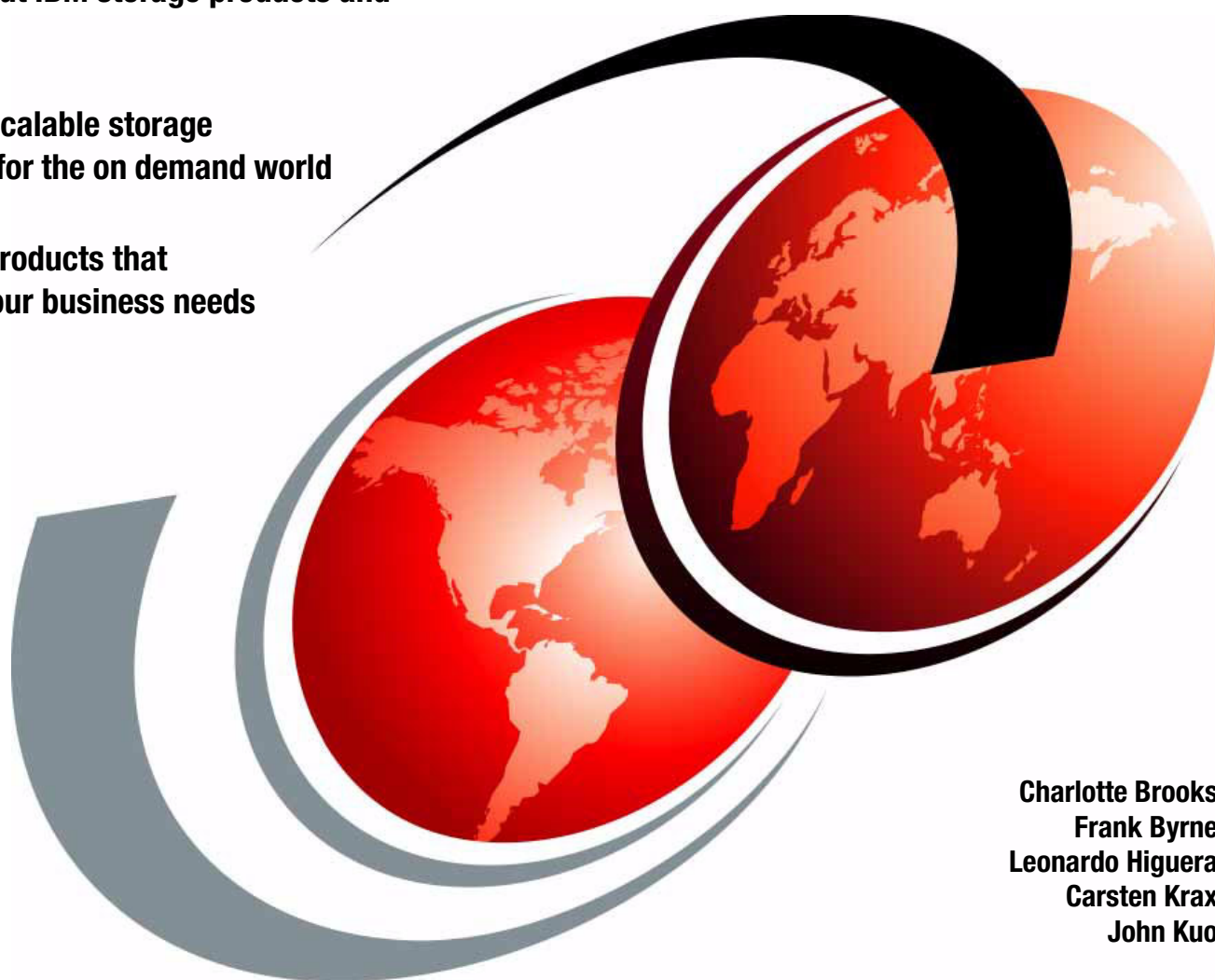


IBM System Storage Solutions Handbook

Learn about IBM storage products and solutions

Flexible, scalable storage solutions for the on demand world

Find the products that answer your business needs



Charlotte Brooks
Frank Byrne
Leonardo Higuera
Carsten Krax
John Kuo

Redbooks



International Technical Support Organization

IBM System Storage Solutions Handbook

October 2006

Note: Before using this information and the product it supports, read the information in “Notices” on page xxiii.

Seventh Edition (October 2006)

This edition applies to the IBM System Storage and TotalStorage products portfolio as of July 2006.

© Copyright International Business Machines Corporation 1999, 2003, 2004, 2005, 2006. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Figures	xv
Notices	xxiii
Trademarks	xxiv
Preface	xxv
The team that wrote this redbook.	xxv
Become a published author	xxvii
Comments welcome.	xxvii
Chapter 1. Introduction: Information on demand with IBM System Storage.	1
1.1 Information on demand	2
1.1.1 Infrastructure simplification	2
1.1.2 Business continuity	3
1.1.3 Information lifecycle management	4
1.2 On demand storage environment	5
1.3 More information	7
Chapter 2. Business continuity solutions.	9
2.1 Introduction to business continuity	10
2.1.1 Three aspects of IT business continuity	10
2.1.2 The objectives of IT business continuity	12
2.2 IT recovery timeline	12
2.3 Selecting business continuity solutions	13
2.3.1 Business continuity is an end-to-end process	14
2.3.2 Leverage infrastructure simplification for business continuity	15
2.4 The tiers of business continuity	15
2.4.1 Definition of the tiers of business continuity	16
2.5 Blending tiers into an optimized solution	18
2.6 Backup and Restore solutions (Tier 1-4)	20
2.6.1 Microsoft Volume Shadow Copy Service (VSS)	20
2.6.2 IBM Tivoli Storage Manager	21
2.6.3 IBM Tape storage systems	21
2.6.4 IBM System Storage N series Snapshot solutions	21
2.7 Rapid Data Recovery solutions (Tier 4-6)	22
2.7.1 Data consistency for database environments	23
2.7.2 Rapid Data Recovery for System z	25
2.7.3 TotalStorage Productivity Center for Replication solutions for UNIX and Windows	26
2.7.4 IBM System Storage SAN Volume Controller	26
2.8 Continuous availability solutions (Tier 7)	26
2.8.1 Geographically Dispersed Parallel Sysplex (GDPS)	27
2.8.2 HACMP/XD for AIX	31
2.9 More information	33
Chapter 3. Information lifecycle management	35
3.1 Overview	36
3.2 Standards and organizations	36
3.2.1 Storage Networking Industry Association (SNIA)	37

3.3	Why ILM is needed	38
3.4	ILM elements.	39
3.4.1	Tiered storage management.	39
3.4.2	Long-term data retention.	41
3.4.3	Data lifecycle management.	42
3.4.4	Policy-based archive management.	44
3.5	IBM ILM solutions	45
3.5.1	Application and database archive solutions	45
3.5.2	E-mail archive solutions	46
3.5.3	Data lifecycle management solutions	47
3.5.4	Enterprise content management solutions	47
3.6	Services and consulting	48
3.7	More information	48
	Chapter 4. Storage infrastructure simplification	49
4.1	Overview	50
4.2	Physical storage consolidation solutions.	50
4.2.1	Key benefits	50
4.2.2	Key offerings	50
4.3	Virtualization solutions	51
4.3.1	Key benefits	51
4.3.2	Key offerings	51
4.4	Automation solutions.	52
4.4.1	Key benefits	52
4.4.2	Key offerings	52
4.5	Storage services	53
4.5.1	Global Services Center for Infrastructure Services	53
4.5.2	Global Services Customer Solution Center Network	53
4.5.3	IBM High Availability Services - for resilient infrastructure	54
4.5.4	IT consolidation.	54
4.5.5	Implementation Services.	54
4.5.6	Infrastructure Resource Management.	55
4.5.7	On demand IT infrastructure strategy	55
4.5.8	Resilient business and infrastructure solutions	55
4.5.9	Resilient business and infrastructure assessment	56
4.5.10	IT optimization solution	56
4.5.11	IT consolidation.	56
	Part 1. Disk systems	57
	IBM System Storage point-in-time copy comparison	59
	FlashCopy function definitions	60
	IBM System Storage disk mirroring comparison	60
	Disk mirroring function definitions	61
	Chapter 5. Entry-level disk storage	63
5.1	Overview	64
5.2	Specifications	66
5.3	IBM TotalStorage DS300	66
5.4	IBM TotalStorage DS400	68
5.5	ServeRAID Manager	69
5.6	Premium features	70
5.6.1	EXP400 expansion units.	70
5.6.2	FlashCopy.	71
5.7	Advanced software features	71

5.8 More information	72
Chapter 6. Expandable storage	73
6.1 Overview and key features of the EXP24	74
6.1.1 High availability	74
6.1.2 Flexible and scalable storage	75
6.2 Overview and key features of the EXP Plus	75
6.2.1 High availability	76
6.2.2 Flexible and scalable storage	76
6.3 More information	77
Chapter 7. Mid-range disk systems	79
7.1 Highlights - what's new	80
7.2 Overview	80
7.3 IBM System Storage DS4100	82
7.3.1 Features	82
7.3.2 Specifications	84
7.4 IBM System Storage DS4300	85
7.4.1 Features	85
7.4.2 Specifications	86
7.5 IBM System Storage DS4300 Turbo	87
7.5.1 Features	87
7.5.2 Specifications	88
7.6 IBM System Storage DS4500	88
7.6.1 Features	88
7.6.2 Specifications	89
7.6.3 Specifications	90
7.7 IBM System Storage DS4700 Express	90
7.7.1 Features	91
7.7.2 Specifications	92
7.8 IBM System Storage DS4800	93
7.8.1 Features	94
7.8.2 Specifications	95
7.9 IBM System Storage DS4000 EXP100 Expansion	96
7.9.1 Features	96
7.9.2 Hardware description	96
7.9.3 Specifications	97
7.9.4 SATA versus FC disks	98
7.10 IBM System Storage DS4000 EXP710 Expansion Unit	99
7.10.1 Features	99
7.10.2 Specifications	100
7.11 IBM System Storage DS4000 EXP810 Expansion Unit	100
7.11.1 Features	101
7.11.2 Specifications	101
7.12 DS4000 Storage Manager	102
7.13 Copy services	104
7.13.1 FlashCopy	105
7.13.2 VolumeCopy	106
7.13.3 Enhanced Remote Mirroring	108
7.14 DS4000 Service Alert	110
7.15 DS4000 Alert Manager	111
7.16 More information	111
Chapter 8. Enterprise Disk Systems	113

8.1 Overview	114
8.1.1 Infrastructure simplification	115
8.1.2 Business continuity	115
8.1.3 Performance	116
8.1.4 System z and zSeries performance	117
8.2 DS6000 and DS8000 technology	120
8.2.1 Enhancements of the DS6000 and DS8000 series	121
8.3 The DS6000 series	123
8.3.1 The DS6800 controller enclosure	124
8.3.2 The DS6000 expansion enclosure	126
8.3.3 DS6800 Model 1750-511	128
8.3.4 DS6000 Model 1750-EX1	129
8.3.5 More information	129
8.4 The DS8000 series	130
8.4.1 The DS8000 design enhancements	130
8.4.2 The DS8000 architecture	132
8.4.3 DS8000 models	138
8.4.4 IBM Standby Capacity on Demand offering for the DS8000	140
8.4.5 More information	140
8.5 The Enterprise Storage Server Family	140
8.6 Storage management	142
8.6.1 DS Management Console	142
8.6.2 IBM TotalStorage Productivity Center	144
8.7 Copy Services	145
8.7.1 IBM System Storage FlashCopy (Point-in-time Copy feature)	145
8.7.2 Remote Mirror and Copy feature	146
8.8 Comparison of products	151
8.9 Additional information	152
Chapter 9. IBM System Storage N series	155
9.1 N series hardware overview	156
9.2 System Storage N3700	157
9.2.1 N3700 hardware features	157
9.2.2 N3700 specifications	158
9.3 System Storage N5000 series	158
9.3.1 System Storage N5000 series highlights	159
9.3.2 Differences between appliance and gateway models	159
9.3.3 N5000 hardware features	161
9.3.4 N5000 specifications	161
9.4 N series expansion units	163
9.4.1 EXN1000 expansion unit	163
9.4.2 EXN2000 expansion unit	164
9.5 Data ONTAP software overview	164
9.6 N series software solutions	165
9.6.1 Data availability	165
9.6.2 Flexible data management and instant data replication	166
9.6.3 Mixed environment support	167
9.6.4 Server and storage consolidation with multistore	168
9.6.5 Tiered storage	168
9.6.6 Advanced data protection with SnapShot technology	169
9.6.7 Local and remote mirroring solutions	170
9.6.8 Improving data availability with clustered failover	173
9.6.9 Optimizing workloads with SnapMover	174

9.6.10	Automated data backup with SnapVault	175
9.6.11	Disk-based backups for Open Systems	176
9.6.12	Backing up N series via NDMP and Tivoli Storage Manager	177
9.6.13	Storage virtualization and simplified management with SnapDrive	178
9.6.14	Rapid backups and restores with SnapManager	179
9.6.15	Single mailbox recovery for Exchange	180
9.6.16	End-to-end data validation for Oracle databases	181
9.6.17	Data permanence solutions	181
9.6.18	N series management.	183
9.6.19	Obliterating data with disk sanitization	186
9.7	More information	186
 Chapter 10. IBM System Storage DR550.		187
10.1	Data retention and compliance	188
10.1.1	Characteristics of retention managed data	188
10.1.2	IBM strategy and positioning.	188
10.2	Positioning of the DR550	189
10.3	IBM System Storage DR550 overview	190
10.3.1	Hardware	190
10.3.2	Software	192
10.4	IBM System Storage DR550 Express	192
10.5	Data retention features	193
10.5.1	Methods of archive retention.	194
10.6	Document Management System.	196
10.7	More information	197
 Part 2. Tape systems		199
 Chapter 11. IBM TotalStorage and System Storage tape drives.		201
11.1	Technology overview	202
11.1.1	LTO technology.	202
11.1.2	The IBM LTO Ultrium 3 technology.	206
11.1.3	3592 technology	210
11.1.4	3590 technology	212
11.2	IBM LTO tape drives	215
11.2.1	IBM System Storage TS1030 Tape Drive.	215
11.2.2	IBM TotalStorage 3580 Tape Drive.	216
11.3	TotalStorage 359x tape drive	216
11.3.1	IBM TotalStorage 3592 Tape Drive (3592 J1A)	217
11.3.2	IBM System Storage TS1120 Tape Drive.	218
11.3.3	IBM TotalStorage 3590 Tape Drive.	219
11.3.4	IBM TotalStorage 3592 Tape Controller Model J70	220
11.3.5	IBM System Storage TS1120 Tape Controller Model C06	221
11.4	More information	222
11.5	Tape media	222
11.5.1	LTO cartridge	222
11.5.2	IBM 3592 cartridge	223
11.5.3	IBM 3590 cartridge	224
 Chapter 12. IBM Tape Libraries		225
12.1	Overview	226
12.2	IBM System Storage TS3100 Tape Library Express Model	227
12.2.1	Features	228
12.2.2	Specifications	228

12.3 IBM System Storage TS3200 Tape Library	229
12.3.1 Features	229
12.3.2 Specifications	230
12.4 IBM System Storage TS3310 Tape Library	231
12.4.1 Features	232
12.4.2 Specifications	233
12.5 IBM System Storage TS3500 Tape Library	234
12.5.1 Structure of the IBM System Storage TS3500 Tape Library.....	236
12.5.2 Dual accessors and service bays in the TS3500	241
12.5.3 Advanced Library Management System (ALMS)	242
12.5.4 Tape drives supported by the IBM System Storage TS3500 Tape Library....	243
12.6 IBM TotalStorage 3953 Tape System for System z	243
12.6.1 Features	244
12.7 IBM TotalStorage 3494 Tape Library	245
12.7.1 Consolidation and multiplatform attachment.....	246
12.7.2 Performance	246
12.7.3 Enterprise tape drive support	247
12.7.4 Storage capacity	247
12.7.5 Tape Library Base Frame	247
12.7.6 Tape Drive Expansion Frame	248
12.7.7 Tape Storage Frame.....	248
12.7.8 High Availability Frame	248
12.7.9 IBM 3494 Library Manager	249
12.7.10 More information	249
12.8 Enterprise Tape System 359x family	250
12.9 Other tape libraries	250
12.9.1 IBM TotalStorage 3581 Tape Autoloader	250
12.9.2 IBM TotalStorage 3582 Tape Library	251
12.9.3 IBM TotalStorage 3583 Tape Library	253
Chapter 13. IBM Tape Virtualization products	255
13.1 IBM Virtualization Engine TS7510	256
13.1.1 Introduction to open tape virtualization	256
13.1.2 IBM Virtualization Engine TS7510 overview	258
13.1.3 IBM Virtualization Engine TS7510 (3954 CV5).....	260
13.1.4 IBM System Storage 3952 Tape Frame Model F05	261
13.1.5 The TS7510 Cache Controller (3955 SV5).....	261
13.1.6 The TS7510 Cache Module (3955 Model SX5)	261
13.1.7 IBM Virtualization Engine TS7510 Management Console	262
13.1.8 TS7510 Data management and advanced features	263
13.1.9 IBM Virtualization Engine TS7510 solution areas.....	268
13.2 IBM TotalStorage 3494 Virtual Tape Server	269
13.2.1 Introduction to IBM TotalStorage 3494 Virtual Tape Server	269
13.2.2 Virtual Tape Server logical components and functions.....	271
13.2.3 IBM TotalStorage 3494 Virtual Tape Server Model B10 and B20.....	273
13.2.4 IBM TotalStorage Virtual Tape Frame 3494 Model CX1.....	276
13.2.5 IBM TotalStorage 3494 Peer-to-Peer Virtual Tape Server	276
13.3 Enterprise Removable Media Manager.....	280
13.3.1 Introduction to Enterprise Removable Media Manager.....	281
13.3.2 eRMM central reporting and monitoring	282
13.3.3 eRMM logical components and functions	283
13.3.4 eRMM control flow	286
13.3.5 Supported features (eRMM 1.2.4).....	289

13.3.6 Supported platforms (eRMM 1.2.4)	289
13.3.7 Strategic fit and positioning	290
13.4 More Information	290
Chapter 14. Complementary storage products	293
14.1 Overview	294
14.2 IBM 7204 external disk drive	296
14.3 IBM 7205 SDLT tape drive	296
14.4 IBM 7206 tape drive	297
14.4.1 Model 336	297
14.4.2 Model VX2	298
14.4.3 Model VX3	298
14.5 IBM 7207 SLR tape drive	298
14.5.1 Model 122	298
14.5.2 Model 330	298
14.6 IBM 7210 DVD-RAM drive	299
14.6.1 Model 025	299
14.6.2 Model 030	299
14.7 IBM 7212 storage device enclosure	300
14.8 IBM 7212-312 VXA-2 autoloader	300
Part 3. Storage networking	301
Chapter 15. Introduction to storage networking	303
15.1 Overview	304
15.2 Storage Area Network (SAN)	304
15.2.1 Fibre Channel	305
15.2.2 SAN topologies	306
15.2.3 Physical components of the SAN infrastructure	311
15.2.4 Naming and addressing	312
15.2.5 FICON	313
15.3 IP Storage Networking technologies	313
15.3.1 iSCSI	314
15.3.2 FCIP	315
15.3.3 iFCP	316
15.3.4 Comparison iSCSI - FCIP - iFCP	317
15.4 Network Attached Storage (NAS)	318
15.4.1 NAS gateways	319
15.5 Wavelength Division Multiplexing	320
15.6 Selecting the best alternative	320
15.7 More information	321
Chapter 16. Storage Area Network products	323
16.1 SAN switches and connectivity products	324
16.1.1 Common characteristics	324
16.1.2 Other switch features	324
16.2 Entry SAN switches	326
16.2.1 IBM TotalStorage SAN16B-2	326
16.2.2 IBM TotalStorage SAN16M-2	327
16.2.3 Cisco MDS 9020 Fabric Switch	328
16.3 Mid-range SAN switches	329
16.3.1 IBM TotalStorage SAN32B-2	330
16.3.2 IBM TotalStorage SAN32M-2	331
16.3.3 IBM TotalStorage SAN64B-2	332

16.3.4	Cisco MDS 9120 and 9140 Multilayer Switches	333
16.3.5	Cisco MDS 9216A Fabric Switch	335
16.4	Enterprise SAN switches.	335
16.4.1	IBM TotalStorage SAN256B	336
16.4.2	IBM TotalStorage SAN140M.	337
16.4.3	IBM TotalStorage SAN256M.	338
16.4.4	Cisco MDS 9506 Multilayer Director	339
16.4.5	Cisco MDS 9509 Multilayer Director	340
16.4.6	Cisco MDS 9513 Multilayer Director	341
16.5	Multiprotocol routers	342
16.5.1	IBM System Storage SAN18B-R multiprotocol router.	343
16.5.2	IBM TotalStorage SAN16B-R multiprotocol router	345
16.5.3	IBM TotalStorage SAN16M-R multiprotocol SAN router.	346
16.5.4	IBM TotalStorage SAN04M-R.	347
16.5.5	Cisco MDS 9216i Fabric Switch	348
16.6	SAN solutions	349
16.6.1	Infrastructure simplification solutions	349
16.6.2	Business continuity solutions	351

Part 4. IBM System Storage open software 355

Chapter 17. IBM System Storage virtualization	357
17.1 What storage “virtualization” is	358
17.1.1 Levels of storage virtualization	359
17.2 IBM System Storage approach to virtualization	361
17.3 IBM System Storage virtualization family	362
17.4 More information	363

Chapter 18. SAN Volume Controller	365
18.1 The need for storage virtualization	366
18.2 IBM System Storage SAN Volume Controller (SVC)	367
18.2.1 Overview	367
18.2.2 SVC commonly used terms.	369
18.3 SVC architecture	370
18.3.1 SVC virtualization	372
18.3.2 SVC software licensing.	375
18.4 SVC solutions	377
18.4.1 Business continuity	377
18.4.2 Infrastructure simplification	378
18.5 SVC compatibility	379
18.6 More information	380

Chapter 19. IBM TotalStorage Productivity Center	381
19.1 Overview	382
19.2 IBM TotalStorage Productivity Center Version 3.1 enhancements	383
19.3 IBM TotalStorage Productivity Center Limited Edition	385
19.4 IBM TotalStorage Productivity Center Standard Edition	385
19.5 IBM TotalStorage Productivity Center for Replication.	385
19.6 SMI-S	386
19.7 More Information	387

Chapter 20. IBM TotalStorage Productivity Center for Data	389
20.1 TotalStorage Productivity Center for Data overview	390
20.1.1 TPC for Data and data management	390

20.1.2 TPC for Data product highlights	391
20.2 TPC for Data architecture	392
20.3 TPC for Data functions	393
20.3.1 Dashboard	393
20.3.2 Discover and monitor information	395
20.3.3 Alerts	397
20.3.4 Reporting	398
20.3.5 Data Manager for Databases	400
20.3.6 Data Manager for Chargeback	400
20.4 Supported platforms	402
20.5 More information	403
Chapter 21. IBM TotalStorage Productivity Center for Fabric	405
21.1 TPC for Fabric - overview	406
21.1.1 TPC for Fabric - components	406
21.2 SAN management	407
21.2.1 Standards and SAN management tools	407
21.2.2 Discovery	408
21.2.3 Zones and zone control	409
21.2.4 Topology views	409
21.2.5 SNMP events	410
21.3 Functions	410
21.3.1 Viewing topology	410
21.3.2 Semantic Zooming	413
21.3.3 Zone and zonesets	416
21.3.4 Monitoring the fabric	418
21.4 More information	420
Chapter 22. TotalStorage Productivity Center for Disk	421
22.1 Overview	422
22.2 TPC for Disk - what's new in the V3.1	422
22.3 TPC support for SAN Volume Controller	423
22.4 Monitoring	423
22.5 Alerting	423
22.6 Reporting	424
22.7 More information	426
Chapter 23. TotalStorage Productivity Center for Replication	429
23.1 TPC for Replication overview	430
23.2 Consistency management	431
23.3 Replication management	432
23.3.1 Health Overview	433
23.3.2 Storage subsystems	433
23.3.3 Logical paths	434
23.3.4 Session and copy sets	434
23.4 More information	436
Chapter 24. IBM Tivoli Storage Manager	437
24.1 IBM Tivoli Storage Manager family of products	438
24.1.1 IBM Tivoli Storage Manager core functionality	438
24.1.2 IBM Tivoli Storage Manager key attributes	441
24.2 How Tivoli Storage Manager works	445
24.2.1 Tivoli Storage Manager backup and archive concepts	445
24.2.2 Progressive incremental backups	446

24.2.3	Tivoli Storage Manager Administration Center	446
24.2.4	IBM Tivoli Storage Manager backup-archive client.	451
24.2.5	IBM Tivoli Storage Manager storage agent.	452
24.2.6	Tivoli Storage Manager security	453
24.2.7	Tivoli Storage Manager API	453
24.3	Client backup and restore operations	454
24.3.1	Traditional LAN and WAN backup topology	454
24.3.2	SAN (LAN-free) backup topology	455
24.3.3	Split-mirror/point-in-time copy backup using SAN.	456
24.3.4	NAS and N Series backup and restore	457
24.3.5	Image backup	459
24.4	IBM System Storage Archive Manager	460
24.4.1	Overview	460
24.5	IBM Tivoli Storage Manager disaster recovery	461
24.6	Tivoli Storage Manager server-to-server communications	463
24.6.1	Using server-to-server virtual volumes for disaster recovery	464
24.7	Tivoli Storage Manager and high availability clustering	464
24.7.1	IBM High Availability Cluster Multi-Processing (HACMP).	465
24.7.2	Backup-archive client support with HACMP	466
24.7.3	Tivoli Storage Manager server and Microsoft Cluster Server (MSCS)	467
24.7.4	Tivoli Storage Manager backup-archive client support with MSCS	468
24.8	Tivoli Storage Manager and tape vaulting.	469
24.8.1	Electronic tape vaulting.	469
24.9	Remote disk mirroring and tape vaulting solutions	470
24.9.1	Collocation considerations for off-site vaulting	472
24.9.2	Reclamation considerations for off-site vaulting	472
24.10	More information	474
	Chapter 25. TotalStorage Expert	475
25.1	Overview	476
25.2	Key features	476
25.2.1	Performance management	476
25.2.2	Asset management.	477
25.2.3	Capacity management	477
25.3	How IBM TotalStorage Expert works	477
25.4	Planning and implementation considerations	478
25.5	More information	479
	Part 5. z/OS storage management.	481
	Chapter 26. Data Management with DFSMS	483
26.1	Managing Enterprise data	484
26.2	The components of DFSMS	484
26.3	DFSMSdftp (Data Facility Product)	484
26.3.1	Organizing and accessing data with DFSMSdftp.	484
26.3.2	Locating data sets.	486
26.4	System Managed Storage.	486
26.4.1	Parallel Sysplex	487
26.4.2	Naming conventions	487
26.4.3	Extended Format data sets.	487
26.4.4	Constructs.	487
26.4.5	ISMF	489
26.4.6	Data management utilities	490
26.5	DFSMSdsss (Data Set Services)	491

26.5.1 Overview	491
26.5.2 Interface to other services	491
26.6 DFSMShsm (Hierarchical Storage Manager)	491
26.6.1 Overview	491
26.6.2 Space management	492
26.6.3 Availability management	492
26.6.4 Tape environment	493
26.7 DFSMSrmm (Removable Media Manager)	494
26.7.1 Overview	494
26.7.2 Policies for retention and movement	495
26.7.3 Control Data Set	495
26.8 More information	495
Chapter 27. SORT products for the System z.	497
27.1 Overview	498
27.2 Performance considerations	498
27.3 DFSORT performance	498
27.3.1 Extended Format data sets	498
27.3.2 ICETPEX	499
27.3.3 Memory sorting	499
27.3.4 ICEGENER utility	500
27.4 Symbols	500
27.5 DFSORT/VSE memory sorting	501
27.6 Sort features	501
27.7 User exits	502
27.8 ICETOOL utility	502
27.9 More information	502
Part 6. Appendixes	503
Appendix A. 7133 serial disk system	505
Overview	506
Serial Storage Architecture	506
More information	507
Appendix B. Introduction to SMI-S and CIM/WBEM.	509
Appendix C. Redundant Arrays of Independent Disks (RAID)	513
Glossary	519
Abbreviations and acronyms	525
Related publications	529
IBM Redbooks	529
Other publications	530
How to get IBM Redbooks	530
Help from IBM	530
Index	531

Figures

1-1	IS, BC, and ILM	2
1-2	Business Continuity	4
1-3	Information lifecycle management	4
1-4	On demand storage environment	6
2-1	Three aspects of Business Continuity	11
2-2	Timeline of an IT recovery	12
2-3	Historical challenge in selecting business continuity solutions	14
2-4	The five components of IT business continuity	15
2-5	Tiers of business continuity	16
2-6	Three business continuity solution bands	19
2-7	Phases of an IT BC implementation	20
2-8	Software replication versus Hardware replication	23
2-9	Sample update sequence in a database	24
2-10	Sample Rolling Disaster	25
2-11	GDPS/PPRC topology	28
2-12	GDPS/XRC topology	30
2-13	HACMP/XD	32
3-1	Information lifecycle management	36
3-2	SNIA vision for ILM	37
3-3	ILM elements	39
3-4	Traditional non-tiered storage environment	40
3-5	Multi-tiered storage environment	41
3-6	ILM policies	43
3-7	Information value changes	44
3-8	Value of information and archive/retrieve management	45
4-1	Heterogeneous and complex IT environment	49
5-1	DS300	64
5-2	Cost-effective iSCSI SAN using DS300	64
5-3	SAN connection options for DS400	65
5-4	DS300 (1701-1RL) storage subsystem rear view	67
5-5	DS300 (1701-1RS) storage subsystem rear view	67
5-6	DS300 storage subsystem model 1701-2RD rear view	68
5-7	DS400 (1700-2RD) storage subsystem rear view	69
5-8	ServeRAID Manager	70
6-1	IBM EXP24 models D24	74
6-2	IBM 2104 expandable storage plus 320 models DS4 and TS4	76
7-1	IBM System Storage DS4100	82
7-2	DS4100 Back view	83
7-3	DS4100 Controller	84
7-4	IBM System Storage DS4300	85
7-5	DS4300 front view	86
7-6	DS4500 front view	89
7-7	DS4500 Rear view	89
7-8	DS4700 Express front view	91
7-9	DS4700 mod 72x rear view	92
7-10	IBM System Storage DS4800	93
7-11	DS4800 front and back views	94
7-12	EXP100 storage expansion unit	96

7-13	DS4000 EXP100 Back view	97
7-14	Storage data in a networks storage environment	98
7-15	EXP710 expansion unit	100
7-16	DS4000 EXP810 Expansion Unit	100
7-17	DS4000 EXP810 rear components	101
7-18	Storage Manager 9.16	102
7-19	How DS4000 FlashCopy works	106
7-20	VolumeCopy	107
7-21	VolumeCopy integration with FlashCopy	108
7-22	Enhanced Remote Mirroring	109
8-1	Enterprise Disk Systems	114
8-2	The server based architecture rings continuous performance improvements	116
8-3	Scaling options.	120
8-4	Disk enclosure switched connections	121
8-5	Four paths to each disk drive.	122
8-6	Grouping of volumes in LSSs	123
8-7	The DS6000 series	124
8-8	DS6800 controller card	124
8-9	DS6000 disk architecture.	125
8-10	DS6800 controller architecture	126
8-11	SBOD controller card.	127
8-12	Example of expansion enclosure cabling.	127
8-13	DS6800 Model 1750-511 rear view	128
8-14	DS6000 with 2 expansion enclosures	129
8-15	The DS8000 series	130
8-16	Logical partitions (LPARs)	132
8-17	DS8000 - Primary Frame	133
8-18	Array across loops	134
8-19	The DS8000 architecture.	135
8-20	RIO-G design.	136
8-21	DS8000 Frames.	137
8-22	DS8100	138
8-23	Model 922 and 9A2	139
8-24	ESS Family	141
8-25	DS6000 Simulated Manager	143
8-26	DS8000 Real-time Manager	143
8-27	DS CLI	144
8-28	FlashCopy	145
8-29	Establish FlashCopy on existing Remote Mirror and Copy Primary	146
8-30	Metro Mirror	147
8-31	Global Copy	148
8-32	Global Mirror	149
8-33	z/OS Global Mirror	150
8-34	z/OS Metro/Global Mirror	151
8-35	DS Storage Manager feature activation.	153
9-1	N series hardware overview.	156
9-2	System Storage N3700	157
9-3	System Storage N5000	159
9-4	Gateway provides functionality of N series technology to other vendor storage sub-systems	160
9-5	RAID4 and RAID-DP	165
9-6	Traditional volume provisioning	166
9-7	N series FlexVol technology	166

9-8	N series FlexClone technology	167
9-9	N series multiprotocol support	167
9-10	N series MultiStore	168
9-11	Tiered storage	169
9-12	N series SnapShot - How it works	170
9-13	N series SyncMirror	171
9-14	N series SnapMirror	171
9-15	SnapMirror consolidation	172
9-16	N series cluster failover	173
9-17	Clustered SyncMirror	174
9-18	N series SnapMover	175
9-19	N series SnapVault	176
9-20	N series mixed SnapVault and Open Systems SnapVault environment	177
9-21	NDMP backup with Tivoli Storage Manager Extended Edition	178
9-22	N series SnapDrive	179
9-23	N series SnapValidator	181
9-24	System Storage N series with SnapLock	183
9-25	N series FilerView	184
9-26	N series Data Fabric Manager	185
9-27	N series disk sanitization	186
10-1	Representation of components of the DR550	189
10-2	DR550 with console kit	191
10-3	IBM System Storage DR550 Express	193
10-4	Chronological retention policy	195
10-5	Event based retention policy	196
10-6	Deletion hold and release	196
11-1	Magnified servo band showing a pair of servo bursts	203
11-2	Encoding bits using the servo stripes within the servo bursts	204
11-3	Four data bands written between the servo tracks	205
11-4	Shingling - writing to tape generation 1 versus generation 2	206
11-5	Ultrium 1, 2 and 3 compatibility	210
11-6	3590 tape cartridge	213
11-7	IBM 3590 recording	213
11-8	IBM System Storage TS1030 Tape Drive	215
11-9	The IBM TotalStorage 3580 Tape Drive	216
11-10	IBM 3592 J1A Tape Drive	217
11-11	The IBM 3590 and IBM 3592 Tape Drive Model J1A	218
11-12	IBM TS1120 tape drive	219
11-13	IBM TotalStorage Enterprise Tape Drive 3590 in a rack	220
11-14	IBM 3592 model J70 tape controller	221
11-15	IBM TS1120 Tape Controller	221
11-16	Ultrium cartridge view from top and rear	222
11-17	Ultrium cartridge view from top and front	223
12-1	IBM System Storage TS3100 Tape Library	227
12-2	IBM System Storage TS3200 Tape Library	229
12-3	TS3200 internal view	230
12-4	TS3200 Library station I/O in the left magazine	230
12-5	IBM System Storage TS3310 Tape Library Tape Library - Base and Expansion	232
12-6	TS3310 Base unit rear and back view	233
12-7	IBM System Storage TS3500 Tape Library	235
12-8	TS3500 Frame L23/53	237
12-9	TS3500 Frame D23 or D53	238
12-10	IBM Multipath host with two HBAs to the Library System	240

12-11 IBM Multipath fail over from Host Side	240
12-12 TS3500 Performance single accessor versus Dual accessor	242
12-13 IBM TotalStorage 3953 Tape System for zSeries diagram	244
12-14 IBM TotalStorage 3953 Tape System for System z	245
12-15 IBM TotalStorage 3494 Tape Library	246
12-16 IBM TotalStorage 3581 Tape Autoloader	250
12-17 IBM TotalStorage 3582 Tape Library	252
12-18 IBM TotalStorage 3583 Tape Library	254
13-1 Traditional backup	256
13-2 Backup to disk buffer	257
13-3 Tape virtualization with IIBM Virtualization Engine TS7510.	258
13-4 IBM Virtualization Engine TS7510 hardware overview	259
13-5 TS7510 Management Console	262
13-6 TS7510 OnDemand Virtual Tape Space Allocation.	264
13-7 TS7510 Import/Export function	265
13-8 TS7510 Network Replication function	266
13-9 TS7510 Auto Replication function	267
13-10 TS7510 Remote Copy function	268
13-11 IBM TotalStorage 3494 Virtual Tape Server	269
13-12 VTS virtual drives and hard disk caching.	270
13-13 VTS Virtual Drives emulation.	271
13-14 Data in the TVC and on physical tape	272
13-15 VTS distance to IBM 3584 and 3953	275
13-16 Hardware components of the VTS transparent to the host system	275
13-17	276
13-18 PtP VTS physical and logical view.	277
13-19 Remote PtP VTS implementation	278
13-20 Customer issues with tape.	281
13-21 eRMM tape library hardware virtualization.	282
13-22 eRMM drive utilization	283
13-23 eRMM drive histogram.	283
13-24 eRMM Groups and Pools.	285
13-25 eRMM logical architecture	286
13-26 eRMM Control Flow - Steps 1 to 2.	287
13-27 eRMM Control Flow - Steps 3 to 6.	288
13-28 eRMM Control Flow - Steps 7 to 9.	289
14-1 IBM 7204 external disk drive	296
14-2 IBM 7205 SDLT tape drive.	297
14-3 IBM 7206 4 mm tape drive.	297
14-4 IBM 7207 SLR tape drive.	298
14-5 IBM 7210 DVD-RAM drive.	299
14-6 IBM 7212 storage device enclosure.	300
14-7 IBM 7212-312 VXA-2 autoloader.	300
15-1 A picture of SAN	305
15-2 Fibre Channel protocol layers	306
15-3 Fibre Channel point-to-point topology	307
15-4 Fibre Channel loop topology	308
15-5 Fibre Channel switched topology.	309
15-6 Fibre Channel single switched topology	310
15-7 Fibre Channel switched topology (redundant switches)	311
15-8 LC and SC connectors.	311
15-9 SFP (Small Form Factor Pluggable Media) and GBIC (Gigabit Interface Converter)	312
15-10 Host Bus Adapter.	312

15-11 IP Storage networking technology	314
15-12 iSCSI protocol layering model	315
15-13 FCIP with IP tunneling	315
15-14 FCIP Encapsulation model	316
15-15 iFCP Gateway architecture	317
15-16 iSCSI, FCIP and iFCP protocol stacks.	318
15-17 A picture of NAS appliance	319
15-18 Wave Division Multiplexer concept	320
16-1 Sample Configuration for SAN16B-2	327
16-2 IBM TotalStorage SAN16B-2.	327
16-3 IBM TotalStorage SAN16M-2	328
16-4 One disk storage array sample	329
16-5 Cisco MDS 9020 Fabric Switch	329
16-6 IBM TotalStorage SAN32B-2.	330
16-7 High-availability server clustering solution.	331
16-8 IBM TotalStorage SAN32M-2	332
16-9 Business Continuity Solution with SAN32M-2	332
16-10 IBM TotalStorage SAN64B-2.	333
16-11 Cisco MDS 9120 and 9140 Fabric Switches	334
16-12 Cisco MDS 9216A Fabric Switch.	335
16-13 IBM TotalStorage SAN256B	337
16-14 IBM TotalStorage SAN 140M.	338
16-15 IBM TotalStorage SAN256M	339
16-16 Cisco MDS 9506 Multilayer Director	340
16-17 Cisco MDS 9509 Multilayer Director	341
16-18 Cisco MDS 9513 Multilayer Director	342
16-19 SAN multiprotocol routing	343
16-20 IBM System Storage SAN18B-R	344
16-21 Inter-SAN routing	344
16-22 IBM TotalStorage SAN16B-R multiprotocol router.	346
16-23 IBM TotalStorage SAN16M-R multiprotocol SAN router	346
16-24 IBM TotalStorage SAN04M-R	347
16-25 Cisco MDS 9216i Fabric Switch.	348
16-26 Infrastructure simplification with high-availability server solution.	350
16-27 Infrastructure simplification with 4 Gbps solution.	351
16-28 Typical backup site solution.	352
16-29 FC over IP	352
16-30 FC over DWDM or CWDM.	352
16-31 Remote tape backup solution	353
17-1 SAN with virtualization	358
17-2 Levels of application of storage virtualization	359
17-3 In-band and out-of-band virtualization models.	360
17-4 SNIA storage model.	362
18-1 Today's environment	366
18-2 SVC components	371
18-3 How to zone for correct SVC disk access	373
18-4 Extents being used to create a VDisk	373
18-5 SVC logical view	374
18-6 SVC managed storage-based licensing.	375
18-7 FlashCopy license illustration	376
18-8 Metro Mirror and Global Mirror license illustration	376
18-9 Business continuity with SVC Global Mirror.	377
18-10 Data migration using SAN	378

18-11	Data migration using SVC	379
19-1	IBM TotalStorage Productivity Center infrastructure	382
19-2	TPC Master Console	385
19-3	IBM TotalStorage Productivity Center Products	386
20-1	Types of data typically found in an enterprise and what do with them.	391
20-2	TPC for Data - dashboard	393
20-3	Availability report - ping	395
20-4	Capacity report - file system used space.	396
20-5	Asset report - discovered LUNs.	397
20-6	TPC 3.1 Alert Log	398
20-7	Asset report - By computer	399
20-8	Data Manager for Databases DB2 free space report.	400
20-9	Data Manager for Chargeback.	401
21-1	Topology Viewer	410
21-2	Topology Viewer for Fabric	411
21-3	Topology Viewer Panel Layout	412
21-4	Topology Viewer MiniMap	413
21-5	Fabric-L0 (F0)	414
21-6	Fabric-L1 (F1)	415
21-7	Fabric-L2 (F2) Switch.	416
21-8	Zoning example graphical view	416
21-9	Zoning tabular view	417
21-10	Entry panel for Zone Configuration	417
21-11	Existing zone configuration	418
21-12	The active zone-definition panel	418
21-13	Fabric performance monitor.	419
22-1	TotalStorage Productivity Center for Disk (Disk Manager)	422
22-2	TPC for Disk configuring alerts	424
22-3	TotalStorage Productivity Center for Disk Reporting by Array.	426
22-4	TPC for Disk Volume HBA assignment report.	426
23-1	TPC for Replication architecture	430
23-2	TPC for Replication freeze command	431
23-3	TPC for Replication heartbeat functionality	432
23-4	TPC for Replication Health Overview	433
23-5	TPC for Replication adding storage subsystem.	434
23-6	TPC for Replication configuring logical paths	434
23-7	TPC for Replication copy set	435
23-8	TPC for Replication sessions.	435
23-9	TPC for Replication Add Copy Sets wizard	436
24-1	Tivoli Storage Manager Extended Edition client, server, application, and database support.	442
24-2	Data movement with Tivoli Storage Manager and the storage hierarchy	445
24-3	Integrated Solutions Console Administration Center	447
24-4	Health Monitor	448
24-5	Enterprise Management.	448
24-6	Storage devices	449
24-7	Policy Domains	449
24-8	Server maintenance.	450
24-9	Reporting	450
24-10	Disaster Recovery Manager	451
24-11	The Tivoli Storage Manager client interface	452
24-12	How the API fits into the DR550 solution.	454
24-13	Tivoli Storage Manager LAN and WAN backup.	455

24-14	Tivoli Storage Manager LAN-free backup	456
24-15	Tivoli Storage Manager split-mirror/point-in-time copy backup	457
24-16	Tape library attached only to NAS server	458
24-17	Tivoli Storage Manager and NDMP backup.	458
24-18	Tivoli Storage Manager image backup	460
24-19	Tivoli Storage Manager Disaster Recovery (DRM) functions	462
24-20	Tivoli Storage Manager server-to-server communications	464
24-21	HACMP and Tivoli Storage Manager server high availability configuration.	465
24-22	HACMP and Tivoli Storage Manager client high availability configuration	466
24-23	MSCS and Tivoli Storage Manager active/active configuration.	467
24-24	MSCS and Tivoli Storage Manager client high availability configuration.	468
24-25	Tivoli Storage Manager and electronic tape vaulting	470
25-1	TIBM TotalStorage Expert sample configuration	478
26-1	DFSMS environment	489
26-2	Example of an ISMF Data Entry Panel	490
A-1	7133 Models D40 and T40.	506
B-1	Component interaction in CIM environment.	511
C-1	RAID 0	514
C-2	RAID 1	515
C-3	RAID 5	516
C-4	RAID-00	517
C-5	RAID-10	518
C-6	RAID-50	518

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law. INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.


This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

e-business on demand®	DS6000™	Power PC®
eServer™	DS8000™	PowerPC®
ibm.com®	Enterprise Storage Server®	Predictive Failure Analysis®
iSeries™	Enterprise Systems Connection	POWER4™
i5/OS®	Architecture®	POWER5™
pSeries®	Express Storage™	Redbooks™
xSeries®	ESCON®	Redbooks (logo)  ™
z/OS®	FlashCopy®	RACF®
z/VM®	FICON®	RS/6000®
z/VSE™	Geographically Dispersed Parallel	S/390®
zSeries®	Sysplex™	ServeRAID™
z9™	GDPS®	System i™
AFS®	Hiperspace™	System p™
AIX 5L™	HyperSwap™	System x™
AIX®	HACMP™	System z™
AS/400®	Informix®	System z9™
BladeCenter®	IBM®	System Storage™
CICS®	Lotus Notes®	System Storage DS™
Domino®	Lotus®	SANergy®
DB2 Universal Database™	Magstar®	Tivoli Enterprise™
DB2®	Micro Channel®	Tivoli Enterprise Console®
DFS™	Netfinity®	Tivoli®
DFSMSdfp™	NetView®	TotalStorage®
DFSMSdss™	Notes®	Virtualization Engine™
DFSMSHsm™	OpenPower™	VM/ESA®
DFSMSrmm™	OS/390®	VSE/ESA™
DFSORT™	OS/400®	WebSphere®
DS4000™	Parallel Sysplex®	

The following terms are trademarks of other companies:

eXchange, IPX, Java, JRE, JVM, Nearline, ONC, RSM, Solaris, StorageTek, Streamline, Sun, Sun Microsystems, Sun StorageTek, Ultra, VSM, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Active Directory, Internet Explorer, Microsoft, Outlook, Windows NT, Windows Server, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

Success in the on demand world depends on a company's ability to leverage information. Greater dependency on information means greater dependency on storage. This dependency as well as the explosion of information, that is created by e-business, is making storage a strategic investment priority for companies of all sizes.

This IBM® Redbook provides overviews and pointers for information about the most current IBM System Storage™ and TotalStorage® products, showing how IBM delivers the right mix of products for nearly every aspect of business continuance and business efficiency. IBM System Storage products can help you store, safeguard, retrieve, and share your data.

We cover these topics:

- ▶ First we introduce the basic storage solutions areas: information lifecycle management, infrastructure simplification, and business continuity.
- ▶ Part 1 describes disk products, including the entire Disk Storage (DS) Series, from entry-level offerings, such as the DS300 and DS400, to mid-range with the DS4000™ family and DS6000™, to high end with the DS8000™. It also includes the DR550.
- ▶ Part 2 is an overview of tape products, covering tape drives, tape libraries, and virtualization products, including LTO and 3590, and 3592 technology.
- ▶ Part 3 starts by describing storage networking infrastructure and protocols, and then presents the switches and directors to form SAN solutions.
- ▶ Part 4 discusses the IBM System Storage software portfolio for open systems and includes Storage Virtualization products, such as the SAN Volume Controller, TotalStorage Productivity Center, Tivoli® Storage Manager, and ETL Expert.
- ▶ Part 5 describes the z/OS® storage management software: DFSMS and DFSORT™.
- ▶ Part 6 is the appendixes, which cover standards, such as Storage Management Initiative Specification (SMI-S), Common Information Model (CIM), and Web Based Enterprise Management (WBEM); also discussed is redundant array of independent disks (RAID).

This redbook is intended as a reference for basic and comprehensive information about the IBM System Storage and TotalStorage products portfolio. This book provides a starting point when establishing your own enterprise storage environment.

The team that wrote this redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Charlotte Brooks is an IBM Certified IT Specialist and Project Leader for Storage Solutions at the International Technical Support Organization, San Jose Center. She has 15 years of experience with IBM in storage hardware and software support, deployment, and management. She has written many Redbooks™, and has developed and taught IBM classes in all areas of storage and storage management. Before joining the ITS0 in 2000, she was the Technical Support Manager for Tivoli Storage Manager in the Asia Pacific Region.

Frank Byrne is a Senior IT Specialist in the United Kingdom, providing day-to-day guidance and advice on the usage of z/OS and associated products. He has 39 years of experience in

large systems support. His areas of expertise include Parallel Sysplex® implementation, and he has written extensively on Parallel Sysplex usage. He is the author of several Redbooks, including *z/OS V1R3 DFSMS Technical Guide*, SG24-6569.

Leonardo Higuera (Leo) is a Hardware Country Specialist in IBM Colombia. He holds a degree in Systems Engineering and is also a Graduate in Management of Projects. He has worked at IBM for 11 years. He has several years of experience working with Intel®-based system and storage solutions. His areas of expertise include High-Availability, Virtualization, Consolidation, design, deployment, and implementation of complex solution, management of critical situations, and customer satisfaction.

Carsten Krax is a Storage Specialist within the IBM Systems Sales Organization in Germany. He is an expert on Business Continuity and Infrastructure Simplification solutions for open systems. His areas of expertise include Disk systems, Tape systems, SAN, NAS, and Storage Software. Before joining IBM in early 2001, Carsten worked from 1999 to 2001 as Storage Specialist for an IBM Business Partner, and from 1995 to 1998 as a Systems Administrator for an IBM customer. He holds a degree in Electrical Engineering from the University of Applied Sciences in Dortmund, Germany, and has coauthored several IBM Redbooks on tape.

John Kuo is an I/T specialist for Open System responsible for Networking and Storage services support in IBM Taiwan. He has 17 years of IT experience and has previously been a zSeries® System Service Representative. His current major responsibilities include DS4000, DS6000, and SVC country support. As well as regular product problem determination, problem solving and fault isolation tasks, he is also takes part in many Business Continuity projects delivering Metro Mirror and Global Mirror functions on ESS and DS8000.



Figure 1 The team: Carsten, Leo, Charlotte, John, and Frank

Thanks to the following people for their contributions to this project:

Gustavo Castets, Bertrand Dufrasne, Babette Haeusser, Emma Jacobs, Mary Lovelace, Alex Osuna, Deanna Polm, Sangam Racherla, Jon Tate, Gabrielle Velez
International Technical Support Organization

Muhammad Dahroug, Rick Higgens, Joseph Hill, Thomas Luther, Reena Master, Stefan Neff, Harold Pike, John Sing, Peter Thurston, Ulf Troppens, Matthias Werner, Abbe Woodcock
IBM product development, support, and marketing

Frank Yaeger, DFSORT development team, worked on Chapter 27, "SORT products for the System z" on page 497.

The authors of previous editions of this redbook.

Tom Chang and staff
Garden Inn Hotel, Los Gatos, California

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You will team with IBM technical professionals, Business Partners and/or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you will develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our Redbooks to be as helpful as possible. Send us your comments about this or other Redbooks in one of the following ways:

- ▶ Use the online **Contact us** review redbook form found at:

ibm.com/redbooks

- ▶ Send your comments in an email to:

redbook@us.ibm.com

- ▶ Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400



Introduction: Information on demand with IBM System Storage

When responding to every customer demand, market opportunity, and external threat requires an integration of people, processes, and information. Because data is a critical component of businesses today, efficiently handling the growing amounts of information is one of the keys to success for an on demand business. Information needs to be promptly available, and this is the aim of “*information on demand*”.

Information on demand enables organizations to be flexible and responsive in dealing with their customers and helps companies improve the quality and timeliness of information, whenever it is needed and wherever it resides. To get there implies:

- ▶ Infrastructure simplification
- ▶ Business continuity
- ▶ Information lifecycle management

This chapter discusses these topics and explains how the IBM System Storage portfolio can help businesses build their “*on demand storage environment*” to achieve information on demand.

1.1 Information on demand

In today's ever more competitive and growing business environment, information is an increasingly valuable, but costly organizational asset.

The volume of information is growing very rapidly in most organizations. And with this, the need to protect and manage information also continues to increase. Organizations are seeking to minimize risk, reduce costs, and increase flexibility by aligning IT investments according to information value and business needs. This is a step towards information on demand. With information on demand, business can respond with flexibility and speed to customer requirements and market opportunity. Getting there involves three aspects:

1. Infrastructure simplification (IS): Simplification of the underlying IT infrastructure and its management to lower the cost and complexity.
2. Business continuity (BC): Assuring the security and durability of information.
3. Information lifecycle management (ILM): Efficiently managing information over its lifecycle.

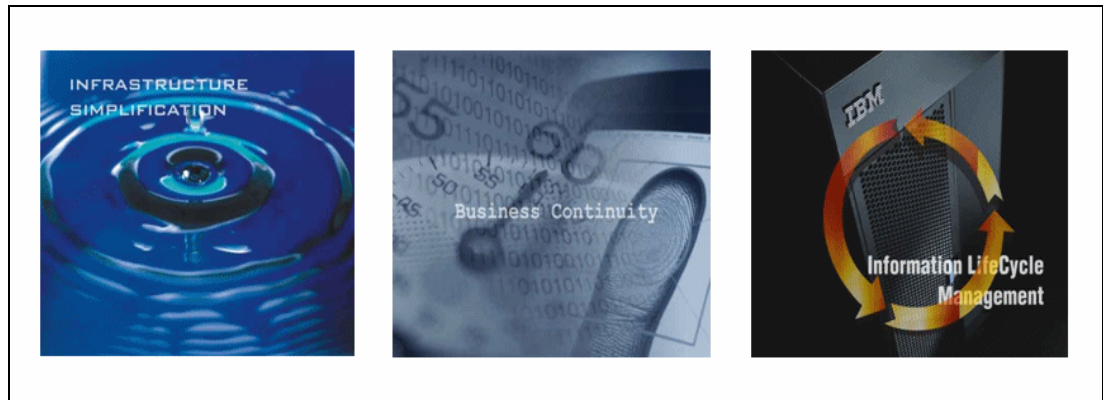


Figure 1-1 IS, BC, and ILM

1.1.1 Infrastructure simplification

Infrastructure simplification is a process by which organizations contain expenses, enable business growth, and reduce operational risks by optimizing IT resources. Simplified infrastructures hold the promise of improved system optimization and Total Cost of Ownership (TCO), higher personnel productivity, and greater application availability through infrastructure resiliency. IBM System Storage products are designed to help clients obtain these benefits through *consolidation, virtualization, and automated management*. Once simplified, the infrastructure can be better managed for lower cost and with fewer errors.

Consolidation

IBM System Storage products can help consolidate the storage environment. Consolidated storage environments have fewer elements to manage which can increase resource utilization and simplify storage management, can provide the ability to share storage servers over extended distances and provide economies of scale for owning disk storage servers. These environments can be more resilient and provide an infrastructure for virtualization and automation. There are four main methods of consolidation: centralization of data centers, physical consolidation, data integration and application integration.

Virtualization

Virtualization of physical resources is an essential part of a simplified storage environment. They are more flexible since the logical files and disks can be manipulated more easily than numerous individual physical disk and tape devices. Virtualization can further improve resource utilization, simplify storage management and accelerate application deployment with improved resource sharing. These capabilities can help organizations enact a tiered storage environment, where information is stored at a cost befitting its value at that point in time.

Automated management

Understanding the environment and managing it cost-effectively is essential and IBM automated management software can help doing that.

IBM storage management products help with understanding how storage is configured and used and in easily making changes to improve the storage network. Many tasks, such as provisioning new storage, can be simplified to a single step by using our advanced automated management software.

Note: For more information about infrastructure simplification, see Chapter 4, “Storage infrastructure simplification” on page 49 and this Web site:

<http://www.ibm.com/servers/storage/solutions/is/>

1.1.2 Business continuity

The business climate in today's on demand era is highly competitive. Customers, employees, suppliers and business partners expect to be able to tap into your information at any hour of the day from any corner of the globe. If you have continuous business operations, then people can get what they need from your business—helping bolster your success and competitive advantage. Thus downtime is unacceptable today. Businesses must also be increasingly sensitive to issues of customer privacy and data security, so that vital information assets are not compromised. To achieve all this, you need a comprehensive Business Continuity plan for your business.

As shown in Figure 1-2 on page 4, business continuity can be achieved with:

- ▶ High availability
- ▶ Continuous operations
- ▶ Disaster recovery

High availability is achieved by means of fault tolerant, failure-resistant infrastructure supporting continuous application processing.

Continuous operations imply that nondisruptive backups and system maintenance coupled with continuous availability of applications.

Disaster recovery means protection against unplanned outages such as natural disasters through reliable and predictable recovery methods.

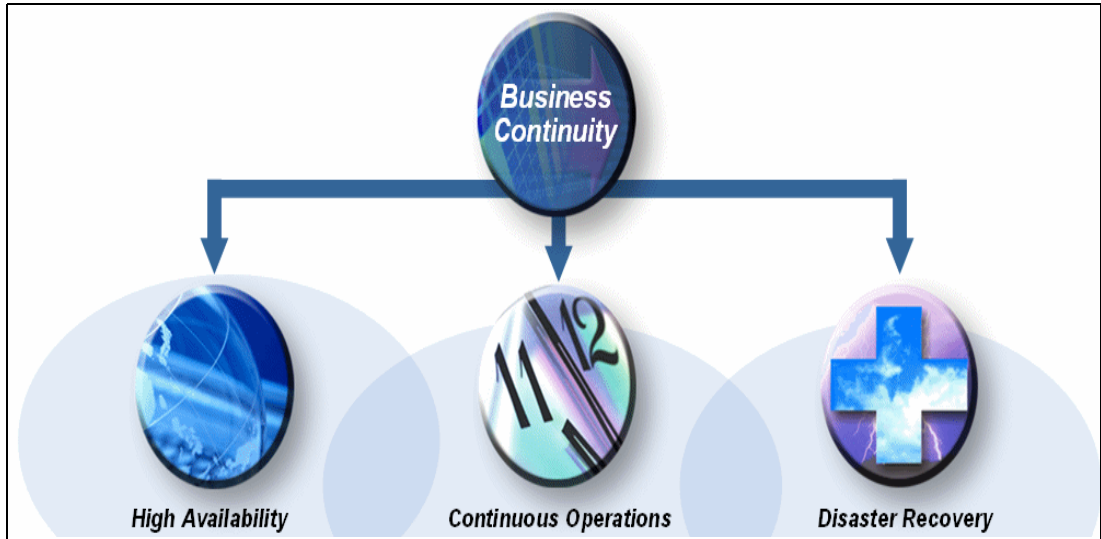


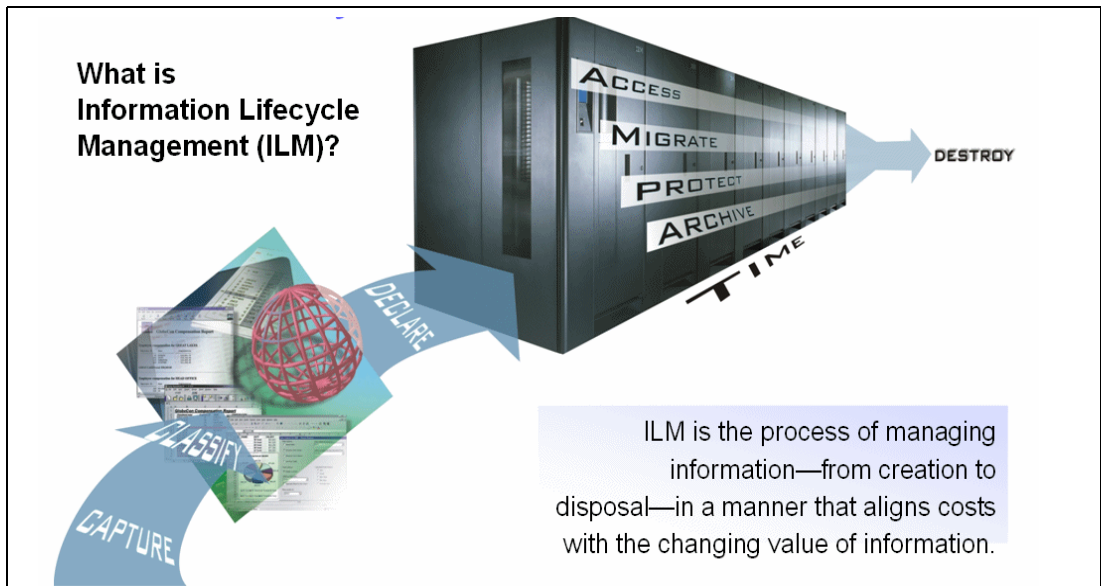
Figure 1-2 Business Continuity

Note: For more information about business continuity, see Chapter 2, “Business continuity solutions” on page 9, and refer to this Web site:

http://www-1.ibm.com/servers/storage/solutions/business_continuity/

1.1.3 Information lifecycle management

Information has become an increasingly valuable asset. But as the quantity of information grows, it becomes increasingly costly and complex to store and manage. Information lifecycle management (ILM) is a process for managing information through its lifecycle, from conception until disposal, in a manner that optimizes storage and access at the lowest cost. The most efficient ILM strategy for a business manages information according to its value. For small and medium-sized enterprises, predicting storage needs and controlling costs can be especially challenging as the business grows.



ILM is the process of managing information—from creation to disposal—in a manner that aligns costs with the changing value of information.

Figure 1-3 Information lifecycle management

IBM unique experiences with the broad range of ILM technologies, and its broad portfolio of offerings and solutions, including offerings in System Storage hardware, System Storage Open Software, and DB2® Content Management, can help provide businesses with the best solutions to manage their information throughout its lifecycle.

Note: For more information about information lifecycle management, see Chapter 3, “Information lifecycle management” on page 35 and this Web site:

<http://www-1.ibm.com/servers/storage/solutions/ilm/>

1.2 On demand storage environment

To achieve information on demand via IS, BC, and ILM as explained above, one needs an on demand storage environment. Here we outline the complete range of IBM System Storage solutions as an on demand storage environment. This environment will ultimately help support the goals of your business.

Figure 1-4 on page 6 shows the building blocks of an on demand storage environment and these are discussed in more detail in the sections below.

Hardware infrastructure

A complete range of IBM storage hardware products provides flexibility in choice of service quality and cost.

IBM has brought together, into one family, a broad range of disk systems to help small and medium size businesses to large size enterprises to select the right solutions for their needs. The IBM System Storage DS family combines the high-performance of the IBM System Storage DS6000 and DS8000 series enterprise servers with IBM System Storage DS4000 series of mid-range systems, with low priced entry systems. This family is complemented by a full range of IBM System Storage capabilities like advanced copy services, management tools and virtualization services to help protect your data.

Note: For more information about IBM disk systems offerings, see Part 1, “Disk systems” on page 57.

IBM also offers a variety of tape storage systems. From a single tape drive to libraries capturing up to petabytes of data, IBM System Storage tape drives, tape libraries and virtual tape servers offer a range of solutions to meet your data management needs.

Note: For more information about IBM tape offerings, see Part 2, “Tape systems” on page 199.

IBM System Storage SAN solutions provide integrated small and medium business and enterprise solutions with multiprotocol local, campus, metropolitan and global storage networking. IBM provides the choice of Brocade, Cisco, and McDATA switches and directors.

Note: For more information about IBM SAN offerings, see Chapter 16, “Storage Area Network products” on page 323.

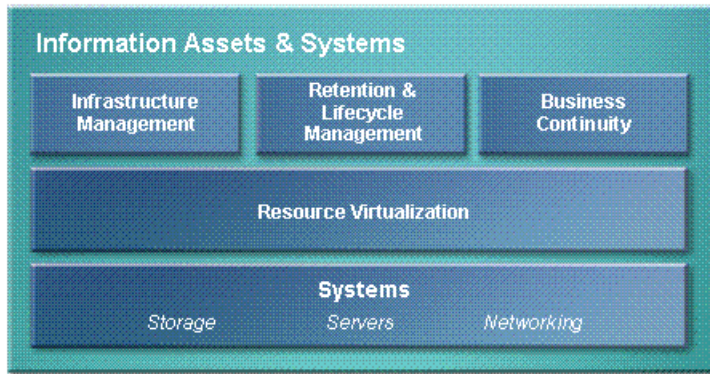


Figure 1-4 On demand storage environment

IBM System Storage open software family

IBM System Storage open software family offers storage software solutions for open systems environment in the area of storage infrastructure management, hierarchical storage management, archive management, recovery management and storage virtualization.

Storage virtualization

Storage virtualization products work to improve the flexibility and utilization of your storage. Virtualization pools your storage into a single logical reservoir of capacity for centralized management. This reservoir can include storage capacity from multiple vendors and platforms in heterogeneous environments. Virtualization also reduces the efforts required for hardware reconfiguration, to help support Business Continuity.

The Virtualization family features the IBM System Storage SAN Volume Controller for virtualization at the block level.

<http://www-1.ibm.com/servers/storage/software/virtualization/svc/index.html>

Storage infrastructure management

Storage infrastructure management is designed to make resource sharing possible across the enterprise, including heterogeneous networks. It helps empower administrators by providing an integrated view of the entire storage environment including software and hardware. Storage infrastructure management provides insight into the historic, operational and predictive analytics of the storage environment that, in turn, can help administrators improve storage capacity and network utilization, and help avoid business outages. It also supports policy-based automation, such as capacity provisioning, performance optimization and data management, helping to provide outstanding business agility.

Here are the products for storage infrastructure management:

- ▶ IBM TotalStorage Productivity Center
<http://www-1.ibm.com/servers/storage/software/center/index.html>
- ▶ IBM TotalStorage Productivity Center with Advanced Provisioning
<http://www-1.ibm.com/servers/storage/software/center/provisioning/index.html>

Hierarchical Storage Management

IBM Hierarchical Storage Management capabilities provide a way to capture low-activity or inactive data and feed it into a hierarchy of lower-cost or tiered storage. This helps you control your data storage growth and costs. Automated, policy-based capabilities determine where data should be stored, based on factors such as how critical it is to your business, how accessible and available you want it to be, and the cost structures of available devices.

Interoperability with IBM Content and Records Management products allows enterprise data to be moved from one medium to another with efficiency while helping avoid disruptions in service.

IBM HSM products are IBM Tivoli Storage Manager for Space Management and Tivoli Storage Manager HSM for Windows®.

<http://www.ibm.com/software/tivoli/products/storage-mgr-space/>
<http://www.ibm.com/software/tivoli/products/storage-mgr-hsm/>

Archive management

Archive management gives complete solutions that are designed to help enterprises archive, retain, and manage data, including e-mail, to help satisfy regulatory, legal, and other business requirements. Archive management products are interoperable with many content management products available in the marketplace, including the IBM DB2 Content Management family. They support a wide variety of storage media, write one read many (WORM) tape and magnetic disks.

The IBM System Storage DR 550 is designed as a preconfigured, integrated solution to help store, retrieve, manage, share and secure regulated and unregulated data. It uses IBM Tivoli Storage Manager for Data Retention to manage and enforce the data-retention policies.

<http://www.ibm.com/software/tivoli/products/storage-mgr-data-reten/>

Recovery management

Recovery management solutions are designed to quickly and reliably recover enterprise data when needed. They use centralized Web-based management, intelligent backup and archiving (with minimal or no impact on application availability), and automated policy-based data migration services.

IBM Tivoli Storage Manager offers you these features for recovery management.

<http://www.ibm.com/software/tivoli/products/storage-mgr/>

Note: For more information about these IBM System Storage open software family, see Part 4, “IBM System Storage open software” on page 355.

Storage orchestration

Storage orchestration is being designed to help automated management and allocation of storage resources in concert with your business goals and policies. This will improve return of IT assets and utilization. Beginning with policy-based, automated work flows, storage orchestration is planned to include optional real-time monitoring and intelligence that is designed to enable autonomic management of the on demand storage environment.

IBM provides the following products in the area of storage orchestration:

- ▶ IBM Tivoli Intelligent Orchestrator

<http://www-306.ibm.com/software/tivoli/products/intell-orch/>

- ▶ IBM Tivoli Provisioning Manager

<http://www.ibm.com/software/tivoli/products/prov-mgr/>

1.3 More information

- ▶ For the latest information about IBM System Storage offerings, refer to:

<http://www-1.ibm.com/servers/storage/index.html>

- ▶ The following link gives an A-Z listing of all IBM System Storage products:

<http://www-1.ibm.com/servers/storage/product/index.html>

- ▶ For more information about virtualization and IBM System Storage virtualization offerings, refer to:

<http://www-1.ibm.com/servers/storage/software/virtualization/>



Business continuity solutions

In this chapter, we discuss the concepts and the solutions for business continuity as they relate to IBM System Storage. In particular, we discuss:

- ▶ Business continuity introduction
- ▶ IT business continuity timeline
- ▶ Business continuity solution selection methodology
- ▶ Tiers of business continuity
- ▶ Segmentation of business continuity solutions
- ▶ Backup and Restore solutions
- ▶ Rapid Data Recovery solutions
- ▶ Continuous availability solutions

2.1 Introduction to business continuity

Today's on demand business climate is highly competitive. Customers, employees, suppliers and business partners expect to be able to use applications and access data at any time from anywhere. Businesses must also be increasingly sensitive to data protection and security. Add regulatory requirements and the inherent demands of participating in the global economy, and the demands of modern IT management become apparent. Business continuity is one of three key storage principles (along with infrastructure simplification and information lifecycle management) that companies can implement to help address these demands.

A 24x7 enterprise needs a comprehensive business continuity plan that supports high availability, continuous operations and disaster recovery. Implementing such a plan can help organizations achieve resiliency by:

- ▶ Supporting competitive efforts through more efficient and cost-effective risk management
- ▶ Helping to achieve critical business application and data availability based on business value
- ▶ Facilitating compliance with government rules and regulations
- ▶ Safeguarding against internal and external threats and providing sustained operations even in the event of a disaster.

Most enterprises cannot afford the cost of downtime due to planned or unplanned system outages. Although the indirect, longer term impacts of downtime - lost market share, decreased productivity, noncompliance with regulations, reduced competitiveness, damaged brand reputation and eroded customer loyalty - are harder to measure, they are equally important. Strengthening the resiliency of your business can help mitigate or avoid them.

2.1.1 Three aspects of IT business continuity

Business continuity for key applications and business processes has three primary facets, as shown in Figure 2-1, by providing:

- ▶ High availability
- ▶ continuous operations
- ▶ Disaster recovery

There is actually considerable overlap between these three solutions areas - some of the solutions that you can implement to prepare for disaster recovery, can also help with high availability and with continuous operations. In this way, your investment in disaster recovery can help your operations even if you never suffer a disaster.

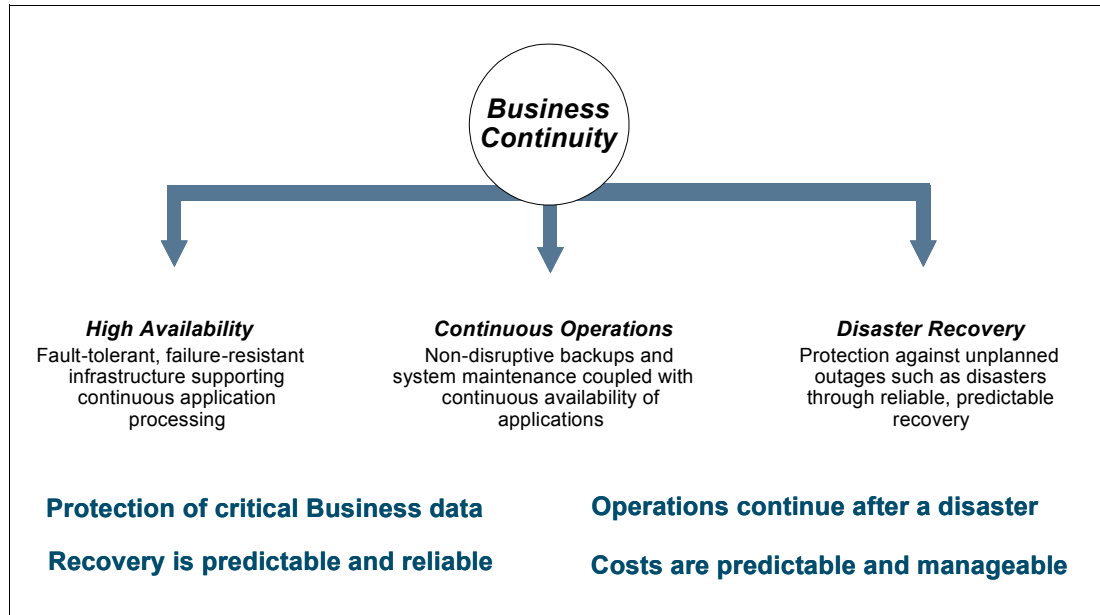


Figure 2-1 Three aspects of Business Continuity

High availability

High availability is the ability to provide access to applications. High availability is often provided by clustering solutions that work with operating systems coupled with a hardware infrastructure that has no single points of failure. If a server that is running an application suffers a failure, the application is picked up by another server in the cluster, with minimal or no interruption to the users. Today's servers and storage systems are also built with fault-tolerant architectures to minimize application outages due to hardware failures.

You can think of high availability as a resilient IT infrastructure that masks failures, and thus continues to provide access to applications.

Continuous operations

Continuous operations is the ability to keep things running under normal operations, for example, where applications can remain online during scheduled backups or planned maintenance. Continuous operations technologies provide the ability to perform repetitive, ongoing, and necessary infrastructure actions, while still maintaining high availability.

Normally, all the components providing continuous operations are situated in the same computer room. The building, therefore, becomes the single point of failure. Thus, a continuous operation setup does not usually of itself provide a disaster recovery solution.

You can think of continuous operations as the ability to keep applications running, during scheduled backups or planned maintenance.

Disaster recovery

Finally, disaster recovery is the ability to recover a datacenter at a different site if a disaster destroys the primary site or otherwise renders it inoperable. In a disaster recovery scenario, the processing resumes at a different site, and on different hardware. A non-disaster problem, such as a corruption of a key customer database, may indeed be a catastrophe for a business, but it is not, by our definition, a disaster, unless processing must be resumed at a different location and on different hardware.

You can think of disaster recovery as the ability to recover at a different site from unplanned outages.

2.1.2 The objectives of IT business continuity

The objectives of business continuity are to protect critical business data, to make key applications available, and to enable operations to continue after a disaster. These must happen in such a way that recovery time is both predictable and reliable, with predictable and manageable costs.

2.2 IT recovery timeline

In Figure 2-2 is a diagram of an IT recovery. Time proceeds from left to right.

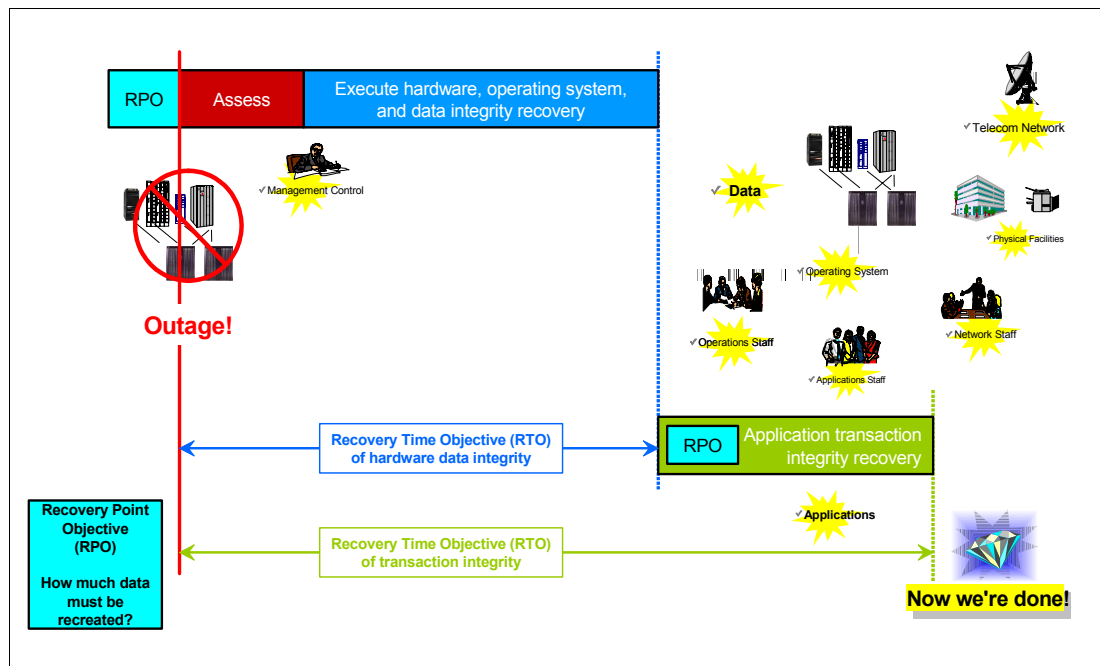


Figure 2-2 Timeline of an IT recovery

Three major components are:

- ▶ Assessment
- ▶ Hardware, operating system recovery
- ▶ Transaction integrity recovery

Assessment

After the outage occurs, the first step is that management must assess the outage (this incurs elapsed time). Because there is a significant capital cost to declaring a disaster and executing a recovery, then management will need to be sure that the situation warrants committing their organization to that expense. Once management has decided to declare a disaster, then they initiate the business continuity process.

Hardware, operating system recovery

The first stage of the business continuity process is to recover the hardware, operating systems, and the data itself. Operations, networking, telecommunications, physical facilities, and associated staff are involved.

At the end of this stage, the operating systems and the data are recovered. The data ideally is accurate and data consistent to a point-in-time prior to the outage. The time duration to this point is the *Recovery Time Objective of hardware data integrity*.

Transaction integrity recovery

Hardware data integrity is not the same as database/application integrity. The storage and servers cannot know what the logical database relationship is between multiple data blocks in the database. Therefore, once the first stage is complete, the transaction integrity recovery must next be performed by the applications staff, on the application and database.

The applications staff performs transaction integrity recovery. Hopefully, this is a database restart and not a database recovery. This process will back out incomplete logical units of work, and restore the database to logical integrity as of the most recent time possible.

When the transaction integrity recovery (rollback or roll forward) is complete, we now have the application and databases ready for user access. This duration is the **Recovery Time Objective of transaction integrity**.

Note: There is a difference in elapsed time between RTO of hardware data integrity, and RTO of transaction integrity. When discussing the Recovery Time Objective (RTO), it is important to distinguish which of the two is being referred to. Operations and application staff can have differing perceptions of the RTO depending on whether the RTO is assumed to be at the hardware recovery level, or at the application recovery level.

Recovery Point Objective (RPO)

Finally, observe how the Recovery Point Objective (RPO) is depicted in Figure 2-2. RPO (which is how much data which must be recreated) is shown as the time offset before the outage occurred.

Note: The RPO data recreation happens in the transaction integrity recovery stage. RPO data recreation cannot happen in the hardware and operating system recovery stage, because the server and storage components do not have knowledge of the logical relationships between multiple application and database blocks of data.

2.3 Selecting business continuity solutions

From an IT infrastructure standpoint, there is a large variety of valid business continuity products. The fundamental challenge is to select the optimum blend of all these business continuity products and technologies.

A common problem in the past has been a tendency to view the business continuity solution as individual product technologies and piece parts; see Figure 2-3. Instead, business continuity solutions need to be viewed as a whole, integrated multi-product solution.

In this section we propose a Business Continuity Solution Selection Methodology that can be used to sort, summarize, and organize the various business requirements in a methodical

way. Then, we methodically use those business requirements to efficiently identify a proper and valid subset of business continuity technologies to address the requirements.

- Each vendor and product area tends to build separate pieces of the solution.
- Insufficient interlocking of the different areas.
- Business Continuance and Disaster Recovery need to be seen as an **integrated product solution**.
- There are many valid technologies, but how to choose among them?

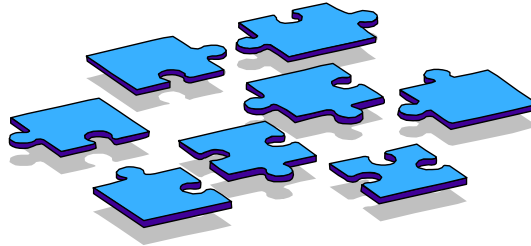


Figure 2-3 Historical challenge in selecting business continuity solutions

2.3.1 Business continuity is an end-to-end process

To combine and properly choose between multiple products, disciplines, and skills to effect a successful IT business continuity solution, we first observe that we can categorize all valid business continuity IT technologies into five component domains:

- ▶ Servers
- ▶ Storage
- ▶ Software and automation
- ▶ Networking and physical infrastructure
- ▶ Skills and services required to implement and operate the above

All IT infrastructure necessary to support the business continuity solution can be categorized as one of these five components; see Figure 2-4.

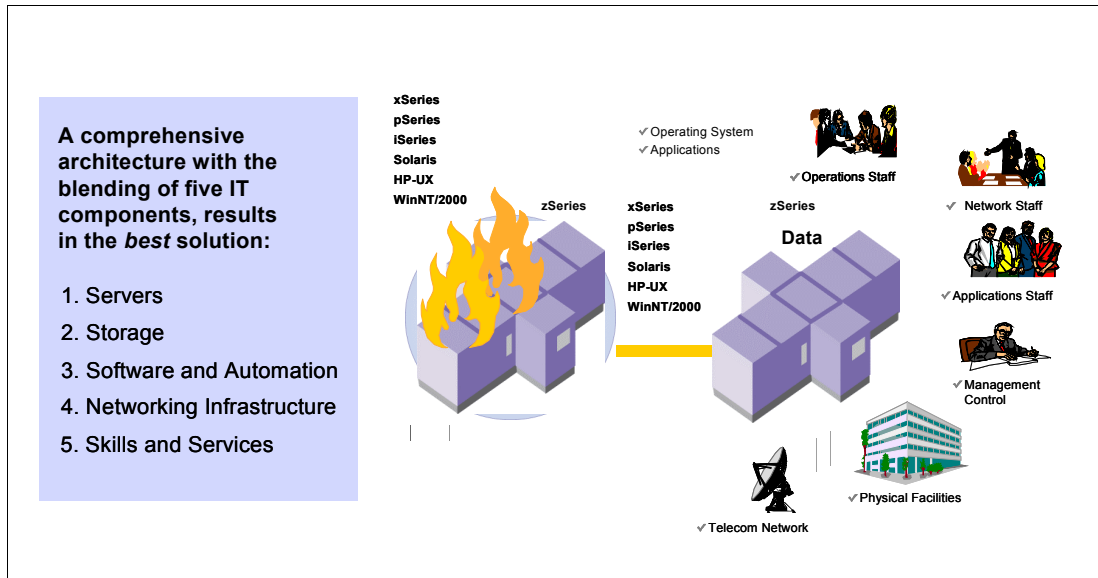


Figure 2-4 The five components of IT business continuity

These five categories provide a framework to organize the various component evaluation skills that will be needed. To gather the proper mix of evaluation skills together, facilitates an effective comparison, contrast, and blend of all five product component areas to arrive at an optimum solution.

2.3.2 Leverage infrastructure simplification for business continuity

Complexity can prevent even the best organization from acting nimbly to meet the ever-changing market and client demands. Solutions for infrastructure simplification can help to improve efficiency, lower total cost of ownership, and reduce time-consuming and costly errors.

From a business continuity perspective, infrastructure simplification can do the following:

- ▶ Introduce common architectural recovery platforms (operating systems, applications, databases, servers, storage systems)
- ▶ Reduce the components to recover
- ▶ Reduce the difficulties in managing dynamic application changes, managing IT environment changes, and managing data and storage allocations

An excellent idea is to do a good job of infrastructure simplification as a prerequisite to implementing improved IT business continuity. This has the effect of introducing consolidation (and cost savings) as a foundation for implementing the new BC functions.

2.4 The tiers of business continuity

By categorization of business continuity technology into the various tiers, we have the capability to more easily match our desired RTO time with the optimum set of technologies. The reason for multiple tiers is that as the RTO time decreases, the optimum business continuity technologies for RTO must change. For any given RTO, there is always a particular set of optimum price/performance business continuity technologies.

The concept of business continuity tiers (Figure 2-5) is a common method used in today's best practices for business continuity solution design. It was originally developed by the IBM US SHARE User Group in 1988. The concept of tiers is powerful, because the tiers concept recognizes that for a given customer Recovery Time Objective, all business continuity products and technologies can be sorted into a RTO solution subset that addresses that particular RTO range.

The tiers concept is flexible. As products and functions change and improve over time, the tier chart only needs to be updated by the addition of that new technology into the appropriate tier and RTO.

The tiers chart below, gives a generalized view of some of today's IBM business continuity technologies by tier. As the recovery time becomes shorter, then more aggressive business continuity technologies must be applied to achieve that RTO (carrying with them their associated increase in value and capital cost).

The tiers also reflect the way an IT organization can incrementally grow and improve their IT BC over time. Each preceding tier provides a foundation for the subsequent higher tier. Notice that implementing a higher tier does not remove the need for the lower tier. In fact, the higher tier can exist because it is based upon the foundation of the tiers below it.

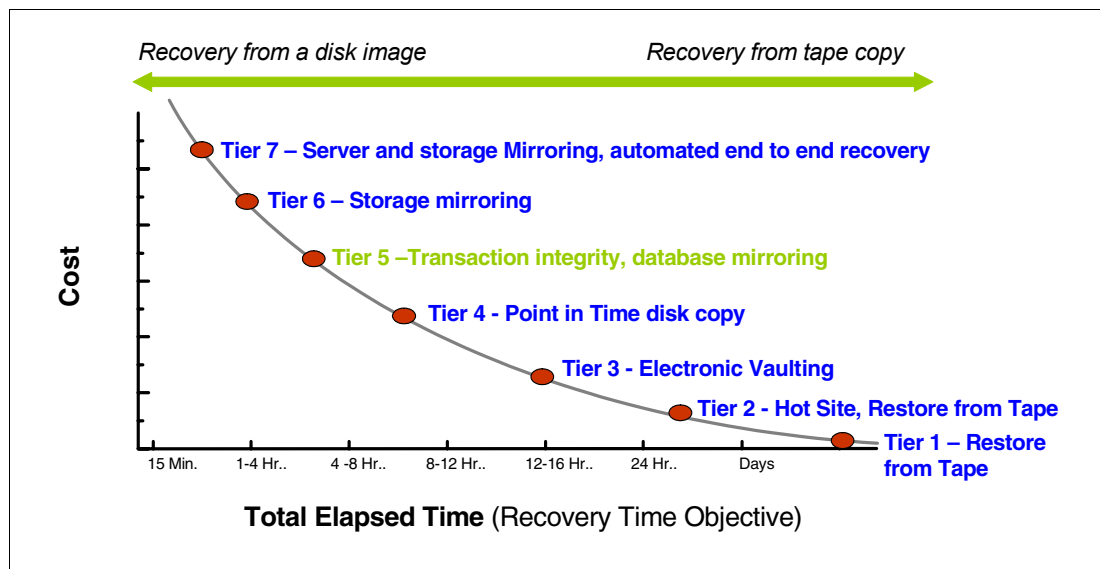


Figure 2-5 Tiers of business continuity

2.4.1 Definition of the tiers of business continuity

In this section we will describe the tiers of business continuity. These tiers are generally accepted examples of today's IT BC tiers. We suggest that you should refine these tiers to create your own specific version of BC tiers chart, specific to your organization, installation, and recovery times.

Tier 0: No off-site data

Businesses with a Tier 0 business continuity solution have no Business Continuity plan.

- There is no saved information, no documentation, no backup hardware, and no contingency plan.

The length of recovery time in this instance is unpredictable. In fact, it may not be possible to recover at all.

Tier 1: Data backup with no hotsite

Businesses that use Tier 1 business continuity solutions back up their data at an off-site facility. Depending on how often backups are made, they are prepared to accept several days to weeks of data loss, but their backups are secure off-site. However, this tier lacks the systems on which to restore data. Sample solutions include:

- ▶ Pickup Truck Access Method (PTAM)
- ▶ Disk Subsystem or Tape based mirroring to locations without processors
- ▶ IBM Tivoli Storage Manager

Tier 2: Data backup with a hotsite

Businesses using Tier 2 business continuity solutions make regular backups on tape. This is combined with an off-site facility and infrastructure (known as a hotsite) in which to restore systems from those tapes in the event of a disaster. This tier of solution will still result in the need to recreate several hours to days worth of data, but it is more predictable in recovery time. Sample solutions include:

- ▶ PTAM with hotsite available
- ▶ IBM Tivoli Storage Manager

Tier 3: Electronic vaulting

Tier 3 solutions utilize components of Tier 2. Additionally, some mission-critical data is electronically vaulted. This electronically vaulted data is typically more current than that which is shipped via the Pickup Truck Access Method (PTAM). As a result, there is less data recreation or loss after a disaster occurs. Sample solutions include:

- ▶ Electronic Vaulting of Data
- ▶ IBM Tivoli Storage Manager - Disaster Recovery Manager

Tier 4: Point-in-time copies

Tier 4 solutions are used by businesses that require both greater data currency and faster recovery than users of lower tiers. Rather than relying largely on shipping tape, as is common on the lower tiers, Tier 4 solutions begin to incorporate advanced core technologies, such as point-in-time disk copy, basic disk mirroring without data consistency, forwarding of database log files and journals, and so on. Several hours of data loss is typical, but these tiers of solutions use methodologies such as disk point-in-time (PIT) copies, managed with software such as Tivoli Storage Manager, to effect a solution with better currency of data than can typically be achieved through tape based solutions. Sample solutions include:

- ▶ Peer-to-Peer Virtual Tape Server
- ▶ Global Copy
- ▶ FlashCopy® Backup/Restore for SAP Databases
- ▶ IBM TotalStorage Productivity Center (TPC) for Replication

Tier 5: Transaction integrity

Tier 5 has traditionally been reserved for database-specific solutions in which transaction integrity-based database replication solutions are employed. These solutions are used by businesses with a requirement for database transaction and synchronization of data between one production and one or more remote data centers. These solutions can offer little or no data loss, and they can provide methodologies that employ deferred application of data at the remote site to provide protection against propagation of logical data corruption. The implementation of these solutions is entirely dependent on the software application in use.

Sample solutions include:

- ▶ Two-phase commit, such as DB2 remote replication

- ▶ WebSphere® Information Integrator Q replication
- ▶ Oracle Data-Guard

Tier 6: Storage Mirroring with zero or little data loss

Tier 6 business continuity solutions are defined as storage mirroring solutions. In other words, the storage subsystem (whether it is disk or tape) maintains a very high level of data currency, with built-in data integrity. This tier is used by businesses, with little or no tolerance for data loss, which need to restore data to applications rapidly. These solutions provide power-outage data consistency, and have no dependence on the applications. Sample solutions include:

- ▶ Metro Mirror
- ▶ Global Mirror
- ▶ z/OS Global Mirror
- ▶ GDPS® HyperSwap™ Manager
- ▶ Peer-to-Peer VTS with synchronous write
- ▶ IBM TotalStorage Productivity Center (TPC) for Replication
- ▶ HACMP/XD with Logical Volume Mirroring

Tier 7: Highly automated business integrated solution

Tier 7 solutions include all the major components being used for a Tier 6 solution with the additional integration of automation of all server, storage, software, and networking components. This allows a Tier 7 solution to provide near continuous availability, with automated recovery of the applications. Sample solutions include:

- ▶ Geographically Dispersed Parallel Sysplex™ (GDPS) for System z™
- ▶ HACMP/XD with Metro Mirror

2.5 Blending tiers into an optimized solution

Best practice today in designing a business continuity solution is to further use the tiers concept to derive a blended business continuity solution for the entire enterprise. The most common result, from an enterprise standpoint, is a strategic architecture of three bands in a blended business continuity solution. Three bands generally appear as an optimum number, because at the enterprise level, two bands generally are insufficiently optimized (in other words, overkill at some point and underkill at others), and four bands are more complex but generally do not provide enough additional strategic benefit.

To use the tiers to derive a blended, optimized enterprise business continuity architecture, we suggest these steps:

1. Categorize the business' entire set of applications into three bands: low tolerance to outage, somewhat tolerant to outage, and very tolerant to outage. Some applications that are not in and of themselves critical, actually feed the critical applications. Therefore, those applications would need to be included in the higher band.
2. Within each band, there are tiers. The individual tiers represent the major business continuity technology choices for that band. It is not necessary to use all the tiers, and of course, it is not necessary to use all the technologies.
3. Once we have segmented the applications (as best we can) into the three bands, we usually select one best strategic business continuity methodology for that band. The contents of the tiers are the *candidate technologies* from which the strategic methodology is chosen.

Business impact analysis, risk assessments, and program assessments are excellent methodologies and essential tools to assist in defining your application segmentation. These three bands are strategic objectives, which the organization by necessity will implement over time.

Note that the IT business continuity technologies chosen to service each band, become the strategic, consolidated technology platform standards upon which this band's IT business continuity is based.

IBM business continuity solutions in the System Storage Resiliency Portfolio have been segmented into the continuous availability band, the Rapid Data Recovery band, and backup/restore band, shown in Figure 2-6.

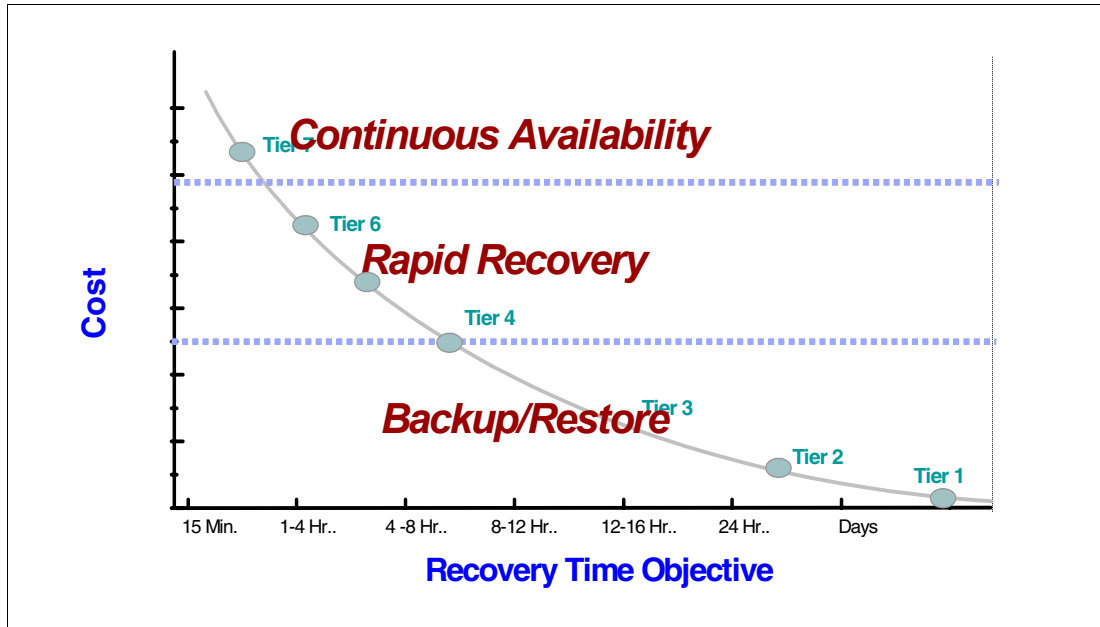


Figure 2-6 Three business continuity solution bands

The implementation of an end-to-end IT business continuity solution does not need to be done all at once. You should plan for a step-by-step incremental IT business continuity project, and start with the current environment. From there you can incrementally build towards the final objective. Each stage provides the necessary foundation for the next step; see Figure 2-7.

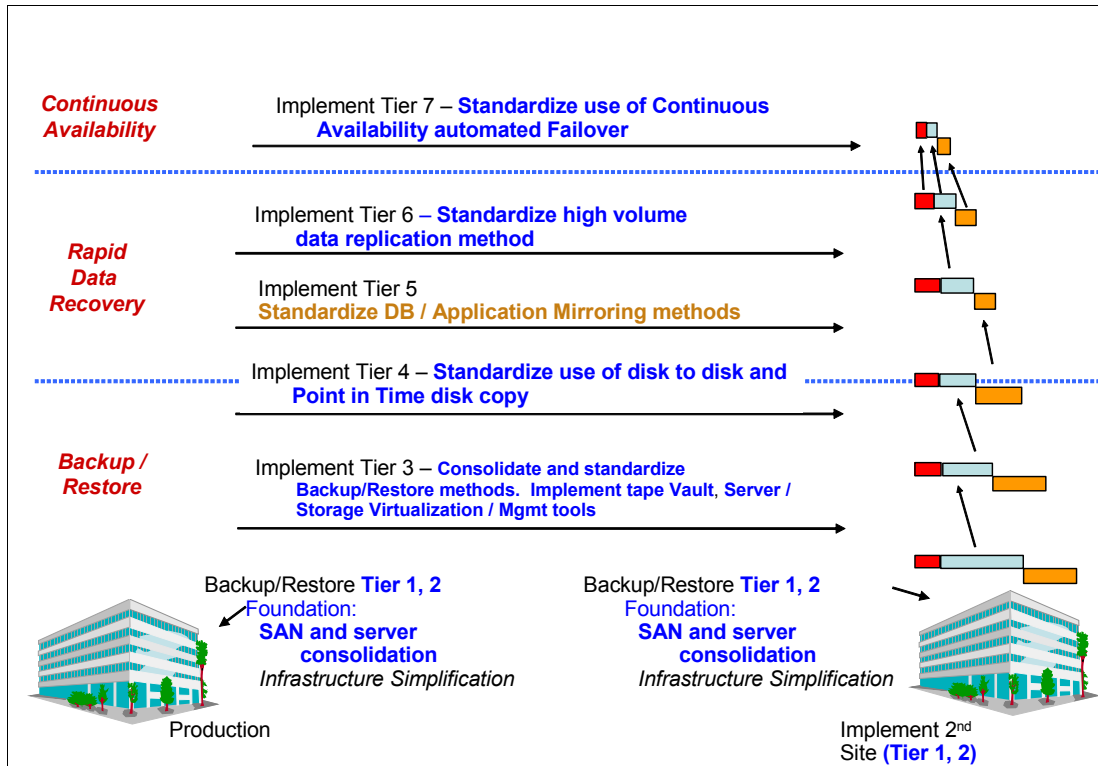


Figure 2-7 Phases of an IT BC implementation

Now that we have introduced the three solution segment bands Backup/Restore, Rapid Data Recovery, and Continuous Availability we will further describe each segment band.

2.6 Backup and Restore solutions (Tier 1-4)

Backup and Restore is the most simple and basic solution to protect and recover your data from failure by creating another copy of data from the production system. The second copy of data allows you to restore data to the time of the data backup.

Backup and Restore spans Tier 1 through Tier 4. This section describes the following solutions:

- ▶ Volume Shadow Copy Services (VSS)
- ▶ IBM Tivoli Storage Manager
- ▶ IBM Tape storage systems
- ▶ IBM System Storage N series Snapshot solutions

2.6.1 Microsoft Volume Shadow Copy Service (VSS)

The Microsoft Volume Shadow Copy Service (VSS) is a storage management interface included with Microsoft Windows Server® 2003. VSS enables a storage array to interact with third-party applications that use the VSS Application Programming Interface (API).

Volume Shadow Copy Services falls under Tier 4 which is similar to products supporting FlashCopy services.

Common uses of VSS are:

- ▶ Creating consistent backups of open files and applications
- ▶ Creating shadow copies for shared folders

Hardware providers are available from IBM for various disk systems; the hardware provider interacts with the Microsoft VSS service to create and maintain the shadow copies, using the FlashCopy disk function. Hardware providers for the IBM System Storage DS8000, DS6000, DS4000, and SAN Volume Controller can be downloaded.

2.6.2 IBM Tivoli Storage Manager

IBM Tivoli Storage Manager is software that enables you to protect your organization's data from failures and other errors by storing backup, archive, space management, and bare-metal restore data, as well as compliance and disaster-recovery data in a hierarchy of offline storage. Tivoli Storage Manager family of software products is designed to provide centralized, automated data protection that can help reduce the risks associated with data loss while also helping to reduce complexity, manage costs, and address compliance with regulatory data retention requirements.

Since it is designed to protect a company's important business information and data in case of Disaster, the Tivoli Storage Manager server should be one of the main production systems that is available and ready to run for recovery of your business data and applications.

There are six solutions to achieve each Disaster Recovery tier:

- ▶ Tier 1 - IBM Tivoli Storage Manager manual off-site vaulting
- ▶ Tier 2 - IBM Tivoli Storage Manager manual off-site vaulting with a hot site
- ▶ Tier 3 - IBM Tivoli Storage Manager electronic vaulting
- ▶ Tier 4 - IBM Tivoli Storage Manager with SAN attached duplicates
- ▶ Tier 5 - IBM Tivoli Storage Manager clustering
- ▶ Tier 6 - IBM Tivoli Storage Manager running in a duplicate site

Further information can be found in Chapter 24, "IBM Tivoli Storage Manager" on page 437.

2.6.3 IBM Tape storage systems

IBM tape drives, libraries and virtualization products can be used to build solutions for data archiving, backup and disaster recovery. They can provide an important component of a comprehensive business continuity strategy that supports high availability, near continuous operations and disaster recovery.

Further information can be found Part 2, "Tape systems" on page 199.

2.6.4 IBM System Storage N series Snapshot solutions

IBM System Storage N series offers a comprehensive set of Tier 4 business continuity solutions based on Snapshot technology:

- ▶ **Snapshot:** A SnapShot is a "frozen," read-only view of a WAFL (Write Anywhere File Layout) volume that provides easy access to old versions of files, directory hierarchies, and/or LUNs. Users can directly access up to 255 SnapShot copies per WAFL volume to recover from accidental deletions, corruptions, or modifications of their data.
- ▶ **SnapRestore:** SnapRestore software can help recover data in amounts as small as an individual file up to a multi terabyte volume so that operations can quickly resume.
- ▶ **SnapVault:** SnapVault technology provides a centralized, disk-based backup solution for IBM N series systems. Storing backup data in multiple Snapshot copies on a SnapVault

secondary storage system can let enterprises keep multiple backups online over a period of time for faster restoration.

- ▶ **SnapManager for Microsoft SQL Server:** SnapManager supports rapid SQL Server backup times, and makes each backup a complete and consistent copy of the original. It allows back up or restore of several databases simultaneously, as well as volume expansion.
- ▶ **SnapManager for Microsoft Exchange:** SnapManager software can provide near-instantaneous hot backups and rapid restores for Exchange environments.
- ▶ **Single Mailbox Recovery for Microsoft Exchange (SMBR):** The combination SnapManager for Microsoft Exchange and SMBR allows rapid recoveries of Exchange databases at almost any level of granularity-storage group, database, folder, single mailbox, or single message.

Further information can be found Chapter 9, “IBM System Storage N series” on page 155.

2.7 Rapid Data Recovery solutions (Tier 4-6)

Rapid Data Recovery is based on maintaining a second copy of data that is consistent at a point-in-time as close to the time of a failure as possible. This consistent set of data allows for the restart of systems and applications without having to restore data and re-applying updates that have occurred since the time of the data backup. It is possible that there may be a loss of a minimal number of in-flight transactions.

Rapid Data Recovery spans Tier 4 through Tier 6 and is different from Continuous Availability, because it does not have the end-to-end server, storage, software, and networking automation that is required to be a Tier 7 solution.

Rapid Data Recovery solutions based on replication technology, as shown in Figure 2-8, can be implemented on three different levels:

- ▶ Application/Database replication
 - Requires less bandwidth
 - Span of consistency is the application or database only
 - More complex implementation, must implement for each application
- ▶ Server replication
 - Can be less complex than application implementation, application independent
 - Uses server cycles, span limited to that operating system
- ▶ Storage replication
 - Implementation is largely platform and application independent, mirror logical disks, supports multiple heterogeneous systems
 - Requires more bandwidth

Application/Database and Server replication is implemented in software, whereas Storage replication exploits the hardware directly.

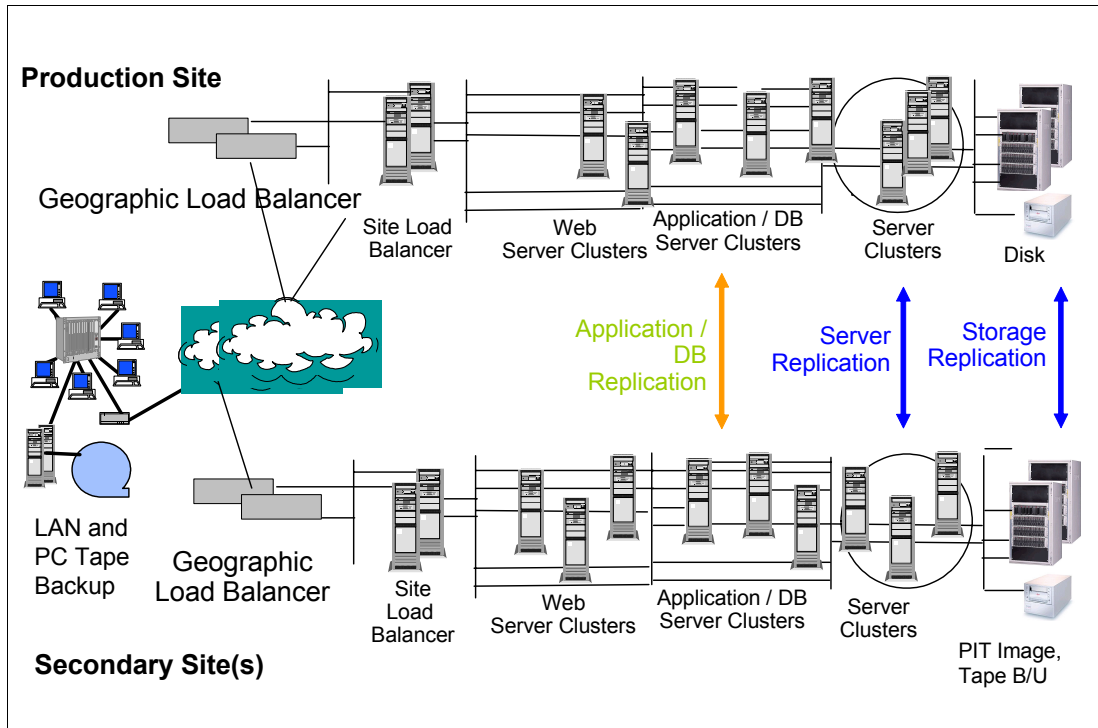


Figure 2-8 Software replication versus Hardware replication

Here we will focus on Rapid Data Recovery solutions based on storage replication. We will describe the following environments:

- ▶ Rapid Data Recovery for System z (GDPS/PPRC HyperSwap Manager)
- ▶ TotalStorage Productivity Center for Replication solutions for UNIX® and Windows
- ▶ IBM System Storage SAN Volume Controller

Before we discuss the particular solutions, we first need to explain some aspects of data consistency when using disk-based mirroring techniques in database environments. We will introduce some important terms, including the Consistency Group.

2.7.1 Data consistency for database environments

In a disaster recovery solution using disk remote mirroring, we want to restart a database application following an outage without having to restore and recover the database. This process has to be consistent, repeatable and fast, measurable in minutes. Restoring the database means to restore the last set of image copy tapes and then apply the log changes to bring the database up to the point of failure. This can take many hours, which would not constitute a Tier 4 or higher solution.

Moreover, actual disasters (fire, explosion, earthquake) are messy. You cannot expect your entire complex to fail simultaneously. Failures will be intermittent and gradual, and the disaster will occur over many seconds or even minutes. This is known as the *Rolling Disaster*: A viable disk mirroring disaster recovery solution must be designed to avoid data corruption that is caused during a Rolling Disaster.

In any operating system, the sequence in which updates are being made is what maintains the integrity of the data. If that sequence is changed, data corruption will occur. The correct sequence must be maintained within a volume, across volumes, and across multiple storage devices. For example, in Figure 2-9, we show the relationship between a database and its log,

which demonstrates the requirement for maintaining I/O integrity. Data consistency across the storage enterprise must be maintained to ensure data integrity.

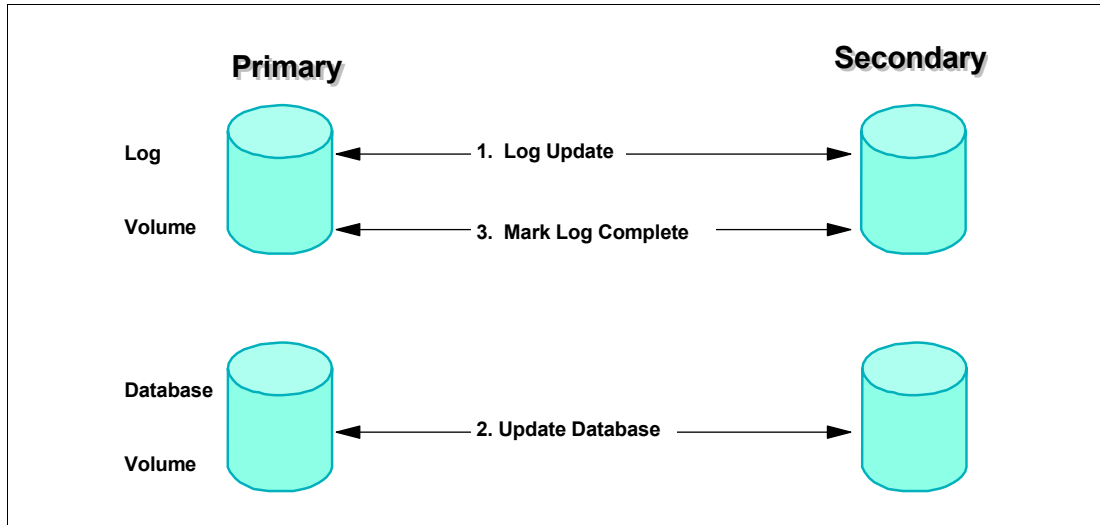


Figure 2-9 Sample update sequence in a database

The order of *Dependent Writes* across volumes must be maintained at remote locations. Failure to do so results in data corruption and introduction of data inconsistency.

In Figure 2-10, we illustrate this concept with an example. The intention to update the database is logged in the database log files at both the primary and secondary volumes (step 1). The database data file is updated at the primary volume, but the update does not reach the remote volume that contains the mirrored data file. The primary location is not aware of the write failure to the secondary volume (step 2). The database update is marked complete in the log files at both the primary and remote locations (step 3). The result is that the secondary site log files say the update was done, but the updated data is not in the database at the secondary location. There is no way to know that the data was corrupted.

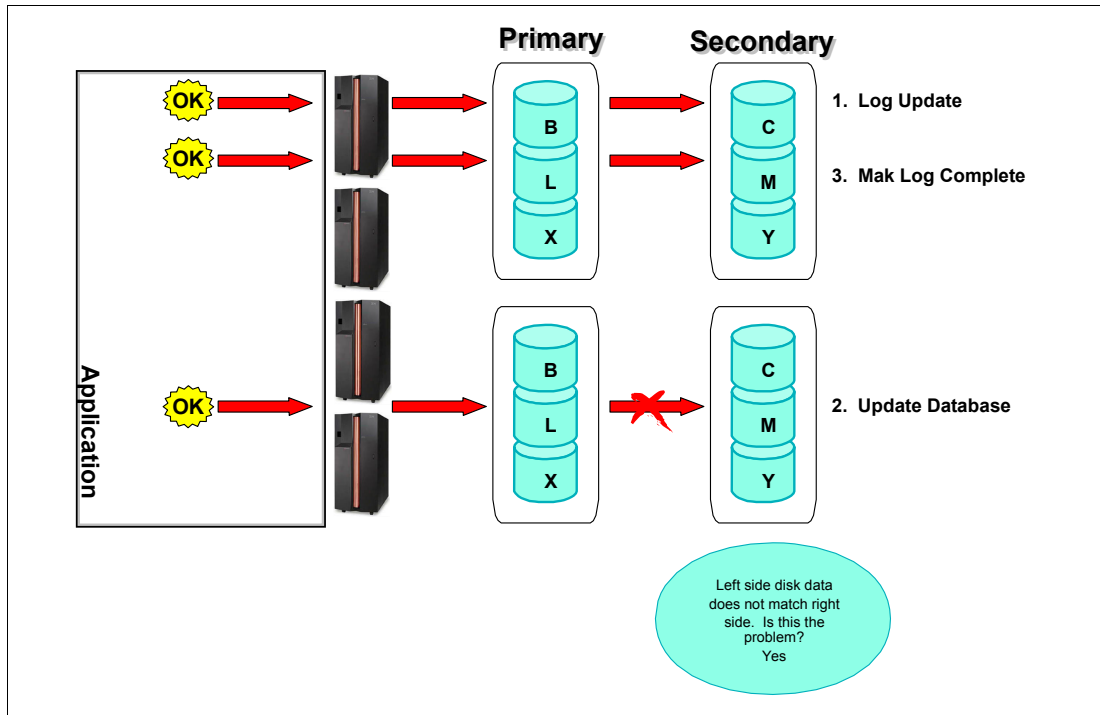


Figure 2-10 Sample Rolling Disaster

So, the issue is the disk subsystem cannot by itself perform the steps required to avoid the *Rolling Disaster* system problem.

For this reason, it is strongly recommended, at a minimum to implement *Consistency Groups* in any mirroring solution. Consistency Groups are an implementation of technology that assists with the consistency of application data capable of dependent writes. To guarantee a fully consistent remote copy, multiple volumes require a Consistency Group functionality. ESS, DS6000, and DS8000 Metro/Global Mirror microcode already has the Consistency Group function for both System z and open systems. SAN Volume Controller Metro/Global Mirror's and DS4000 Global Mirror's microcode has a Consistency Group function for open systems.

If any volume within a Consistency Group cannot complete a write to its counterpart in the remote mirror relationship, an *Extended Long Busy* (ELB) for mainframe environments or a *SCSI Queue Full* condition for open systems will be issued, preventing further writes to any of the volumes within the Consistency Group. This wait period is the perfect time to issue a freeze to all volumes involved to maintain consistency. If a write cannot complete, the storage system will not back out incomplete transactions on its own. Instead, the application will need to recognize that the transaction was incomplete and take the appropriate actions. Once the storage system pauses the application I/O to the affected primary volumes, the write dependent mechanism of the application prevents the Metro/Global Mirror secondary volumes from becoming inconsistent.

2.7.2 Rapid Data Recovery for System z

When used in a two site implementation, the GDPS/PPRC HyperSwap Manager can be a Tier 6 Rapid Data Recovery solution for disaster recovery situations. It is not a Tier 7 solution because it lacks the recovery automation provided by a full GDPS/PPRC implementation. 2.8.1, "Geographically Dispersed Parallel Sysplex (GDPS)" on page 27 provides more information about general GDPS.

Rapid Data Recovery for System z is provided by an IBM Global Services service offering, Geographically Dispersed Parallel Sysplex (GDPS) HyperSwap Manager, in the GDPS suite of offerings. It uses IBM System Storage Metro Mirror (previously known as Synchronous PPRC) to mirror the data between disk subsystems. Metro Mirror is a hardware-based mirroring and remote copying solution for the IBM System Storage DS6000, DS8000, and SVC systems.

Further information can be found on the GDPS Web site:

<http://www.ibm.com/systems/z/gdps/>

2.7.3 TotalStorage Productivity Center for Replication solutions for UNIX and Windows

Rapid data recovery for UNIX and Windows solutions are based on TotalStorage Productivity Center (TPC) for Replication, including Two Site Business Continuity, together with DS6000, DS8000, and SVC Copy Services. These products provide dynamic failover to a secondary storage device. TPC for Replication provides a single point of control for automated failover. It manages and protects data for both planned and unplanned exception conditions in a heterogeneous open systems environment. TPC for Replication can bring up a remote site with just a Database restart and data consistency is guaranteed and can be adapted to user needs.

By the use of Consistency Groups, this solution freezes the environment at a known point instead of mirroring literally hundreds of time-offset failures in a short amount of time. So, it protects data in the event of a rolling disaster.

Further information can be found at the following Web sites.

IBM TotalStorage Productivity Center for Replication:

<http://www-03.ibm.com/servers/storage/software/center/replication/index.html>

2.7.4 IBM System Storage SAN Volume Controller

IBM System Storage SAN Volume Controller (SVC) creates a virtual pool of storage, so that it appears as one logical device to centrally manage and to allocate capacity as needed. It also provides one solution to help achieve the most effective on demand use of your key storage resources. The SVC addresses the increasing costs and complexity in data storage management by shifting the storage management intelligence from individual SAN controllers into the network by using virtualization.

SVC falls under Tier 4 for FlashCopy and Tier 6 for Metro Mirror.

Further information can be found Chapter 18, "SAN Volume Controller" on page 365.

2.8 Continuous availability solutions (Tier 7)

This section briefly describes continuous availability solutions in the following environments:

- ▶ Continuous availability for IBM System zSeries (GDPS)
- ▶ HACMP™/XD

Tier 7 solutions are distinguished by their built-in automation capabilities.

2.8.1 Geographically Dispersed Parallel Sysplex (GDPS)

GDPS is a family of IBM Global Services offerings for a single or multi-site environment, which provides an integrated, end-to-end solution for enterprise IT Business Continuity, integrating software automation, servers, storage, and networking.

GDPS automation provides the capability to manage the remote copy configuration and storage subsystems, automate System z operational tasks, manage and automate planned reconfigurations, and do failure recovery from a single point of control. GDPS offerings include Tier 7 and Tier 6 recovery capability.

The GDPS family of System z Business Continuity solutions consists of two major offering categories, with sub-offerings in each category. They are:

- ▶ GDPS/PPRC solutions, based on IBM System Storage Metro Mirror (formerly PPRC)-, including:
 - GDPS/PPRC, a Tier 7 solution
 - GDPS/PPRC HyperSwap Manager, a Tier 6 solution
 - RCMF/PPRC, a remote copy management solution for PPRC
- ▶ GDPS/XRC solutions, based on System Storage z/OS Global Mirror -(formerly XRC)-, including:
 - GDPS/XRC, a Tier 7 solution
 - RCMF/XRC, a remote copy management solution for XRC

GDPS/PPRC overview

GDPS/PPRC is designed as a *continuous availability and disaster recovery solution*. Metro Mirror (PPRC) hardware disk mirroring synchronously mirrors data that reside on a set of disk volumes, called the primary volumes, to secondary disk volumes in a second system.

The physical topology of GDPS/PPRC, shown in Figure 2-11, consists of a System z base or Parallel Sysplex cluster spread across two sites that are separated by up to 100 km/62 miles of fiber – with one or more z/OS systems at each site.

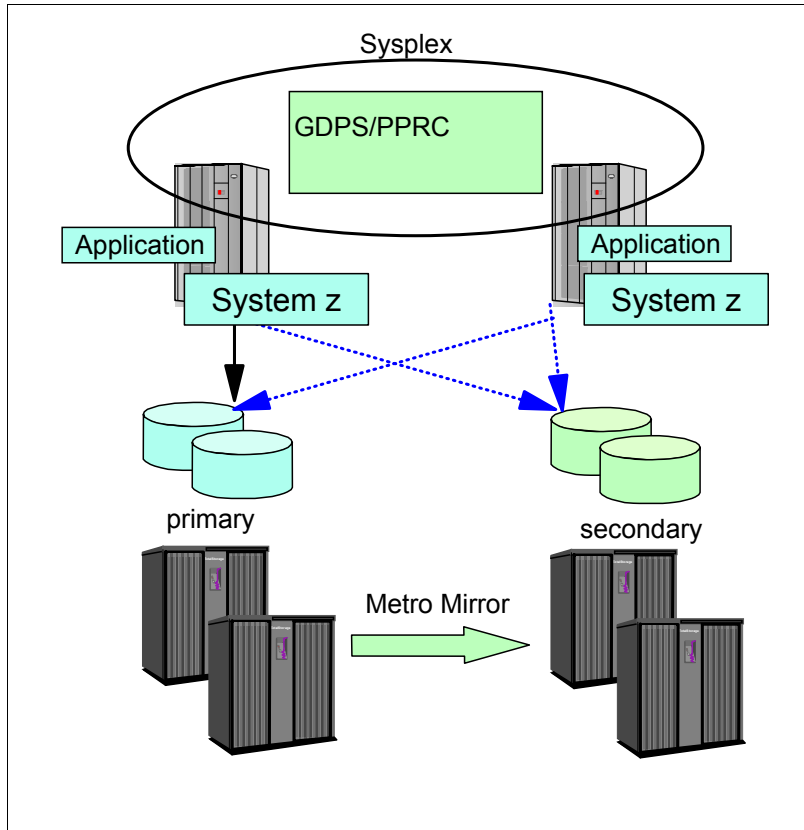


Figure 2-11 GDPS/PPRC topology

With GDPS/PPRC, you can perform a controlled site switch for both planned and unplanned site outages, with no or minimal data loss, maintaining full data integrity across multiple volumes and storage subsystems and the ability to perform a normal DBMS restart – not DBMS recovery – in the second site. GDPS/PPRC is application independent, and therefore can cover the complete application environment.

Near continuous availability of data with HyperSwap

GDPS in the Metro Mirror environment provides the HyperSwap functionality. The HyperSwap function can help significantly reduce the time needed to switch to the secondary set of disks while keeping the z/OS systems active, together with their applications. In this case, HyperSwap broadens the near continuous availability attributes of GDPS/PPRC by extending the Parallel Sysplex redundancy to disk subsystems.

GDPS/PPRC has a major sub-offering named GDPS/PPRC HyperSwap Manager, described in “GDPS/PPRC HyperSwap Manager overview” on page 29. This product extends Parallel Sysplex availability to disk subsystems, even if multiple sites are not available and the Parallel Sysplex is configured in only one site.

GDPS/PPRC management of zSeries operating systems

In addition to managing images within the base or Parallel Sysplex cluster, GDPS can manage a client's other zSeries and System z production operating systems; including z/OS, Linux® for zSeries, z/VM®, and VSE/ESA™. For example, if the volumes associated with the Linux for zSeries images are mirrored using Metro Mirror, GDPS can restart these images as part of a planned or unplanned site reconfiguration. The Linux for zSeries images can either run as a logical partition (LPAR) or as a guest under z/VM.

GDPS/PPRC management for open systems LUNs (Logical Unit Number)

GDPS/PPRC technology has been extended to manage a heterogeneous environment of z/OS and Open Systems data. If installations share their disk subsystems between the z/OS and Open Systems platforms, GDPS/PPRC can manage a common Metro Mirror Consistency Group (described in 2.7.1, “Data consistency for database environments” on page 23) for both System z and open systems storage, thus providing data consistency across both z/OS and Open Systems data. This allows GDPS to be a single point of control to manage business resiliency across multiple tiers in the infrastructure, improving cross-platform system management and business processes.

GDPS/PPRC Multi-Platform Resiliency for zSeries

GDPS/PPRC has been enhanced to provide a new function called *GDPS/PPRC Multi-Platform Resiliency for zSeries*. This function is especially valuable for clients who share data and storage subsystems between z/OS and z/VM Linux guests on zSeries, for example, an application server running on Linux on zSeries and a database server running on z/OS. GDPS/PPRC provides the reconfiguration capabilities for the Linux on zSeries servers and data in the same manner as for z/OS systems and data.

In summary, GDPS/PPRC is capable of the following attributes:

- ▶ Near continuous availability solution for z/OS
- ▶ Near transparent business continuity solution for z/OS
- ▶ Common point of control for recovery of a mixed zSeries and Open Systems business continuity environment
- ▶ Recovery Time Objective (RTO) less than an hour
- ▶ Recovery Point Objective (RPO) of zero (optional)
- ▶ Protects against localized area disasters (distance between sites limited to 100 km fiber)

GDPS/PPRC HyperSwap Manager overview

The GDPS/PPRC HyperSwap Manager solution is a subset of the full GDPS/PPRC solution, designed to provide a affordable entry point to the full family of GDPS/PPRC offerings, by providing a Rapid Data Recovery solution for enterprise disk-resident data.

GDPS/PPRC HyperSwap Manager (GDPS/PPRC HM) does this by offering a subset of the full GDPS/PPRC, specifically, the HyperSwap management and Metro Mirror management capabilities.

GDPS/PPRC HyperSwap Manager can provide either of the following two configurations:

1. Near continuous availability of data within a single site

GDPS/PPRC HyperSwap Manager is designed to provide Continuous Availability of data by masking disk outages that are caused by disk maintenance or failures. For example, if normal processing is suddenly interrupted when one of the disk subsystems experiences a hard failure, thanks to GDPS, the applications are masked from this error because GDPS detects the failure and autonomically invokes HyperSwap. The production systems continue to use data from the mirrored secondary volumes. Disk maintenance can also be similarly performed without application impact by executing the HyperSwap command.

2. Near continuous availability of data and D/R solution at metro distances

In addition to the single site capabilities, in a two site configuration, GDPS/PPRC HyperSwap Manager provides an entry-level disaster recovery capability at the recovery site, including the ability to provide a consistent copy of data at the recovery site from which production applications can be restarted. The ability to simply restart applications helps eliminate the need for lengthy database recovery actions.

The GDPS HyperSwap Manager offering features specially priced, limited function Tivoli System Automation and NetView® software pricing, thus enabling a more affordable entry point into the full GDPS automation software.

GDPS/PPRC HyperSwap Manager can be upgraded to full GDPS/PPRC at a later time, preserving the implementation investment and GDPS skills and procedures that have already been developed.

GDPS/XRC overview

GDPS/XRC has the attributes of a *disaster recovery* solution. z/OS Global Mirror (XRC) is a combined hardware-asynchronous and software-asynchronous remote copy solution. The application I/O is signaled completed when the data update to the primary storage is completed. Subsequently, a DFSMSdftp™ component called System Data Mover (SDM) that is typically running in the recovery site, asynchronously offloads data from the primary storage subsystem's cache and updates the secondary disk volumes.

GDPS/XRC has the following attributes:

- ▶ Disaster recovery solution
- ▶ RTO between an hour to two hours
- ▶ RPO less than two minutes, typically 3-5 seconds
- ▶ Protects against localized as well as regional disasters (distance between sites is unlimited)
- ▶ Minimal remote copy performance impact

GDPS/XRC topology

The physical topology of a GDPS/XRC configuration, as shown in Figure 2-12, consists of production systems in one site. The production systems could be a single system, multiple systems that are sharing a disk, or a base or Parallel Sysplex cluster. The recovery site can be located at a virtually unlimited distance from the production site.

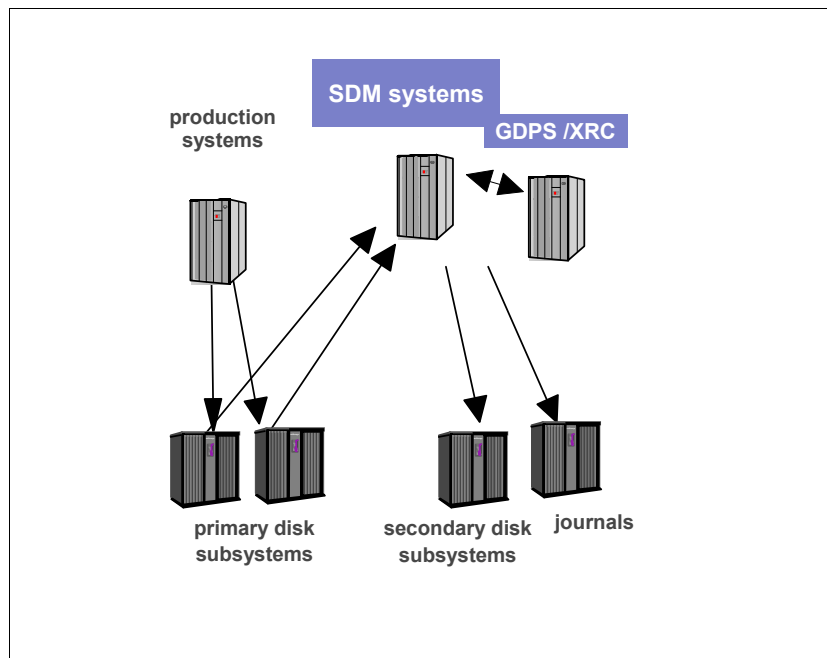


Figure 2-12 GDPS/XRC topology

In GDPS/XRC, the production systems located in the production site can be a single system, multiple systems that are sharing disks, or a base or Parallel Sysplex cluster. GDPS/XRC

provides a single, automated solution designed to dynamically manage storage subsystem mirroring (disk and tape) that allows a business to attain “near transparent” Disaster Recovery with minimal data loss. With GDPS/XRC, you can perform a controlled site switch for an unplanned site outage, and maintain data integrity across multiple volumes and storage subsystems. You then only need to perform a normal DBMS restart in the recovery site, rather than a DBMS recovery. GDPS/XRC is application independent and therefore can cover the client's complete application environment.

Additional GDPS information

For additional information about GDPS solutions or GDPS solution components, refer to these Web sites:

- ▶ GDPS home page:
<http://www.ibm.com/servers/eserver/zseries/gdps/>
- ▶ zSeries Business Resiliency Web site:
<http://www.ibm.com/systems/z/resiliency/>

Also, refer to these IBM manuals and Redbooks:

- ▶ *IBM eServer zSeries: Device Support Facilities User's Guide and Reference R17, GC35-0033*
- ▶ *IBM eServer zSeries: z/OS V1R6.0 DFSMS Advanced Copy Services, SC35-0428*
- ▶ *GDPS Family - An Introduction to Concepts and Capabilities, SG24-6374*

2.8.2 HACMP/XD for AIX

HACMP/XD for AIX solution integrates High-Availability Cluster Multiprocessing (HACMP) on IBM eSystem p™ systems, with the IBM System Storage DS™ Family of disk storage systems, IBM SAN Volume Controller, and Metro Mirror (PPRC) functionality to support high availability and disaster recovery.

This is a Tier 7 solution when HACMP Extended Distance (XD) is used in combination with the Metro Mirror feature. It can be considered an automated Tier 4 solution when HACMP/XD is used in combination with Global Copy.

HACMP helps to provide rapid recovery of application services by automatically moving a workload running on a host server to a recovery server after a failure. HACMP Extended Distance (XD), an optional HACMP feature that helps manage data mirroring, allows HACMP to automate disk-failover management. In a single-site HACMP environment, all cluster nodes sharing volume groups have physical connections to the same disk set. In an HACMP/XD environment, the cluster nodes access the same shared volume groups, but the nodes at each site access the volume groups from different storage systems, shown in Figure 2-13.

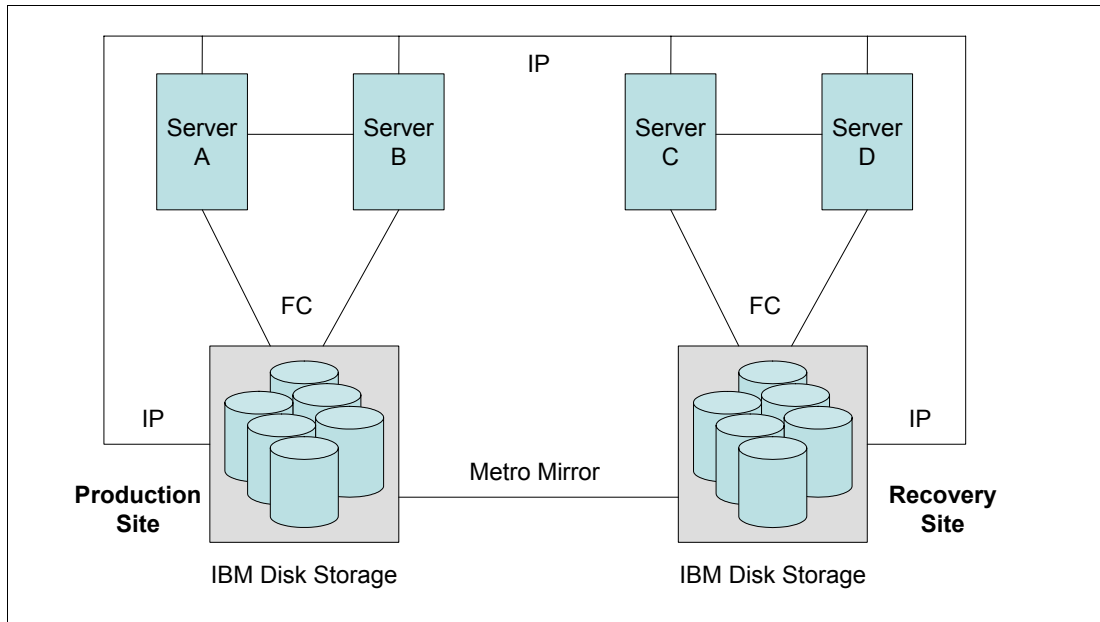


Figure 2-13 HACMP/XD

When the application is active on a primary-site server, HACMP/XD replicates application data updates to the disk system at a secondary site. HACMP/XD supports Metro Mirror functionality on the IBM ESS, DS8000, DS6000, or SVC disk systems. Metro Mirror constantly maintains an up-to-date copy of the primary location data at a remote site within the metropolitan area (typically up to 300 km away using dense wavelength division multiplexing). During a planned or unplanned outage, the solution can automatically switch workload to a remote backup system and then quickly restart your critical applications. Because mirrored data is a time-consistent image of the original data, this can help avoid a long and complicated data recovery process before restoring business operations.

A typical customer environment (Figure 2-13) running Metro Mirror and HACMP/XD would have a four-node, wide-area high availability cluster consisting of two local servers, server 1 and server 2, that are sharing the primary IBM disk or SVC at the primary site. Two remote servers, server 3 and server 4, share the secondary IBM storage at the recovery site. Server 1 is the primary server running an application; server 2 is the primary backup (at the primary site). Servers 3 and 4 are configured as backup servers at the secondary site. The application data is stored on the shared primary IBM diskstorage array. The shared application data is replicated through Metro Mirror over Fibre Channel links to the remote disk system at the secondary site. If cluster server 1 fails and server 2 remains healthy, HACMP can migrate the applications to server 2. All application data updates on server 2 can continue to be replicated from the primary to the secondary disk system.

If server 2 fails or if a local disaster simultaneously disables 1 and 2, HACMP automatically migrates the applications to server 3, typically with minimal service interruption to the application users. HACMP/XD automatically manages the failover to the secondary disk at this time without user intervention. It is also manages the clustered environment, starting the application at the remote site using the mirrored data. Server 3 can start accessing the data at the secondary storage system. When the primary systems become operational, HACMP has the ability to automatically reverse the data replication from the secondary disk to the primary disk with little operations impact.

Additional information

You can find further information at:

http://www.ibm.com/servers/storage/solutions/business_continuity/index.html

For further details on HACMP/XD with PPRC, we suggest the *HACMP/XD: ESS PPRC Planning and Administration Guide*, SC23-4863.

2.9 More information

For a more detailed understanding of Business Continuity Solutions, refer to these Redbooks and Redpapers:

- ▶ *IBM System Storage Business Continuity: Part 1 Planning Guide*, SG24-6547
- ▶ *IBM System Storage Business Continuity: Part 2 Solutions Guide*, SG24-6548
- ▶ *IBM System Storage Business Continuity Solutions Overview*, SG24-6684
- ▶ *Disaster Recovery Strategies with Tivoli Storage Management*, SG24-6844
- ▶ *IBM System Storage Business Continuity Solution Selection Methodology*, REDP-4062
- ▶ *IBM System Storage: Planning for Heterogeneous IT Business Continuity*, REDP-4063



Information lifecycle management

Information is essential to any business. Organizations have the challenge to efficiently manage information throughout its lifecycle, related to its business value. The quantity of information and its value changes over time, and becomes increasingly costly and complex to store and manage.

Explosive data growth coupled with years of decentralized IT management practices have allowed customers' storage environments to grow out of control. They have evolved into expensive, complex systems with fragmented data and perhaps outdated management processes.

In today's IT environments:

- ▶ Storage now typically accounts in excess of 15% of total IT budgets.
- ▶ Data growth is now estimated at over 50% annually - many very large companies have close to 150 TB of storage at the time of writing and industries such as health care and life sciences are growing their data at one TB a day.
- ▶ And, in most cases, disk utilization is under 50%, with up to 40% of data being redundant.

This chapter discusses the importance of information lifecycle management (ILM), its benefits, and introduces you to the elements of data lifecycle management.

3.1 Overview

Information lifecycle management (ILM) is a process for managing information through its lifecycle, from creation until disposal, in a manner that optimizes storage and access at the lowest cost. ILM consists of the policies, processes, practices, and tools used to align the business value of information with the most cost-effective IT infrastructure from the time information is conceived through its final disposition. Information is aligned with business processes through management of service levels associated with applications, metadata, information, and data.

ILM is not just hardware or software—it includes processes and policies to manage the information. It is designed upon the recognition that different types of information can have different values at different points in their lifecycle. Predicting storage needs and controlling costs can be especially challenging as the business grows.

The overall objectives of managing information with information lifecycle management are to help reduce the total cost of ownership (TCO) and help implement data retention and compliance policies. In order to effectively implement ILM, owners of the data need to determine how information is created, how it ages, how it is modified, and if/when it can safely be deleted. ILM segments data according to value, which can help create an economical balance and sustainable strategy to align storage costs with businesses objectives and information value. The adoption of ILM technologies and processes, as shown in Figure 3-1, turns that strategy into a business reality.

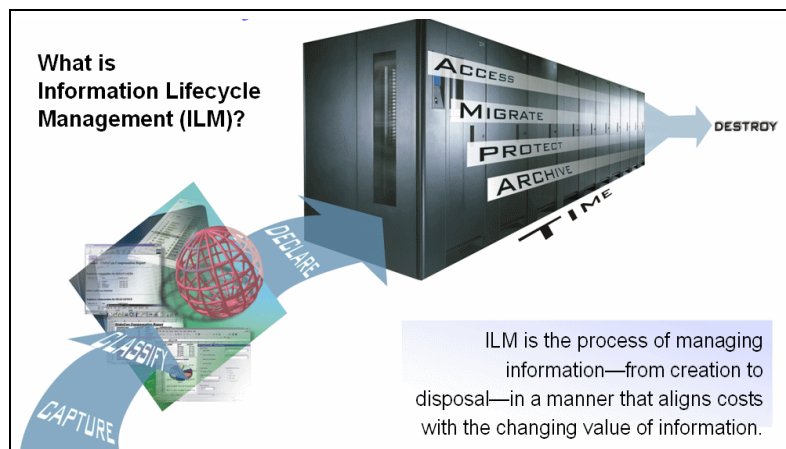


Figure 3-1 Information lifecycle management

3.2 Standards and organizations

The success and adoption of any new technology, and any improvement to existing technology, are greatly influenced by standards. Standards are the basis for the interoperability of hardware and software from different, and often rival, vendors. Although standards bodies and organizations such as the Internet Engineering Task Force (IETF), American National Standards Institute (ANSI), and International Organization for Standardization (ISO) publish these formal standards, other organizations and industry associations, such as the Storage Networking Industry Association (SNIA), play a significant role in defining the standards and market development and direction.

3.2.1 Storage Networking Industry Association (SNIA)

The Storage Networking Industry Association is an international computer system industry forum of developers, integrators, and IT professionals who evolve and promote storage networking technology and solutions. SNIA was formed to ensure that storage networks become efficient, complete, and trusted solutions across the IT community. IBM is one of the founding members of this organization. SNIA is uniquely committed to networking solutions into a broader market. SNIA is using its Storage Management Initiative (SMI) and its Storage Management Initiative Specification (SMI-S) to create and promote adoption of a highly functional interoperable management interface for multivendor storage networking products. SMI-S makes multivendor storage networks simpler to implement and easier to manage. IBM has led the industry in not only supporting the SMI-S initiative, but also using it across its hardware and software product lines. The specification covers fundamental operations of communications between management console clients and devices, auto-discovery, access, security, the ability to provision volumes and disk resources, LUN mapping and masking, and other management operations.

Data Management Forum

SNIA has formed the Data Management Forum (DMF) to focus on defining, implementing, qualifying, and teaching improved methods for the protection, retention, and lifecycle management of data.

Vision for ILM by SNIA and DMF

The Data Management Forum defines ILM as a new management practice for the datacenter. ILM is not a specific product, nor is it just about storage and data movement to low-cost disk. It is a standards-based approach to automating datacenter operations by using business requirements, business processes, and the value of information to set policies and service level objectives for how the supporting storage, compute, and network infrastructure operates.

The key question that flows from this *vision* of ILM is *How do we get there?*, because these capabilities do not fully exist today. This is the work of SNIA and the Data Management Forum: To unify the industry towards a common goal, to develop the relevant standards, to facilitate interoperability, and to conduct market education around ILM. Figure 3-2 illustrates the SNIA vision for ILM.

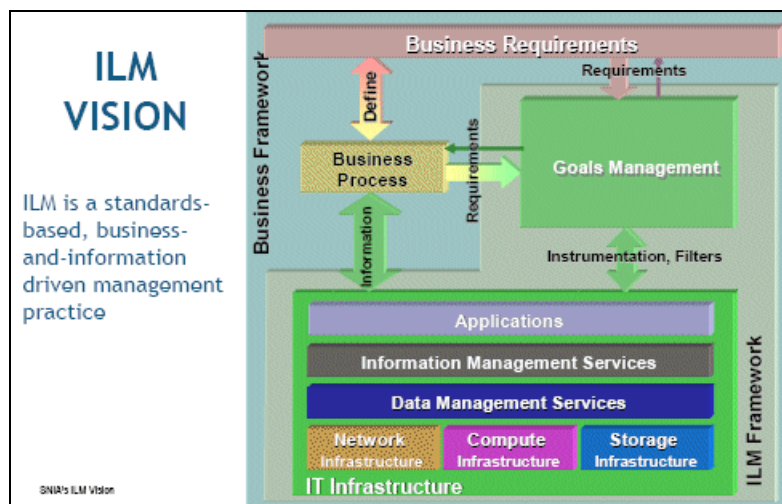


Figure 3-2 SNIA vision for ILM

For additional information about the various activities of SNIA and DMF see its Web site at:
<http://www.snia.org>

3.3 Why ILM is needed

Business, in order to run efficiently, need fast access to the stored data. But today's business environment faces increasing challenges: The explosion of the sheer volume of digital information, the increasing cost of storage management, tight regulatory requirements for data retention, and manual business and IT processes that are increasingly complex and error prone. Although the total value of stored information has increased overall, historically, not all data is created equal, and the value of that data to business operations fluctuates over time. This is commonly referred to as the *data lifecycle*. The existence of the data lifecycle means that all data cannot be treated the same.

Most frequently, the value of data decreases over time, for example, a critical e-mail which you receive today, once replied to and acted upon, becomes less valuable. Business application code actually becomes more valuable as it moves through the development cycle to deployment, then becomes less important as the applications are replaced or phased out. The rate of decline is different for different types of data. However, infrequently accessed or inactive data can become suddenly valuable again as events occur, or as new business initiatives or projects are taken on. Historically, the need to retain information has resulted in a "buy more storage" mentality. However, this approach has only served to increase overall storage management costs and complexity, and has increased the demand for hard-to-find qualified personnel.

Executives today are tasked with reducing overall spending while supporting an ever-increasing number of service and application demands. While support and management tasks increase, IT departments are being asked to justify their position by demonstrating business value to the enterprise. IT must also develop and enhance the infrastructure in order to support business initiatives while facing some or all of these data storage issues:

- ▶ Information and data growing faster than the storage budget.
- ▶ What data can I delete and when? What to keep and for how long?
- ▶ Disk dedicated to specific applications - inhibits sharing.
- ▶ Duplicated copies of files and other data. Where are they and how much space do they use?
- ▶ No mapping of the value of the data to the value of the hardware on which it is stored.
- ▶ Longer time required to backup data, but the window keeps shrinking.
- ▶ Storage performance does not meet requirements.
- ▶ Manual processes causing potential business risk due to errors.
- ▶ Regulatory requirements dictate long-term retention for certain data.
- ▶ Inability to achieve backup/recovery/accessibility objectives for critical data.
- ▶ Inability to grow the support staff to keep up with the demand for storage management in an increasingly complex environment.
- ▶ Multiple backup and restore approaches and processes.
- ▶ Storage management requirements not well defined.

In response to these, it is necessary to define specific objectives to support and improve information management:

- ▶ Control demand for storage and create policies for allocation.
- ▶ Reduce hardware, software, and storage personnel costs.
- ▶ Improve personnel efficiency, optimizing system and productivity.
- ▶ Define and enforce policies to manage the lifecycle of data.
- ▶ Define and implement the appropriate storage strategy to address current and future business requirements.

3.4 ILM elements

To manage the data lifecycle and make your business ready for on demand, there are four main elements that can address your business to an ILM structured environment, as shown in Figure 3-3. They are: Tiered storage management, Long-term data retention, Data lifecycle management and Policy-based archive management.

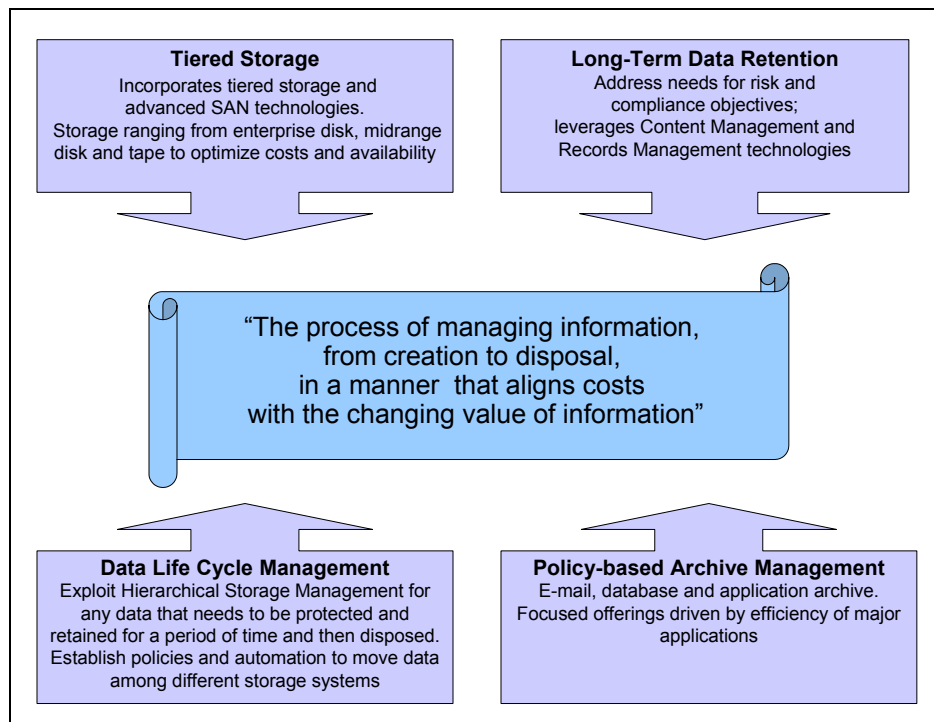


Figure 3-3 ILM elements

3.4.1 Tiered storage management

Most organizations today seek a storage solution that can help them manage data more efficiently. They want to reduce the costs of storing large and growing amounts of data and files and maintain business continuity. Through tiered storage, you can reduce overall disk-storage costs, by providing benefits like:

- ▶ Reducing overall disk-storage costs by allocating the most recent and most critical business data to higher performance disk storage, while moving older and less critical business data to lower cost disk storage.
- ▶ Speeding business processes by providing high-performance access to most recent and most frequently accessed data.

- ▶ Reducing administrative tasks and human errors. Older data can be moved to lower cost disk storage automatically and transparently.

Typical storage environment

Storage environments typically have multiple tiers of *data value*, such as application data that is needed daily and archive data that is accessed infrequently. But typical storage configurations offer only a single tier of storage, as in Figure 3-4, which limits the ability to optimize cost and performance.

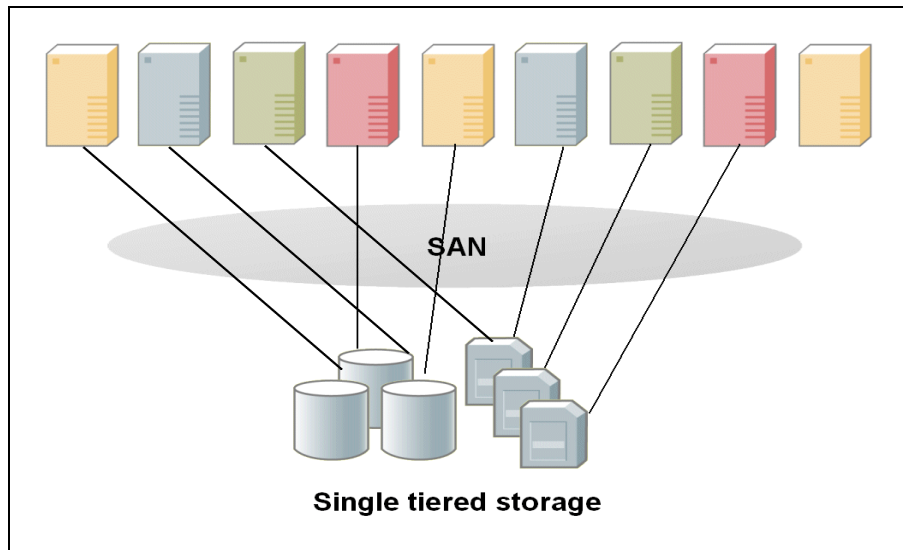


Figure 3-4 Traditional non-tiered storage environment

Multi-tiered storage environment

A tiered storage environment is the infrastructure necessary to align storage cost with the changing value of information. The tiers will be related to data value. The most critical data is allocated to higher performance disk storage, while less critical business data is allocated to lower cost disk storage. Each storage tier will provide different performance metrics and disaster recovery capabilities. Creating classes and storage device groups is an important step to configure a tiered storage ILM environment.

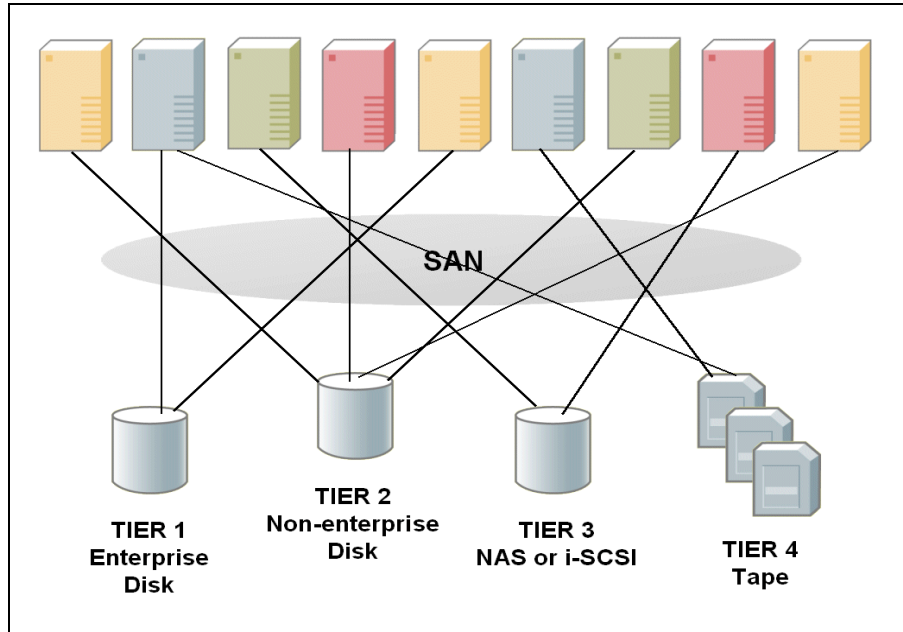


Figure 3-5 Multi-tiered storage environment

An IBM ILM solution in a tiered storage environment is designed to:

- ▶ Reduce the total cost of ownership (TCO) of managing information. It can help optimize data costs and management, freeing expensive disk storage for the most valuable information.
- ▶ Segment data according to value. This can help create an economical balance and sustainable strategy to align storage costs with business objectives and information value.
- ▶ Help make decisions about moving, retaining, and deleting data, because ILM solutions are closely tied to applications.
- ▶ Manage information and determine how it should be managed based on content, rather than migrating data based on technical specifications. This approach can help result in more responsive management, and offers you the ability to retain or delete information in accordance with business rules.
- ▶ Provide the framework for a comprehensive enterprise content management strategy.

Key products of IBM for tiered storage solutions and storage virtualization solutions are:

- ▶ IBM System Storage SAN Volume Controller (SVC)
- ▶ IBM System Storage DS family of disk storage - DS4000, DS6000, and DS8000
- ▶ IBM System Storage N series (NAS)
- ▶ IBM System Storage tape drives, tape libraries, and virtual tape solutions

3.4.2 Long-term data retention

There is a rapidly growing class of data that is best described by the way in which it is managed rather than the arrangement of its bits. The most important attribute of this kind of data is its retention period, hence it is called *retention-managed data*, and it is typically kept in an archive or a repository. In the past it has been variously known as *archive data*, *fixed content data*, *reference data*, *unstructured data*, and other terms implying its read-only nature. It is often measured in terabytes and is kept for long periods of time, sometimes forever.

In addition to the sheer growth of data, laws and regulations governing the storage and secure retention of business and client information are increasingly becoming part of the business landscape, making data retention a major challenge to any institution. An example of these is the Sarbanes-Oxley Act in the US, of 2002, and there are many others.

Businesses must comply with these laws and regulations. Regulated information can include e-mail, instant messages, business transactions, accounting records, contracts, or insurance claims processing, all of which can have different retention periods, for example, for 2 years, for 7 years, or retained forever. Moreover, some data must be kept just long enough and no longer. Indeed, content is an asset when it needs to be kept; however, data kept past its mandated retention period could also become a liability. Furthermore, the retention period can change due to factors such as litigation. All these factors mandate tight coordination and the need for ILM.

IBM ILM data retention strategy

Regulations and other business imperatives, stress the need for an Information Lifecycle Management process and tools to be in place. The unique experience of IBM with the broad range of ILM technologies, and its broad portfolio of offerings and solutions, can help businesses address this particular need and provide them with the best solutions to manage their information throughout its lifecycle. IBM provides a comprehensive and open set of solutions to help.

IBM has products that provide content management, data retention management, and sophisticated storage management, along with the storage systems to house the data. Key products of IBM for data retention and compliance solutions are:

- ▶ IBM Tivoli Storage Manager, including IBM System Storage Archive Manager
- ▶ IBM DB2 Content Manager Family, which includes DB2 Content Manager, Content Manager OnDemand, CommonStore for Exchange Server, CommonStore for Lotus® Domino®, and CommonStore for SAP
- ▶ IBM DB2 Records Manager
- ▶ IBM System Storage DS4000 with SATA disks
- ▶ IBM System Storage DR550
- ▶ IBM TotalStorage and System Storage tape (including WORM) products

Important: The IBM offerings are intended to help clients address the numerous and complex issues relating to data retention in regulated and unregulated business environments. Nevertheless, each client's situation is unique, and laws, regulations, and business considerations impacting data retention policies and practices are constantly evolving. Clients remain responsible for ensuring that their information technology systems and data retention practices comply with applicable laws and regulations, and IBM encourages clients to seek appropriate legal counsel to ensure their compliance with those requirements. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law.

3.4.3 Data lifecycle management

At its core, the process of ILM moves data up and down a path of tiered storage resources, including high-performance, high-capacity disk arrays, lower-cost disk arrays such as serial ATA (SATA), tape libraries, and permanent archival media where appropriate. Yet ILM involves more than just data movement; it encompasses scheduled deletion and regulatory compliance as well. Because decisions about moving, retaining, and deleting data are closely tied to application use of data, ILM solutions are usually closely tied to applications.

ILM has the potential to provide the framework for a comprehensive information-management strategy, and helps ensure that information is stored on the most cost-effective media. This helps enable administrators to make use of tiered and virtual storage, as well as process automation. By migrating unused data off of more costly, high-performance disks, ILM is designed to help:

- ▶ Reduce costs to manage and retain data.
- ▶ Improve application performance.
- ▶ Reduce backup windows and ease system upgrades.
- ▶ Streamline data management.
- ▶ Allow the enterprise to respond to demand—in real-time.
- ▶ Support a sustainable storage management strategy.
- ▶ Scale as the business grows.

ILM is designed to recognize that different types of information can have different value at different points in their lifecycle. As shown in Figure 3-6, data can be allocated to a specific storage level aligned to its cost, with policies defining when and where data will be moved.

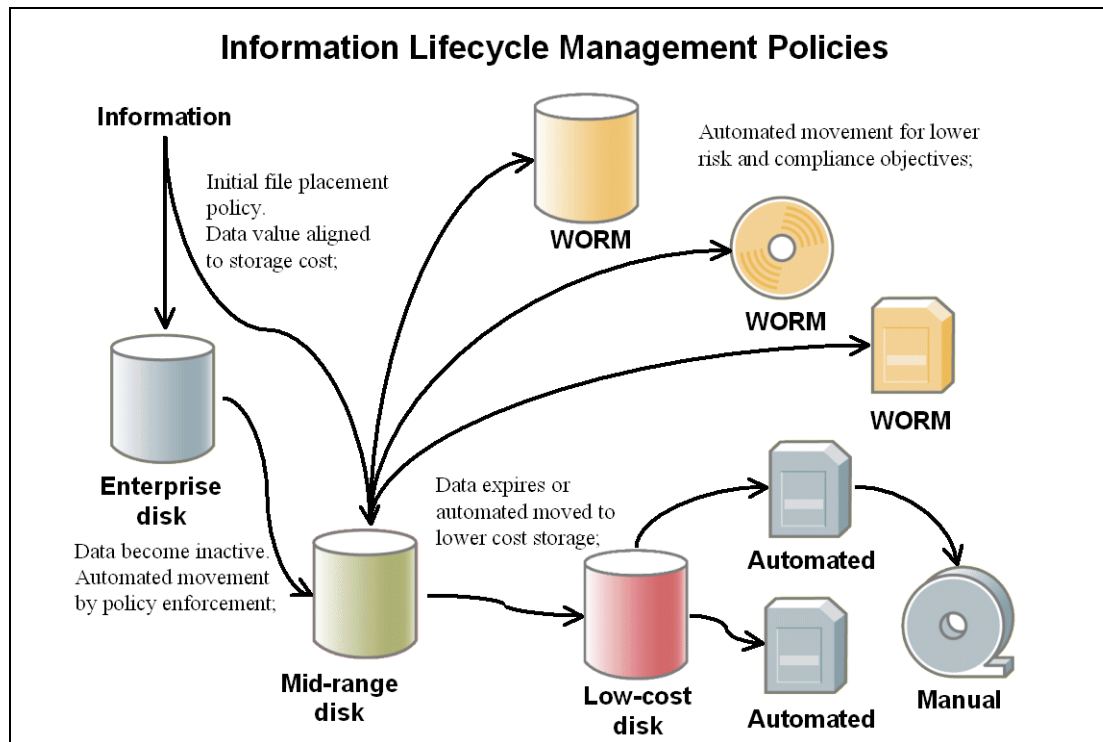


Figure 3-6 ILM policies

The value of a piece of information may change sometimes, and data that was previously inactive, and was migrated to a lower-cost storage, now could be needed and should be processed in a high-performance disk. A data lifecycle management policy can be defined to move the information back to enterprise storage, making the storage cost aligned to data value, as illustrated in Figure 3-7 on page 44.

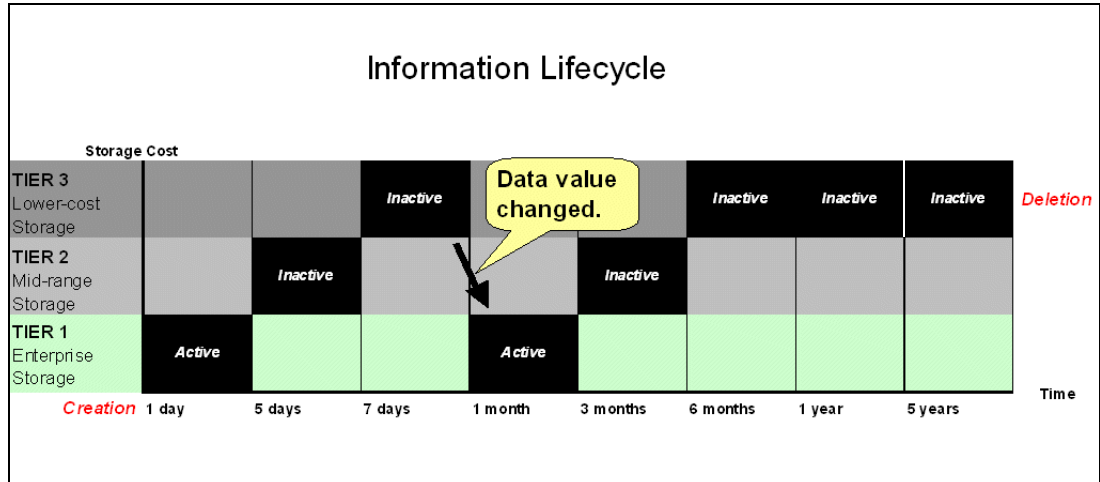


Figure 3-7 Information value changes

Key products of IBM for lifecycle management are:

- ▶ IBM TotalStorage Productivity Center
- ▶ IBM System Storage SAN Volume Controller (SVC)
- ▶ IBM Tivoli Storage Manager including IBM System Storage Archive Manager
- ▶ IBM Tivoli Storage Manager HSM

3.4.4 Policy-based archive management

As businesses of all sizes migrate to e-business solutions and a new way of doing business, they already have mountains of data and content that have been captured, stored, and distributed across the enterprise. This wealth of information provides a unique opportunity. By incorporating these assets into e-business solutions, and at the same time delivering newly generated information media to their employees and clients, a business can reduce costs and information redundancy and leverage the potential profit-making aspects of their information assets.

Growth of information in corporate databases such as Enterprise Resource Planning (ERP) systems and e-mail systems makes organizations think about moving unused data off the high-cost disks. They need to:

- ▶ Identify database data that is no longer being regularly accessed and move it to an archive where it remains available.
- ▶ Define and manage what to archive, when to archive, and how to archive from the mail system or database system to the back-end archive management system.

Database archive solutions can help improve performance for online databases, reduce backup times, and improve application upgrade times. E-mail archiving solutions are designed to reduce the size of corporate e-mail systems by moving e-mail attachments and/or messages to an archive from which they can easily be recovered if needed. This action helps reduce the need for end-user management of e-mail, improves the performance of e-mail systems, and supports the retention and deletion of e-mail, as shown in Figure 3-8 on page 45.

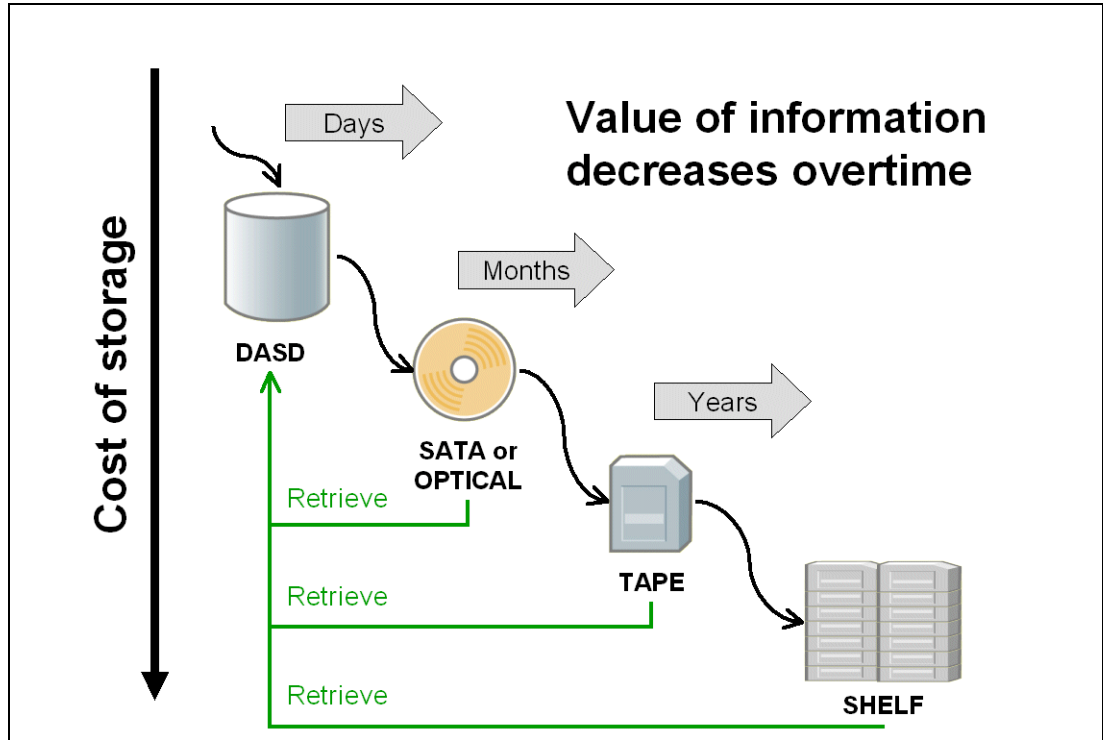


Figure 3-8 Value of information and archive/retrieve management

The way to do this is to migrate and store all information assets into an e-business-enabled content manager. ERP databases and e-mail solutions generate large volumes of information and data objects that can be stored in content management archives. An archive solution allows you to free system resources, while maintaining access to the stored objects for later reference. Allowing it to manage and migrate data objects gives a solution the ability to have ready access to newly created information that carries a higher value, while at the same time still being able to retrieve data that has been archived on less expensive media.

Key products of IBM for archive management are:

- ▶ IBM Tivoli Storage Manager including IBM System Storage Archive Manager
- ▶ IBM DB2 Content Manager family of products
- ▶ IBM DB2 CommonStore family of products

3.5 IBM ILM solutions

IBM has experience with the broad range of ILM technologies, its relationships with leading ISVs, and its broad portfolio of offerings and solutions - including offerings in IBM System Storage disk and tape systems, IBM System Storage Open Software, and IBM DB2 Content Management software - can help your business select the best solutions to manage your information throughout its lifecycle.

3.5.1 Application and database archive solutions

Over time, databases, such as enterprise resource planning (ERP) and customer relationship management (CRM) systems, can start to fill with older information that is used infrequently. Some of this information might be retained only for reference purposes. This information can

needlessly increase demands on high-performance disk systems, slow application performance, and increase the time required to back up databases and upgrade applications.

Application and database archive solutions are designed to identify database data that are no longer being accessed regularly and move the data to an archive, where it remains available if needed. Because applications can span multiple database systems and even multiple platforms, select application and database archive solutions can recognize related data from multiple systems and archive it together.

Key offerings

- ▶ IBM DB2 CommonStore for SAP
IBM DB2 CommonStore for SAP helps you archive information from operational SAP databases and related files. It supports any SAP operational database, including IBM DB2 Universal Database™, IBM Informix® and Oracle. The offering is SAP AG certified for all current releases: SAP R/3 Enterprise, mySAP.com, and SAP NetWeaver framework.
- ▶ Princeton Softech Active Archive Solutions software
Princeton Softech Active Archive Solutions software, working together with IBM System Storage systems, enables you to selectively archive and remove rarely accessed data from a wide range of production databases running on UNIX, Windows, and IBM eServer™ System z platforms to alternative storage, including the IBM System Storage DR550 retention-managed storage offering.
- ▶ EASY Software Enterprise x
The EASY for mySAP application covers archiving scenarios for SAP users. EASY for mySAP, working with IBM System Storage systems, supports full SAP Archive Link functionality, and delivers a variety of additional functionality that can help improve business processes and efficiency. EASY ENTERPRISE.x complements the IBM System Storage DR550 system and can enable businesses to manage the lifecycle of data assets from creation to retirement.
- ▶ IBM Express Storage™ Platform
The IBM Express Storage Platform is designed to provide an affordable, flexible, and prepackaged storage offering to address the unique storage demands of small and medium-sized businesses and large departmental users.

3.5.2 E-mail archive solutions

For many companies, e-mail increasingly consumes storage. And dealing with e-mail poses a problem for many end users. IBM e-mail archive solutions are designed to provide comprehensive solutions for managing e-mail.

IBM offers e-mail archive solutions that can help reduce the size of e-mail systems, reduce management and improve performance. They can help you:

- ▶ Improve productivity by reducing the need for end-user management of e-mail
- ▶ Improve performance of e-mail systems
- ▶ Address compliance requirements for e-mail
- ▶ Reduce e-mail storage costs

Key offerings

- ▶ IBM DB2 CommonStore for Exchange Server and IBM DB2 CommonStore for Lotus Domino
IBM DB2 CommonStore for Exchange Server and IBM DB2 CommonStore for Lotus Domino are designed to manage e-mail archiving and retrieval. They can help trim the

size of your e-mail system to reduce storage costs, improve e-mail system performance and provide virtually unlimited mailbox space for each user.

- ▶ IBM System Storage DR550 system

The IBM System Storage DR550 system is designed to help businesses address the growing challenge of managing and securing retention managed data and other critical information assets with operational efficiency. The DR550 is an integral component of IBM extensive content and data retention offerings and helps companies become on demand businesses while responding with flexibility and speed to new regulatory requirements.

3.5.3 Data lifecycle management solutions

Storage management costs are a top concern for businesses. One way to save is to move less-frequently-accessed or less-critical data to lower-cost devices. Data lifecycle management solutions provide a way to identify such data and position it within a tiered-storage infrastructure.

Data lifecycle management can help eliminate unneeded points of management and reduce the number of physical devices in your infrastructure. It can help you:

- ▶ Control your data storage and costs
- ▶ Automatically position data correctly at creation time to meet anticipated needs
- ▶ Increase flexibility through variable-cost IT infrastructure
- ▶ Increase productivity through automated, policy-based capabilities

Key offerings

- ▶ IBM System Storage SAN Volume Controller (SVC)

The SAN Volume Controller (SVC) simplifies the storage infrastructure by enabling changes to the physical storage with minimal or no disruption to applications. This will be great helpful while you doing data migration from tiered storage environment.

- ▶ IBM Tivoli Storage Manager

Helps protect an organization's data from failures and other errors by storing backup, archive, space management and bare-metal restore data, as well as compliance and disaster-recovery data in a hierarchy of offline storage.

- ▶ TotalStorage Productivity Center for Data

IBM TotalStorage Productivity Center (TPC) for Data helps discover, monitor, and create enterprise policies for disks, storage volumes, file systems, files, and databases. Knowing where all your storage is located and the properties of your data places you in a better position to act intelligently on your data.

- ▶ IBM Express Storage Platform

The IBM Express Storage Platform is designed to provide an affordable, flexible, and prepackaged storage offering to address the unique storage demands of small and medium-sized businesses and large departmental users.

3.5.4 Enterprise content management solutions

IBM content management helps integrate and deliver critical business information on demand. IBM software and solutions support challenging information types, such as images, documents, e-mail, Web content, e-records, multimedia and computer report output.

IBM comprehensive content management portfolio can assist your organization in improving productivity, enhancing responsiveness and addressing with regulatory demands.

Key offerings

- ▶ IBM DB2 Content Manager

At the heart of the IBM content management portfolio, IBM DB2 Content Manager software is designed to store and manage a broad range of business content, including documents, e-mail, Web content and more. DB2 Content Manager can help you quickly and easily access information across diverse formats to respond to a customer or a partner's request and securely deliver relevant information.

- ▶ IBM System Storage DR550 system

IBM System Storage DR550 data system can help your business address the growing challenge of managing and securing retention managed data and other critical information assets with operational efficiency. The DR550 system is an integral component of IBM extensive content and data retention offerings. It can help you become an on demand business, responding with flexibility and speed to new regulatory requirements.

3.6 Services and consulting

ILM is a complex area with many variables to consider. It can be difficult to know just where to start. In general, businesses are recommended to implement a phased approach to ensure effectiveness and ROI.

IBM offers an extensive set of planning and solutions services to assist businesses in developing their ILM strategies, providing assessments, and helping businesses meet the challenges of regulatory compliance and efficient data management. These services range from a quick data assessment to comprehensive planning, architectural definition, and infrastructure implementation.

3.7 More information

For more information about IBM ILM offerings and capabilities, see these Web sites and publications:

- ▶ <http://www.ibm.com/servers/storage/solutions/ilm/index.html>
- ▶ http://www.ibm.com/servers/storage/services/solutions_ilm.html
- ▶ <http://www.ibm.com/software/info/contentmanagement>
- ▶ *ILM Library: Techniques with Tivoli Storage and IBM TotalStorage Products*, SG24-7030

Storage infrastructure simplification

The growth of IT, including applications and processes, means that many customers find themselves in an environment with many administrators, many servers, many different types of disk and tape storage, many networks, many and different processes by area, many communication devices, and many and different type of management systems. Day by day the infrastructure is growing and changing, placing more demands on the human and technology resources. Perhaps Figure 4-1 strikes a chord with you.

Infrastructure simplification is all about helping your business meet its objectives. The IT challenge is clear - simplify and optimize the infrastructure so that you can efficiently respond in a cost-effective way. IBM can provide the technology, business and industry expertise for simplifying the infrastructure environment in real business environments.

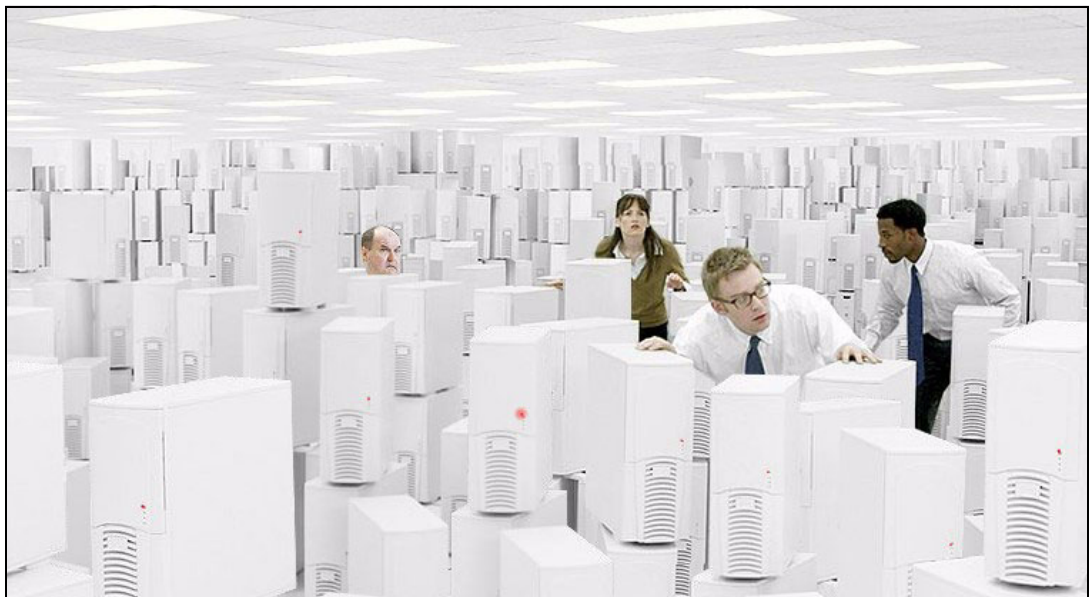


Figure 4-1 Heterogeneous and complex IT environment

4.1 Overview

This chapter covers products, services and offerings which provide infrastructure simplification with IBM System Storage.

IBM solutions for infrastructure simplification help improve efficiency; lower total cost of data ownership, storage, and management; and reduce time-consuming and costly errors.

IBM solutions for infrastructure simplification focus on the following key areas:

- ▶ Physical storage consolidation
- ▶ Virtual storage consolidation
- ▶ Automated data management
- ▶ Storage Services

Note: For a detailed explanation on Infrastructure Simplification, we recommend the redbook *Introduction to Storage Infrastructure Simplification*, SG24-7114.

4.2 Physical storage consolidation solutions

Physical storage consolidation eliminates the islands in your infrastructure. By consolidating heterogeneous open and enterprise storage in the same solution, you will save space, reduce administration costs, and increase availability, performance and quality of services. Physical storage consolidation should decrease the long-term costs associated with storing and managing data.

4.2.1 Key benefits

- ▶ Consolidation can help eliminate unneeded points of management and reduce the number of physical devices in the infrastructure.
- ▶ Centralize the physical storage environment with scale up technologies
- ▶ Consolidate data with scale out technologies
- ▶ Increase resource utilization to save on storage costs
- ▶ Simplify storage management to improve productivity
- ▶ Achieve economies of scale for acquiring and operating disk, tape and network resources

4.2.2 Key offerings

IBM offers consolidation products for disk, tape, and storage networking.

Disk storage consolidation

IBM System Storage DS family systems (DS4000, DS6000, and DS8000) either alone or using the IBM System Storage SAN Volume Controller, can potentially lower long-term costs and complexity. These disk storage systems can help improve management efficiencies and flexibility, reduce floor space, and provide point-in-time and mirroring performance and functional advantages. The DS family helps consolidate storage by providing wide ranging attachment to different platforms, and offering high density storage in a resource-efficient footprint.

Tape storage consolidation

IBM tape technology has high-density and high-performance media (LTO and 3592), as well as highly scalable libraries which can attach to and are sharable by multiple server platforms and applications. This helps simplify the number and type of tape devices which are required.

Storage area networking

IBM System Storage SAN products and solutions provide integrated SAN solutions with multiprotocol local, campus, metropolitan and global storage networking.

4.3 Virtualization solutions

Storage virtualization provides abstraction of physical storage devices so they can be accessed as a pool of logical resources. This offers the advantage of separating or buffering the physical storage of data and devices from the management and use of data. IBM has been a leader in offering the advantages of virtualization with more than twenty years experience, beginning with mainframe technology.

IBM virtualization solutions include:

- ▶ **Disk drive virtualization**
Disk firmware abstracts the physical cylinder, head and sector into a single logical block address. This feature is inherent in all disk drives.
- ▶ **Storage system partitioning**
The storage system creates multiple partitions from a single storage resource. IBM System Storage DS8000 systems provide this type of virtualization.
- ▶ **Block virtualization**
Software abstracts multiple disk arrays into a single storage resource. IBM System Storage SAN Volume Controller provides block virtualization.
- ▶ **Tape virtualization**
Software abstracts both tape drives and tape media onto a single disk resource. This type of virtualization is provided by the IBM TotalStorage 3494 Virtual Tape Server and IBM Virtualization Engine TS75100.

4.3.1 Key benefits

- ▶ Enable a nondisruptive, tiered storage environment
- ▶ Use one point of management for disk storage
- ▶ Streamline storage management by individually managing physical devices
- ▶ Reduce the cost and improve the flexibility of replication services
- ▶ Accelerate application deployment with improved resource sharing
- ▶ Reduce total cost of ownership
- ▶ Improve application availability

4.3.2 Key offerings

- ▶ **IBM System Storage SAN Volume Controller:** Increase the flexibility of the storage infrastructure. It combines storage capacity from multiple vendors into a single reservoir of capacity, that can be easily deployed and re-deployed to multiple servers. So-called “virtual disks” mask the underlying infrastructure and any changes to it, from the attached servers, and can be easily increased and moved to different hardware.

- ▶ **IBM TotalStorage 3494 Virtual Tape Server**
- ▶ **IBM Virtualization Engine TS75100:** These solutions emulate tape drives, libraries, and media using fast disk-based hardware. Using these solutions allows customers to seamlessly upgrade to faster technology, while preserving underlying application compatibility. Tape virtualization solutions also provide scalability and availability.

4.4 Automation solutions

Understanding the IT environment and managing it cost-effectively are essential to success in today's highly competitive, global economy, and automated management software can help. IBM storage management solutions help improve the efficiency of IT administration by eliminating manual sources of error and time-consuming repetitive tasks, while providing a continual understanding of the ways the infrastructure is being used. IBM storage management software helps show how storage is being consumed and how to manage storage more cost-effectively in the future. Advanced IBM automated management software can help simplify certain tasks, such as provisioning new storage, to a single step.

4.4.1 Key benefits

- ▶ Use continual monitoring to better understand, predict and manage your storage environment
- ▶ Change manual tasks to automated tasks
- ▶ Simplify the IT environment and reduce failures
- ▶ Improve administrator productivity
- ▶ Automatically deploy additional capacity, preventing out-of-resource conditions and improving application availability
- ▶ Reduce human errors during routine tasks
- ▶ Accommodate changes more easily without disruption
- ▶ Reduce user errors that can disrupt operations by automating routine, error-prone tasks
- ▶ Reduce management efforts and accomplish management tasks by using less-experienced staff, once automated management is established

4.4.2 Key offerings

IBM TotalStorage Productivity Center

IBM TotalStorage Productivity Center is an open solution which reduces the effort of managing complex storage infrastructures, improves storage capacity utilization and improves administrative efficiency. TPC helps simplify the management of traditional and virtualized SAN environments. It consists of:

- ▶ IBM TotalStorage Productivity Center for Data
- ▶ IBM TotalStorage Productivity Center for Fabric
- ▶ IBM TotalStorage Productivity Center for Disk
- ▶ IBM TotalStorage Productivity Center for Replication

IBM Tivoli Provisioning Manager

IBM Tivoli Provisioning Manager enables on demand computing across the entire datacenter through software, server, storage and network automation. IBM Tivoli Provisioning Manager automates the manual tasks of provisioning and configuring servers and virtual servers,

operating systems, middleware, applications, storage and network devices acting as routers, switches, firewalls and load balancers.

4.5 Storage services

In addition to a broad range of products from workgroup disk arrays to petabytes-capable file systems, IBM offers a wide range of services for every storage need. Some of these services include: data migration offerings to help move from older storage to modern IBM storage, storage cost and value analyses to understand the current storage environment and what capabilities it needs to grow into, and in depth onsite analysis of storage ownership costs. See this Web site for more information

http://www.ibm.com/servers/storage/services/solutions_is.html

4.5.1 Global Services Center for Infrastructure Services

Our infrastructure has been designed specifically to deliver applications at enhanced levels of security, performance and availability.

Highlights

- ▶ Speed - helps speed, deployment made possible by IBM security-rich reference builds and workflow automations
- ▶ Reliability - helps increase reliability, which is made possible, in part by a well-managed and security-enhanced infrastructure
- ▶ Economic Benefit - increases the potential to reduce cost through improved utilization of shared services

The infrastructure services include the data center, security, storage, systems, and monitoring and support through our Applications Management Center (AMC). The services can also optionally include disaster recovery and high availability.

4.5.2 Global Services Customer Solution Center Network

Planning and implementing in-house, multisite workstation rollouts and refreshes can be challenging - especially when geographically diverse locations are involved. Without a single point of responsibility, prudent resource management, careful planning and an experienced team, project deadlines can spin out of control - wasting valuable time and resources. Further, costs related to setting up, integrating, testing and eventually delivering a multisite workstation rollout can be excessive and unpredictable. The IBM Global Services Customer Solution Center Network can help by providing a single vendor to manage large-scale or small-scale integration and redistribution services needs, including receiving, storing, integrating, and shipping rollout assets, so that budgets and critical deadlines are met.

Highlights

- ▶ An integration and rollout solution specific to individual requirements
- ▶ Predictable costs and schedules
- ▶ A single source for hardware integration and rollout
- ▶ Integration, testing, shipping and redistribution for all elements of the rollout, whether manufactured by IBM or another provider

- ▶ Skilled technical and project administration professionals experienced in large and small rollouts, both domestically and internationally

4.5.3 IBM High Availability Services - for resilient infrastructure

IBM High Availability Services - for resilient infrastructure helps improve critical IT systems availability, moving even closer to a continuous operations environment by reducing unplanned outages and reducing or potentially eliminating planned outages.

Highlights

- ▶ Availability readiness review
- ▶ Comprehensive availability assessment
- ▶ Development/implementation of tailored solutions to meet your business objectives
- ▶ An assigned Availability Manager
- ▶ Service Level Agreements for availability guarantees

4.5.4 IT consolidation

IT consolidation is a logical first step in designing a more rational, efficient and resilient infrastructure that can support the on demand business model. IT Consolidation Services provides more than just the physical combining of assets.

Highlights

- ▶ Higher availability
- ▶ Opportunities for resource realignment
- ▶ Increased performance and simplified server environments
- ▶ Faster application deployments
- ▶ Improved cost avoidance processes
- ▶ Simplified server environments, including Linux systems

4.5.5 Implementation Services

Implementation Services are a few of the many services available. For more information, please contact your IBM representative.

Enterprise application integration services for Microsoft technologies

Enterprise application integration services can help realize the benefits of a tightly-aligned e-business infrastructure, supported by various Microsoft and other enterprise-based technologies - locally or globally. By leveraging existing frameworks, IBM can help integrate complex, disparate systems to create an enterprise-wide information flow solution that is designed to be flexible and cost-effective.

Highlights

- ▶ Analysis and strategy services to help plan and prioritize the effort and costs of IT initiatives to achieve business objectives
- ▶ Architecture and design services to assist in defining an effective strategy for using current and anticipated applications and Microsoft technologies
- ▶ Iterative design, build and test services, including development of a comprehensive plan to help ensure a managed, well-supported implementation
- ▶ Implementation services to help deploy the solution and required infrastructure into the production environment

4.5.6 Infrastructure Resource Management

Whether you are implementing an on demand initiative, IT consolidation or new technologies, IBM offers a variety of capabilities that can help optimize your resources.

Highlights

- ▶ Helps provide a complete solution for deployment, managing and supporting the technology infrastructure from procurement to disposal
- ▶ Provides services necessary from expert design and implementation of infrastructure resource management strategies, to a full range of out-tasking solutions that can augment your staff and complement operations
- ▶ Combines IBM world class skills, best practices, ITIL-aligned solutions, highly refined technologies and services, and industry-leading solution partners

4.5.7 On demand IT infrastructure strategy

An on demand business can be defined as one whose IT infrastructure delivers optimal business value, enhances competitive advantage, and accommodates evolving business goals and new market opportunities. In their efforts to become an effective on demand business, many companies share the same goals: to grow revenue, manage risk, improve productivity and enhance profitability. Working with a wide variety of clients, IBM has identified four highlights that are integral to meeting these goals:

- ▶ An on demand infrastructure that supports your on demand initiatives to help grow revenue
- ▶ A resilient infrastructure that is proactive rather than reactive to help manage and mitigate risk
- ▶ An operationally efficient infrastructure that helps optimize IT performance, reduce spending and drives cost savings
- ▶ An infrastructure that enables improved productivity through strategic technology adoption and resource management

4.5.8 Resilient business and infrastructure solutions

A resilient business infrastructure provides a security-rich, agile, available and recoverable environment that can handle planned and unplanned events, and positive and negative impacts to business. Survival is dependent upon the rapid response of both critical business processes and the supporting infrastructure.

Highlights

- ▶ Identifies business processes and elements that are critical to a resilient, agile enterprise
- ▶ Determines risks (stresses and demands) to the business
- ▶ Evaluates current business and technology infrastructure using a six layer framework
- ▶ Creates vulnerability, responsiveness and prevention indices
- ▶ Prepares next steps and recommendations for mitigating risks

4.5.9 Resilient business and infrastructure assessment

An important aspect of on demand infrastructure is resiliency - having the flexibility to respond rapidly in the event of changes and threats - whether they are computer viruses, earthquakes or sudden spikes in demand for IT resources.

Highlights

- ▶ Examines the many working layers required for an optimal infrastructure implementation, including strategy, organization, business and IT
- ▶ Compares your environment to other similar companies in your industry
- ▶ Provides a definition of potential threats or disruptions, prioritizes the level and impact of risk
- ▶ Recommends areas of the business that need to be addressed to meet business goals
- ▶ Incorporates business and IT components into a single, security-enhanced and competitive resiliency strategy

4.5.10 IT optimization solution

When moving to an on demand environment, IT Optimization Services can help better leverage IT assets and capabilities to improve return on investment by analyzing the IT environment, including data center operations, networks, systems, applications and IT management processes.

Highlights

- ▶ Examines an organization's ability to achieve business value in seven key areas; business/IT alignment, physical infrastructure, applications and data, process, organization, network and the financial environment
- ▶ Prioritizes the resulting initiatives within a balanced scorecard

4.5.11 IT consolidation

IT consolidation is a logical first step in designing a more rational, efficient and resilient infrastructure that can support the on demand business model. IT Consolidation Services provides more than just the physical combining of assets.

Highlights

- ▶ Higher availability
- ▶ Opportunities for resource realignment
- ▶ Increased performance and simplified server environments
- ▶ Faster application deployments
- ▶ Improved cost avoidance processes
- ▶ Simplified server environments, including Linux systems



Part 1

Disk systems

Part 1 covers disk products that either can be directly attached to a server (Direct Attached Storage or DAS), used in a Storage Area Network (SAN) infrastructure, or network attached (NAS).

The IBM System Storage disk products portfolio covers the needs of a wide spectrum of possible implementations, from entry-level to large enterprise.

The IBM System Storage DS family combines the high-performance of the IBM System Storage DS6000 and DS8000 series enterprise servers with the IBM System Storage DS4000 series of mid-range systems, with newly introduced low priced entry systems. The family is further complemented by a range of expansion enclosures to expand the disk storage capacities of individual systems into 100's of terabytes.

Furthermore, a full range of IBM System Storage capabilities such as advanced copy services, management tools, and virtualization services are available to help protect data.

Figure 2 shows the extensive range of DS disk storage systems that are discussed in more detail in subsequent chapters.



Figure 2 IBM Disk Storage systems

Enterprise storage

Under Enterprise disk storage, you find disk systems with the following characteristics:

- ▶ Excellent external storage solution for medium to large scale deployments
- ▶ Medium and large data centers
- ▶ Multi-workload mission critical environments
- ▶ Long distance disaster replication
- ▶ High performance, highly scalable, high function storage systems
- ▶ Designed to avoid any single points of failure for high availability. Highly reliable architecture with hot-swap, redundant components
- ▶ Designed for nondisruptive upgrades in capacity, performance, and cache. Long-term, highly expandable storage solutions to protect investment
- ▶ Flexible management including SMI-S compatible API
- ▶ Advanced business continuance solutions
- ▶ Excellent price/performance/value

Mid-level storage

IBM System Storage mid-level storage disks systems are characterized as:

- ▶ Ideal external storage solution for smaller deployments
- ▶ Often used in remote branch office locations
- ▶ Storage systems designed to offer low cost, scalability, high performance

In particular the DS4000 series incorporates:

- ▶ High reliability with hot-swap redundant power supplies, controllers and drives
- ▶ High performance dual ported 2 Gbps and 4 Gbps FC drives as well as SATA drives
- ▶ Easy to use, common DS Storage Manager

Near-line storage and reference data

For near-line storage and reference data, IBM System Storage disk products in that category offer:

- ▶ Excellent external storage solution for small to large scale deployments for
- ▶ Data Protection – local snapshot and remote mirrors to help support rapid recovery and zero-time backup windows
- ▶ Fixed Content Support – those unchanging files and blocks that have light to moderate access profiles
- ▶ Designed with no single point of failure for high availability. Designed to offer high reliability with hot-swappable redundant components
- ▶ Designed to be nondisruptive for upgrades in capacity and performance, helping to protect investment
- ▶ Flexible management including SMI-S compatible FSM API

Entry-level storage

IBM System Storage entry-level disk systems are:

- ▶ Excellent external storage solution for smaller deployments
 - Smaller data center

- Remote branch office locations
- ▶ Storage subsystems designed to provide low cost, scalability, high performance
- ▶ Designed to offer high reliability with hot-swap redundant power supplies, controllers and drives
- ▶ Easy to use Storage manager helps removes complexity out of SAN deployment and management

Network attached storage

IBM N series (NAS) products provide a wide-range of network attachment capabilities to a broad range of host and client systems. The range includes NAS gateways, providing heterogeneous access to Fibre Channel attached storage arrays, and NAS expandable storage systems, which come with Fibre Channel disks.

IBM System Storage point-in-time copy comparison

The following table provides a quick reference to the point-in-time copy (FlashCopy) functions available on the IBM System Storage SAN Volume Controller, DS4000, and ESS - DS6000 - DS8000.

Each box has one of the following indicators:

- ▶ **Y** – function is currently generally available
- ▶ **N** – function is not available

The table is intended as a guideline only; consult the appropriate product chapters and product documentation for details of the individual product implementations.

As product enhancements are continually made, see your IBM representative for the latest information.

	SAN Volume Controller	IBM DS6000, DS8000, ESS	IBM DS4000
File level	N	Y- z/OS only	N
Block level	Y	Y	Y
Physical Copy	Y	Y	Y
Logical Copy (no copy)	Y	Y	Y
Transition NoCopy -> Copy	Y	Y	N
Logical Copy (space efficient)	N	N	Y
Persistent Copies	Y	Y	N
Multiple Concurrent Copies	N	Y	Y
Source / Target Read / Write Capable	Y	Y	Y
Incremental Copies	N	Y	N
Consistency Groups	Y	Y	N
Across heterogeneous disk vendors	Y	N	N

Figure 3 Point-in-time copy product function comparison

FlashCopy function definitions

File Level: Is the point-in-time copy performed at the file level?

Block / LUN Level: Is the point-in-time copy performed at the block / LUN / volume level?

Physical Copy: Can a physical copy be made of the entire source to the target?

Logical Copy (no background copy): Does the capability exist to not copy the complete source LUN / volume

Transition Nocopy -> Copy: Does the capability exist to change no background copy to a full copy without doing a new point-in-time copy operation?

Logical Copy (space efficient): Does the capability exist to utilize less target space than occupied by the source?

Persistent Copies: Does the source / target relationship remain intact until explicitly withdrawn?

Multiple Concurrent Copies: Can multiple targets be made from one source at the same point in time?

Source / Target Read / Write Capable: Are both the source and target fully read / write capable?

Incremental Copies: Are changes to the source tracked so only changed data is copied to the target?

Consistency Groups: Can associated volumes be treated as one or more groupings and have the point-in-time operation performed across the grouping at the same point in time?

Heterogeneous disk subsystems: Can the point-in-time copy have source and target on different vendor disk subsystems?

IBM System Storage disk mirroring comparison

The following table provides a quick reference to the disk mirroring (Metro Mirror, Global Copy, Global Mirror) functions available on the IBM System Storage SAN Volume Controller, DS4000, and ESS - DS6000 - DS8000.

Each box has one of the following indicators:

- ▶ **Y** – function is currently generally available
- ▶ **N** – function is not available

The table is intended as a guideline only; consult the appropriate product chapters and product documentation for details of the individual product implementations.

As product enhancements are continually made, see your IBM representative for the latest information.

	SAN Volume Controller	IBM DS6000, DS8000, ESS	IBM DS4000
File level	N	N	N
Block level	Y	Y	Y
Metro Mirror	Y	Y	Y
Distance Support	10KM native ~100 KM w/ channel ext.	300KM Sync Unlimited Async	10KM native ~100 KM w/ channel ext.
Global Copy	N	Y	Y
Global Mirror	N	Y	Y
Asynchronous Cascading 3 site	N	Y	N
Failover / Failback Support	Y	Y	N
Consistency Groups	Y	Y	N
Freeze / Thaw support	N	Y	N
Suspend / Resume	Y	Y	Y
Source/Target can be FC Source	Y	Y	Y
Source / Target same cluster	Y	Y	N
Heterogeneous disk Source/Target	Y	N	N

Figure 4 Disk mirroring product function comparison

Disk mirroring function definitions

File Level: Is the point-in-time copy performed at the file level?

Block / LUN Level: Is the point-in-time copy performed at the block / LUN level / volume level?

Metro Mirror synchronous copy - Host notified of write complete after source and target I/O have completed.

Distance Support:

- ▶ SVC – 10 km Fibre Channel - channel extender IP/DWDM based solutions for longer distances.
- ▶ ESS / DS6000 / DS8000 – 103 km over ESCON®, 300 km over Fibre Channel or IP/DWDM based network; unlimited distance supported for asynchronous mirroring.
- ▶ DS4000 – 10 km over FCP; channel extender IP/DWDM based solutions for longer distances.

Global Copy asynchronous copy: Host notified of write complete once I/O to source completed. I/O to target is performed later; data broadcast where data integrity not insured.

Global Mirror asynchronous copy with consistency: Target I/O applied in with data integrity; data suitable for database restart at remote site.

Failover / failback support: Capability to quickly re-establish target -> source relationship in event of a primary site failure, copying only incremental changes back.

Consistency groups: Can associated volumes be treated as one or more groupings such that disk mirroring error triggers cause suspension and keep data integrity across an entire grouping of volumes?

Freeze / thaw (run) support: This function is linked to consistency groups. If an error occurs, I/O stops (freeze) to the consistency group, and an error is posted to the host for automation to act upon, and then host I/O resumes (thaw / run)?

Suspend / resume: Can the remote copy environment be suspended in the event of a planned or unplanned outage and then resumed? This means that only changes while the remote copy environment is suspended are copied to the targets versus entire remote copy environment must be copied in entirety.

Source / target can be FlashCopy source: Can the source or target of a disk mirror be the Source for FlashCopy?

Source / Target same cluster: Can the source and target of a remote copy pair reside in the same cluster / subsystem? This capability is primarily useful for testing purposes.

Heterogeneous disk source / target: Can the source and target of a remote copy pair reside on different vendor disk subsystems?



Entry-level disk storage

This chapter describes the IBM TotalStorage DS300 and DS400 disk systems. These two systems are entry-level, cost-effective workgroup SAN storage for IBM System x and BladeCenter® servers.

The main difference between the two products is that the DS300 is designed to use iSCSI to connect over an existing Ethernet infrastructure, whereas the DS400 is capable of using Fibre Channel connectors. The DS400 is also capable of supporting more disk capacity than the DS300, through the use of EXP400 expansion units, and both support FlashCopy.

Note: The DS300 and DS400 are no longer marketed. They are replaced by the IBM System Storage DS3000 series.

5.1 Overview

The IBM TotalStorage DS300 and DS400 are entry-level, low cost workgroup storage subsystems for System x and BladeCenter servers. The DS300 and DS400 offer a solution for workgroup storage applications, such as file, print and Web serving, as well as remote boot storage for diskless servers.



Figure 5-1 DS300

Leveraging standard Ethernet infrastructure via iSCSI and low cost highly reliable SCSI drives, the DS300 offers a simple, affordable storage area network (SAN) solution that can be used in direct attached or network-attached environments for shared storage as shown in Figure 5-2. For even greater flexibility, the DS300 allows simultaneous support of a range of operating systems for System x and BladeCenter servers—in either single or dual-controller configurations.

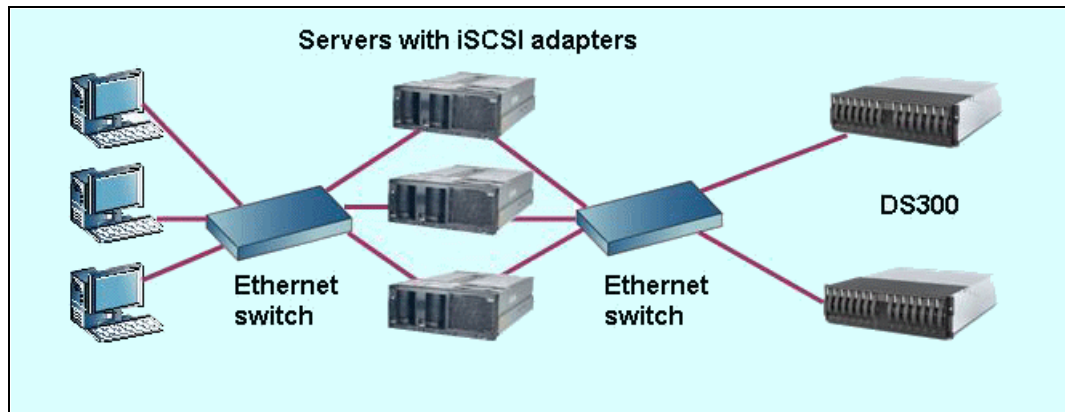


Figure 5-2 Cost-effective iSCSI SAN using DS300

An alternative to iSCSI adapters in the servers is to use iSCSI software initiators - with the disadvantage of less security and increased CPU utilization. Suggested software initiators include:

- ▶ Microsoft iSCSI software initiator
- ▶ Linux iSCSI initiator

Unlike the DS300, the DS400 offers 2 Gbps FC host fabric ports and can therefore be connected to Fibre Channel switches as shown in Figure 5-3 on page 65.

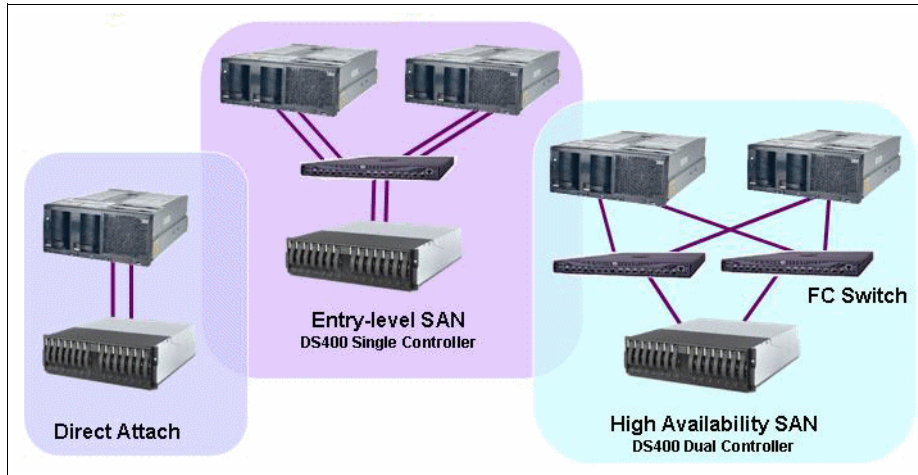


Figure 5-3 SAN connection options for DS400

The modular and scalable design of the DS300 and DS400 can easily be upgraded to meet the growing needs of storage, in addition to advanced software features on the DS300 and DS400, such as access control lists and online array expansion. These features are designed to enable an easy and dynamic storage configuration according to the changing usage needs as well as sharing storage across multiple application servers.

Both units are designed to deliver advanced functionality for business continuity and disaster recovery. Using a rack-mountable 3U enclosure with 14 SCSI drives and redundant design, hot-swap power and cooling modules, the DS300/400 features RAID reliability and high availability software to help maintain operations. The enclosure supports either one or two controllers for high-availability configurations. Each controller is designed to support up to three 1 Gbps Ethernet ports (one for management and two for data paths) so that - even if a line or a controller fails - multiple paths between servers and storage can help maintain storage accessibility.

The DS300/400 also offers optional advanced features such as space-efficient IBM FlashCopy solutions, which are designed to help reduce storage backup windows and improve storage utilization for Microsoft Exchange, Microsoft SQL Server 2000 and Lotus Notes® applications.

The DS300 and DS400 are designed to offer excellent storage management and configuration capabilities through IBM ServeRAID™ Manager (the same software used on the System x servers for RAID management), helping enable clients to manage multiple IBM TotalStorage DS300 and DS400 subsystems, as well as direct-attached internal ServeRAID controllers, all through a single, easy-to-use management console with an extensive online help library. The IBM ServeRAID Manager also provides configuration wizards to simplify the initial storage setup, and supports most common operating environments for workgroup applications, including Microsoft Windows and Linux platforms.

The management software of the DS300/400 is built directly into the client/server workgroup manager IBM Director. IBM Director provides viewing and tracking of the hardware configuration of remote systems in detail and monitors the usage and performance of critical components, such as processors, disks and memory.

5.2 Specifications

Table 5-1 Comparison of DS300 and DS400

Product	DS300	DS400
Machine/Models	1701-1RL, 1701-1RS, 1701-2RD	1700-1RS, 1700-2RD
Platform Support	Windows 2000, Windows 2003, Linux	Windows 2000, Windows 2003, Linux
Host Connectivity	iSCSI	Fibre Channel
SAN Support	Direct, Switched Ethernet	Direct, FC-AL, Switched Fabric
Copy Services	FlashCopy	FlashCopy
Availability Features	Fault Tolerant, RAID, Redundant Hotswap Power, Hotswap drives, Dual controller, dual pathing drivers	Fault Tolerant, RAID, Redundant Hotswap Power, Hotswap drives, Dual controller, dual pathing drivers
Controller	Single or dual active 1 GB iSCSI RAID Controllers	Single or dual Active 2 GB FC RAID Controllers
Cache (min/max)	256 MB (1RL), 256 MB/ 1 GB (Single), 512 MB/ 2 GB (Dual)- Battery Back-up	256 MB / 1 GB (Single), 512 MB / 2 GB (Dual)- Battery Back-up
RAID Support	0, 1, 5, 10, 50	0, 1, 5, 10, 50
Capacity (min,max)	36 GB, 4.2 TB	36 GB, 12 TB with 2 EXP400 Expansion Units
Drive Interface	Ultra320 SCSI	Ultra320 SCSI
Drive Support	36 GB, 73 GB, 146 GB, 300 GB 10,000 RPM Disk Drives; 36 GB, 73 GB, 146 GB 15,000 RPM Disk Drives	73 GB, 146 GB, 300 GB 10,000 RPM Disk Drives; 36 GB, 73 GB, 146 GB 15,000 RPM Disk Drives
Clustering support	Microsoft Windows2003 MSCS	Microsoft Windows 2003 MSCS

5.3 IBM TotalStorage DS300

The IBM TotalStorage DS300 includes easy-to-use IBM ServeRAID Manager, which provides common management software. It scales to 4.2 TB of physical storage capacity using 300 GB Ultra320 SCSI drives.

By utilizing standard System x hot-swap Ultra320 SCSI drives which are designed to offer high mean-time between-failure, the DS300 is ideal for high I/O processor (IOP) performance that allows for future growth. The modular design of the DS300 can easily be upgraded to support increasing volumes of data flexibly and affordably. The DS300 can also be enabled to share storage across multiple application servers.

The IBM TotalStorage DS300 comes in a variety of models.

1. The DS300 storage subsystem 1701-1RL consists of the following components:
 - One iSCSI RAID controller. The controller has one 1 Gbps Ethernet input connector for data and management.

- One power-supply-with-fan unit.
- 7 hard disk drive filler panels.
- LED indicators (8).
- 256 MB raid cache memory.
- 512 MB system memory.
- 4.2 TB maximum storage capacity.
- RAID levels 0, 1, 5, 10 and 50.

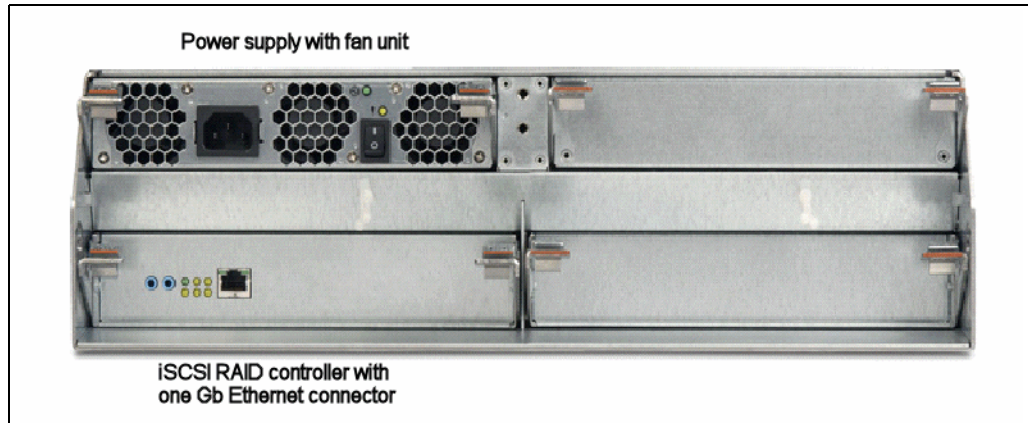


Figure 5-4 DS300 (1701-1RL) storage subsystem rear view

2. The DS300 storage subsystem 1701-1RS comes with the following components:

- One iSCSI RAID controllers. The controller has one 1 Gbps Ethernet input connector for data and management.
- Two power-supply-with-fan units.
- 14 hard disk drive filler panels.
- LED indicators (8).
- 256 MB raid cache memory per controller.
- 512 MB system memory (can be upgraded to 1 GB).
- 4.2 TB maximum storage capacity.
- Raid levels 0, 1, 5, 10 and 50.

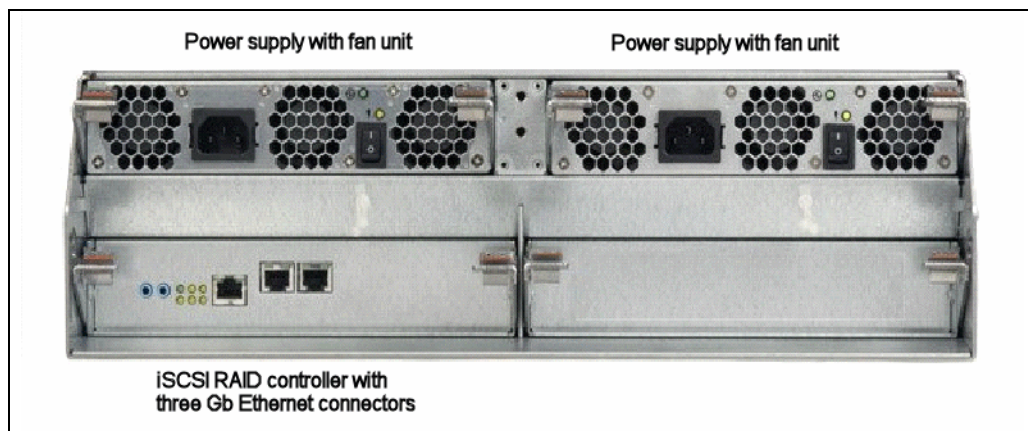


Figure 5-5 DS300 (1701-1RS) storage subsystem rear view

3. The DS300 storage subsystem 1701-2RD contains the following components:

- Two iSCSI RAID controllers. Each controller has one 1 Gbps Ethernet input connector for management, two 1 Gbps Ethernet input connectors for data.
- Two power-supply-with-fan units.
- 14 hard disk drive filler panels.

- LED indicators (12).
- 256 MB raid cache memory per controller.
- 512 MB system memory (can be upgraded to 1 GB).
- 4.2 TB maximum storage capacity.
- RAID levels 0, 1, 5, 10 and 50.

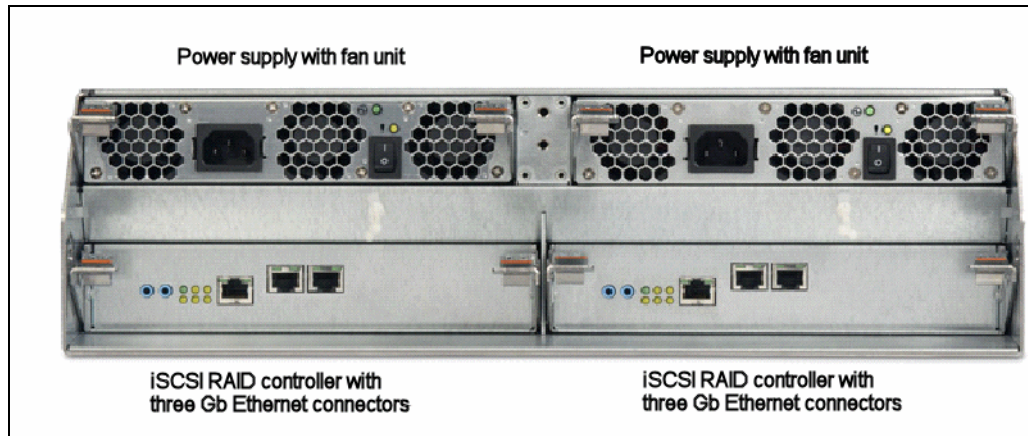


Figure 5-6 DS300 storage subsystem model 1701-2RD rear view

5.4 IBM TotalStorage DS400

The DS400 storage subsystem uses industry-proven 2 GB Fibre Channel connectors and is also capable of supporting more disk capacity than the DS300, through the use of IBM TotalStorage EXP400 expansion for affordable migration from Direct Attached Storage (DAS) to a Storage Area Network (SAN). The expansion allows a total of 40 disk drives or 12 TB of physical capacity, using 300 GB Ultra320 SCSI drives.

The DS400 storage subsystem comes with one (1700-1RS) or two (1700-2RD) Fibre Channel RAID controllers. The IBM TotalStorage DS400 storage subsystem comes with the following components:

- ▶ One 1 Gbps Ethernet connector for management.
- ▶ Two 2 Gbps Fibre Channel host ports.
- ▶ One Fibre Channel RAID controller (2 for 1700-2RD).
- ▶ Two power-supply-with-fan units.
- ▶ 14 hard disk drive filler panels.
- ▶ 256 MB raid cache memory per controller.
- ▶ 512 MB system memory (can be upgraded to 1 GB).
- ▶ 12 TB maximum storage capacity.
- ▶ RAID levels 0, 1, 5, 10 and 50.

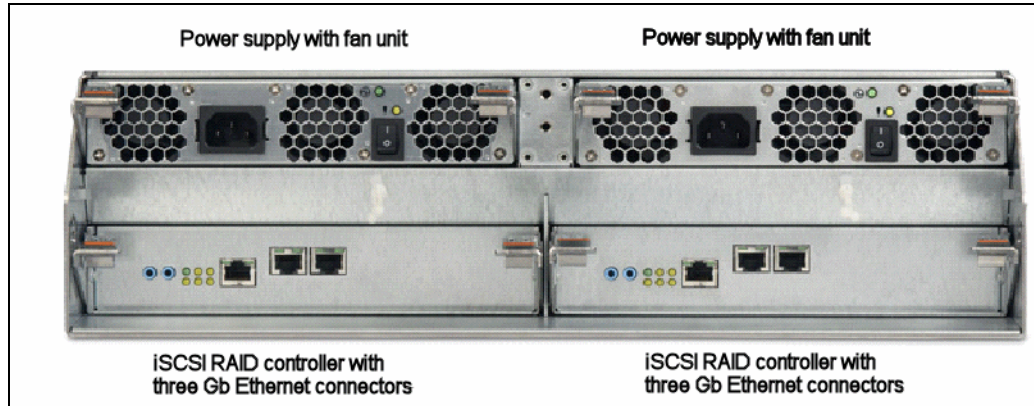


Figure 5-7 DS400 (1700-2RD) storage subsystem rear view

5.5 ServeRAID Manager

The ServeRAID Manager is used to configure the storage subsystem and to monitor its status. The steps to configure the DS300/400 ServeRaid Manager are as follows:

1. Install IBM ServeRAID Manager.
2. Add a ServeRAID agent for the DS300/400 storage subsystem.
3. Update the controller firmware if required.
4. Set the IP addresses for the iSCSI ports on the DS300.
5. Create a RAID array, create a logical drive and configure the access control list.

For monitoring the systems, the program should be run constantly and checked frequently. It offers the best way to diagnose and repair storage-subsystem failures. The program can help to:

- ▶ Determine the nature of the failure.
- ▶ Locate the failed component.
- ▶ Determine the recovery procedures to repair the failure.

After logging in for the first time, you must add the management station before you can begin setting up the network storage. The management station agent is a monitoring agent for network storage. After you add an agent a management station and complete the configuration of the attached storage system enclosures, you can monitor the storage from the ServeRAID Manager console.d.

Figure 5-8 on page 70 shows how IBM ServeRAID Manager displays the Physical and Logical device views, showing the physical devices and logical devices connected to the controller or enclosure.

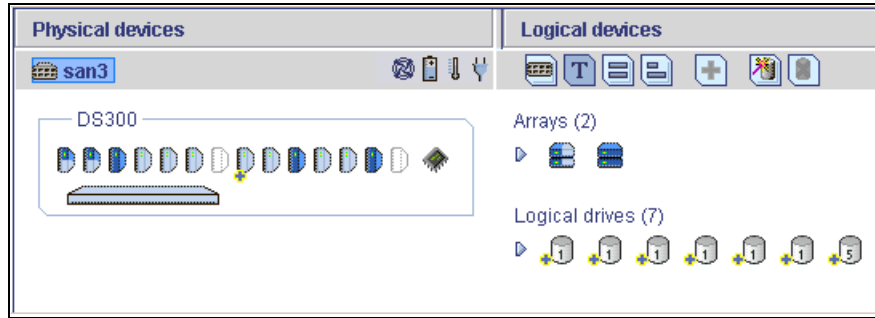


Figure 5-8 ServeRAID Manager

The indicator colors are:

- ▶ Blue - Normal
- ▶ Yellow - Warning
- ▶ Red-Error
- ▶ Gray - Not applicable to the devices

For example, the fan indicator changes to yellow when one fan fails; it changes to red when a second fan fails and cooling is no longer adequate. Although the storage subsystem has fault LEDs, these LEDs do not necessarily indicate which component has failed or which type of recovery procedure must be performed. In some cases, such as loss of redundancy in a fan, the fault LED is not lit. Only the ServeRAID Manager program can detect the failure.

Tip: The *IBM TotalStorage DS400 Problem Determination Guide* and *IBM TotalStorage DS300 Problem Determination Guide* provide examples of typical configuration and problem determination flowcharts and are available online at:

<http://www.ibm.com/pc/support/site.wss/document.do?lndocid=MIGR-56887>

5.6 Premium features

The following premium features are offered, depending on the model.

5.6.1 EXP400 expansion units

The DS400 has two external small computer system interface (SCSI) connectors which can connect two optional EXP400 expansion units. They support up to 14 Ultra320 SCSI hard disk drives on a single or dual bus allowing up to 12 TB of total storage.

The EXP400 expansion unit:

- ▶ Delivers fast and high-volume data access.
- ▶ Provides storage functions across multiple drives and to multiple hosts.
- ▶ Designed for continuous and reliable service
- ▶ Modular design.
- ▶ Provides redundant disk drives.
- ▶ Provides redundant power supply with fan units
- ▶ Provides SCSI Bus Expander and Enclosure Services Module (ESM).
- ▶ Uses hot-swap technology for easy replacement without turning off the expansion unit.

Expansion unit models 1733-1RU and -1RX come with two 500-watt AC power supply with fan units, one ESM (the second ESM is optional), a filler panel to cover the empty ESM bay,

and 14 drive filler panels. Expansion unit model 1733-2RX comes with two -48 volt DC power supply with built-in fan units, two ESMS, and 14 drive filler panels.

Refer to the following documents for more information about the EXP400:

- ▶ *IBM EXP400 Storage Expansion Unit Type 7133 Hardware Maintenance Manual and Troubleshooting Guide*
ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/90p3548.pdf
- ▶ *IBM EXP400 Storage Expansion Unit Installation Guide*
ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/88p9249.pdf
- ▶ *IBM EXP400 Storage Expansion Unit User's Guide*
ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/88p9250.pdf

5.6.2 FlashCopy

FlashCopy provides the ability to make an instantaneous point-in-time copy of a volume. For the DS300/DS400, it is disk space efficient and uses a copy-on-write methodology (only changed data is copied) preserving the primary volume for unchanged data to save space on the disk drives. Only the FlashCopy Management Command-Line Tool can be used to create and manage FlashCopies; however, the Storage Manager can be used to view the properties of a FlashCopy.

FlashCopy can be used for many purposes. Combined with application agents, which are required to quiesce the application to permit an accurate copy, FlashCopy can save a point-in-time copy of a volume for backup purposes. The volume can be mounted as a copy of a volume for use in databases, application development or short term backup and recovery scenarios.

Up to 254 FlashCopy volumes are supported.

5.7 Advanced software features

The following features are available for both DS300 and DS400 systems:

- ▶ Access Control Lists – ability to select which servers can access each storage partition or target (LUN mapping/masking). Access Control Lists allow the user to assign LUNs to servers and designate read/write or modify rights in order to improve security of data.
- ▶ Online LUN expansion.
- ▶ Online RAID level migration.
- ▶ DS300/DS400 ServeRAID™ Integration with IBM Director.
- ▶ Global Hot Spare.

Boot from SAN

DS400 and DS300 can be used to boot from the SAN rather than from local disks on individual servers providing maximum consolidation of IT resources and minimum equipment cost. Boot from SAN is a remote boot technology where the source of the boot disk is on the SAN. The server communicates with the SAN through host bus adapters (HBA). The HBA BIOS contains the instructions that enable the server to find the boot disk on the SAN. Booting from the SAN provides also a rapid Disaster Recovery as all of the boot information and production data are stored on a local SAN, remote SAN or both environments.

DS300 operates on iSCSI and it supports booting from SAN using iSCSI interconnect to the SAN, provided iSCSI HBAs are used to enable the boot process.

Attention: Boot from SAN is not supported using iSCSI software initiator. No sharing of boot images is allowed, as Windows servers cannot currently share a boot image. Each server requires its own dedicated LUN to boot.

5.8 More information

For more information about the IBM TotalStorage entry-level disk systems consult the following documents and Web sites:

<http://www.ibm.com/servers/storage/disk/ds/ds300/index.html>

<http://www.ibm.com/servers/storage/disk/ds/ds400/index.html>

- ▶ *IBM TotalStorage DS300 and DS400 Hardware Maintenance Manual*, at:

ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/25k8190.pdf

- ▶ IBM Redbook *IBM TotalStorage DS300 and DS400 Best Practices Guide*, SG24-7121



Expandable storage

This chapter presents the Expandable Storage Plus (EXP Plus) 2104 and EXP24 family.

The IBM 2104 Expandable Storage Plus and EXP24 provide flexible, scalable, and low-cost disk storage in a compact package for pSeries® servers. These disk enclosures are ideal for enterprises, such as Internet or application service providers, that need high-performance external disk storage in a small footprint.

6.1 Overview and key features of the EXP24

The IBM TotalStorage EXP24 offers the following features:

- ▶ Attachment to IBM eServer pSeries POWER4™, POWER5™, and OpenPower™ servers running AIX and Linux. This includes iSeries™ models 520, 550, 570 or 595 using an AIX 5L™ or Linux partition.
- ▶ Scales up to 7.2 TB of storage capacity
- ▶ Provides attachment for up to eight servers.
- ▶ Incorporates high-performance Ultra320 SCSI disk storage with 320 MBps throughput for the 320 models.
- ▶ Features up to twenty-four 10,000 RPM disk drives, with capacities of 73.4 GB, 146.8 GB, and 300 GB or 15,000 RPM disk drives, with capacities of 36.4 GB, 73.4 GB, and 146.8 GB.
- ▶ Configurable as either 4 groups of 6 drives or 2 groups of 12 drives with either single or dual connection to any group of drives
- ▶ RAID 1, 3, and 5 support
- ▶ Provides high availability to safeguard data access.
- ▶ Provides scalability for fast-growing environments.

The IBM TotalStorage EXP24 family consists of the 7031 models D24 or T24. The Model D24 is a horizontal 4 EIA by 19-inch rack drawer for mounting in equipment racks, and is shown in Figure 6-1. The Model T24 is a vertical tower for floor-standing applications.



Figure 6-1 IBM EXP24 models D24

These models provide industry-standard Ultra320 SCSI interfaces combined with Ultra320 SCSI adapters and high-performance disk storage with 320 MBps throughput.

6.1.1 High availability

To help ensure that information is accessible, the EXP24 is designed for high availability. It features hot-swappable, auto-docking disk drives. It has redundant cooling fans, and an option for dual, hot-pluggable power and power supplies.

Availability is enhanced with the use of RAID options or UNIX software mirroring. SCSI RAID adapters provide RAID 1, 3, and 5 functions, and the operating system can provide mirroring through the operating system.

The EXP24 supports two host SCSI bus configurations. The Single Bus Ultra320 SCSI Repeater Card (#5741) provides a single initiator to each SCSI group allowing each SCSI Drive group to have a single connection between the server and the EXP24. The Dual Bus Ultra320 SCSI Repeater Card (#5742) provides a dual initiator to each SCSI group allowing each SCSI drive group to have dual connections between the server (or servers) and the

TotalStorage EXP24. This feature can provide options for high availability and for multiple servers sharing a single drive group. The Dual Bus Ultra320 SCSI Repeater Card can be used on any or all of the drive groups in the enclosure together with other drive groups in the enclosure using the standard connection option.

Up to four initiator cards can be configured - and they can be any combination of single and dual bus.

The EXP24 has two power supplies so that it can be operated from two independent power sources.

6.1.2 Flexible and scalable storage

The rack-mounted Model DS4 drawer can reside in a variety of 19-inch racks, including IBM 7014-T00 and 7014-T42.

The rack-mounted D24 model and the desk-side tower T24 model can be populated with up to twenty-four disk drives. Available options include 10,000 RPM (73.4 GB, 146.8 GB, and 300 GB) and 15,000 RPM (36.4 GB, 73.4 GB and 146.8 GB) disk drives. Within the EXP24 enclosure, these drive capacities and speeds can be mixed and drives can be added in increments as few as one or as many as twenty-three. Housing multiple drawers in a rack provides the flexibility to build storage capacity from gigabytes to terabytes.

6.2 Overview and key features of the EXP Plus

The IBM 2104 Expandable Storage Plus offers the following features:

- ▶ Provides an ideal package for Internet service providers when combined with an IBM RS/6000® or pSeries server.
- ▶ Scales from up to 4 TB of capacity per drawer or tower to more than 56 TB per rack.
- ▶ Provides a shared storage for all major types of servers.
- ▶ Provides a single or split-bus configuration flexibility to one or two servers.
- ▶ Incorporates high-performance Ultra320 SCSI disk storage with 320 MBps throughput for the 320 models.
- ▶ Features up to fourteen 10,000 RPM disk drives, with capacities of 73.4 GB, 146.8 GB, and 300 GB or 15,000 RPM disk drives, with capacities of 36.4 GB, 73.4 GB, and 146.8 GB.
- ▶ Provides high availability to safeguard data access.
- ▶ Provides scalability for fast-growing environments.

Note: The 2104 models DS4 and TS4 are withdrawn from marketing as of 31 March, 2006. They are replaced by the IBM TotalStorage EXP24.

The IBM 2104 Expandable Storage Plus family consists of the 320 models DS4 or TS4 (Figure 6-2 on page 76).



Figure 6-2 IBM 2104 expandable storage plus 320 models DS4 and TS4

The 320 models provide industry-standard Ultra320 SCSI interfaces combined with Ultra320 SCSI adapters and high-performance disk storage with 320 MBps throughput. The 320 models are designed for high availability.

6.2.1 High availability

To help ensure that information is accessible, the EXP Plus is designed for high availability. It features hot-swappable, auto-docking disk drives, as well as an optional redundant power supply and cooling unit, all designed to minimize downtime.

Availability is enhanced with the use of RAID options or UNIX software mirroring. SCSI RAID adapters provide RAID 0, 1, 1E, 5 and 5E functions and can be used to attach the EXP Plus to RS/6000 and pSeries servers for increased data protection. To further facilitate availability, SCSI enclosure services can monitor and report problems to the server.

A number of SCSI bus configurations are supported by the EXP Plus 320 models. When attached to a single host system, up to 14 disk drives can be configured per host adapter SCSI port in each EXP Plus 320 enclosure to meet large capacity requirements. In dual-host environments that require the optional second U320 Port feature (#2642) be installed, the EXP Plus 320 can be configured as two separate SCSI buses, each bus supporting up to seven disk drives. Each bus is attached to two independent host connections. The third option is in high-availability, dual-host clustered configurations utilizing high availability cluster multi-processing (HACMP) for AIX, where the two host ports (requires feature number 2642) on the EXP Plus 320 are each connected to the clustered hosts on a single SCSI bus, and up to 12 disk drives enable non-concurrent usage with failover support.

The EXP Plus 320 has several options for power. The simplest uses a single AC power supply (#3430) with a fan unit (#3499). For dual redundant power supplies and ability to attach to two power sources, two AC power supplies (#3430) can be ordered.

6.2.2 Flexible and scalable storage

The EXP Plus helps to address the needs of many different types of environments. Four models are available to fit the need for standalone storage in one- or two-server environments, or for rack-mounted storage used by several servers.

The rack-mounted Model DS4 drawer can reside in a variety of 19-inch racks, including IBM 7014-T00 and 7014-T42. The largest rack (7014-T42) can hold up to fourteen EXP Plus drawers for a total physical capacity of over 56 TB.

The standalone tower Model TS4 are ideal for one- or two-server environments, providing up to 4 TB of capacity in a small desk-side tower unit.

The rack-mounted DS4 model and the desk-side tower TS4 model can be populated with up to fourteen disk drives. Available options include 10,000 RPM (73.4 GB, 146.8 GB, and 300 GB) and 15,000 RPM (36.4 GB, 73.4 GB and 146.8 GB) disk drives. Within the EXP Plus 320 enclosure, these drive capacities and speeds can be mixed and drives can be added in increments as few as one or as many as thirteen. Housing multiple drawers in a rack provides the flexibility to build storage capacity from gigabytes to terabytes.

6.3 More information

- ▶ For information about the EXP Plus models, refer to:
<http://www.ibm.com/servers/storage/disk/expplus>
- ▶ For information about the EXP24 models, refer to:
<http://www.ibm.com/servers/storage/disk/exp24/index.html>
- ▶ For the supported adapters, refer to:
<http://www.ibm.com/servers/storage/disk/expplus/pdf/interop-matrix.pdf>
<http://www.ibm.com/servers/storage/disk/exp24/interop.html>



Mid-range disk systems

The DS4000 series systems are RAID controller devices that contain Fibre Channel (FC) interfaces to connect the host systems and the disk drive enclosures.

All DS4000 series systems have hot-swap and redundant power supplies and fans. With the exception of the base entry model, all DS4000 series systems have dual RAID controllers, which are also hot-swappable, therefore providing excellent system availability even if one of the parts should malfunction.

In this chapter, we present the features and major characteristics of the different models that make up the current DS4000 family:

- ▶ IBM System Storage DS4100
- ▶ IBM System Storage DS4300 and DS4300 Turbo
- ▶ IBM System Storage DS4500
- ▶ IBM System Storage DS4700 Express
- ▶ IBM System Storage DS4800
- ▶ IBM System Storage DS4000 EXP100 Expansion Unit
- ▶ IBM System Storage DS4000 EXP710 Expansion Unit
- ▶ IBM System Storage DS4000 EXP810 Expansion Unit

We also briefly discuss the Storage Manager software that is used to manage the DS4000 series systems and show how it has evolved over time.

7.1 Highlights - what's new

This section summarizes the new functions, supported platforms, and features of the DS4000 family of disk storage products, since the last edition of this book in July 2005.

- ▶ Technical specifications
- ▶ Fibre Channel interconnection 4 Gbps technology
- ▶ EXP810 expansion unit
- ▶ DS4700
- ▶ DS4800
- ▶ Storage Manager updated to Version 9.16

7.2 Overview

IBM System Storage and System Storage DS4000 series are disk storage products using redundant array of independent disks (RAID) that contain the Fibre Channel (FC) interface to connect both the host systems and the disk drive enclosures. The DS4000 series of disk storage systems are an IBM solution for mid-range/departmental storage requirements. The IBM System Storage DS4000 series family are:

- ▶ IBM System Storage DS4100

The DS4100 entry-level SATA storage system is available in a single and dual controller configuration. The single controller model provides entry-level storage expandable to 5.6 TB. The dual controller model supports up to 44.8 TB with the DS4000 EXP100 expansion. The single controller model has 2 Fibre Channel host connections and the dual controller provides 4 FC ports.

- ▶ IBM System Storage DS4300 and DS4300 Turbo

The DS4300 is a mid-level, 1 Gbps Fibre Channel storage system available in a single and dual controller configuration. It is suitable for consolidation and clustering applications, and scales to over 16 TB of disk.

The DS4300 Turbo is a mid-level disk system supporting up to 44.8 TB of disk, facilitating storage consolidation for medium-sized customers. It provides an end-to-end 2 Gbps Fibre Channel solution (the host interface on base DS4300 is 2 Gbps, while the turbo auto senses to connect to 1 Gbps or 2 Gbps) and offers up to 65 percent read performance improvement.

Note: IBM System Storage DS4300 Express Model is designed to be easy to acquire, install, and manage, as it features standard configuration packages. For the latest information, see:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4300/express/browse.html>

- ▶ IBM System Storage DS4500

The DS4500 scales up to 67.2 TB of FC disk capacity using 224 300 GB drives with EXP700s or EXP710s, up to 89.6 TB of Serial ATA disk capacity with EXP100. The DS4500 offers advanced replication services to support business continuance and disaster recovery.

- ▶ IBM System Storage DS4700 Express

IBM System Storage DS4700 Express offers high performance 4 Gbps capable Fibre Channel connections and up to 33.6 TB of physical storage capacity. The DS4700 can

expand from workgroup to enterprise-wide capability by attaching up to six DS4000 EXP810 disk enclosures.

► IBM System Storage DS4800

The IBM System Storage DS4800 is the most powerful in the DS4000 Series. It features a 4 Gbps Fibre Channel interface designed for data-intensive applications that demand increased connectivity. The DS4800 can support up to 224 disk drives in EXP810, EXP710, and EXP100 disk units making it a great choice for performance-oriented or capacity-oriented storage requirements. Available in four models: the 80A and 82A with 4 GB of cache, the 84A with 8 GB of cache, and the 88A with 16 GB of cache. All models support over 67.2 TB of Fibre Channel (FC) physical storage capacity and 89.6 TB of Serial ATA (SATA)

The DS4000 Series is a rebranding of the former IBM System Storage/IBM System Storage (including FAStT) and also includes newer models. The mapping between former DS4000/FAStT models and renamed DS4000 models is detailed in Table 7-1.

Table 7-1 DS4000 product renaming

Original Name	New Name
IBM System Storage FAStT Disk System	IBM System Storage DS4000
FAStT	DS4000
FAStT family	DS4000 Series
N/A	DS4800
N/A	DS4700 Express
FAStT900	DS4500
N/A	DS4300 Express
FAStT600 Turbo	DS4300 Turbo
FAStT600	DS4300
FAStT100	DS4100
N/A	DS4000 EXP810 expansion unit
N/A	DS4000 EXP710 Express expansion unit
N/A	DS4000 EXP710 expansion unit
N/A	DS4000 EXP100 Express
EXP100	DS4000 EXP100

Along with the product rebranding, some of the features available with the DS4000 products have been renamed as shown in Table 7-2. The renaming also includes a couple of new mirroring features which are described in 7.13, “Copy services” on page 104.

Table 7-2 DS4000 services renaming

Original Name	New Name
FAStT Storage Manager	DS4000 Storage Manager
FAStT Service Alert	DS4000 Service Alert
N/A	DS4000 Alert Manager

Original Name	New Name
FAST FlashCopy	FlashCopy for DS4000
FAST VolumeCopy	VolumeCopy for DS4000
FAST Remote Volume Mirror (RVM) <ul style="list-style-type: none"> ▶ FAST Synchronous Mirroring ▶ N/A ▶ N/A 	Enhanced Remote Mirroring for DS4000 <ul style="list-style-type: none"> ▶ Metro Mirroring for DS4000 ▶ Global Copy for DS4000 ▶ Global Mirroring for DS4000

Note: Refer to the product interoperability matrix for the latest information at:

<http://www.ibm.com/servers/storage/disk/ds4000/pdf/interop-matrix.pdf>

7.3 IBM System Storage DS4100

The IBM System Storage DS4100 Disk System, shown in Figure 7-1, is a 3U rack-mountable Serial ATA (SATA) RAID controller and disk drive enclosure. It is an entry-level server that can help address storage consolidation and near-line application storage needs. The DS4100 is designed for mid-range and entry-level application environments that need lower cost storage, but do not want to give up the availability, manageability, data integrity, and software/hardware features achieved with Fibre Channel-based storage systems.

A typical use of the DS4100 is as data archive and data backup storage. With up to seven expansion enclosures attached, it can support 112 SATA disk drives for a total storage capacity of 44.8 TB.

7.3.1 Features

There are two models available:

- ▶ Model 1724-100 (Dual Controller)
- ▶ Model 1724-1SC (Single Controller)



Figure 7-1 IBM System Storage DS4100

The DS4100 Dual Controller model has the following major characteristics:

- ▶ Compact 3U rack-mountable enclosure.

- ▶ Dual 2 Gbps hot-swappable RAID controllers with 512 MB of battery backed cache (256 MB per controller) on dual controller models.
- ▶ Redundant, hot-swappable power supplies and cooling on dual controller models.
- ▶ Support for RAID-0, 1, 3, 5 and 10.
- ▶ Supports global hot spares.
- ▶ One expansion port per controller.
- ▶ Dual Controller model supports up to seven IBM System Storage DS4000 EXP100 Expansion Units (a license key for 4th - 7th enclosure is required). This allows for up to 112 disk drives, including the 14 internal drives in the standard enclosure, offering a potential capacity of 44.8 TB.

Note: DS4100 only supports the EXP100 expansion enclosure.

- ▶ There are four 2 Gbps host ports, two for each controller; this enables a direct, redundant attach of two servers (clustering is also supported) without the use of a switch.
- ▶ Support both short and long wave Fibre Channel host attachment, allowing distances of up to 10 km (6.2 miles).
- ▶ Support one storage partition in standard configuration. There is an option to expand up to 4, 8 or 16 storage partitions.
- ▶ FlashCopy (premium feature)
- ▶ Standard configuration only supports Windows connection; Linux, NetWare, and UNIX support needs a host kit option.
- ▶ On demand functions such as Dynamic Volume Expansion, Dynamic Capacity Expansion and Dynamic RAID Level Migration, allowing unused storage to be brought online without stopping operations.
- ▶ DS4000 Service Alert which is capable of automatically alerting IBM if a problem occurs.
- ▶ Each RAID controller fits into the back of the system disk and has the following components: battery, cache memory, two host ports and one expansion port, one Ethernet port, one serial port. Each RAID controller has a unique hardware Ethernet address, which is printed on a label on the front. There is also a label giving the cache battery manufacturing and installation dates on the top of the controller unit. Figure 7-2 shows the parts of the back of the DS4100.

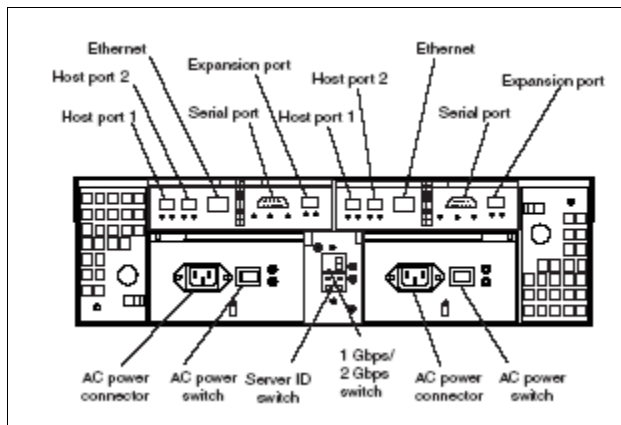


Figure 7-2 DS4100 Back view

Figure 7-3 on page 84 shows a zoom in of the controller card.

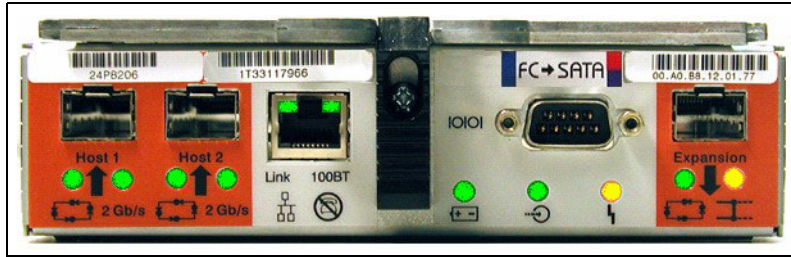


Figure 7-3 DS4100 Controller

The DS4100 Single Controller model has the following differences (or restrictions) compared to the dual controller model:

- ▶ Single 2 Gbps RAID controller with 256 MB of battery backed cache. The single controller can be upgraded to the DS4100 Dual Controller model.
- ▶ Single power supply and redundant cooling fan assemblies.
- ▶ The DS4100 single controller includes as standard three 250 GB SATA disk modules and supports 11 additional internal SATA disk modules.
- ▶ Expandable to over 5.6 TB in the controller enclosure with 14 x 400 GB drives. No expansion units can be attached.
- ▶ Warranty: Three year parts and labor warranty, 9x5 next business day, upgradeable to 24x7 with four hour response

Restriction: The EXP100 expansion enclosure cannot be attached to the single controller DS4100 model.

7.3.2 Specifications

Table 7-3 summarizes the DS4100 specifications.

Table 7-3 IBM DS4100 specifications

Model	1724-1SC or 1724-100
RAID controller	Single or Dual active 2 GB RAID controllers
Cache	256 MB or 512 MB total, battery-backed
Host interface	2/4 Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard
Drive interface	2 Gbps FC-AL
Supported drives	250 GB and 400 GB 7200 rpm
RAID	levels 0, 1, 3, 5, 10
Storage partitions	Standard Host Group, upgradeable to 4, 8 or 16 partitions
Maximum drives supported	Single: 14 drives, Dual: 112 Serial ATA drives (using seven DS4000 EXP100 Expansion Units)
Fans and power supplies	Single/dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack - 3U

For the latest specification information, check:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4100/index.html>

7.4 IBM System Storage DS4300

The IBM System Storage DS4300 is an affordable, scalable disk system for storage consolidation and clustering applications. Its modular architecture, which includes Dynamic Capacity Addition, Dynamic Volume Expansion and a Turbo option can support e-business on demand® environments by helping to enable storage to grow as demands increase. Autonomic features such as online firmware upgrades and DS4000 Service Alert also help enhance the system's usability.

The DS4300 helps consolidate direct-attached storage into a centrally managed, shared or SAN environment.

7.4.1 Features

The DS4300 is a 3U rack-mountable 2 Gbps Fibre Channel RAID storage system and disk drive enclosure housing up to 14 FC drives with capacities from 36.4 GB to 300 GB. There are two models available:

- ▶ Model 1722-60X (Dual Controller)
- ▶ Model 1722-6LU (Single Controller)

The DS4300 is shown in Figure 7-4.



Figure 7-4 IBM System Storage DS4300

The DS4300 Dual Controller model has the following major characteristics:

- ▶ There are four host ports, two for each controller; this enables a cluster solution without the use of a switch.
- ▶ Dual 2 Gbps hot-swappable RAID controllers with 512 MB battery backed up cache in the turbo model.
- ▶ Redundant, hot-swappable power supplies and cooling.
- ▶ Support RAID levels 0, 1, 3, 5 or 10.
- ▶ Supports global hot spares.
- ▶ One expansion port per controller.
- ▶ Internal controller capacity is 14 FC drives up to 300 GB each

- ▶ Supports three IBM System Storage DS4000 EXP700/710 Expansion Units (a license key is required); this allows for up to 56 disk drives, offering a potential capacity of 16.8 TB Fibre Channel disk.
- ▶ Supports eight IBM System Storage DS4000 EXP100 Expansion Units, this allows for up to 112 disk drives, offering a potential capacity of 56 TB of Serial ATA disk storage. Requires DS4000 Storage Manager 9.10 or later.
- ▶ Autonomic functions such as Dynamic Volume Expansion and Dynamic Capacity Addition, allowing unused storage to be brought online without stopping operations, and DS4000 Service Alert, which is capable of automatically alerting IBM if a problem occurs
- ▶ FlashCopy and Volume Copy are available.
- ▶ Supports both short and long wave Fibre Channel host attachment, allowing distances up to 10 km (6.2 miles).
- ▶ Supports one storage partition in standard configuration. There is an option to expand up to 4, 8 or 16 storage partitions.
- ▶ The DS4300 includes licensed IBM Machine Code and support to attach to servers running Microsoft Windows operating systems. For attachment to IBM @server pSeries/AIX, Sun/Solaris™, HP/HP-UX, and other Intel-based operating system servers, optional features must be ordered to obtain the IBM Machine Code licenses and support necessary

Note: The DS4300 can be upgraded to the DS4300 Turbo.

The DS4300 Single Controller model has the following differences (or restrictions) compared to the dual controller model:

- ▶ Single 2 Gbps RAID controller with 256 MB of battery backed cache. The single controller can be upgraded to the DS4300 Dual Controller model and the DS4300 Turbo.
- ▶ Single power supply and redundant cooling fan assemblies.
- ▶ Cannot attach any expansion units - only 14 FC internal drives.

All DS4300 models support hot-swap Fibre Channel hard drive CRUs and cooling fan CRUs, so you can remove and replace these components without turning off the disk system. Figure 7-5 shows the components and controls on the front of the DS4300.

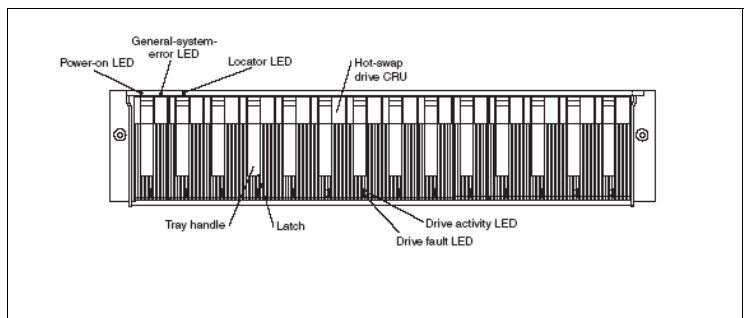


Figure 7-5 DS4300 front view

7.4.2 Specifications

Table 7-4 show the specifications for the DS4300.

Table 7-4 IBM DS4300 Disk System

Model	1722-6LU or 1722-60X
-------	----------------------

RAID controller	Single or Dual active 2 GB RAID controllers
Cache	Base model: 256 MB or 512 MB total, battery-backed
Host interface	4 Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard
Drive interface	Base model: 2 Gbps FC-AL
Supported drives	36.4 GB, 73.4 GB, 146.8 GB, 300 GB 10,000 rpm 18.2 GB, 36.4 GB, 73.4 GB, 146.8 GB 15,000 rpm
RAID	levels 0, 1, 3, 5, 10
Storage partitions	Base model: 1, upgradeable to 4, 8 or 16
Maximum drives supported	14 FC drives 4.2 TB (single controller) or 56 FC drives (using three EXP700/710 Expansion Units) 16.8 TB 112 SATA drives (using 8 EXP100 Expansion Units) 44.8 TB
Fans and power supplies	Single or Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack - 3U

For the latest specification information, check:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4300/index.html>

7.5 IBM System Storage DS4300 Turbo

To meet growing needs for storage capacity and performance, the DS4300 has a Turbo option. The Turbo option extends the high-availability and scalability characteristics of the entry-level DS4300 base model, and may be added during the initial order or later as an upgrade.

7.5.1 Features

The DS4300 Turbo is an upgraded version of the base DS4300 model and has the following characteristics and features in comparison:

- ▶ DS4300 Turbo performs at more than 75,000 cached I/Os per second - a 70 percent improvement and provides up to 768 MBps throughput.
- ▶ Higher scalability over base DS4300, scalable to 33 TB of Fibre Channel disk by attaching up to seven EXP700 or EXP710 units.
- ▶ Supports eight IBM System Storage DS4000 EXP100 Expansion Units, this allows for up to 112 disk drives, offering a potential capacity of 28 TB of SATA disk storage.
- ▶ Increased cache 1 GB per controller on Turbo for a total of 2 GB.
- ▶ Host interface on base DS4300 is 2 Gbps. The DS4300 Turbo auto senses to connect to either 1 Gbps or 2 Gbps.
- ▶ Provides four 2 Gbps Fibre Channel host ports designed to economically support multipath failover for directly attached UNIX and Intel processor-based clustered servers.
- ▶ Up to 64 storage partitions.

- ▶ Enhanced Remote Mirroring is available on the DS4300 Turbo.

Because the DS4300 turbo can scale up to 112 disk drives and 33.6 TB via EXP700/EXP710 (FC) 250 GB, 56 TB via EXP100 (Serial ATA), it can be upgraded from a workgroup SAN to an enterprise network storage system, thereby providing flexibility to grow with your business. In addition, using the DS4000 Storage Manager software, multiple DS4300 Turbo servers can be combined to help address additional performance and capacity requirements further enhancing your scalability options.

7.5.2 Specifications

Table 7-5 shows the specifications of the DS4300 Turbo.

Table 7-5 IBM DS4300 Turbo specifications

Model	1722-60U
RAID controller	Dual active 2 GB RAID controllers
Cache	2 GB total, battery-backed
Host interface	4 Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard
Drive interface	Auto-sensing 1 Gbps/2 Gbps FC-AL
Supported drives	36.4 GB, 73.4 GB, 146.8 GB, 300 GB 10,000 rpm 18.2 GB, 36.4 GB, 73.4 GB, 146.8 GB 15,000 rpm
RAID	levels 0, 1, 3, 5, 10
Storage partitions	8 standard, upgradeable to 16 or 64
Maximum drives supported	112 FC drives (using seven EXP700/710 Expansion Units) 112 SATA drives (using 8 EXP100 Expansion Units)
Fans and power supplies	Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack - 3U

For the latest specification information, check:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4300/index.html>

7.6 IBM System Storage DS4500

The IBM System Storage DS4500 offers investment protection with advanced functions and flexible features. It easily scales from 36 GB to over 67.2 TB of Fibre Channel disk storage or 89.6 TB of SATA disk to support growing storage requirements. DS4500 offers advanced replication services to support business continuance and disaster recovery.

7.6.1 Features

The DS4500 is an enterprise class Disk System designed to provide performance and flexibility for today's demanding data-intensive computing environments. The DS4500 is show in Figure 7-6 on page 89.



Figure 7-6 DS4500 front view

Figure 7-7 shows the DS4500 rear view. It shows two redundant hot-swap 175 watt power supplies, communications/fan module, host-side mini-hubs (two are standard, a maximum of four can be installed), drive-side mini-hubs (two are standard, a maximum of four can be installed), and power supplies.

Four drive modules (mini hubs) may be attached to each controller for an aggregate maximum of 224 drives. The fan/comm module provides serial and Ethernet support and dual redundant fans to cool all the rear-mounted components.

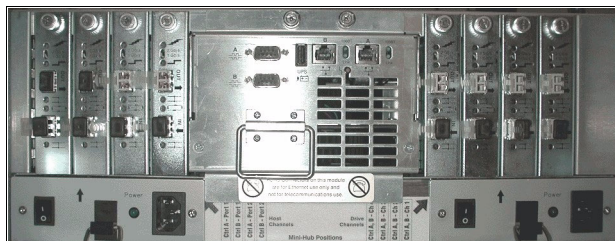


Figure 7-7 DS4500 Rear view

7.6.2 Specifications

- ▶ 2 Gbps FC connectivity and maximum bandwidth of 800 MBps
- ▶ Supports continuous availability through redundant components and mirrored, battery backed-up cache.
- ▶ Up to 64 storage partitions.
- ▶ Supports up to 224 FC drives and provides more than 67 TB using flexible combinations of 18.2, 36.4, 73.4, 146.8 and 300 GB drives in up to 16 DS4000 EXP700 or 710 disk enclosures
- ▶ Supports up to 224 SATA drives and provides more than 89 TB using 400 GB and 250 GB drives in up to 16 DS4000 EXP100 disk enclosures.
- ▶ Includes DS4000 Storage Manager to help manage the storage subsystem.
- ▶ Provides flexibility for multiplatform storage environments by supporting a wide variety of servers, operating systems and cluster technologies. (Support for IBM AIX, Microsoft Windows 2000, Windows 2003, Windows NT®, Novell NetWare, Sun™ Solaris, HP-UX, Red Hat Linux, VMWare, Linux IA64).
- ▶ Offers advanced replication functions such as FlashCopy, VolumeCopy, and Enhanced Remote Mirroring for high data availability and protection.

7.6.3 Specifications

Table 7-6 shows the specifications of the DS4500.

Table 7-6 DS4500 Disk System

Model	1742-90U
RAID controller	Dual active 2 GB RAID controllers
Controller cache	Total of 2 GB, battery-backed
Host interface	Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL)
Connectivity	4 standard, up to 8 via additional host and drive side mini-hubs
Drive interface	FC-AL
Supported drives (internal)	36.4 GB, 73.4 GB, 146.8 GB, 300 GB 10,000 rpm 18.2 GB, 36.4 GB, 73.4 GB, 146.8 GB 15,000 rpm
RAID levels	0, 1, 3, 5 and 10
Storage partitioning	Up to 64
Maximum drives supported	224 (using 16 EXP100 or 16 EXP700/710 Expansion Units)
Fans and power supplies	Dual redundant, hot-swappable
Rack support	Industry standard 19" rack - 4U

For the latest specification information, check:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4500/index.html>

7.7 IBM System Storage DS4700 Express

The IBM System Storage DS4700 Express Model supports a high-performance 4 Gbps FC interface for data-intensive applications that demand increased connectivity. The DS4700 Express 4 Gbps technology is backward compatible with both 2 Gbps and 1 Gbps technologies. Maximum rated bandwidth is 1600 Mbps. The DS4700 Express is available in two models: Model 70 has 2 GB of physical cache memory (1 GB per controller) and the Model 72 has 4 GB of physical cache memory (2 GB per controller). The DS4700 Express supports up to 112 disk drives with the attachment of six DS4000 EXP810 disk enclosures, for over 33.6 TB of physical storage capacity. DS4700 Express supports up to 16 hard disk drive internally in the standard controller. Figure 7-8 on page 91 shows the DS4700 with 16 hard disk drives installed.

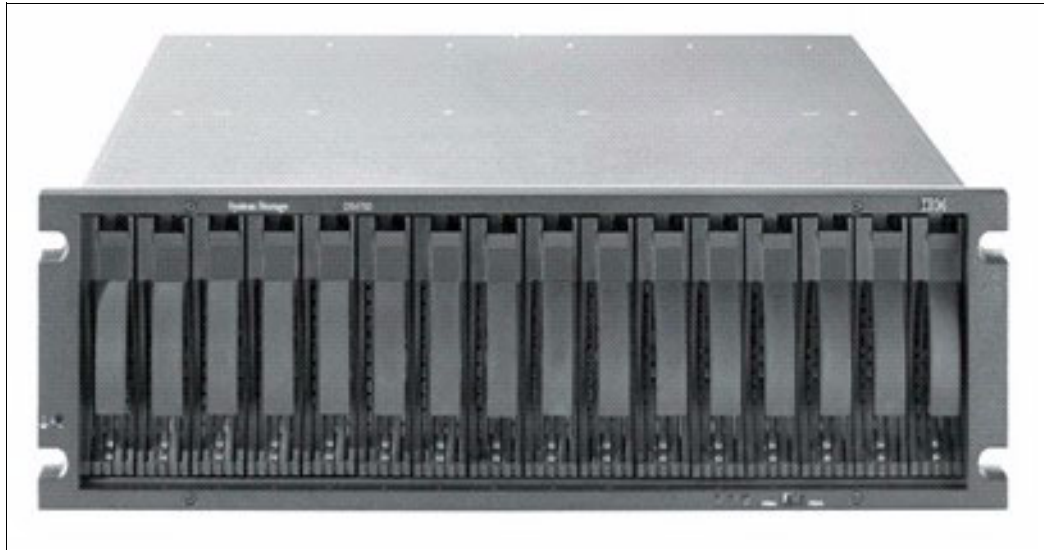


Figure 7-8 DS4700 Express front view

7.7.1 Features

- ▶ Dual RAID controllers – Dual 4 Gbps capable, redundant and hot-swappable, battery backed, write cache mirroring
- ▶ Model 72A supports
 - 4 GB cache - battery backed cache is designed to be protected for at least three days
 - Eight Host/SAN connections side and four drive side ports for two controllers
- ▶ Model 70A supports
 - 2 GB cache - battery backed cache is designed to be protected for at least three days
 - Four Host/SAN connections side and four drive side ports for two controllers
- ▶ Supporting 2 dual redundant FC disk loops
- ▶ Eight or four host ports provides a cluster solution without using a switch.
- ▶ Auto negotiates 1, 2, or 4 Gbps host connection speeds
- ▶ Supported up to 16 disk drives inside the controller, of the following type:
 - 2 Gbps FC: 15K rpm, 146 GB/73 GB/36 GB (E-DDM)
 - 2 Gbps FC: 10K rpm, 300 GB/146 GB/73 GB (E-DDM)
 - 4 Gbps FC: 15K rpm, 146 GB/73 GB/36 GB (E-DDM)
- ▶ Hot pluggable drives - Add new drives for extra capacity without booting server
- ▶ Supports up to 112 FC drives with 6 DS4000 EXP810 Expansion Units
- ▶ Includes DS4000 Storage Manager to help manage the storage subsystem
- ▶ Provides flexibility for multiplatform storage environments by supporting a wide variety of servers, operating systems and cluster technologies. (Support for IBM AIX, Microsoft Windows 2000, Windows NT, Novell NetWare, Sun Solaris, HP-UX, Red Hat Linux, Funware, Linux IA64).
- ▶ Offers advanced replication functions such as FlashCopy and Remote Mirroring for high data availability and protection

Figure 7-9 on page 92 shows the DS4700 model 72x rear view. Each controller has two Ethernet ports, independent slots battery and two modules for power and cooling.

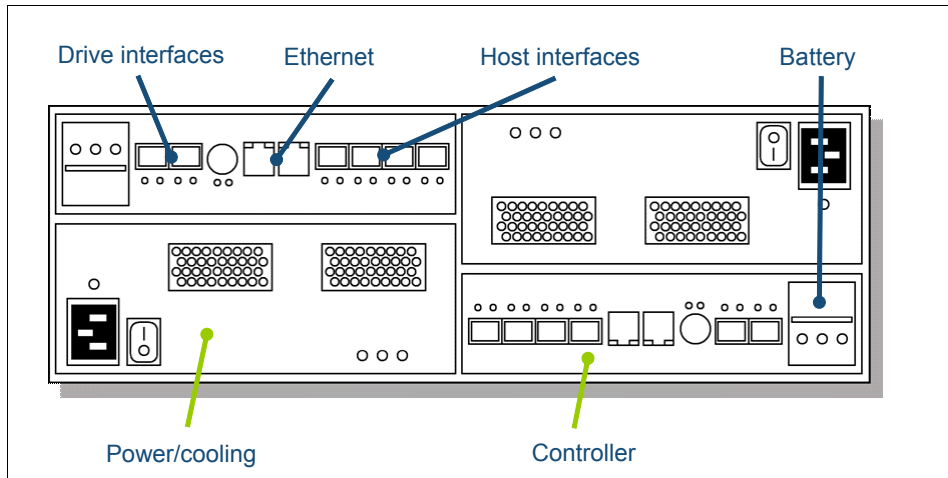


Figure 7-9 DS4700 mod 72x rear view

7.7.2 Specifications

Table 7-7 System Storage DS4700 Express specifications

Model	1814-72H/1814-72A	1814-70H/1814-70A
RAID controller	Dual active 4 GB RAID controllers	Dual active 4 GB RAID controllers
Capacity of base enclosure	16 HDD	16 HDD
Controller cache	Total of 4 GB, battery-backed	Total of 2 GB, battery-backed
Host interface	FC-AL	FC-AL
Connectivity	Eight host side and four drive side ports	Four host side and four drive side ports
Drive interface	FC-AL	FC-AL
Supported drives	2 Gbps FC: 15K rpm, 146GB/73 GB/36GB (E-DDM) 2 Gbps FC: 10K rpm, 300GB/146GB/73GB (E-DDM) 4 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)	2 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM) 2 Gbps FC: 10K rpm, 300GB/146GB/73GB (E-DDM) 4 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)
RAID levels	0, 1, 3, 5 and 10	0, 1, 3, 5 and 10
Storage partitioning	Up to 64	Up to 64
Maximum drives supported	Up to 112 FC or SATA drives	Up to 112 FC or SATA drives
Fans and power supplies	Dual redundant, hot-swappable	Dual redundant, hot-swappable
Rack support	Industry standard 19" rack - 3U	Industry standard 19" rack - 3U

For the latest specification information, check:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4700/index.html>

7.8 IBM System Storage DS4800

The IBM System Storage DS4800 delivers high performance, advanced function, and high availability, modular and scalable storage capacity, with SAN attached 4 Gbps FC connectivity. A 4U rack-mountable enclosure houses the DS4800 redundant, dual-active intelligent RAID controllers equipped with sixteen 4 Gbps ports (eight ports per controller), for attachment of host servers and DS4000 Storage Expansion Unit enclosures. This disk system supports RAID levels 0, 1, 3, 5, and 10. The DS4800 represents the sixth-generation architecture of the DS4000 series family. The DS4800 attaches up to sixteen DS4000 Storage Expansion Units which can be either FC or SATA, or a combination of both types by use of the optional DS4800 FC/SATA Enclosure Intermix feature. Advanced DS4000 storage management, copy service options, and optional advanced disaster recovery functions are available for the DS4800, including FlashCopy, VolumeCopy, and Enhanced Remote Mirroring. The DS4800 RAID controller cache size can be either 4 GB, 8 GB or 16 GB, depending on the DS4800 model. Depending on the model, the DS4800 either ships with the 8-storage-partition premium feature or can be ordered with 8-, 16-, or 64-storage-partition premium features. The DS4800 storage system's eight 4 GBps drive connections support up to 224 disk drives with the attachment of 14 EXP810, 16 EXP710, or 16 EXP100 disk enclosures. Each of the four models support over 67.2 TB FC physical storage capacity and 89.6 TB of SATA physical storage capacity. Figure 7-10 shows the DS 4800 unit.



Figure 7-10 IBM System Storage DS4800

The IBM System Storage DS4800 contains two controllers, two power supply-fans, one interconnect-battery unit. The RAID controllers are located at the back of the unit. The two power supply-fans and the interconnect-battery unit are located behind the front bezel. Figure 7-11 on page 94 shows the front bezel, a front view without the cover, and a back view, as well as the locations of the components of DS4800.

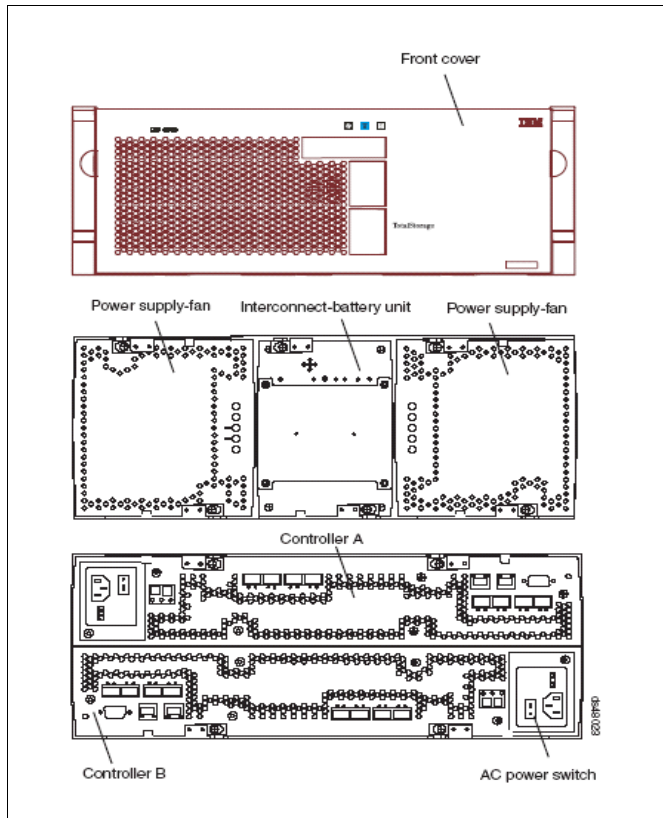


Figure 7-11 DS4800 front and back views

7.8.1 Features

- ▶ Dual RAID controllers – Dual 4 Gbps capable, redundant and hot-swappable, battery backed, write cache mirroring
- ▶ Model 80A supports 4 GB cache
- ▶ Model 82A supports 4 GB cache
- ▶ Model 84A supports 8 GB cache
- ▶ Model 88A supports 16 GB cache
- ▶ Supporting 2 dual redundant FC disk loops
- ▶ Host interface 8 host ports—Fibre Channel (FC) Switched and FC-AL standard,
- ▶ Auto negotiates 1, 2, or 4 Gbps host connection speeds
- ▶ Supported drives
 - 36.4 GB, 73.4 GB, 146.8 GB and 300 GB 10,000 rpm (Fibre Channel)
 - 36.4 GB, 73.4 GB and 146.8 GB 15,000 rpm (Fibre Channel)
 - 400 GB and 250 GB 7,200 rpm (SATA)
 - 2 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)
 - 2 Gbps FC: 10K rpm, 300GB/146GB/73GB (E-DDM)
 - 4 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)
- ▶ RAID levels 0, 1, 3, 5, 10
- ▶ Storage partitions 8, 16 or 64 storage partitions—choice of 8, 16 or 64 required. Upgrade: 8–16, 16–64

- ▶ Maximum drives supported by all models: 224 FC drives (using 14 DS4000 EXP810, 16 EXP710 Expansion Units)
- ▶ Dual redundant, hot-swappable fans and power supplies
- ▶ Provides flexibility for multiplatform storage environments by supporting a wide variety of servers, operating systems and cluster technologies. (Support for IBM AIX, Microsoft Windows 2000, Windows NT, Novell NetWare, Sun Solaris, HP-UX, Red Hat Linux, VMWare, Linux IA64).
- ▶ Offers advanced replication functions such as FlashCopy, Remote Mirroring, Enhanced Remote Copy for high data availability and protection

7.8.2 Specifications

Table 7-8 shows the DS4800 specifications

Table 7-8 IBM System Storage DS4800 specifications

Model	1815-80A, 1815-82A, 1815-84A, 1815-88A
RAID controller	Dual active
Cache	Model 80A: 4 GB battery-backed Model 82A: 4 GB battery-backed Model 84A: 8 GB battery-backed Model 88A: 16 GB battery-backed
Host interface	8 host ports—Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard, auto-sensing 1 GBps/2 GBps/4 GBps
Drive interface	8 drive ports—Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard, auto-sensing 2 GBps/4 GBps
Supported drives	36.4 GB, 73.4 GB, 146.8 GB and 300 GB 10,000 rpm (Fibre Channel) 36.4 GB, 73.4 GB and 146.8 GB 15,000 rpm (Fibre Channel) 400 GB and 250 GB 7,200 rpm (SATA) 2 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM) 2 Gbps FC: 10K rpm, 300GB/146GB/73GB (E-DDM) 4 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)
RAID	levels 0, 1, 3, 5, 10
Storage partitions	8, 16 or 64 storage partitions—choice of 8, 16 or 64 required Upgrade: 8–16, 16–64
Maximum drives supported	All models: 224 FC drives (using 14 DS4000 EXP810, or 16 EXP710 Expansion Units)
Fans and power supplies	Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack - 4U

For the latest specification information, check:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4800/index.html>

7.9 IBM System Storage DS4000 EXP100 Expansion

DS4000 EXP100 Storage Expansion Unit provides SATA disks to the DS4000 product line, including the IBM DS4100, IBM DS4300, IBM DS4400, and IBM DS4500.

The EXP100 connects to the DS4000 disk system using redundant 2 Gbps Fibre Channel connections, supporting increased reliability and performance as compared to older, non-redundant parallel Advanced Technology Attachment (ATA) products.

Note: IBM System Storage DS4000 EXP100 Express Model Expansion Unit has standard configurations, and is designed to be easy to acquire, install, and manage. For the latest information, see:

<http://www.ibm.com/servers/storage/disk/ds4000/exp100/express/index.html>

7.9.1 Features

The EXP100 is designed for mid-range and enterprise application environments that need lower cost storage, but with the availability, manageability, data integrity and software/hardware features achieved with Fibre Channel based storage systems.

- ▶ For data archival, data reference, and near-line storage applications
- ▶ Redundant, hot-swappable power and cooling components
- ▶ 3U, 19 inch rack-mountable enclosure supporting up to 14 disks per enclosure
- ▶ Support for AIX, Windows, and Linux environments
- ▶ Autonomic functions such as Dynamic Volume Expansion and Dynamic Capacity Addition, allowing unused storage to be brought online without stopping operations, and DS4000 Service Alert, which is capable of automatically alerting IBM if a problem occurs

Figure 7-12 shows the EXP100.



Figure 7-12 EXP100 storage expansion unit

7.9.2 Hardware description

The EXP100 SATA disk enclosure supports up to fourteen 250 GB, 7200 RPM SATA disk drives or/and 400 GB, 7200 RPM SATA disk drives, offering up to 3.5 TB per 3U enclosure. It attaches to the DS4500, DS4300 and DS4100 Disk Systems via dual-active Fibre Channel drive loops.

Current SATA disk drives are single-ported and lack true multipathing and multi-initiator capabilities. But the SATA Interface Card, provides the drives in the EXP100 the dual-link capability and switching functionality, needed to achieve redundant I/O paths. The SATA

Interface Card also provides the SATA drives with basic Fibre Channel-like functionality not otherwise available, enabling them to be managed like Fibre Channel drives.

Note: The 250 GB 7200 rpm SATA drives were withdrawn from marketing on December 16, 2005 so are no longer available for new orders.

The DS4000 EXP100 has several removable components. These components, called Customer Replaceable Units (CRUs), are accessible from the front or back of the DS4000 EXP100. The components are:

- ▶ Up to fourteen SATA hard disk drives.
- ▶ Two Environmental Services Monitors (ESMs).
- ▶ Two power supplies with two fans.

These hot-swap features of the DS4000 EXP100 enable the removal and replacement of the SATA hard disk drives, power supplies, ESMs, and fan units without turning off the expansion unit, maintaining the availability of the system while a hot-swap device is removed, installed, or replaced. The EXP100 ESMs and user controls are shown in Figure 7-13

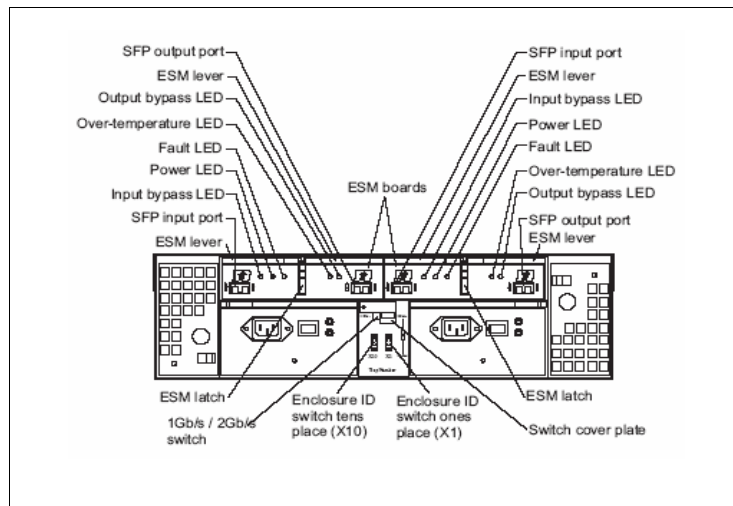


Figure 7-13 DS4000 EXP100 Back view

7.9.3 Specifications

Table 7-9 shows the specifications of the EXP100

Table 7-9 IBM DS4000 EXP100 expansion unit specifications

Model	1710-10U
Number of drives supported	Up to 14 SATA drives (250 GB, 400 GB)
Fans and power supplies	Dual redundant, hot-swappable
Dimensions	(with bezel) 13.23 cm H x 48.18 cm W x 59.74 cm D (5.21 in x 18.97 in x 23.52 in)
Weight	Fully configured weight (with 14 disk drives installed): 40.2 kg (88.4 lb)

7.9.4 SATA versus FC disks

Without any doubt, technical characteristics and performance of FC disks remain for now superior to those of SATA disks. However, not all storage applications require the superior features of Fibre Channel. When used for the appropriate enterprise applications, SATA disks offers a tremendous cost advantage over FC. First, SATA drives are cheaper to manufacture and because of their larger individual capacity, SATA drives are on average sixty percent cheaper per gigabyte than FC disks. The fact is that in large capacity systems, the drives themselves account for the vast majority of the cost of the system. Using SATA disks will substantially reduce the TCO of the storage system.

Classes of storage

Storage data can reside at three different locations within the network storage hierarchy. This is also known as tiered storage, shown in Figure 7-14.

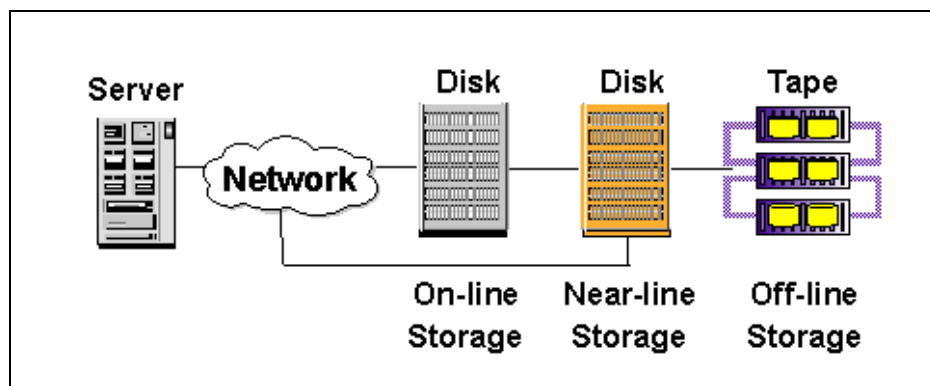


Figure 7-14 Storage data in a networks storage environment

Particular data types are suitable for storage at the various levels.

► **Online (primary) storage**

Best suited for *business critical* applications that require constant instantaneous access to data, such as databases and frequently accessed user data. This data requires continuous availability and typically has high performance requirements. Business-critical data will be stored on Fibre Channel disk implemented in enterprise-class storage solutions.

► **Near-Line (secondary) storage**

Used for *business important* applications that require quicker access compared with offline storage (as tape), but do not require the continuous, instantaneous access provided by online storage. Secondary storage represents a large percentage of a company's data and is an ideal fit for SATA technology.

► **Offline (archival) storage**

Used for applications where infrequent serial access is required, such as backup for long-term storage. For this type of storage, tape remains the most economical solution.

Conclusion: Data storage implementations best suited to utilize SATA technology, reside at the “near-line” or secondary location within the network storage hierarchy, and offer a cost-effective alternative to FC disks at that location.

General characteristics for primary, secondary and archival storage in traditional IT environments are summarized in Table 7-10 on page 99.

Table 7-10 Storage classes in traditional IT environments

Class of Storage	On-line	Near-line	Off-line
Primary media	FC disk	SATA disk	Tape
Price	Highest	Low cost-per-GB	Lowest
IOPS performance	Highest	Minimal	N/A
MBps performance	Highest	High	Lowest
Time to data	Immediate	Close to immediate	Drive mount time
Media reliability	Highest	Good	Good - lower
Uptime	24/7	< 24/7	< 24/7
Typical applications	ERP/RDBMS	Fixed content	Archive retrieval

7.10 IBM System Storage DS4000 EXP710 Expansion Unit

The EXP710 (Figure 7-15 on page 100) is a rack-mountable storage expansion enclosure which is an enhanced *switched bunch of disks* (SBOD) version of the EXP700.

Many SANs use FC-AL within the storage systems to connect the individual storage devices (disks or tapes) to the controllers. FC-AL can daisy-chain up to 125 drives in a loop (or on a hub) on one controller. This architecture is commonly referred to as a *Just a Bunch of Disks* (JBOD) or spanning. However, this architecture can cause performance problems as the number of drives increases and can also make it difficult to diagnose problems. To improve performance, a FC-AL switch can be used in place of the FC-AL hub and this allows the controller to establish point-to-point connections between each storage device and the controller. The result is higher bandwidth, better performance and full isolation of the drives, eliminating the risk of a single drive disrupting the loop and causing other drives to fail. It is commonly referred to as a *Switched Bunch of Disks* (SBOD).

Note: IBM System Storage DS4000 EXP710 Express Model Expansion Unit features standard configurations, designed to be easy to acquire, install. For the latest information, see:

<http://www.ibm.com/servers/storage/disk/ds4000/exp710/express/>

7.10.1 Features

The EXP710 supports the 2 Gbps Fibre Channel connection and holds 14 slim-line FC disks. It provides 4.2 TB of disk storage capacity per 3U rack-mountable enclosure. It can be attached to the DS4800, DS4500, DS4400, DS4300 Turbo, and DS4300 disk systems.

The main difference between the EXP700 and the EXP710 is the physical ESMs, and the firmware that is installed on them. In fact, as the two expansion units share so many components, the EXP700 is easily upgradeable to the EXP710. To do this, upgrade the Disk System to V9.1 or later, take the EXP700's offline, replace the ESMs, re-cable the expansion unit, then power on.



Figure 7-15 EXP710 expansion unit

7.10.2 Specifications

Table 7-11 IBM DS4000 EXP710 specifications

Model	1740-710
Number of drives supported	Up to 14 FC drives (18.2, 36.4, 73.4, 146.8 or 300 GB)
Fans and power supplies	Dual redundant, hot-swappable
Dimensions	(with bezel) 13.23 cm H x 48.18 cm W x 59.74 cm D (5.21 in x 18.97 in x 23.52 in)

7.11 IBM System Storage DS4000 EXP810 Expansion Unit

The IBM System Storage DS4000 EXP810 storage expansion enclosure (Machine Type 1812, Models 81A and 81H) implements high-capacity, Fibre Channel disk storage. The expansion unit delivers fast, high-volume data transfer, retrieval, and storage functions for multiple drives to multiple hosts. The EXP810 allows continuous, reliable service, using hot-swap technology for easy replacement without shutting down the system and support redundant, dual-loop configurations. The EXP810 is a rack-mountable enclosure that supports up to 16 2 Gbps fibre channel 300 GB disk drive modules, with up to 4.8 TB of capacity per enclosure. The EXP810 has redundant 2 or 4 Gbps Fibre Channel connections, and is supported on the DS4800 and DS4700 Express. The EXP810 supports redundant, dual-loop configurations with the DS4000s and other EXP810 and EXP710 enclosures. The storage expansion enclosure provides continuous, reliable service, and uses hot-swap technology for easy replacement without the need to shut down the system. Up to seven EXP810 units can be connected together in a fibre channel loop, providing connections to a maximum of 112 hard drives. Figure 7-16 shows the EXP810.

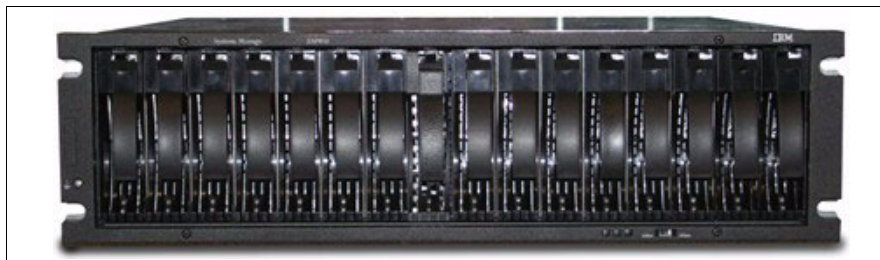


Figure 7-16 DS4000 EXP810 Expansion Unit

7.11.1 Features

- ▶ Machine type:1812, Model number:81A/81H
- ▶ 2 Gbps and 4 Gbps Fibre Channel drives supported
- ▶ Supported drives 2 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)
- ▶ Supported drives 2 Gbps FC: 10K rpm, 300GB/146GB/73GB (E-DDM)
- ▶ Supported drives 4 Gbps FC: 15K rpm, 146GB/73GB/36GB (E-DDM)
- ▶ Supported on the DS4800 or DS4700 Express at V6.16 firmware and higher
- ▶ Can be intermixed with EXP710
- ▶ New design supporting up to 16 drives in 3U enclosure
- ▶ Up to 4.8 TB physical capacity per expansion unit using sixteen 300 GB disk drives
- ▶ Dual Environmental Service Modules (ESM) - 4 Gb FC
- ▶ Field replaceable components
- ▶ Planned NEBS Level 3 certification
- ▶ Approximately 14% storage density increase over EXP710 in the same drive enclosure
- ▶ Figure 7-17 shows the rear components of DS4000 EXP810

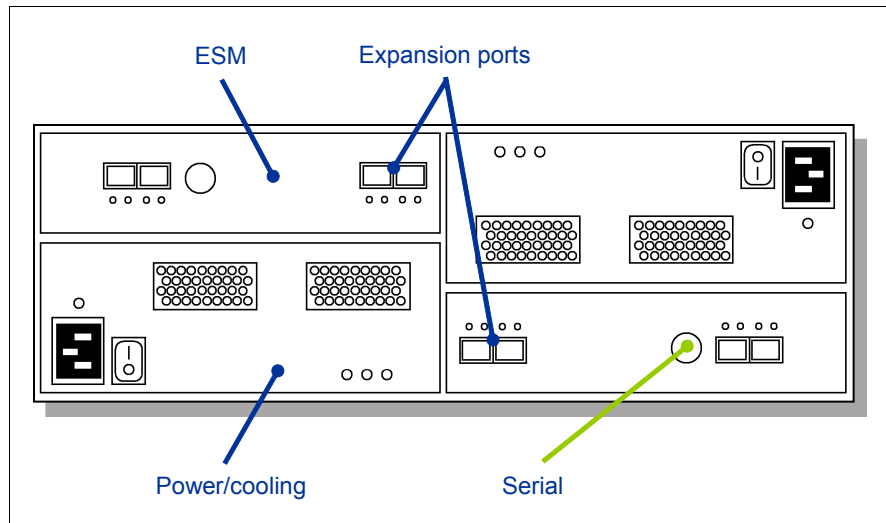


Figure 7-17 DS4000 EXP810 rear components

7.11.2 Specifications

Table 7-12 IBM System Storage DS4000 EXP810 Expansion Unit

Model	1812,-81A/81H
Number of drives supported	Up to 16 FC drives
Fans and power supplies	Dual redundant, hot-swappable
Dimensions	(with bezel) 12.95 cm H x 48.26 cm W x 57.15 cm D (5.1 in x 19 in x 22.5 in)
Weight	40 kg (88 lb)

7.12 DS4000 Storage Manager

The IBM DS4000 Storage Manager software is used to configure, manage and troubleshoot the DS4000 Disk System. Its front-end component, the Storage Manager client, is a Java™ based management tool that is available for various operating systems.

It is used primarily to configure RAID arrays and logical drives, assign logical drives into storage partitions, replace and rebuild failed disk drives, expand the size of the arrays and logical drives and convert from one RAID level to another. It allows troubleshooting and management tasks, like checking the status of the Disk System components, updating the firmware of the RAID controllers and managing the Disk System. Finally, it offers advanced functions such as FlashCopy, Enhanced Remote Mirroring and VolumeCopy (these are premium features that need to be purchased).

DS4000 Storage Manager 9.16 is the most recent version available at the time of writing this book. SM 9.16 supports controller firmware versions 4.01.xx.xx - 06.16.xx.xx. DS4000 Storage Manager 9.16 (with controller firmware 06.16.xx.xx) provides support for attachment of the DS4000 EXP710 storage expansion enclosure and the new EXP810 storage expansion enclosures to DS4800 storage subsystems. SM 9.16 supports automatic ESM firmware synchronization with the new EXP810 storage expansion enclosure. When you install a new ESM into an existing EXP810, in a DS4000 storage subsystem that supports automatic ESM firmware synchronization, this feature resolves firmware mismatch conditions by automatically synchronizing the firmware in the new ESM with the firmware in the existing ESM.

Figure 7-18 shows the most recent version of Storage Manager with a DS4800

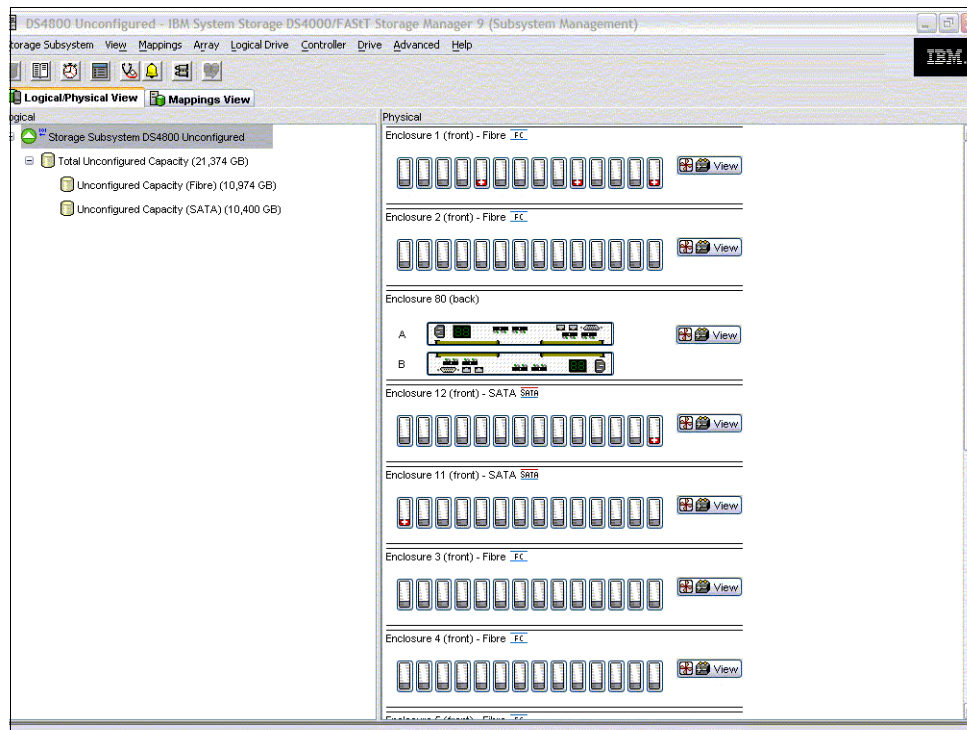


Figure 7-18 Storage Manager 9.16

Latest enhancements to the Storage Manager include:

▶ **Support for mixed drive configurations**

It is now possible to intermix Fibre Channel (FC) drives and SATA drives in one Disk System. The flexibility to intermix different drive types (FC and SATA) in one Disk System gives you the ability to use the advantages of both drive technologies. For example, it is now possible to have the primary (production) storage on FC drives and the secondary (near line or archiving) storage on SATA drives without the need of having different, separate Disk Systems. Or the FlashCopy or VolumeCopy of an array made of FC drives can be created on cheaper SATA drives.

▶ **Support for Microsoft VDS/VSS**

Microsoft has defined an interface layer for the purpose of integrating storage system functions with the Windows Server 2003 operating system. The ultimate aim is to integrate snapshot and capacity management functions into applications, such as Microsoft Exchange, within their applications. Storage Manager 9.12 or higher supports VDS/VSS.

– Volume Shadow Copy Service (VSS)

- Provides Point-in-Time (PiT) copies
- Works with applications to suspend for a few seconds.
- Uses snapshot (FlashCopy) services from the storage subsystem.

– Virtual Disk Service (VDS)

- Common disk interface for block storage
- Common storage management across vendors
- API for third party management tools

▶ **Task assistants**

The task assistants provide a task-oriented starting point for carrying out common storage management operations - helping to simplify storage administration and reduce TCO.

With the Enterprise Management Window Task Assistant, you can add/name/rename storage subsystems, set up an alert destination, and manage a storage subsystem. This last option launches the Subsystem Management Window Task Assistant. With this assistant, you can configure a storage subsystem, define hosts, create new storage partitions, map additional logical drives, and other common tasks.

▶ **New RDAC driver architecture**

The new RDAC driver, sometimes referred to as the Multi-Path Proxy (MPP) driver is a new multipath driver architecture that is used in the new versions of multipathing drivers (a.k.a. RDAC) for Windows 2000/2003 and Linux. The driver was completely rewritten from scratch to include some improvements and to be able to accommodate future planned enhancements in the device driver.

▶ **Full command line interface**

Any task that can be performed with the Storage Manager GUI can now also be performed through the Command Line Interface.

▶ **Controller service mode**

In previous versions of Storage Manager it was not possible to do any problem determination on a “live” controller, as the controller was either online (processing IO) or failed (not processing IO). Now, if a controller is suspected as the root cause of a problem it can be set into service mode and IBM support can run low-level shell commands to further analyze the errors.

► **Support for VERITAS ASL 4.0**

Previous versions of the DS4000 Storage Manager support a component called the VERITAS Array Support Library (ASL), which is essentially a vendor-supplied plug-in for the VERITAS Volume Manager (a.k.a. VxVM) Dynamic Multipathing Driver (DMP). VxVM 4.0 has defined some new capabilities of the ASL which are supported in Storage Manager 9.10 and later by the feature known as “VERITAS ASL 4.0”.

Other important features of Storage Manager include:

► **Dynamic Volume Expansion**

Dynamic Volume Expansion (DVE) is a modification operation that you can use to increase the capacity of a standard logical drive or a FlashCopy repository logical drive. To increase the capacity, a DVE operation uses the free capacity that is available on the logical drive group of the logical drive. The operation is considered to be dynamic because you can continuously access data on logical drive groups, logical drives, and disk drives during the operation.

► **Dynamic Capacity Expansion**

The ability to increase the available free capacity on an array Dynamic Capacity Expansion (DCE) without needing to restart the host system is a very important feature. In today’s IT environment, the need for storage space grows constantly. Many customers exhaust their existing space sooner or later and have to expand their storage capacity. It is essential that this process be nondisruptive and not cause any downtime. With DS4000 Storage Manager, you can simply add new disk drives to the Disk System and start the expansion procedure while the system remains fully operational. Once the procedure starts, you cannot stop it. Be aware that you may see some performance impact because the expansion process competes with normal disk access.

Further dynamic volume/array reconfiguration such as changing the logical drive segment size (DSS) and changing the RAID level (DRM) are discussed in more detail in the redbook *IBM System Storage DS4000 Series, Storage Manager and Copy Services*, SG24-7010.

7.13 Copy services

Table 7-13 shows the support for FlashCopy, VolumeCopy, and the Enhanced Remote Mirror Option on the DS4100, DS4300, DS4300 Turbo, DS4400, DS4500, DS4700, and DS4800 Storage Subsystems. This support matrix requires that the DS4000 Storage Subsystem controller firmware be at version 06.1x.xx.xx and the Storage Manager software be at version 9.16.

Table 7-13 DS4000 Copy Services f

Feature	DS4100 Base	DS4300	DS4300 Turbo	DS4400 DS4500 DS4700 DS4800
FlashCopy	Supported	Supported	Supported	Supported
Maximum number of logical drives that can be defined	1024	1024	1024	2048
Maximum total FlashCopy logical drives	512	512	512	1024

Feature	DS4100 Base	DS4300	DS4300 Turbo	DS4400 DS4500 DS4700 DS4800
Maximum FlashCopy logical drives per base RAID logical drive	4	4	4	4
VolumeCopy	Supported	Not supported	Supported	Supported
Maximum number of VolumeCopy target logical drives for a given source logical drive	1023	1023	1023	2047
Maximum copy pairs per array	1023	1023	1023	2047
Maximum running copies per array	8	8	8	8
Remote Mirror Option	Supported	Not supported	Supported	Supported
Total Remote Mirror logical drives	32	N/A	32	64

Refer to the *Storage Manager Version 9 Copy Services User's Guide*, GC26-7707, for a more detailed explanation of copy services for the DS4000 family of servers.

7.13.1 FlashCopy

A FlashCopy logical drive is a point-in-time image of a logical drive. It is the logical equivalent of a complete physical copy, but it is created much more quickly, and requires less disk space. In DS4000 Storage Manager, the logical drive from which you are basing the FlashCopy, called the *base logical drive*, must be a standard logical drive in the storage subsystem. Typically, you create a FlashCopy so that an application (e.g., backup software) can access the FlashCopy and read the data while the base logical drive remains online and user-accessible. When the backup completes, the FlashCopy logical drive is no longer needed.

When you take a FlashCopy, the controller suspends I/O to the base logical drive for only a few seconds. At the same time, it creates a new logical drive called the *FlashCopy repository logical drive* where it stores FlashCopy metadata and copy-on-write data. When the controller finishes creating the FlashCopy repository logical drive, I/O write requests to the base logical drive can resume. However, before a data block on the base logical drive is modified, a copy-on-write occurs, copying the contents of blocks that are to be modified into the FlashCopy repository logical drive, for safekeeping. Since the FlashCopy repository logical drive stores copies of the original data in those data blocks, further changes to those data blocks write directly to the base logical drive without another copy-on-write. And, since the only data blocks that are physically stored in the FlashCopy repository logical drive are those that have changed since the time of the FlashCopy, the FlashCopy technology uses less disk space than a full physical copy.

Figure 7-19 on page 106 shows how this works.

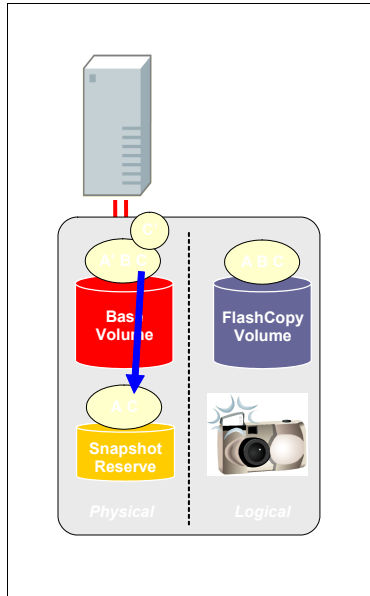


Figure 7-19 How DS4000 FlashCopy works

The Storage Management software issues a warning message when the FlashCopy repository logical drive nears a user-specified threshold (a percentage of its full capacity; the default is 20 percent). When this condition occurs, you can use the Storage Manager software to expand the capacity of your FlashCopy repository logical drive from free capacity on the array.

7.13.2 VolumeCopy

The VolumeCopy feature is a firmware-based mechanism for replicating logical drive data within a storage subsystem. This feature is designed as a system management tool for tasks such as relocating data to other drives for hardware upgrades or performance management, data backup, and restoring snapshot logical drive data.

A VolumeCopy creates a complete physical replication of one logical drive (source) to another (target) within the same storage subsystem. The target logical drive is an exact copy or *clone* of the source logical drive.

Because VolumeCopy is a full point-in-time replication, it allows for analysis, mining, and testing without any degradation of the production logical drive performance. It also brings improvements to backup and restore operations, making them faster and eliminating I/O contention on the primary (source) logical drive.

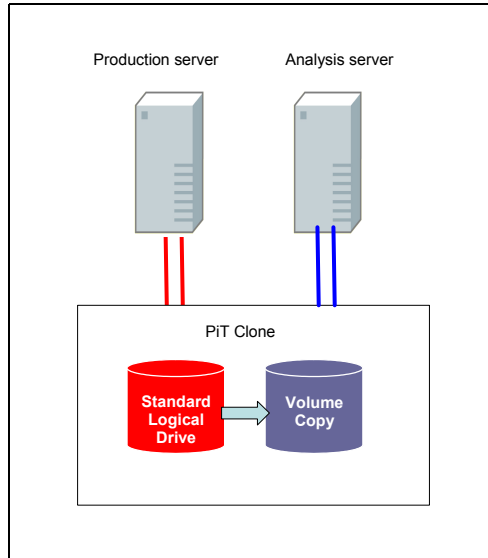


Figure 7-20 VolumeCopy

The VolumeCopy premium feature must be enabled by purchasing a feature key. For efficient use of VolumeCopy, FlashCopy must be installed as well, so VolumeCopy is only available as a bundle which includes a FlashCopy license.

When a background process is reading the source logical drive and writing the data to the target logical drive, host I/O access to the source and target logical drives are restricted:

- ▶ All write requests to a source logical drive of a copy request at the time of copying are rejected.
- ▶ All write and read requests to the target logical drive while the copying is in progress are rejected.

These restrictions are necessary to ensure the integrity of the point-in-time copy. If the logical drive being copied is large, this can result in an extended period of time without the ability to make updates or changes to the data. As shown in Figure 7-21 on page 108, FlashCopy which allows a point-in-time copy to be made while maintaining read/write access, enables a complete copy to be created without interrupting the I/O activity of the production logical drive.

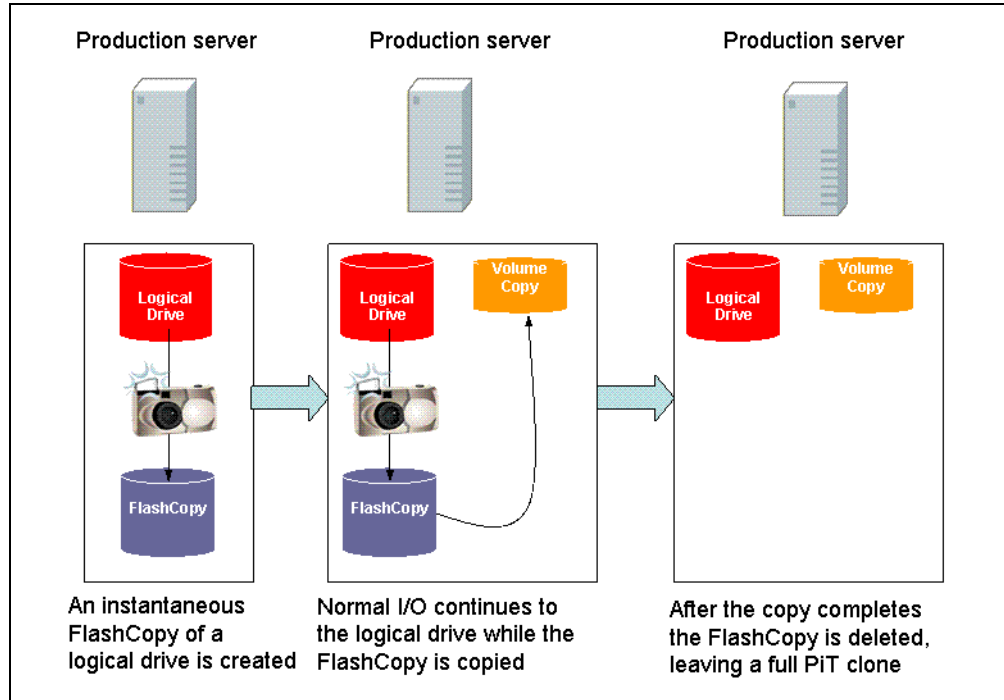


Figure 7-21 VolumeCopy integration with FlashCopy

7.13.3 Enhanced Remote Mirroring

The Enhanced Remote Mirroring (ERM) option is a premium feature of the IBM DS4000 Storage Manager 9.12 or higher software and can be enabled by purchasing a premium feature key.

The ERM option is used for online, real-time replication of data between storage subsystems over a remote distance (Figure 7-22 on page 109). In the event of disaster or unrecoverable error at one storage subsystem you can promote the second storage subsystem to take over responsibility for normal I/O operations. Currently, 64 mirror relationships are supported per DS4000 system.

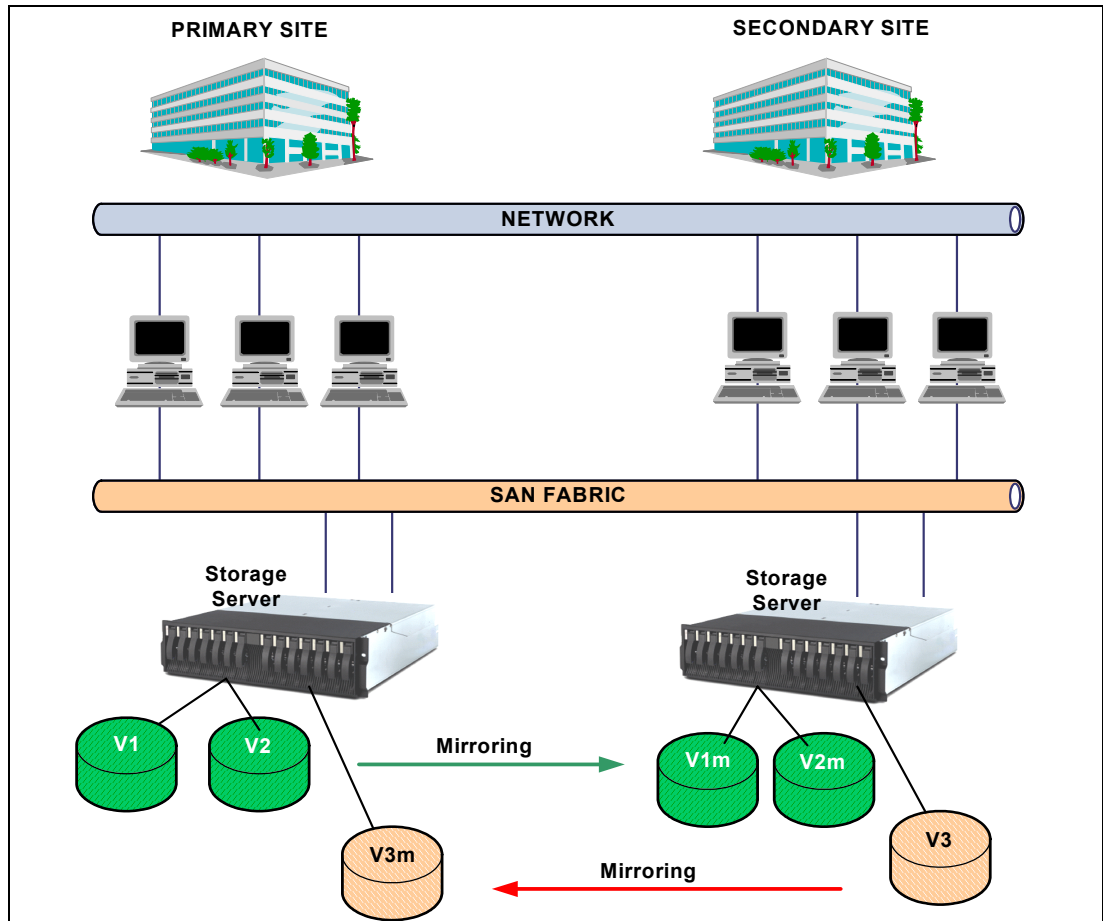


Figure 7-22 Enhanced Remote Mirroring

The Enhanced Remote Mirroring is a redesign of the former Remote Volume Mirroring (RVM) and offers three different operating modes:

► **Metro Mirror**

Metro Mirroring is a synchronous mirroring mode. Any host write request are written to the primary (local) storage subsystem and then transferred to the secondary (remote) storage subsystem. The remote storage controller reports the result of the write request operation to the local storage controller which reports it to the host. This mode is called synchronous, because the host application does not get the write request result until the write request has been executed on both (local and remote) storage controllers. This mode corresponds to the former RVM functionality.

► **Global Copy**

Global Copy is an asynchronous write mode. All write requests from a host are written to the primary (local) storage subsystem and immediately reported as completed to the host system. Regardless of when data was copied to the remote storage subsystem, the application does not wait for the I/O commit from the remote site. However, Global Copy does not ensure that write requests performed to multiple drives on the primary site are later processed in the same order on the remote site. As such, it is also referred to as Asynchronous Mirroring without Consistency Group. This contrasts with the previous Remote Volume Mirroring feature which only supported synchronous data transfers, where the write I/O had to complete on both the local and remote subsystems before it was acknowledged as complete.

► **Global Mirror**

Global Mirroring is an asynchronous write that ensures that the write requests are carried out in the same order at the remote site. This mode is also referred to as Asynchronous Mirroring with Consistency Group.

The new Enhanced Remote Mirroring has also been equipped with new functions for better Business Continuity and easier maintenance tasks. These functions include:

Secondary logical drive accessibility

With the former RVM, a secondary logical drive was not accessible to any host, even if you mapped the drive to the host. With the new ERM, this behavior is changed and a read-only access is now possible.

Suspend Mirror and Resume Mirror capability

With the previous version of Remote Mirroring, any sort of communication failure, planned or unplanned, required that the mirror be reestablished block-by-block from the beginning. Enhanced Remote Mirroring available with version 9.10 and later firmware (06.10) supports a more robust resume operation, also referred to as a delta re synchronization, where only the data written to the primary logical drive while the mirror was suspended, or broken, will be sent to the remote site once the mirror is restored. This includes a mirror that was either manually suspended or stopped due to an unplanned communication loss. Since the data on the secondary logical drive is not changing, you can access the “frozen” secondary logical volume as read-only and use it for test purposes or to back it up. The new Enhanced Remote Mirroring has a delta log which tracks updates to the primary logical drive during the planned or unplanned communication interruption. Once communications is re-established, you can invoke the Resume Mirror function. It will re synchronize the changed data between the primary and the secondary logical drives. No full synchronization takes place.

Change Write Mode option

You can switch among the different mirroring modes at any time, for an established mirror relationship. This is called *Dynamic Mode Switching*. You can switch between:

- Metro Mirroring (synchronous write mode)
- Global Copy (asynchronous write mode without Consistency Groups)
- Global Mirroring (asynchronous write mode with Consistency Groups)

Test Mirror communication function

After the mirroring relationship between the primary and the secondary logical drive is established, it is possible and easy to test the mirror communication status. Using the SM GUI client, you can perform the test with only one mouse click, and the result is graphically displayed (green or red traffic light picture).

Enhanced Hardware Compatibility for ERM

ERM is available for the DS4800, DS4700, DS4500, DS4400, and DS4300 Turbo. On the DS4300 Turbo it supports 32 mirror pairs while on the other models, it supports 64 mirror pairs.

7.14 DS4000 Service Alert

DS4000 Service Alert (hereafter called Service Alert) is a feature of the IBM System Storage DS4000 Storage Manager that monitors system health and automatically notifies the IBM Support Center when problems occur. Service Alert sends an e-mail to a call management center that identifies your system and captures any error information that can identify the

problem. The IBM support center analyzes the contents of the e-mail alert and contacts you with the appropriate service action.

Service Alert complements but does not replace the basic hardware maintenance agreement in place for the DS4000 storage subsystems. With Service Alert activated, the IBM support center will monitor Service Alert e-mails with the same coverage being provided in the basic hardware maintenance agreement. The service is available worldwide to all current and new DS4000 Disk System customers.

DS4000 Service Alert requires a services contract. Please contact your IBM representative for further information regarding the service offering.

7.15 DS4000 Alert Manager

IBM System Storage DS4000 Alert Manager, a services solution designed to support the remote monitoring of installed DS4000s, expands the capability beyond the existing DS4000 Service Alert. DS4000 Alert Manager is designed to automatically notify IBM Service and Support of problems. DS4000 Alert Manager enables IBM Service and Support to dial in to the DS4000 information logs to aid problem determination.

DS4000 Alert Manager allows two-way communication between the DS4000 controller and IBM remote support. This “appliance” is an integrated solution designed to support the remote monitoring of installed DS4000s. The appliance is designed to automatically notify IBM Service when the DS4000 issues an alert. IBM Service can electronically contact the appliance (including DS4000 event log files) through a modem connection and request the appliance software to obtain information about the alert. Event log files and additional information are transmitted back to IBM Service, from the appliance, to assist IBM Service with problem determination and problem source identification. The appliance provides a heartbeat function that allows IBM Service to periodically contact the appliance and determine whether it is still functioning.

Highlights of DS4000 Alert Manager:

- ▶ Designed to help prevent or reduce possible downtime
- ▶ Provide alert notification and remote service access for DS4000s
- ▶ Monitor appliance and automatically notifies IBM Service of problems
- ▶ Allows service to dial in to obtain information and logs to aid problem determination
- ▶ Provide a heartbeat function to determine whether monitoring appliance is functional
- ▶ Capable of supporting up to four DS4000s if within fifty feet of appliance

7.16 More information

More detailed information about the DS4000 family can be found in the following publications and Web sites:

<http://www.ibm.com/servers/storage/disk/ds4000/index.html>

- ▶ IBM Redbook *Fibre Array Storage Technology A FASTT Introduction*, SG24-6246
- ▶ IBM Redbook *IBM System Storage DS4000 Series, Storage Manager and Copy Services*, SG24-7010
- ▶ Redpaper *Introducing IBM TotalStorage FASTT EXP100 with SATA Disks*, REDP-3794
- ▶ Publication *Storage Manager Version 9 Copy Services User's Guide*, GC26-7707

- ▶ Interoperability matrix:

<http://www.ibm.com/servers/storage/disk/ds4000/interop-matrix.html>



Enterprise Disk Systems

The family of IBM Enterprise Disk Systems offers a broad range of scalable solutions to address various enterprise storage needs. By leveraging IBM leading technology the Enterprise Disk Systems provide a significant choice in functionality, performance and resiliency.

This chapter discusses the features and major characteristics of the following enterprise disk storage products:

- ▶ IBM System Storage DS6000 Series
- ▶ IBM System Storage DS8000 Series
- ▶ IBM TotalStorage Enterprise Storage Server (ESS)

8.1 Overview

The IBM Enterprise Disk Systems deliver high-performance, high-availability storage, with flexible configuration for different business requirements. The IBM System Storage DS6000 and DS8000 deliver an enterprise storage continuum of systems with shared replication services and common management interfaces. The DS6000 and DS8000 series systems are designed to help simplify the storage infrastructure, support business continuity and optimize information lifecycle management. See Figure 8-1 for an overview of the family.

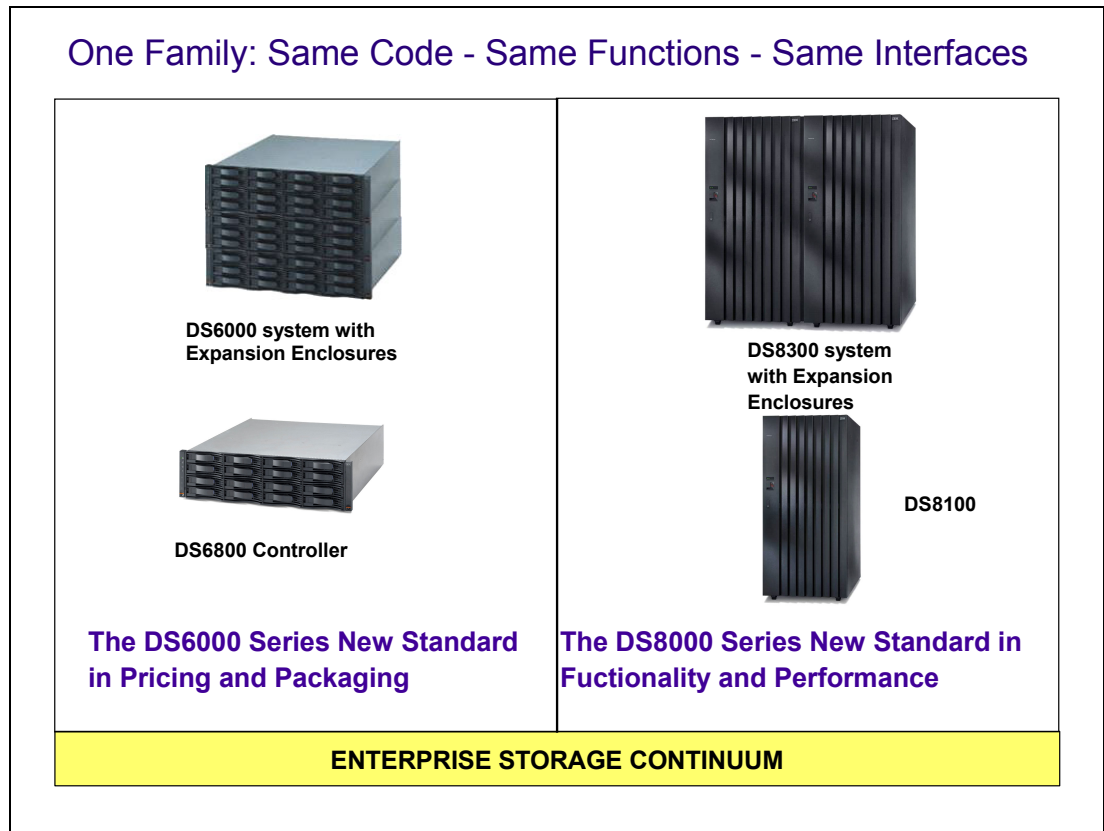


Figure 8-1 Enterprise Disk Systems

The DS6000 Series

The DS6000 offers true enterprise-class functionality with modular design and an attractive entry-level pricing for medium and large businesses. The DS6000 series helps simplify data management and enables easy scalability, which allows accommodation of the continuing exponential data growth. By leveraging proven software function of the Enterprise Storage Server, the DS6000 series brings proven enterprise class technology to a modular package.

The DS8000 Series

The IBM System Storage DS8000 series is designed for the most demanding, mission critical environments requiring the highest level of availability. The DS8000 series is designed to set an entirely new industry standard for high-performance, high-capacity by delivering a dramatic leap in performance and scalability.

8.1.1 Infrastructure simplification

The IBM System Storage DS Family offers the opportunity to simplify your IT infrastructure through consolidation and streamlined storage management.

Consolidate storage assets

Consolidation begins with interoperability. The IBM System Storage DS Family can be connected across a broad range of server environment, including IBM z/OS, z/VM, OS/400®, i5/OS® and AIX, as well as Linux, HP-UX, Sun SOLARIS, Novell NetWare, UNIX and Microsoft Windows. You can easily split storage capacity among the attached servers, reducing the total number of storage systems required.

Streamline storage management

The IBM System Storage DS Family incorporates streamlined management tools with easy-to-use, and straightforward GUI based on open SMI-S (Storage Management Initiative Specification) interfaces. The GUI allows users to manage multiple subsystems and controllers, perform logical configurations and administer copy service management functions, all via a Web browser.

The DS6000 series, the DS8000 series, and even the ESS storage subsystems share a common set of advanced functions, including FlashCopy, Metro Mirror, Global Copy, and Global Mirror. So there is only one set of skills necessary to manage the whole enterprise disk storage systems.

There is also a set of common functions for storage management, including the IBM System Storage DS Command-Line Interface (DS CLI) and the IBM System Storage DS open application programming interface (API).

Advanced copy functions and storage management tools are described later in this chapter.

8.1.2 Business continuity

The IBM System Storage DS Family offers the capability to increase your IT continual operation tolerance by data protection and compatible copying service.

Data protection

Many design characteristics and advanced functions of the IBM Enterprise Disk Systems contribute to protect data effectively.

- ▶ Fault-tolerant

The IBM System Storage DS8000 series are designed with no single point of failure. It is a fault-tolerant storage subsystem, which can be maintained and upgraded concurrently with user operations. The DS6000 series is also designed and implemented with component redundancy to help reduce and avoid many potential single points of failure.

- ▶ RAID protected storage

The IBM System Storage Enterprise Disk Systems support RAID-5, RAID-10 configurations, or a combination of both. This gives more flexibility when selecting the redundancy technique for data protection.

Copying service

IBM FlashCopy point-in-time copy functions support higher application availability and continuity of operations because they are designed to shrink backup window time. Metro and

Global Mirror functions allow for creation of duplicate copies of application data at remote sites for rapid recovery purposes.

The IBM System Storage DS Family also supports open standards. As a result, you can mirror a DS8000 series system with a DS6000 series system or an ESS Model 750 or 800 to help lower the total cost of the disaster recovery solution.

8.1.3 Performance

The IBM Enterprise Disk Systems are designed for high performance that takes advantage of IBM leading technologies. In today's world, enterprises need business solutions that can deliver high levels of performance continuously every day, day after day. They also need a solution that can handle different workloads simultaneously, so they can run business intelligence models, large databases for enterprise resource planning (ERP), and online and Internet transactions alongside each other. Some of the unique features that contribute to the overall high-performance design of the IBM Enterprise Disk Systems are:

Server-based design

The design decision to use processor memory as I/O cache is a key element of the IBM storage architecture. Performance improvements can be traced to the capabilities of the processor speeds, the L1/L2 cache sizes and speeds, the memory bandwidth and response time, and the PCI bus performance. (See Figure 8-2)

With the DS6000 and DS8000 series, the cache access has been accelerated further by making the non-volatile storage (NVS) a part of the main memory. Some part of the memory is used for the operating system and another part in each controller card acts as non-volatile storage (NVS), but most of the memory is used as cache. This design to use processor memory makes cache accesses very fast.

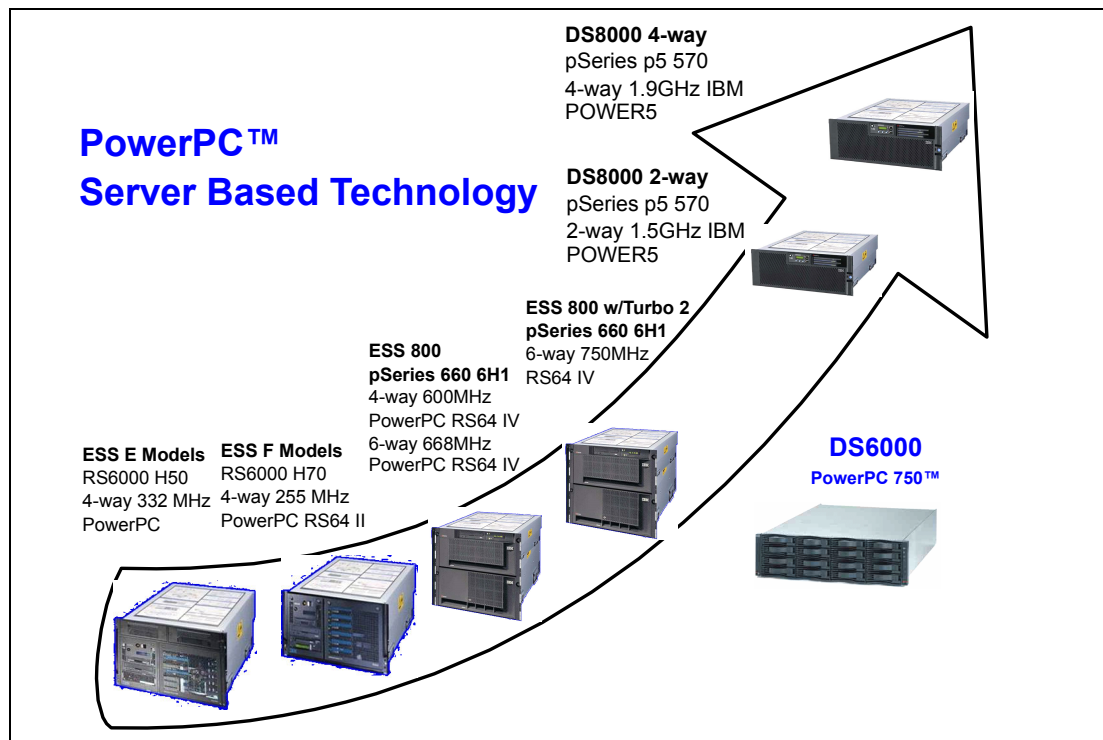


Figure 8-2 The server based architecture rings continuous performance improvements

Sequential Prefetching in Adaptive Replacement Cache (SARC)

Another performance enhancer is the new self-learning cache algorithm. This is described further in 8.2.1, “Enhancements of the DS6000 and DS8000 series” on page 121.

IBM multipathing software

IBM Multipath Subsystem Device Driver (SDD) provides load balancing and enhanced data availability in configurations with more than one I/O path between the host server and the storage server. Most vendors’ priced multipathing software selects the preferred path at the time of initial request. IBM free of charge preferred path multipathing software dynamically selects the most efficient and optimum path to use at each data interchange during read and write operations.

Command Tag Queuing

Command Tag Queuing provides Multiple AIX/UNIX I/O commands which may be queued to the DS6000 which improve performance through autonomic storage management versus the server queuing one I/O request at a time.

8.1.4 System z and zSeries performance

The zSeries hardware and software has evolved over the last forty years from the 360 system, introduced in the mid 1960s. While this has the positive aspect that applications have had continuity of support, the 360 architecture imposes constraints on the performance of both single and multi system environments. There has also been a massive increase in the amount of data stored and processed on the platform which has made the management of the data more complex and time-consuming. To address these issues has required a series of design changes to the disk systems themselves, as well as the System z (including zSeries) hardware and software.

This section discusses:

- ▶ The historical background to the performance constraints
- ▶ How data is stored and accessed
- ▶ The solutions

Hardware design

The architecture of the 360 introduced the concept of devices attached to a control unit, which had a cable (known as the channel) connecting it to the processor.

The channel and control unit

The use of the channel to transfer data between a device and processor storage is governed by a strict protocol which has to be adhered to by the processor and control unit. By having the protocol in a control unit any design changes to the DASD could be made transparent to the processor and the software running on it.

DASD devices

The geometry of the Direct Access Storage Device (DASD) was defined at the same time as the 360 architecture, with a series of platters (disks) mounted in parallel on a spindle. Data reads and writes were performed by a set of heads which were mounted on an assembly to match the number of platter surfaces. A DASD device (in software terms, a *volume*) was subdivided into *tracks* and *cylinders* for the purposes of storing data. A *track* is the area covered by a head during a complete revolution of a platter and a *cylinder* the number of tracks in a vertical plane.

Because the heads move together, they can only be positioned over one cylinder at a time - to access data on another cylinder the whole head assembly had to move (known as a seek). Only one head can transfer data at a time so concurrent operations are not possible. Also the channel can only process one Input or Output (I/O) operation at a time,

Software limitations

The original software design was made on the assumption that DASD was not capable of concurrent operations.

Data sets

The available space on a DASD volume is allocated, in tracks or cylinders, to data sets which have a unique name. Having larger DASD volumes meant less floor space in the machine room and reduced the number of volumes to manage.

Software bottleneck

A Unit Control Block (UCB) is created to represent each device connected to a System z processor. As soon as an I/O operation is started to a device, the UCB is marked as busy and any further requests are queued. With this design it is not possible to start additional I/O operations even if the DASD has the capability. Additionally data sets with response times that were critical to online system response times would need to be monitored and might need to be split across multiple volumes.

Large volumes

As stated in "DASD devices" on page 117 a volume consists of a number of tracks and cylinders. The amount of data that can be written on a track depends on the technology of the heads and reached its maximum with the last of the pre-RAID DASD, the 3390. There were four models in this range, 3390-1, 3390-2, 3390-3 and 3390-9, all of which had the same track capacity and number of heads (15). The difference between the models was the number of cylinders. The model suffix defined the number of cylinders relative to the 3390-1 which had 1113, therefore a 3390-3 had 3339.

The advantage of the 3390-9 was that it occupied the same floor area as a 3390-1 but had nine times the capacity. However this also meant that the demands for the data on the volume increased, which caused queuing for I/O operations and relatively poor response times. This made the 3390-9 of little use for online systems and the 3390-3 became the general choice as a compromise between capacity and performance.

Parallel Sysplex

In response to the increasing requirements, largely Internet driven, for data to be available on a 24 x 7 basis, the Parallel Sysplex was introduced. In simple terms this is a collection of z/OS systems which have access to the same data and can run more than one iteration of an application. This allows for capacity to be increased to meet demand and to allow systems to be taken out for upgrade or maintenance while maintaining a service. However, from a performance point of view there is a potential problem if more than one system tries to access the same volume concurrently. If an I/O operation is started to a volume then that volume has allegiance to the channel that started the I/O and cannot be used by another system.

Solutions

With the introduction of RAID subsystems, DASD volumes became logical entities and there was no longer any reason why multiple I/O operations could not run concurrently either on single or multi-system environments. The only consideration is whether there is a potential conflict between reading and updating in the same data set. The solutions to the performance issues described in the previous sections have been addressed by modifications to both the ESS and the z/OS software. As the volumes seen by z/OS are emulated, the particular model

of the subsystem is not significant and the term ESS is used as a generic to refer to IBM disk systems, including the DS6000 and DS8000.

Workload Manager

The Workload Manager (WLM) is a component of z/OS that distributes systems resources in order that workloads, such as online applications, are able to achieve predefined targets (goals), for example terminal response times. The resources managed include central storage, CPU and I/O devices.

Multiple Allegiance

As mentioned in “Parallel Sysplex” on page 118, the allegiance between a DASD volume and a channel prevents I/O from different systems running concurrently, this restriction has been lifted by the introduction of Multiple Allegiance (MA). This is a combination of features in the z/OS software and the ESS hardware. The I/O operations that z/OS starts to a DASD volume are prefixed by information which indicates the area of the volume of interest (usually a data set) and the nature of the operation to be performed, that is either reading or writing. The ESS maintains tables of each concurrent request and allows them to start providing that no request has an integrity exposure, that is a read and write to the same data set.

Parallel Access Volumes (PAV)

As the UCB design prevents more than one I/O operation being started to a volume, the solution is to use one or more additional UCBs to represent a volume. To achieve this, the ESS is defined with a number of addresses of true volumes, known as Base addresses, and an additional set, known as Aliases. If WLM detects that a goal is not being achieved, due to the queuing of I/O to a base address, it can request that the ESS associates one of the Alias addresses with the Base. Once this is done, the queued I/O requests will be split across the two addresses. If this does not relieve the queuing problem then further requests for Aliases will be made. Once the delay has been relieved, the Aliases will be returned to a free state and can be used for association with another Base address as required.

Emulated Large Volumes

As was stated earlier the 3390-9 met the criteria of reducing the number of volumes to manage, but did not necessarily give good response times. The introduction of PAVs and MA addressed the response time issue, so there was no reason why volumes larger than a 3390-9 could not be defined. This was done in two stages, the first stage was to support the definition of a volume up to 27 times the size of a 3390-1, that is approximately 30,000 cylinders. The second stage was to increase the maximum size to 54 times that of a 3390-1, that is approximately 60,000 cylinders.

These values are another indication of the history of the z/Series software in which the values of 32K and 64K have a significance to some of the instructions in the architecture and are often limits on the value of fields. So the first stage was using the existing code and the second stage was altering the code.

To address the customer requirement for reduced management of volumes, the ESS and z/OS software are capable of supporting a 3390-54. However this means that there is potentially a vast amount of data on a single volume and the queuing for I/O operations means a relatively poor response times. This being the case large volumes are likely to be of little use for online systems.

Priority I/O Queuing

If there is resource contention in an ESS then I/O requests will be queued on a first in first out (FIFO) basis. If workloads are being delayed by I/O response time WLM can pass a priority to the ESS which will use that to decide the position in the queue and not FIFO.

FICON

The latest channel protocol available on the z/series machines Fibre Connectivity (FICON®) which supports a 2 Gigabit connection. This protocol using a multiplexing technique which delivers high bandwidth and is particularly effective with sustained transfer of large data blocks (such as taking backups).

8.2 DS6000 and DS8000 technology

The maximum storage capability of the DS68000 controller is 4.8 TB. Optional DS6000 expansion enclosures provide a maximum storage capability of 38.4 TB. The physical storage capacity of the DS8000 series systems can range from 1.1 TB to 192 TB and it has an architecture designed to scale up to a petabyte. The DS8000 series allows additions and upgrades from one model to another to adapt to changing business requirements.

With this choice, you can build very cost efficient storage systems by adding expansion enclosures to the DS6800 controller or also grow horizontally by adding other DS68000 controllers. You have the option to easily grow into the DS8000 series by adding DS8000 systems to your environment or by replacing DS6000 systems. The expandability is shown in Figure 8-3:

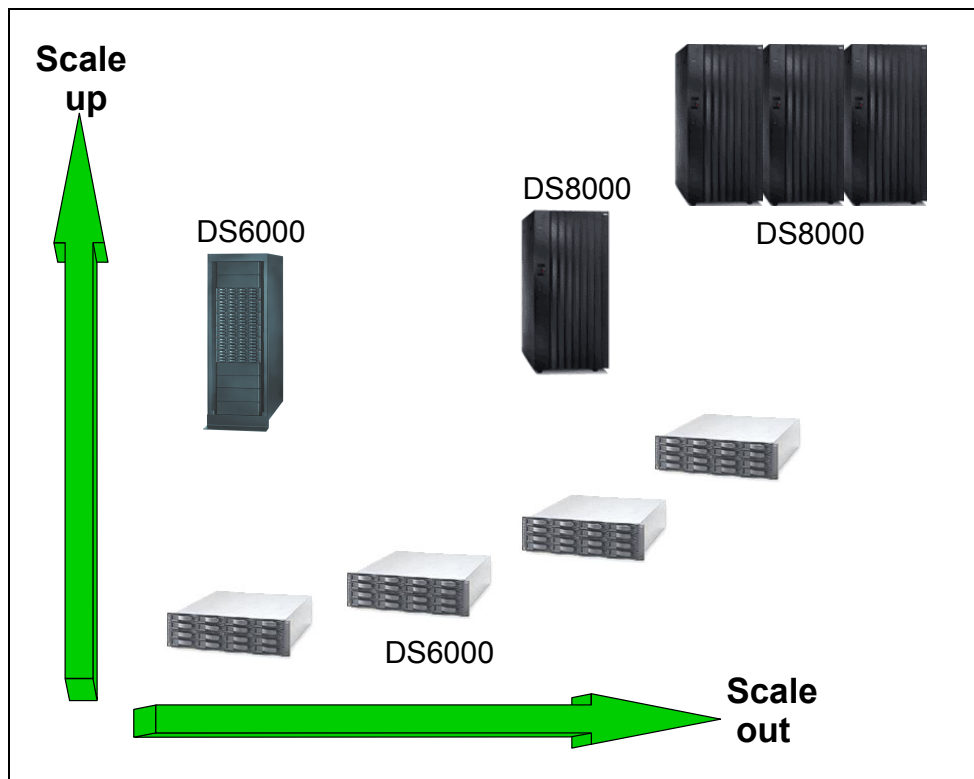


Figure 8-3 Scaling options

Multi-tiered storage

The DS6000/DS8000 series is ideally suited to a multi-tiered storage environment. This helps minimize storage costs by retaining frequently accessed or high-value data on higher performance DS8000 storage servers and archiving less frequently accessed or less valuable information on less-costly DS6000 systems.

8.2.1 Enhancements of the DS6000 and DS8000 series

The DS6000 and DS8000 series incorporate the following performance, resiliency and scalability innovations:

Switched Fibre Channel Arbitrated Loop (FC-AL)

The disk interconnection has changed over the previous ESS. Instead of SSA loops there is now a switched FC-AL implementation (see Figure 8-4). This switching technology uses a point-to-point connection to each disk drive and adapter, which allows maximum bandwidth for data movement, eliminates the bottlenecks of loop designs and allows for specific disk drive fault indication and isolation in the same time. Each disk is attached to both switches. Whenever the device adapter connects to a disk, it uses a switched connection to transfer data. This means that all data travels via the shortest possible path.

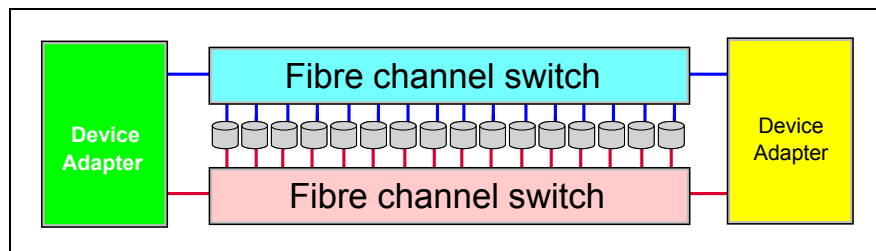


Figure 8-4 Disk enclosure switched connections

Four path switched drive subsystem

There are four paths from the device adapters to each disk drive to provide greater data availability in the event of multiple failures along the data path (see Figure 8-5 on page 122). The four paths provide two FC-AL device interfaces, each with two paths such that either path can be used to communicate with any disk drive on that device interface (in other words, the paths are redundant). One device interface from each device adapter connects to a set of FC-AL devices such that either device adapter has access to any disk drive through two independent switched fabrics (in other words, the device adapters and switches are redundant). In normal operation, however, disk drives are typically accessed by one device adapter and one server. Each path on each device adapter can be active concurrently, but the set of eight paths on the two device adapters can all be concurrently accessing independent disk drives. This avoids any contention between the two device adapters for access to the same disk, and means that all eight ports on the two device adapters can be concurrently communicating with independent disk drives.

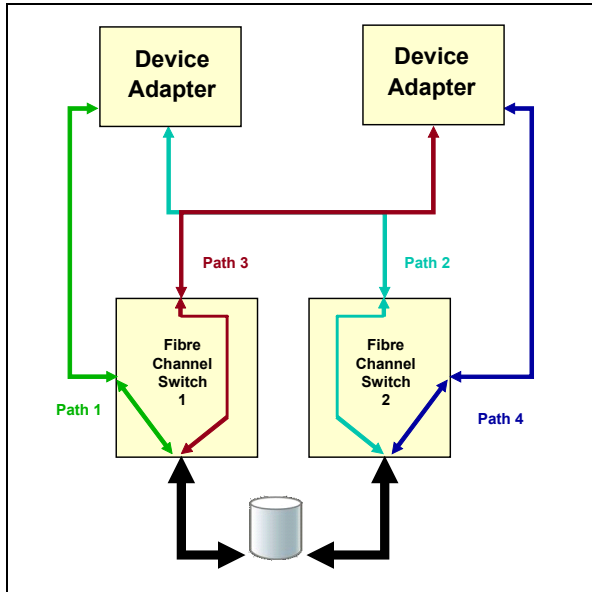


Figure 8-5 Four paths to each disk drive

Self-learning cache algorithms - SARC

The DS6800 and DS8000 use the patent-pending Sequential Prefetching in Adaptive Replacement Cache (SARC) algorithm, developed by IBM Storage Development in partnership with IBM Research. It is a self-tuning, self-optimizing solution for a wide range of workloads with a varying mix of sequential and random I/O streams. SARC is inspired by the Adaptive Replacement Cache (ARC) algorithm and inherits many features from it.

When a host performs a read I/O, the DS6000 and DS8000 fetch the data from the disk arrays via the high performance switched disk architecture. The data is then cached in volatile memory in case it is required again. The DS6000 and DS8000 attempt to anticipate future reads by the SARC algorithm. Data is held in cache as long as possible using this smart algorithm. If a cache hit occurs where requested data is already in cache, then the host does not have to wait for it to be read from the disks.

SARC provides the following:

- ▶ Sophisticated, patented algorithms to determine what data should be stored in cache based upon the recent access and frequency needs of the hosts.
- ▶ Pre-fetching, which anticipates data prior to a host request and loads it into cache.

Predictive Failure Analysis

The DS Enterprise Family uses Predictive Failure Analysis® (PFA) to monitor disk drive operations. PFA takes pre-emptive and automatic actions before critical drive failures occur. The disk drives can anticipate certain forms of failures by keeping internal statistics of read and write errors. If the error rates exceed predetermined threshold values, the drive will be nominated for replacement. Because the drive has not yet failed, data can be copied directly to a spare drive. This avoids using RAID recovery to reconstruct all of the data onto the spare drive.

Configuration flexibility

The DS6000 and DS8000 use virtualization techniques to separate the logical view of hosts from the underlying physical layer. On an ESS there was a fixed association between logical

subsystems (LSS) and device adapters. With the DS6000 and the DS8000, these limitations no longer apply - see Figure 8-6.

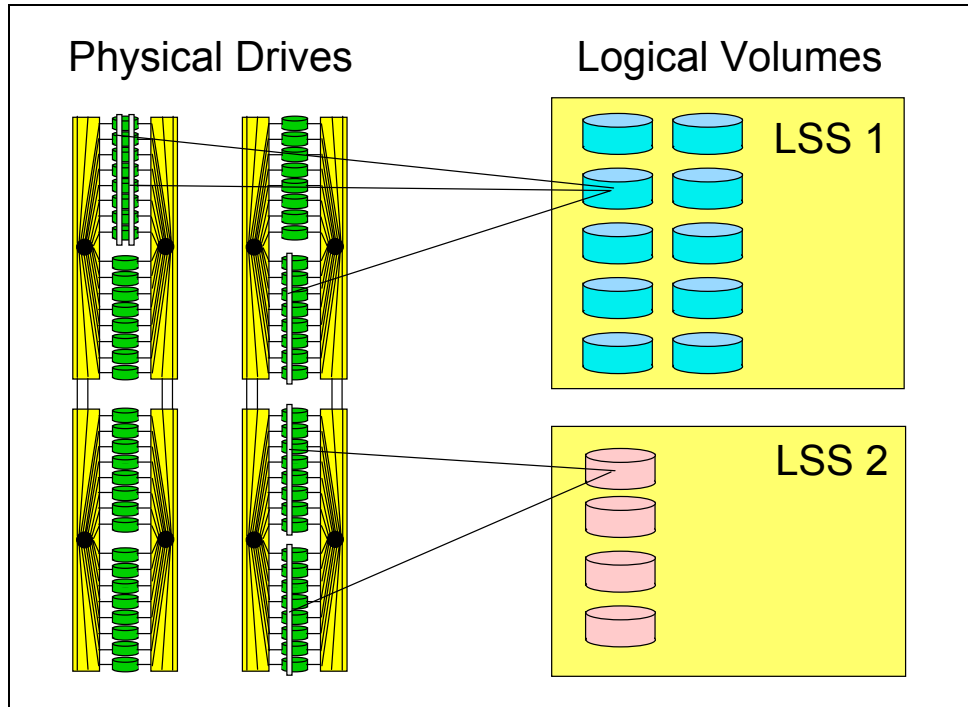


Figure 8-6 Grouping of volumes in LSSs

The DS6000 and DS8000 series provide a high configuration flexibility with the following virtualization techniques:

- ▶ **Dynamic LUN/Volume creation and deletion**
LUNs can be created and deleted without having to reformat a whole array.
- ▶ **Large LUN and large CKD volume support**
LUNs and volumes can be configured to span arrays, which allows for large LUN sizes.
- ▶ **Flexible LUN to LSS association**
There is no predefined association of arrays to logical subsystems (LSS).
- ▶ **Simplified LUN masking**
The access to LUNs by the host systems is controlled via volume groups. Hosts or disks in the same volume group share access to data.

8.3 The DS6000 series

The DS6000 (see Figure 8-7 on page 124) series offers high reliability and enterprise-class functionality for medium and large businesses. The DS6000 series supports connectivity with open systems and mainframe hosts. Copy services are equivalent and interoperable with the DS8000 and the ESS. With the DS6000, enterprise class resiliency is now available in an incredibly small, modular, affordable package. The DS6000 series can grow up to 38.4 TB physical storage.



Figure 8-7 The DS6000 series

The DS6000 consists of the *DS6800 controller enclosure (1750-511)* which has dual Fibre Channel RAID controllers. Capacity can be increased by adding up to 7 *DS6000 expansion enclosures (1750-EX1)*.

8.3.1 The DS6800 controller enclosure

The DS6000 series is based on IBM industry-leading PowerPC® architecture which accelerates data response time and offers impressive performance. At the core of the DS6800 controller enclosure are two active/active RAID controllers (See Figure 8-8).

The controller cards

Each controller card contains an integrated four-port host adapter and a high performance device adapter chipset. The device adapter ASIC on the DS6800 controller provides two disk expansion and two disk control ports. These ports are used to connect the expansion enclosures. Host adapter ports are used for host or SAN attachment.



Figure 8-8 DS6800 controller card

The RAID controller cards are the heart and soul of the system. Each card is the equivalent of a cluster node in an ESS. IBM has leveraged its extensive development of the ESS host adapter and device adapter function to create a total repackaging. It actually uses DS8000 host adapter and device adapter logic, which allows almost complete commonality of function and code between the two series (DS6000 and DS8000).

The processors

The DS6800 utilizes two 64-bit PowerPC 750GX 1 GHz processors for the storage server and the host adapters, respectively, and another PowerPC 750FX 500 MHz processor for the device adapter on each controller card. The DS6800 has 2 GB memory in each controller card, for a total of 4 GB.

The disk subsystem

For the disk subsystem, each controller card has an integrated four-port FC-AL (Fibre Channel Arbitrated Loop) device adapter that connects the controller card to two separate Fibre Channel loops. Each switched loop attaches disk enclosures that each contain up to 16 disks. Each enclosure contains two 22-port Fibre Channel switches. Of these 22 ports, 16 are used to attach to the 16 disks in the enclosure and four are used to interconnect with other enclosures. The remaining two are reserved for internal use. Shown Figure 8-9:

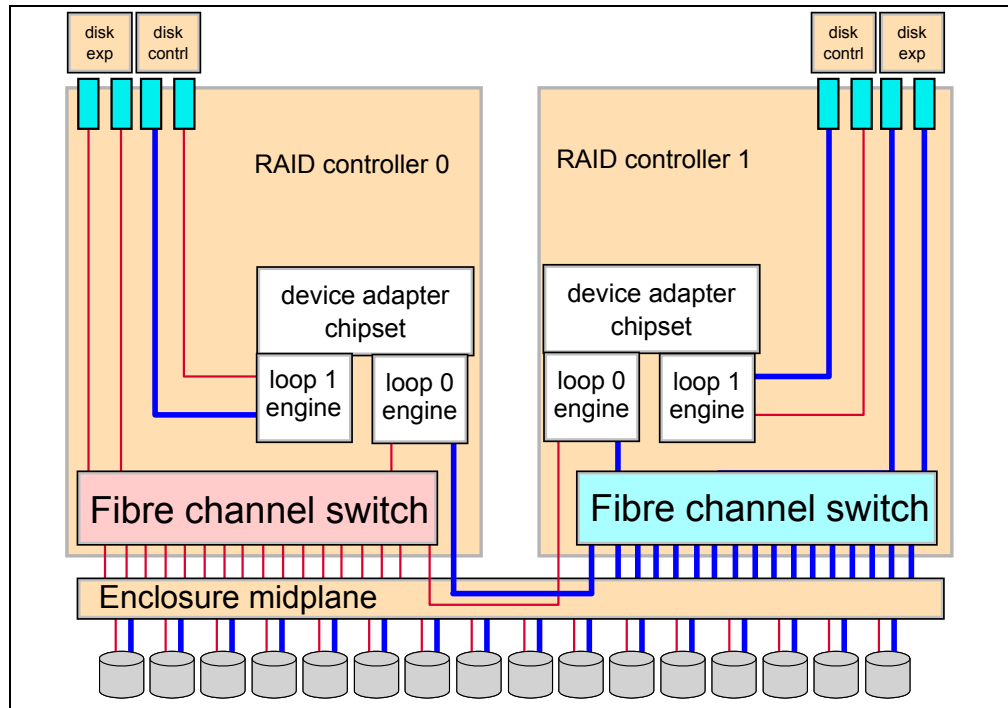


Figure 8-9 DS6000 disk architecture

The configuration process when forming RAID-5 or RAID-10 arrays requires that two global spares be defined in the DS6800 controller enclosure. If you have expansion enclosures, the first enclosure needs another two global spares. More spares could be assigned when drive groups with larger capacity drives are added.

Host adapters

The DS6800 has eight 2 Gbps Fibre Channel host ports. Each host port can be configured individually to operate in Fibre Channel or FICON mode. A host port cannot be both FICON and FCP simultaneously, but it can be changed as required. Host servers should have paths to each of the two RAID controllers of the DS6800 to achieve the best reliability and performance.

The attached hosts interact with microcode running on a Power PC® chipset to access data on logical volumes. The microcode manages all read and write requests to the logical volumes on the disk arrays. For write I/O operations, the controllers use fast-write, which writes the data to volatile memory on one controller and to persistent memory on the other controller. The DS6800 then reports to the host that the write is complete before it has actually been written to disk. This provides much faster write performance. Persistent memory is also called NVS or non-volatile storage. See Figure 8-10 on page 126:

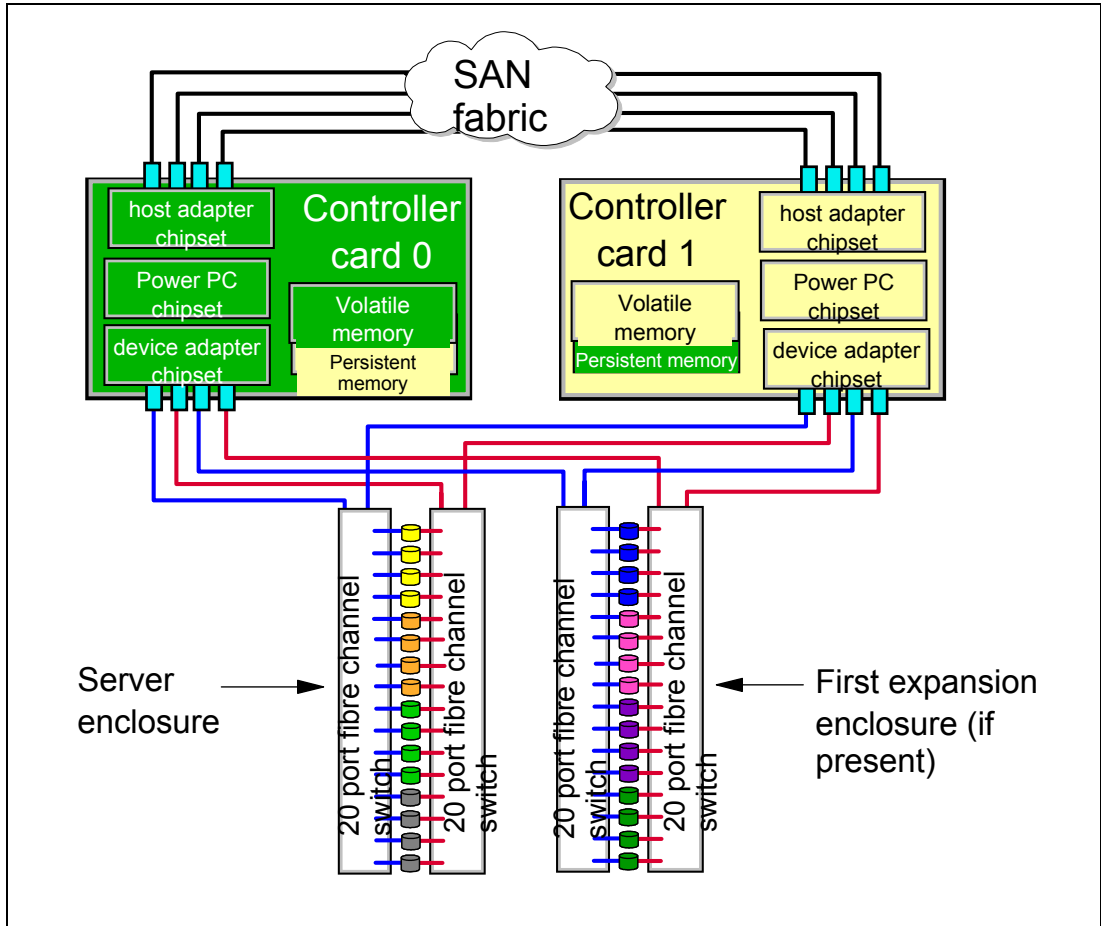


Figure 8-10 DS6800 controller architecture

Power subsystem

The power subsystem of the DS6800 has two redundant power supplies and two battery backup units (BBUs). DS6000 expansion enclosures contain power supplies but not BBUs. The DS6000 power supplies are hot-swappable and a single power supply can support the power requirements of an entire enclosure.

Dense packaging

Calibrated Vectors Cooling technology used in System x and BladeCenter to achieve dense space saving packaging is also used in the DS6800.

8.3.2 The DS6000 expansion enclosure

The DS6000 expansion enclosure is used to add capacity to a DS6800 server enclosure. Up to 7 DS6000 expansion enclosures can be added to a DS6800 controller enclosure.

Expansion enclosure SBOD controller card

The DS6000 SBOD controller card (see Figure 8-11 on page 127) is only found in the expansion enclosure. Each SBOD controller card contains an independent 22-port Fibre Channel switch. Of these 22 ports, 16 are used to attach to the 16 disks in the expansion enclosure. Four more are used to interconnect with other enclosures, with the remaining two ports reserved for internal use.



Figure 8-11 SBOD controller card

The DS6800 supports two dual redundant switched loops. The first loop is for the DS6800 and up to three DS6000 expansion enclosures. The second switched loop is for up to four expansion enclosures. For connections to the previous and next enclosure, four inbound and four outbound 2 Gbps Fibre Channel ports are available.

The disks in the server enclosure are on the first disk loop (loop 0). The first two expansion enclosures are attached to the DISK CONTRL ports to start the second disk loop (loop 1). The DISK EXP ports are used to attach the third expansion enclosure. It joins the same switched loop as the disks in the server enclosure (loop 0), as shown in Figure 8-12:

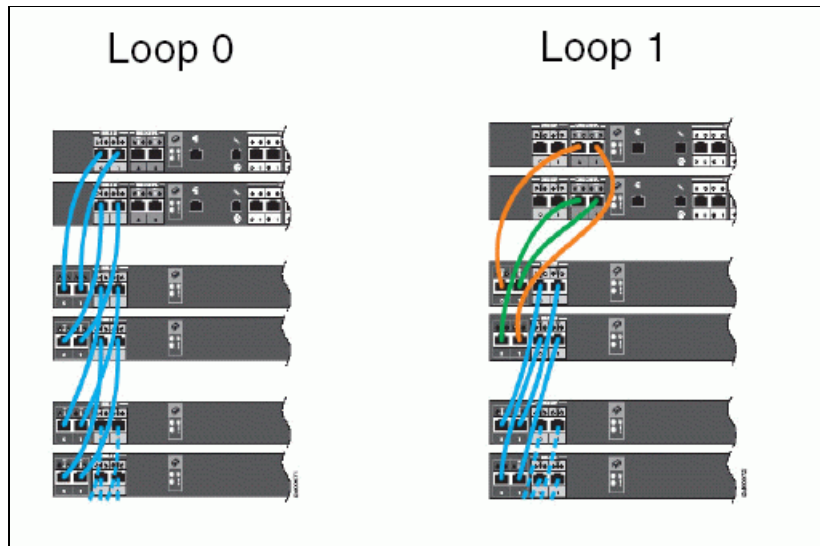


Figure 8-12 Example of expansion enclosure cabling

To distribute the storage enclosures equally on both storage loops, we recommend you follow the install sequence shown in Table 8-1:

Table 8-1 Install sequence of expansion enclosures

Expansion enclosure	Loop
1	1
2	1
3	0
4	1
5	0
6	1
7	0

8.3.3 DS6800 Model 1750-511

The 1750-511 (Figure 8-13) model contains control unit functions as well as a rich set of advanced functions and holds up to 16 disk drive modules (DDMs). It provides a minimum capacity of 584 GB with 8 DDMs and 73 GB per DDM. The maximum storage capacity with 300 GB DDMs is 4.8 TB with 16 DDMs.



Figure 8-13 DS6800 Model 1750-511 rear view

The 1750-511 model has the following features:

- ▶ Two RAID controller cards
- ▶ PowerPC 750GX 1 GHz processors and one PowerPC 750FX processor on each RAID controller card
- ▶ Dual active controllers to provide continuous operations and back up the other controller in case of controller maintenance or an unplanned outage of a controller
- ▶ 4 GB of cache memory (2 GB in each controller card)
- ▶ Battery backed mirrored cache
- ▶ Two battery backup units (one per controller card)
- ▶ Two AC/DC power supplies with imbedded enclosure cooling units
- ▶ Eight Fibre Channel host ports that can be configured as a pairs of FCP or FICON host ports. The hosts ports auto-negotiate to either 2 Gbps or 1 Gbps link speeds
- ▶ Two Fibre Channel switches for disk drive connectivity
- ▶ Eight 2 Gbps device ports (for additional DS6000 expansion enclosures)
- ▶ Attachment to up to seven DS6000 Model 1750-EX1 expansion enclosures
- ▶ Choice of 2 Gbps Fibre Channel disk drives including:
 - 73 GB 15k RPM
 - 146 GB 15k RPM
 - 146 GB 10k RPM
 - 300 GB 10k RPM
- ▶ 3U in height and mountable in a standard 19-inch rack

8.3.4 DS6000 Model 1750-EX1

DS86000 can expand with the expansion enclosure model 1750-EX1 (Figure 8-14). Up to 7 expansion drawers build up to the maximum configuration of 128 DDMs which is comprised of 8 drawers x 16 DDMs.



Figure 8-14 DS6000 with 2 expansion enclosures

Each expansion enclosure contains the following features:

- ▶ Two expansion controller cards, each with:
 - Two 2 Gbps inbound ports
 - Two 2 Gbps outbound ports
 - One Fibre Channel switch
- ▶ Disk enclosure which holds up to 16 Fibre Channel DDMs
- ▶ Two AC/DC power supplies with embedded enclosure cooling units
- ▶ Supports attachment to DS6800 Model 1750-511
- ▶ The DS6800 Model 1750-EX1 can also be mounted in a standard 19-inch rack

8.3.5 More information

For more information about the DS6000 series, see the Redbooks:

- ▶ *The IBM TotalStorage DS6000 Series: Concepts and Architecture*, SG24-6471:
- ▶ *The IBM TotalStorage DS6000 Series: Implementation*, SG24-6781

- ▶ *The IBM TotalStorage DS6000 Series: Copy Services with IBM eServer zSeries*, SG24-6782
- ▶ *The IBM TotalStorage DS6000 Series: Copy Services in Open Environments*, SG24-6783
- ▶ *iSeries and IBM TotalStorage: A Guide to Implementing External Disk on eServer i5*, SG24-7120
- ▶ *IBM TotalStorage DS6000 Series: Performance Monitoring and Tuning*, SG24-7145

8.4 The DS8000 series

The DS8000 series (Figure 8-15) is designed to *break through* to a new dimension of on demand storage. Its innovative design incorporates a high-bandwidth and fault-tolerant Fibre Channel disk technology and hardware-based partitioning implementation. The DS8000 delivers robust, flexible, and cost-effective disk storage for mission-critical workloads and helps to ensure exceptionally high system availability for continuous operations in 24x7 environments.



Figure 8-15 The DS8000 series

The hardware has been optimized to provide enhancements in performance, connectivity, and reliability. The DS8000 series features several models in a new, higher-density footprint, while re-using much of the fundamental architecture of the previous ESS models. This ensures that the DS8000 leverages a very stable and well-proven operating environment.

8.4.1 The DS8000 design enhancements

Here are some key architectural enhancements introduced in the DS8000:

POWER5 processor technology

The DS8000 series exploits the IBM POWER5 technology, which is the foundation of the storage system LPARs. The DS8000 uses 64-bit POWER5 microprocessors in dual 2-way or dual 4-way processor complexes, with up to 256 GB of cache, 4 times as much as the ESS. The maximum host I/O operations per second of the DS8300 models is up to six times the maximum of the ESS Model 800.

Internal fabric

The DS8000 comes with a high bandwidth, fault tolerant internal interconnection, called RIO-G (Remote I/O) which is also used in the System p. The interconnection can operate at speeds up to 1 GHz and offers a 2 Gbps sustained bandwidth per link, with exceptional performance and reliability.

Storage Hardware Management Console (S-HMC)

The DS8000 offers a new integrated management console, which is the service and configuration portal for up to eight DS8000s in the future. Initially there will be one management console for one DS8000 storage subsystem. The S-HMC is the focal point for configuration and Copy Services management, and can run at the integrated display or remotely via a Web browser.

The S-HMC is also the interface for remote services (call home and call back). Remote connections can be configured to meet customer requirements. It is possible to allow one or more of the following: call on error (machine detected), connection for a few days (customer initiated), and remote error investigation (service initiated). The remote connection between the management console and the IBM service organization is via a VPN point-to-point connection over the Internet or modem.

Storage system logical partitions (LPARs)

The DS8000 series provides storage subsystem LPARs as a first in the industry. Logical partitioning introduces a new level of virtualization.

Storage LPAR technology (see Figure 8-16 on page 132) enables the creation of two or potentially more completely separate storage subsystems running the same or different versions of microcode. With these separate resources, each storage system LPAR can be used for completely separate production, test or other unique storage environments within a single physical DS8000 system. Each partition can be established to support specific performance requirements. You can run independent, virtual storage images with differing workloads, and with different operating system environments. The DS8000 series isolates and protects LPARs via hardware and the POWER5 Hypervisor firmware.

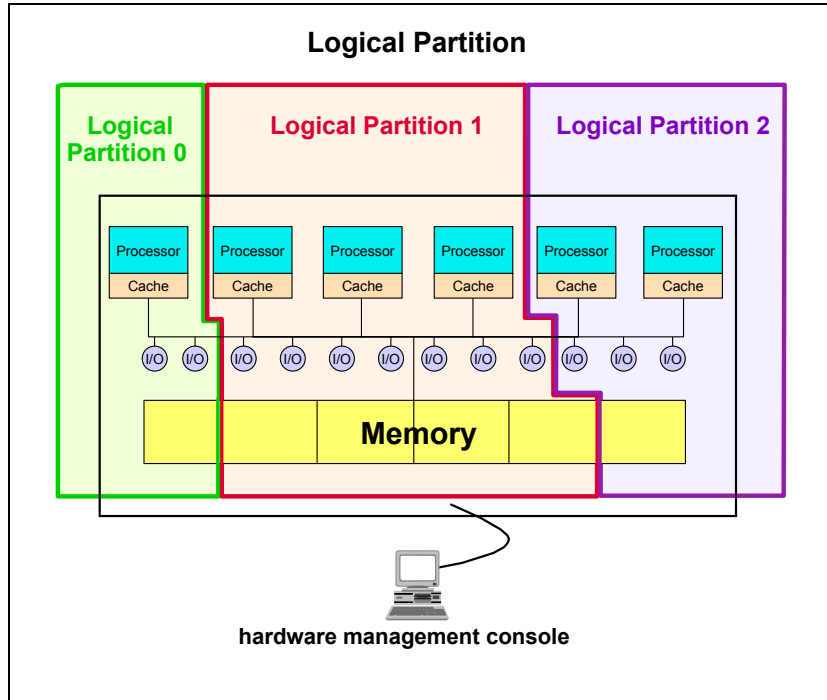


Figure 8-16 Logical partitions (LPARs)

The LPAR functionality is available in the DS8300 Model 9A2, and partitions the subsystem into two virtual storage system images. Initially each storage system LPAR has access to:

- ▶ 50 percent of the processors
- ▶ 50 percent of the processor memory
- ▶ Up to 16 host adapters
- ▶ Up to 320 disk drives (up to 96 TB of capacity)

8.4.2 The DS8000 architecture

This section describes the architecture of the DS8000 systems and provides information about the following components:

- ▶ Processor complexes
- ▶ Disk subsystems (switched FC-AL loops)
- ▶ I/O enclosures (device and host adapters)
- ▶ Power and cooling
- ▶ Battery backup units
- ▶ Hardware Management Console (S-HMC)
- ▶ Network switches
- ▶ Frames and expansion frames

Figure 8-17 on page 133 shows the components used to create the DS8000 Primary Frame.

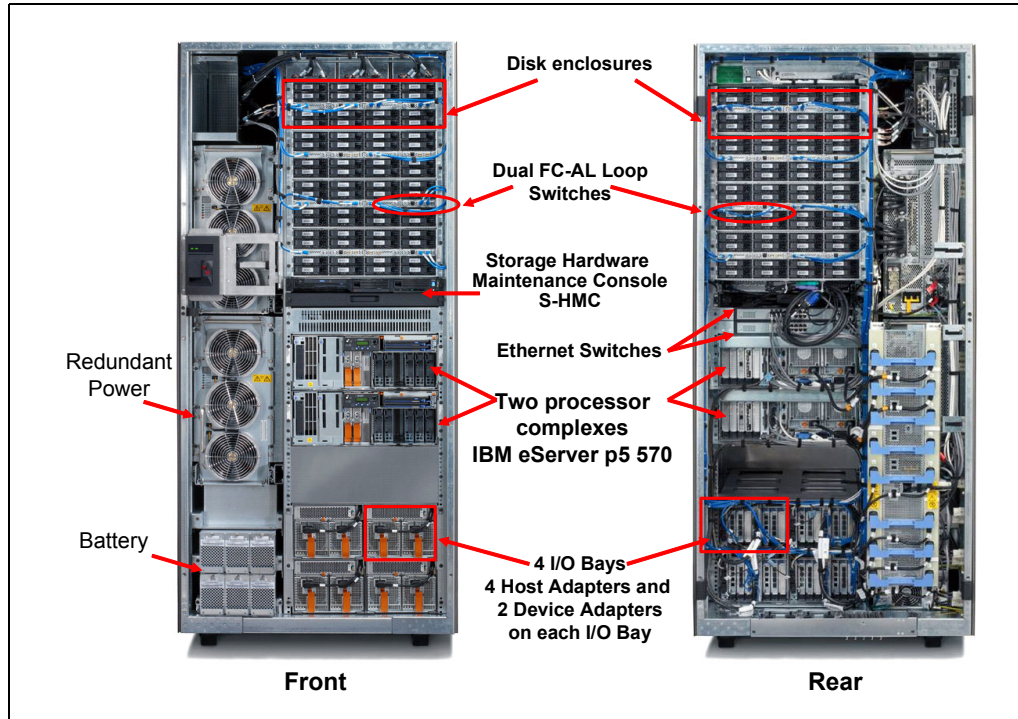


Figure 8-17 DS8000 - Primary Frame

The processor complexes

The DS8000 base frame contains two IBM p5 570 processor complexes, which contain the processor and memory that drive all functions in the DS8000. The symmetric multiprocessor (SMP) p5 570 system features 2-way or 4-way, copper-based, SOI-based POWER5 microprocessors running at 1.5 GHz or 1.9 GHz with 36 MB off-chip Level 3 cache configurations. The p5 570 also has I/O expansion capability using the RIO-G interconnect.

The Model 921 has 2-way processors while the Model 922 and Model 9A2 have 4-way processors.

Each processor complex can be logically partitioned into two LPARs, each of which is the equivalent of an ESS cluster.

The processor memory

The DS8100 Model 921 offers up to 128 GB of processor memory and the DS8300 Models 922 and 9A2 up to 256 GB. Half the memory is located in each processor complex. In addition, the Non-Volatile Storage (NVS) scales to the processor memory size selected, which can also help optimize performance.

Service processor and SPCN

The service processor (SP) performs predictive failure analysis based on any recoverable processor errors. The SP can monitor the operation of the firmware during the boot process, and it can monitor the operating system for loss of control. This enables the service processor to take appropriate action. The SPCN is the system power control network that is used to control the power of the attached I/O subsystem.

The disk subsystems

Each complex uses several four-port FC-AL device adapters to access the disk subsystem. A DS8000 can potentially have up to sixteen of these adapters arranged in eight pairs. Each adapter connects the complex to two separate switched Fibre Channel networks. Each switched network attaches disk enclosures that each contain up to 16 disks. Each enclosure contains two 20 port Fibre Channel switches. Of these 20 ports, 16 are used to attach to the 16 disks in the enclosure and the remaining four are used to interconnect either with other enclosures or to the device adapters.

The DS8000 implements the concept of Array Across Loops (AAL) which splits each array site into two halves. Half of the site is located on the first disk loop of a device adapter (DA) pair and the other half is located on the second disk loop of that DA pair. It is implemented primarily to maximize performance; however it also provides higher level of redundancy with RAID-10 configurations. This is shown in Figure 8-18.

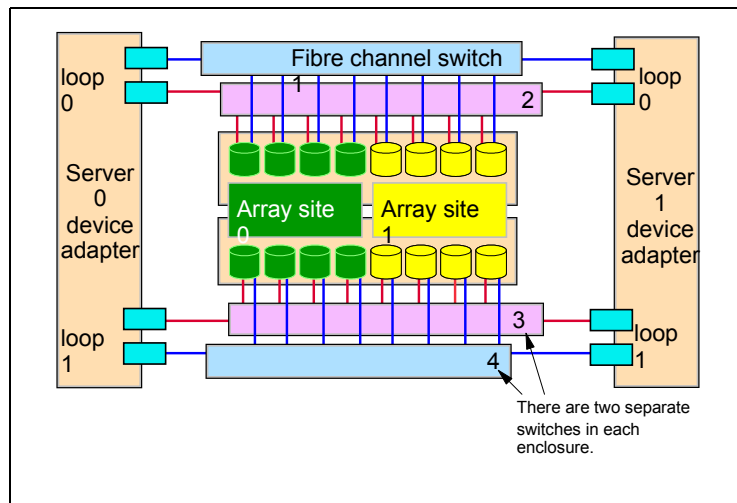


Figure 8-18 Array across loops

The host adapters

The DS8000 offers enhanced connectivity with the availability of four-port Fibre Channel/FICON host adapters (see Figure 8-19 on page 135). Each port can be independently configured as either Fibre Channel protocol (FCP) port for open systems host connection or PPRC FCP links, or as a FICON port to connect to System z hosts. ESCON host connection is also supported, but then a host adapter contains only two ESCON ports. The mix of ESCON ports and FCP ports on the same adapter is not possible. A DS8000 can have both ESCON adapters and FCP/FICON adapters at the same time.

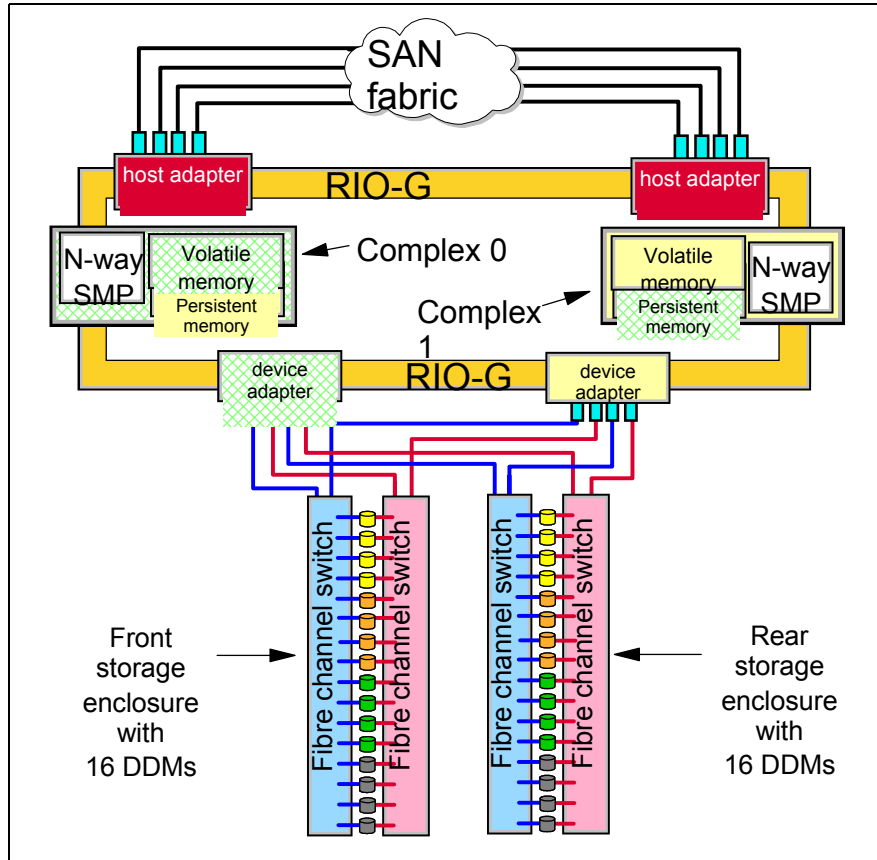


Figure 8-19 The DS8000 architecture

The attached hosts interact with software running on the complexes to access data on logical volumes. Each complex hosts at least one instance of this software (which is called a *server*), running in an LPAR. The servers manage all read and write requests to the logical volumes on the disk arrays. During write requests, the servers use fast-write where the data is written to volatile memory on one complex and persistent memory on the other complex. The server then reports the write as complete before it has been written to disk. This provides much faster write performance.

RIO-G

Both the device and host adapters operate on a high bandwidth fault-tolerant interconnect known as the RIO-G (shown Figure 8-20 on page 136).

The RIO-G ports are used for I/O expansion to external I/O drawers. The RIO-G is evolved from earlier versions of the RIO interconnect.

Each RIO-G port can operate at 1 GHz in bidirectional mode and can pass data in each direction on each cycle of the port. It is designed as a high performance self-healing interconnect. The p5 570 provides two external RIO-G ports, and an adapter card adds two more. Two ports on each processor complex form a loop.

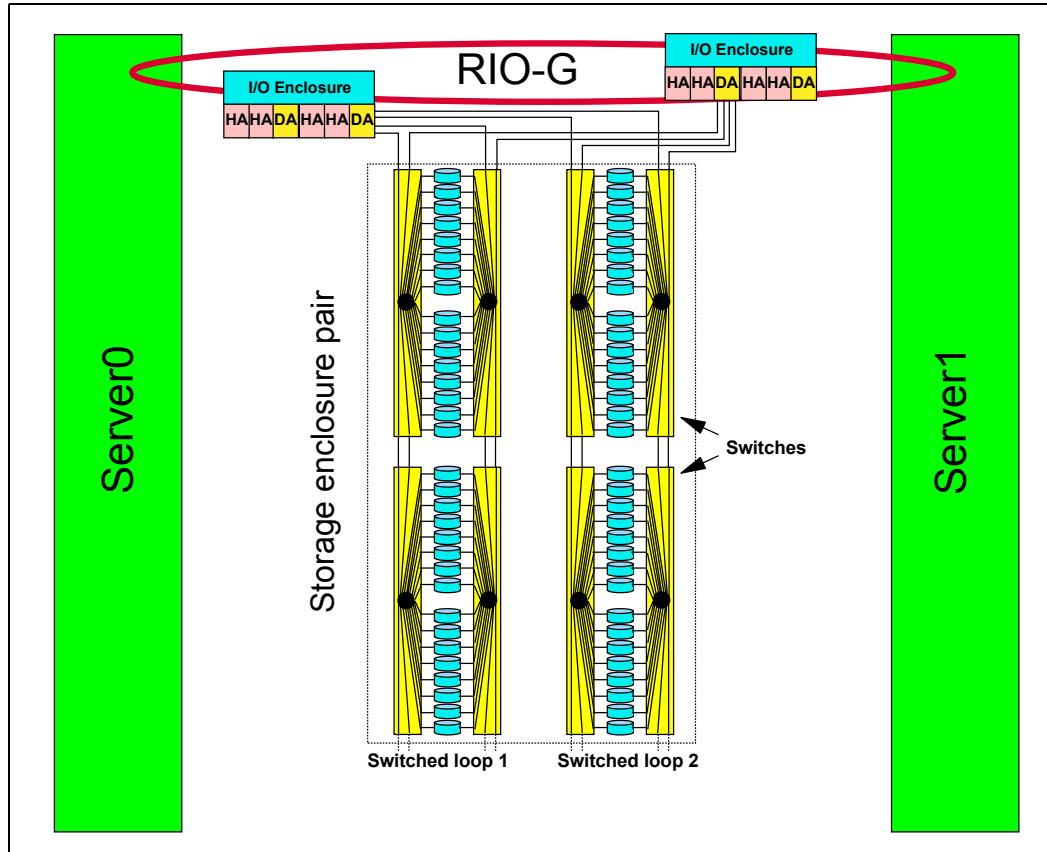


Figure 8-20 RIO-G design

Storage Hardware Management Console (S-HMC)

The DS8000 offers a new integrated management console. This console can serve as the service and configuration portal for up to eight DS8000s in the future. Initially there is one management console for each DS8000 storage subsystem.

It is a focal point with multiple functions such as:

- ▶ Storage configuration
- ▶ LPAR management
- ▶ Advanced Copy Services invocations
- ▶ Interface for local service personnel
- ▶ Remote service and support

S-HMC functions can be managed from the integrated keyboard display or remotely via a Web browser.

Ethernet Switches

Two 16-port gigabit Ethernet switches connect the disk enclosures and the processor complexes, which allow the creation of a fully redundant network. Each processor complex has multiple connections to each switch so that each server can access each switch. These switches are dedicated for use only by the DS8000.

Power and cooling

The DS8000 power and cooling system is highly redundant. There are two redundant primary power supplies (PPSs) in each frame of the DS8000, and each PPS can power the frame by itself. Furthermore, each processor and I/O enclosure has dual redundant power supplies.

The disk enclosures do not have separate power supplies since they draw power directly from the PPSs. They do however have cooling fans located in a plenum above the enclosures.

Battery backup assemblies

The backup battery assemblies help protect data in case of loss of external power. The model 921 contains two battery backup assemblies while the model 922 and 9A2 contain three of them (to support the 4-way processors). If input AC power is completely lost, the battery assemblies are used to allow the contents of NVS memory to be written to a number of DDMs internal to the processor complex, before power off. The FC-AL DDMs are not protected from power loss unless the extended power line disturbance feature has been purchased.

DS800 frames and enclosures

The base frame can contain up to eight disk enclosures, each can contain up to 16 disk drives. They are described as 16-packs because each enclosure can hold 16 disks. In a maximum configuration, the base frame can hold 128 disk drives. Each expansion frame can hold up to 16 disk enclosures which contain the disk drives. In a maximum configuration, an expansion frame can hold 256 disk drives.

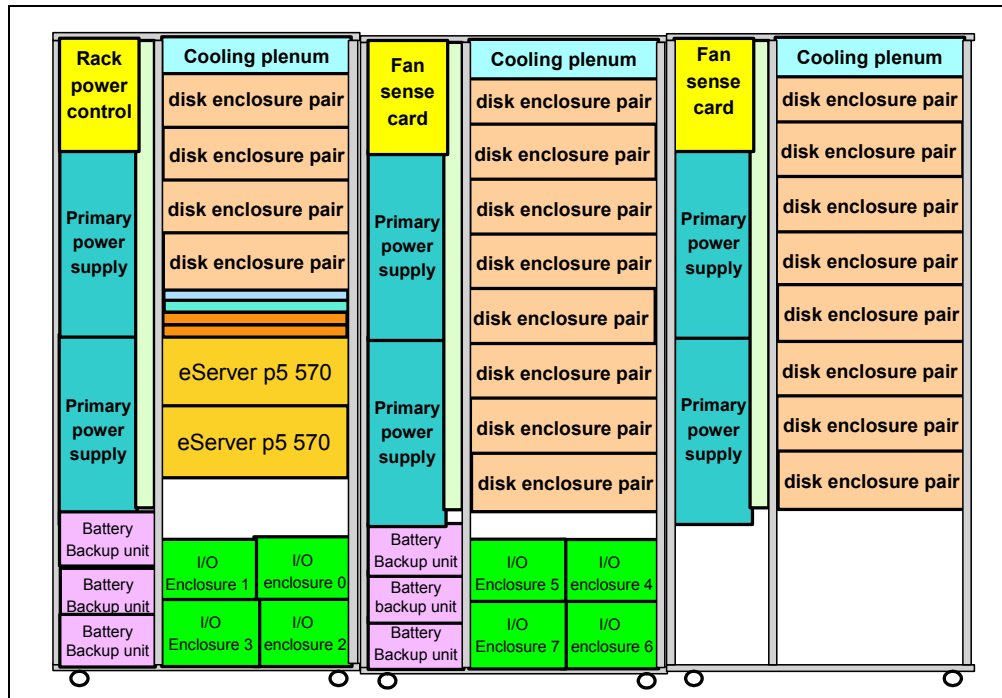


Figure 8-21 DS8000 Frames

Expansion frames can contain I/O enclosures and adapters if they are the first expansion frame attached to either a model 922 or a model 9A2. The second expansion frame in a model 922 or 9A2 configuration cannot have I/O enclosures and adapters (see Figure 8-21), nor can any expansion frame that is attached to a model 921. If the expansion frame contains I/O enclosures, the enclosures provide connectivity between the adapters and the

processors. The adapters contained in the I/O enclosures can be either device or host adapters.

8.4.3 DS8000 models

One of the advantages of the DS8000 series is its linear scalability for capacity and performance. If your business grows rapidly, you may need much more storage capacity, faster storage performance, or both. The DS8000 series can meet these demands within a single storage unit.

DS8100 Model 921

The DS8100 Model 921 (see Figure 8-22) has the following features:

- ▶ Two processor complexes with pSeries POWER5 1.5 GHz two-way CEC each
- ▶ Up to 128 DDMs for a maximum disk storage capacity of 38.4 TB with 300 GB DDMs
- ▶ Up to 128 GB of processor memory which used to be referred as cache with the ESS 800
- ▶ Up to 16 host adapters (four-port 2 Gb FCP/FICON or two port ESCON adapters)

The DS8100 model can support one expansion frame, Model 92E which can expand the disk storage capacity by up to 384 disk drives, for a maximum capacity of 115.2 TB.

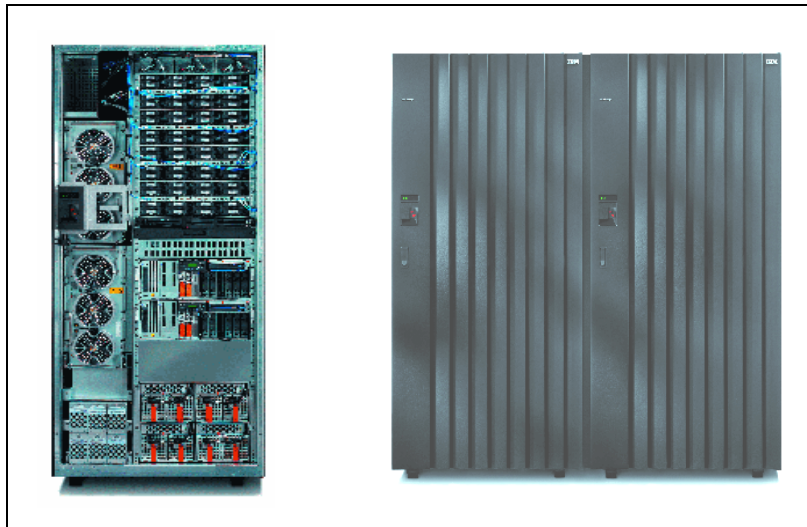


Figure 8-22 DS8100

Note: The Model 921 is field upgradeable to either a Model 922 or Model 9A2.

DS8300 Models 922 and 9A2

The DS8300 Model 922 and 9A2 (shown in Figure 8-23 on page 139) offer higher capacity and performance than the DS8100.

The Model 9A2 provides two storage images through two storage system LPARs within the same physical storage unit.



Figure 8-23 Model 922 and 9A2

Both models provide the following features:

- ▶ Two processor complexes with pSeries POWER5 1.9 GHz four-way CEC each
- ▶ Up to 128 DDMs for a maximum of 38.4 TB with 300 GB DDMs
- ▶ Up to 256 GB of processor memory which used to be referred as cache with the ESS 800
- ▶ Up to 16 host adapters (four-port 2 Gb FCP/FICON or two port ESCON adapters)

The DS8300 models can support either one or two expansion frames. With expansion frame 92E or 9AE, you can expand the disk storage capacity and number of adapters as follows:

- ▶ With one expansion frame, you can support the following expanded capacity and number of adapters:
 - Up to 384 disk drives, for a maximum capacity of 115.2 TB
 - Up to 32 Fibre Channel/FICON or ESCON host adapters
- ▶ With two expansion frames, you can support the following expanded capacity:
 - Up to 640 disk drives, for a maximum capacity of 192 TB

Table 8-2 summarizes the model capabilities.

Table 8-2 Comparison of models for capacity

	2-way base frame only	2-way + expansion frame	4-way base frame only	4-way + expansion frame	4-way + two expansion frames
Disk Drives: <ul style="list-style-type: none"> ▶ 73 GB 15k RPM ▶ 146 GB 10k RPM ▶ 300 GB 10k RPM 	16 to 128 (increment of 16)	16 to 384 (increment of 16)	16 to 128 (increment of 16)	16 to 384 (increment of 16)	16 to 640 (increment of 16)

	2-way base frame only	2-way + expansion frame	4-way base frame only	4-way + expansion frame	4-way + two expansion frames
Maximum Capacity	38.4 TB	115.2 TB	38.4 TB	115.2 TB	192 TB

8.4.4 IBM Standby Capacity on Demand offering for the DS8000

Standby Capacity on Demand (Standby CoD) provides standby on demand storage for the DS8000 and allows you to access the extra storage capacity whenever the need arises. With Standby CoD, IBM installs up to 64 drives (in quantities of 16) in your DS8000. At any time, you can logically configure your Standby CoD capacity for use. It is a nondestructive activity that does not require intervention from IBM. Upon logical configuration, you will be charged for the capacity.

8.4.5 More information

For more information about the DS8000 series, refer to these Redbooks:

- ▶ *The IBM TotalStorage DS8000 Series: Concepts and Architecture*, SG24-6452
- ▶ *The IBM TotalStorage DS8000 Series: Implementation*, SG24-6786
- ▶ *The IBM TotalStorage DS8000 Series: Copy Services with IBM eServer zSeries*, SG24-6787
- ▶ *IBM TotalStorage DS8000 Series: Copy Services in Open Environments*, SG24-6788
- ▶ *IBM TotalStorage DS8000 Series: Performance Monitoring and Tuning*, SG24-7146

8.5 The Enterprise Storage Server Family

The ESS was first made available in 1999 as an enterprise storage system with heterogeneous server connectivity and an extensive suite of advanced functions to support users' mission-critical, high-availability, multi-platform environments. The ESS set a new standard for storage servers at its release, and subsequently was enhanced with second and third generation models. The most current ESS model is the 800.

Note: The ESS Model 750 was withdrawn from marketing on April 28, 2006.

Figure 8-24 on page 141 provides general descriptions of each model and their design points.

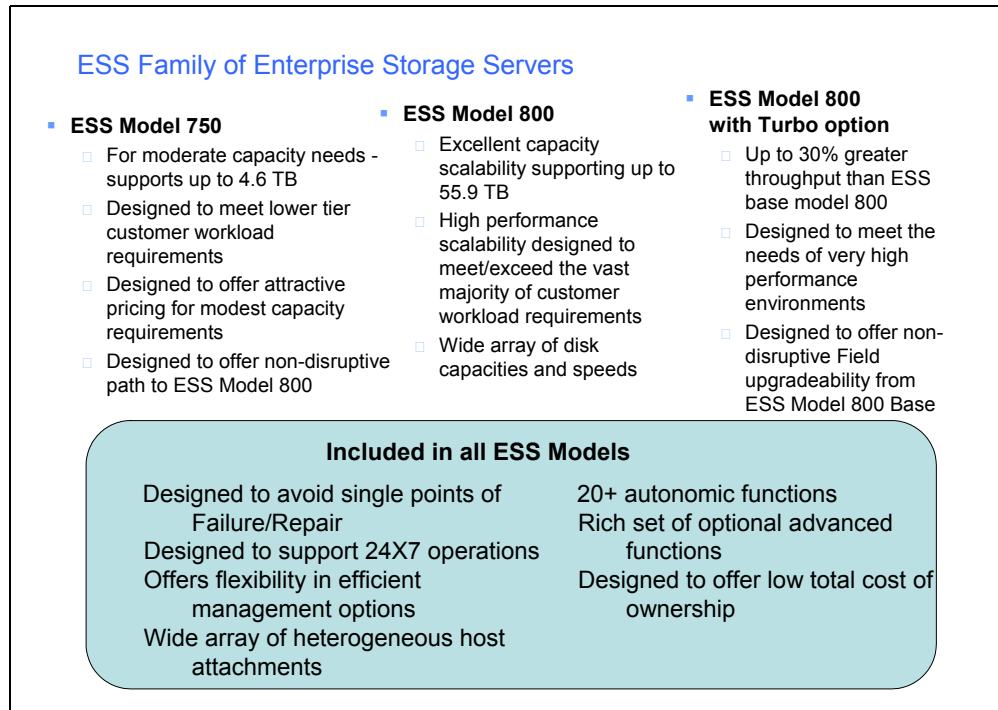


Figure 8-24 ESS Family

Key features that characterize the ESS Family are:

- ▶ SSA (Serial Storage Architecture) technology.
- ▶ A fault-tolerant system, dual active processing clusters with switching, hot spares, hot-swappable disk drives, mirrored write cache and redundant power and cooling.
- ▶ Efficient cache management and powerful back-end with high-performance SSA disk adapters.
- ▶ Array Across Loops (AAL) option, whereby disk arrays are spread across two loops on the SSA device adapter pair.

ESS Model 800

The ESS Model 800 is designed to support 24/7 operations and different type of workloads even when mixing dissimilar workload demands. For the zSeries workload, it has the option of a large cache and efficient cache algorithms.

The ESS Model 800 provides the following features:

- ▶ Two clusters, each contains RISC SMP processors providing two options:
 - Standard processor feature (can be upgraded to Turbo II processor)
 - Turbo II processor feature
- ▶ Up to 64 GB cache size
- ▶ 2 GB non-volatile storage (NVS)
- ▶ Capacity from 582 GB to 55.9 TB
- ▶ Disk drive capacities include:
 - 18.2 GB, 36.4 GB, 72.8 GB and 145.6 GB 10k rpm disks
 - 18.2 GB, 36.4 GB and 72.8 GB 15k rpm disks
- ▶ Up to 384 disk drives

- ▶ RAID-5 and either RAID-10 or non-raid disk groups. Arrays can also be configured across loops.
- ▶ Up to 32 SCSI or ESCON host ports, up to 16 FCP/FICON ports, and intermix configurations

ESS Model 750

The ESS Model 750 is well suited for clients with mid-range capacity and performance needs all in an affordable package. It includes many of the functions of the ESS Model 800 and all of its reliability. It is designed to meet the high availability requirements of mainframe and open systems environments and is an especially good fit for the IBM @server zSeries 8xx servers.

The ESS Model 800 provides the following features:

- ▶ Dual server-based using 2-way processors
- ▶ 8 or 16 GB cache size
- ▶ 2 GB non-volatile storage (NVS)
- ▶ Capacity from 1.1 TB to 4.6 TB using either 72.8 GB 10k rpm or 145.6 GB 10k rpm disk drives
- ▶ RAID-5 and RAID-10 are available and intermixable
- ▶ 64 non-arbitrated, pipelined paths to disks
- ▶ Up to 6 FCP/FICON (1 port each) or ESCON (2 port each) host adapters

Note: Array across loops (AAL) is not supported on the ESS Model 750.

The ESS Model 750 can be upgraded nondisruptively to the ESS Model 800.

8.6 Storage management

This section provides you information about configuration and management capabilities of the IBM System Storage Enterprise Disk Systems.

8.6.1 DS Management Console

The IBM DS6000 and DS8000 series offer new management tools and interfaces which are applicable to both products.

IBM System Storage DS Storage Manager GUI

The DS Storage Manager is a Web based GUI that is used to perform logical configurations and Copy Services management functions. It accessed via a Web browser.

You have the following options to use the DS Storage Manager:

- ▶ **Simulated (Offline) Configuration**

You can create or modify logical configurations when disconnected from the network. After creating the configuration, you can save it and then apply it to the storage system later when connected to the network. Shown as Figure 8-25 on page 143:



Figure 8-25 DS6000 Simulated Manager

► **Real-time (Online) Configuration**

This provides real-time management support for logical configuration and Copy Services features for a network attached storage system. Shown as Figure 8-26:

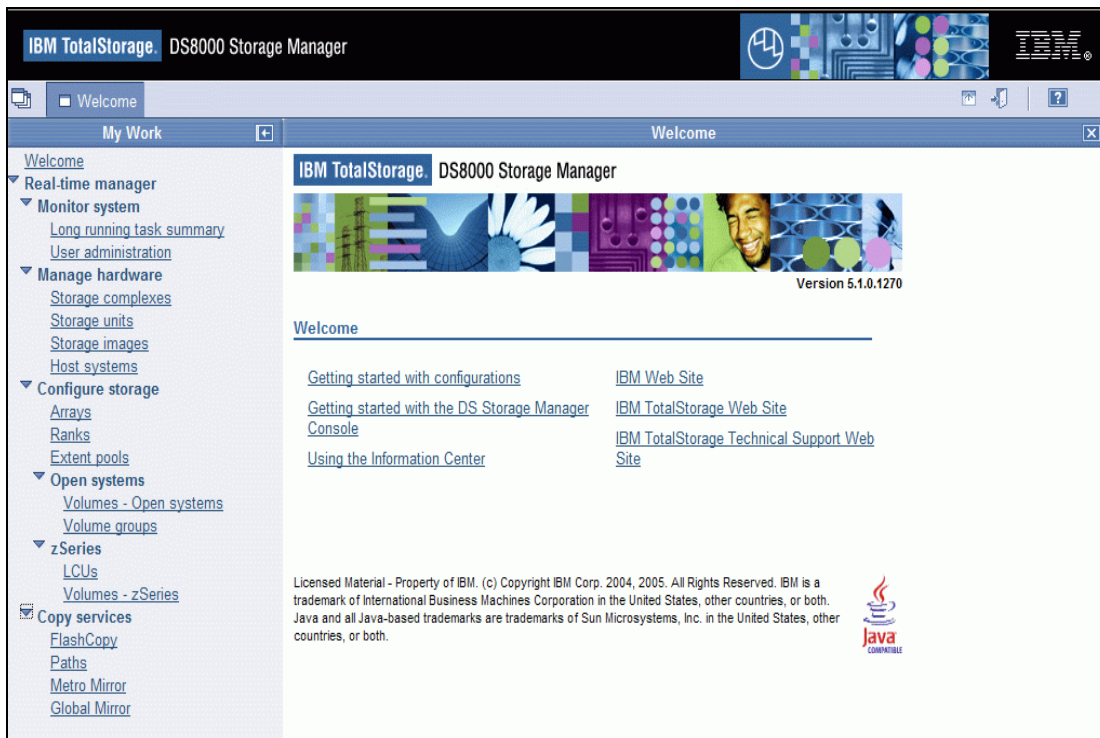


Figure 8-26 DS8000 Real-time Manager

The DS8000 Storage Manager contains more diagnostic features than the DS6000 GUI.

IBM System Storage DS Command-Line Interface (CLI)

The DS CLI is a single CLI which can perform a full set of commands for logical configuration and/or Copy Services activities. The DS CLI can also issue Copy Services commands to an ESS Model 750 or Model 800. IDS CLI commands can also be saved in a script, which can enhance productivity since it eliminates the previous (on ESS) requirement for to create and save a task using the GUI. Shown in Example 8-1:

The following list highlights a few of the specific types of functions that you can perform with the DS Command-Line Interface:

- ▶ Check and verify your storage unit configuration
- ▶ Check the current Copy Services configuration that is used by the storage unit
- ▶ Create new logical storage and Copy Services configuration settings
- ▶ Modify or delete logical storage and Copy Services configuration settings

Example 8-1 DS CLI

```
dsccli> lssi
Date/Time: 27 June, 2005 9:47:23 IBM DSCLI Version: 5.1.0.276
Name ID          Storage Unit      Model WWNN          State  ESSNet
=====
IBM.2107-75ABTV1 IBM.2107-75ABTV0 9A2   5005076303FFC663 Online Enabled
IBM.2107-75ABTV2 IBM.2107-75ABTV0 9A2   5005076303FFCE63 Online Enabled
dsccli> lsarray
Date/Time: 27 June 2006 9:48:27 IBM DSCLI Version: 5.1.0.276 DS: IBM.2107-7ABTV1
Array State Data RAIDtype  arsite Rank DA Pair DDMcap (10^9B)
=====
A0 Assigned Normal 5 (6+P+S) S1    R0   0           146.0
A1 Assigned Normal 5 (6+P+S) S2    R1   0           146.0
dsccli>
```

Figure 8-27 DS CLI

DS Open application programming interface (API)

The DS Open application programming interface (API) is a non-proprietary storage management client application that supports routine LUN management activities, such as LUN creation, mapping and masking, and the creation or deletion of RAID-5 and RAID-10 volume spaces. The DS Open API also enables Copy Services functions such as FlashCopy and Remote Mirror and Copy.

8.6.2 IBM TotalStorage Productivity Center

The IBM TotalStorage Productivity Center is an open storage management solution that helps to reduce the effort of managing complex storage infrastructures, to increase storage capacity utilization and to improve administrative efficiency. It is designed to enable an agile storage infrastructure that can respond to on demand storage needs.

TPC is the integration point for storage and fabric management and replication. It provides a launchpad for the following IBM TotalStorage Open Software Family and Tivoli products:

- ▶ IBM TotalStorage Productivity Center for Disk
- ▶ IBM TotalStorage Productivity Center for Data
- ▶ IBM TotalStorage Productivity Center for Fabric

- ▶ IBM TotalStorage Productivity Center for Replication

For more information about the IBM TotalStorage Productivity Center refer to Chapter 19, “IBM TotalStorage Productivity Center” on page 381.

8.7 Copy Services

Advanced Copy Services are enterprise-level, leading-edge functions which address an organization’s needs for disaster recovery, data migration, and data duplication.

8.7.1 IBM System Storage FlashCopy (Point-in-time Copy feature)

The IBM System Storage FlashCopy (see Figure 8-28) helps reduce or eliminate planned outages for critical applications. FlashCopy enables data to be copied in the background while making both source and copied data available to users almost immediately. The point-in-time copy created by FlashCopy is typically used when a copy of production data must be produced with minimal application downtime. It can be used for online backup, testing of new applications, or for copying a database for data mining purposes.

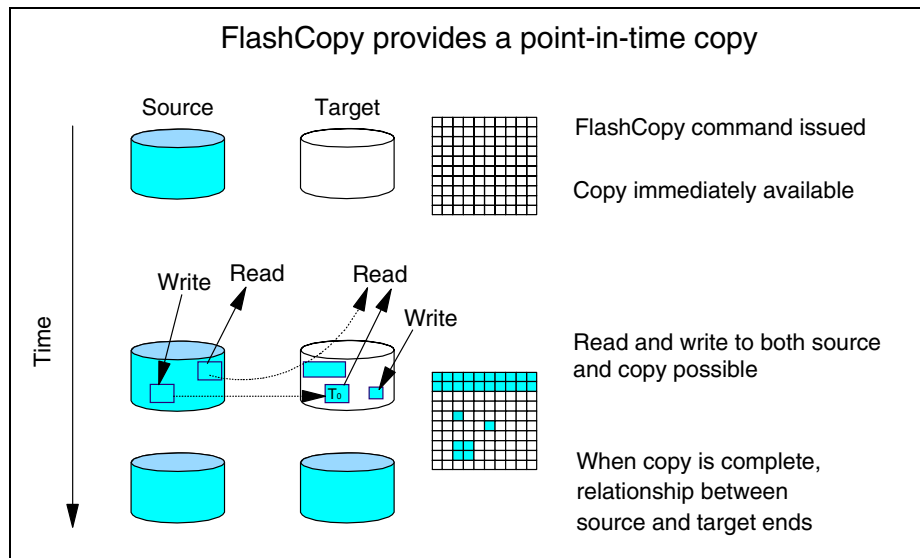


Figure 8-28 FlashCopy

FlashCopy is an additional charged feature. You have to order the Point-in-time Copy feature which includes FlashCopy. FlashCopy is interoperable with the ESS

FlashCopy supports many advanced capabilities, including:

- ▶ **Data Set FlashCopy**
Allows a FlashCopy of a data set in a System z environment.
- ▶ **Multiple Relationship FlashCopy**
Allows a source volume to have multiple targets simultaneously.
- ▶ **Incremental FlashCopy**
Allows update of a FlashCopy target without having to recopy the entire volume.

► **Consistency Group function**

Used to help create a consistent point-in-time copy across multiple LUNs or volumes, and even across multiple DS6800 systems as well as across DS8000 series, ESS 800 and ESS 750 systems.

► **Inband commands over remote mirror link**

In a remote mirror environment, commands to manage FlashCopy at the remote site can be issued from the local or intermediate site and transmitted over the remote mirror Fibre Channel links. This eliminates the need for a network connection to the remote site solely for the management of FlashCopy.

► **FlashCopy to a Remote Mirror Primary**

Allows a FlashCopy target volume to be used also as a remote mirror primary volume. This process allows you to create a point-in-time copy and then make a copy of that data at a remote site with Metro Mirror or Global Copy. Shown as Figure 8-29:

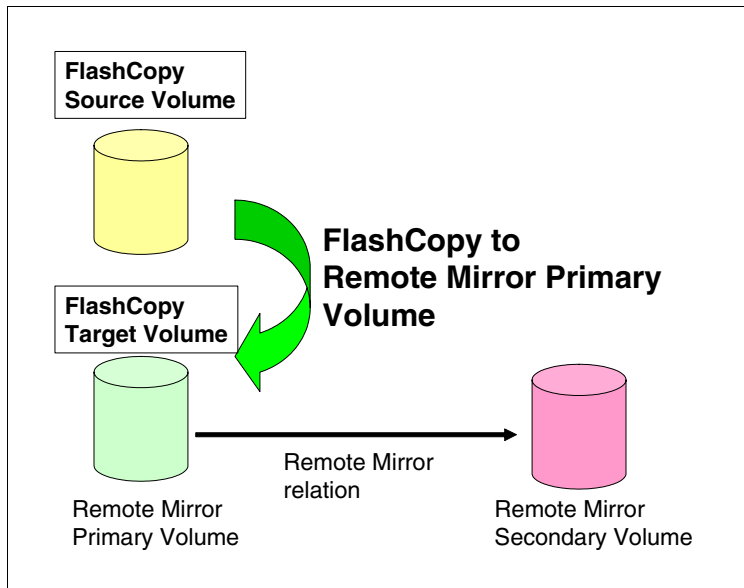


Figure 8-29 Establish FlashCopy on existing Remote Mirror and Copy Primary

8.7.2 Remote Mirror and Copy feature

Remote Mirror and Copy is another separately orderable priced feature which includes Metro Mirror, Global Copy and Global Mirror. The local and remote storage systems must have a Fibre Channel connection between them. The Fibre Channel ports used for Remote Mirror and Copy can be configured either as a dedicated remote mirror link or as a shared port between remote mirroring and Fibre Channel Protocol (FCP) data traffic.

The DS6000 and DS8000 series systems can participate in remote mirror and copy solutions with each other as well as with the ESS Model 800 and Model 750 systems.

IBM supports the following remote mirror and copy solutions:

IBM System Storage Metro Mirror (Synchronous PPRC)

Metro Mirror is a remote-mirroring technique for all supported servers, including z/OS and open systems. It is designed to constantly maintain an up to date copy of the local application data at a remote site which is within the metropolitan area (typically up to 300 km away using DWDM). With synchronous mirroring techniques, data currency is maintained between sites,

though the distance can have some impact on performance. Metro Mirror is used primarily as part of a business continuance solution for protecting data against disk storage system loss or complete site failure. Shown in Figure 8-30:

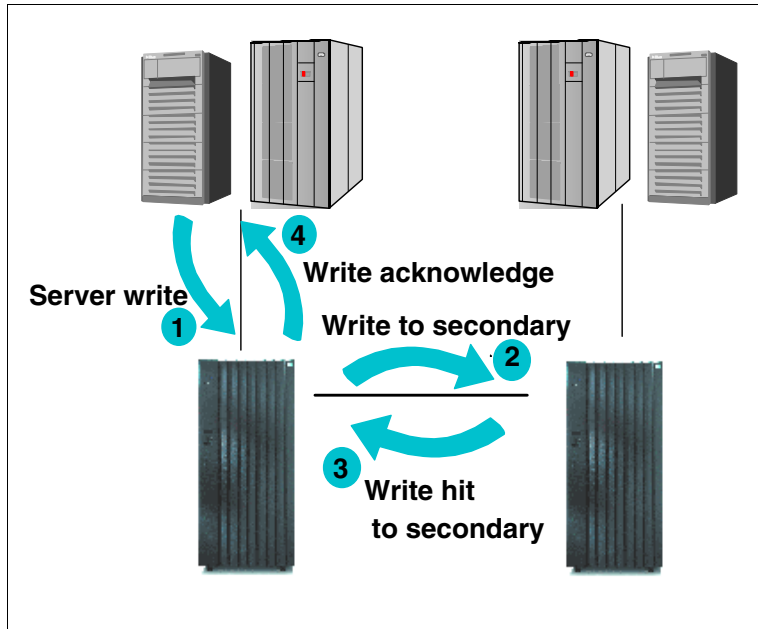


Figure 8-30 Metro Mirror

IBM System Storage Global Copy (Asynchronous PPRC-XD)

Global Copy (see Figure 8-31 on page 148) is an asynchronous remote copy function for z/OS and open systems for longer distances than are possible with Metro Mirror. With Global Copy, write operations complete on the primary storage system before they are received by the secondary system. This capability is designed to prevent the primary system's performance from being affected by wait-time from writes on the secondary system. Therefore, the primary and secondary copies can be separated by any distance. This function is appropriate for remote data migration, off-site backups and transmission of inactive database logs at virtually unlimited distances.

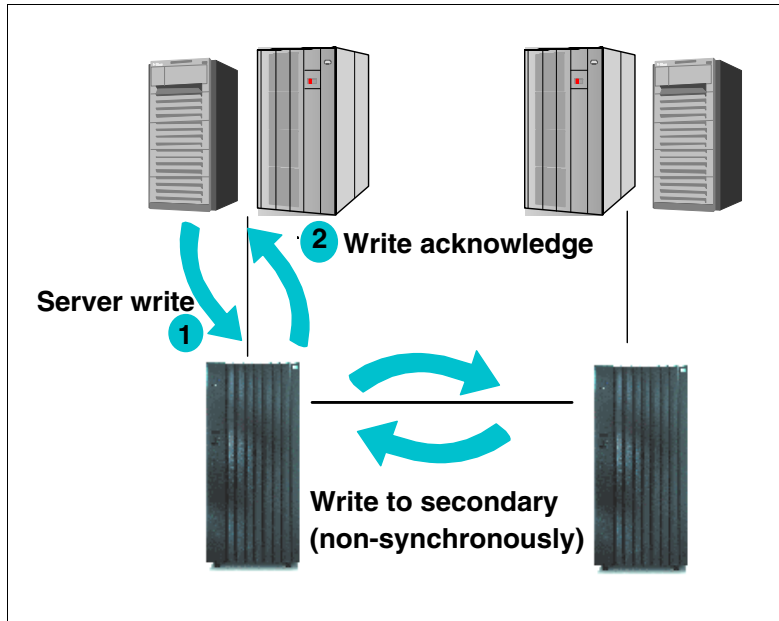


Figure 8-31 Global Copy

IBM System Storage Global Mirror (Asynchronous PPRC)

Global Mirror (Figure 8-32 on page 149) copying provides a two-site extended distance remote mirroring function for z/OS and open systems servers. With Global Mirror, the data that the host writes to the storage unit at the local site is asynchronously shadowed to the storage unit at the remote site. A consistent copy of the data is then automatically maintained on the storage unit at the remote site. This two site data mirroring function is designed to provide a high-performance, cost-effective global distance data replication and disaster recovery solution.

Global Mirror provide the following benefits:

- ▶ Support for virtually unlimited distances between the local and remote sites. The distances are typically limited only by the capabilities of the network and channel extension products.
- ▶ A consistent and restartable copy of the data at the remote site, created with minimal impact to applications at the local site.
- ▶ Dynamic selection of the desired recovery point objectives
- ▶ Efficient synchronization of the local and remote sites with support for failover and failback modes.

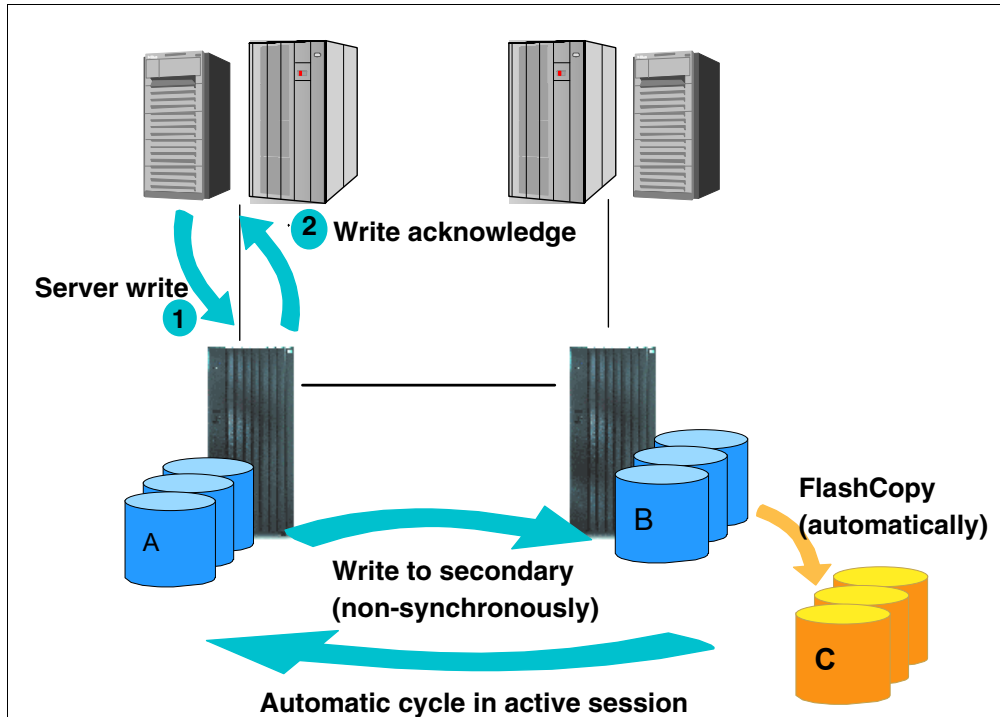


Figure 8-32 Global Mirror

IBM System Storage z/OS Global Mirror (Extended Remote Copy XRC)

z/OS Global Mirror is a remote data mirroring function available for the z/OS and OS/390 operating systems. It maintains a copy of the data asynchronously at a remote location over unlimited distances. z/OS Global Mirror is well suited for large zSeries server workloads and can be used for business continuance solutions, workload movement and data migration.

Note: The DS6000 series systems can only be used as a target system in z/OS Global Mirror operations. z/OS Global Mirror is not supported on the ESS Model 750.

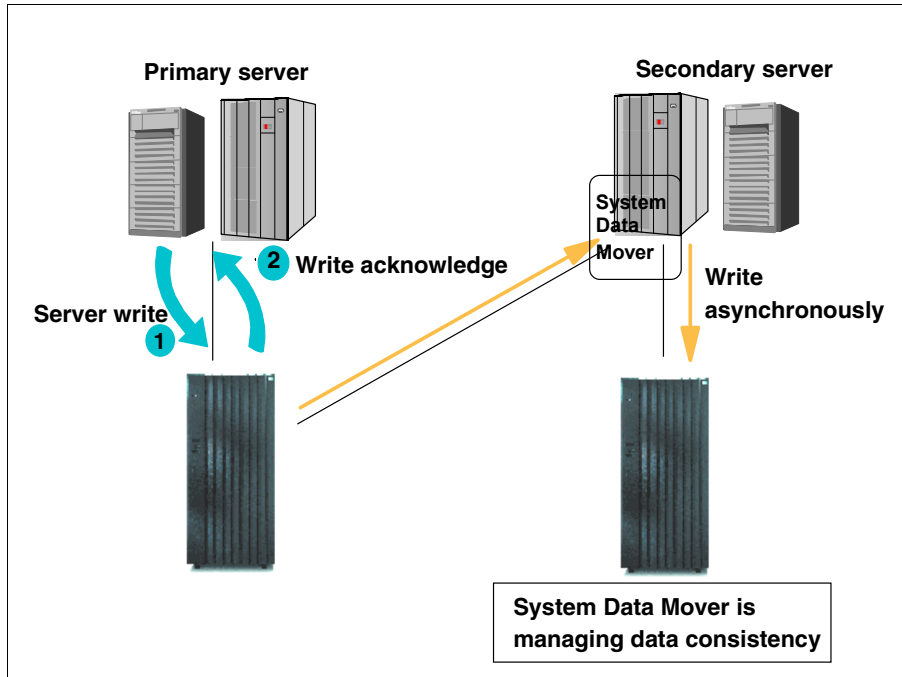


Figure 8-33 z/OS Global Mirror

IBM System Storage z/OS Metro/Global Mirror

This mirroring capability utilizes z/OS Global Mirror to mirror primary site data to a location that is a long distance away and also uses Metro Mirror to mirror primary site data to a location within the metropolitan area. This enables a 3-site high availability and disaster recovery z/OS solution for even greater protection from unplanned outages. Shown as Figure 8-34 on page 151:

Note: z/OS Metro/Global Mirror is not supported on DS6000 series.

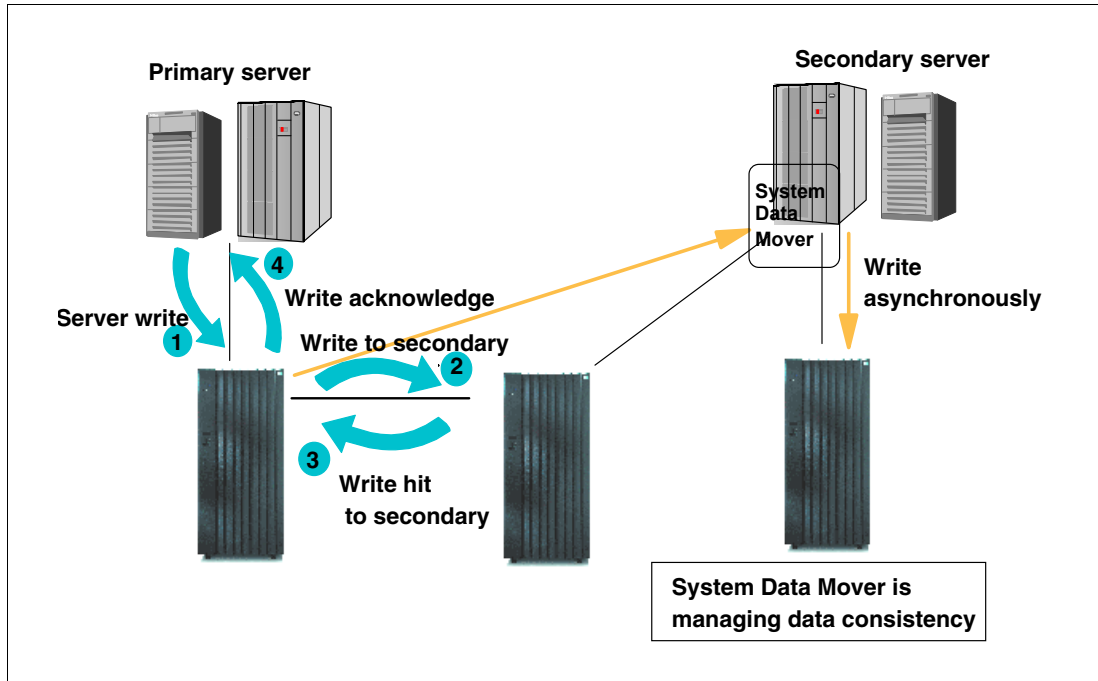


Figure 8-34 z/OS Metro/Global Mirror

8.8 Comparison of products

Here we compare the DS6000 and the DS8000 to each other and to the ESS which is the original enterprise storage server for z and open systems. We also provide comparison of the DS6000 to DS4000.

DS6000 series compared to DS4000

Those familiar with the DS4000 family will find substantial difference between the DS4000 series and DS6000 series of products. Both product families have about the same footprint and capacity, but their functions differ. DS6000 offers higher performance compared to the DS4000.

The DS6000 offers enterprise capabilities not found in mid-range offerings, since it originates from the ESS roots and provides support for System z and System i™ servers as well as open systems.

The implementation of FlashCopy on the DS4000 is different compared to the DS6000 series. For example, the target system on a DS4000 cannot be used for production, while it can on the DS6000.

While the DS4000 also offers remote copy solutions, these functions are not compatible with the DS6000.

DS6000 series compared to ESS

IT sites with ESS will find it very easy to replace their old systems with a DS6000. All functions (with the exception of cascading Metro/Global Copy and z/OS Global Mirror), are the same as on the ESS and are also available on a DS6000.

If you want to keep your ESS and if it is a Model 800 or 750 with Fibre Channel adapters, you can use your old ESS, for example, as a secondary for remote copy. With the ESS at the appropriate LIC level, scripts or CLI commands written for Copy Services will work for both the ESS and the DS6800.

For most environments the DS6800 performs much better than an ESS. You might even replace two ESS 800s with one DS6800. The sequential performance of the DS6800 is excellent. However, when you plan to replace an ESS with a large cache (let's say more than 16 GB) with a DS6800 (which comes with 4 GB cache) and you currently get the benefit of a high cache hit rate, your cache hit rate on the DS6800 will drop down. This is because of the smaller cache. z/OS benefits from large cache, so for transaction-oriented workloads with high read cache hits, careful planning is required.

DS6000 series compared to DS8000 series

You can think of the DS6000 as the smaller sibling of the DS8000. The DS8000 and the DS6000 can work with each other perfectly because of their common switched FC-AL architecture.

All Copy Services (with the exception of z/OS Global Mirror) are available on both systems. You can do Metro Mirror, Global Mirror, and Global Copy between the two series. The CLI commands and the DS Storage Manager GUI look the same for both systems.

So it is very easy to have a mixed environment with DS8000 and DS6000 systems to optimize the cost-effectiveness of your storage solution, while providing the cost efficiencies of common skills and management functions.

The DS8000 can deliver a higher throughput and a dramatic leap in performance. It provides high I/O bandwidth and self-healing interconnect with RIO-G architecture which is only available on the DS8000. At the same time, the DS8000 offers extensive virtualization capabilities with the industry's first implementation of storage logical partitioning. Storage system LPARs are not available on the DS6000.

DS8000 series compared to ESS

The DS8000 is the next generation of the (ESS), so all functions which are available in the ESS are also available in the DS8000 (with the exception of Metro/Global Copy). From a consolidation perspective, it is now possible to replace four ESS Models 800 with one DS8300. And with the LPAR implementation you get an additional consolidation opportunity, because you get two storage system logical partitions in one physical machine.

Since the mirror solutions are compatible between the ESS and the DS8000, it is possible to think about a set up for a disaster recovery solution with the high-performance system DS8000 at the primary site and the ESS at the secondary site, where the same performance is not required.

8.9 Additional information

This section gives complementary information about the support and requirements of the IBM System Storage Enterprise Family products.

Licensed Machine Code (LIC)

The IBM System Storage Enterprise Systems are shipped with IBM Licensed Internal Code (LIC) that is licensed for use by a customer on a specific machine, designated by serial number, under the terms and conditions of the IBM Customer Agreement or the IBM

Agreement for Licensed Internal Code. New releases of LIC offer improvements in both function and reliability. Customers are strongly recommended to obtain the latest level of LIC.

All features and functions are available either as a standard or as an optional feature and include the following:

- ▶ Operating environment and assigned capacity
- ▶ Host attachment capabilities
- ▶ Parallel access volumes on zSeries
- ▶ LPAR and S-HMC capabilities on the DS8000 series
- ▶ Point-in-time Copy features
- ▶ Remote mirror and copy features

You can activate and manage licensed features from the System Storage DS Manager GUI on DS6000 and DS8000 series. Shown as Figure 8-35:

The screenshot shows the IBM TotalStorage DS Manager GUI. The main content area is titled "View machine summary" for an "IBM TotalStorage DS8300". It provides instructions on how to verify functions and lists actions like "Assign a function authorization" and "Manage an activation". Below this, there are several tables showing feature activation details for different models and serial numbers.

IBM 2107 Model 922 Serial number 13-11060			
Feature code	Description		
0700	OEL indicator		
0720	PTC indicator		
0740	RMC indicator		
0760	RMZ indicator		
0780	PAV indicator		
0900	Non-standby CoD		

IBM 2244 Model OEL Serial number 13-000F1			
Description	Total license	Assigned	Unassigned
Operating environment	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model PAV Serial number 13-000F2			
Description	Total license	Assigned	Unassigned
Parallel access volumes	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model PTC Serial number 13-000F3			
Description	Total license	Assigned	Unassigned
Point in time copy	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model RMC Serial number 13-000F4			
Description	Total license	Assigned	Unassigned
Remote mirror and copy	25.0 TB	25.0 TB	0.0 TB

IBM 2244 Model RMZ Serial number 13-000F5			
Description	Total license	Assigned	Unassigned
Remote mirror for z/OS	25.0 TB	25.0 TB	0.0 TB

Figure 8-35 DS Storage Manager feature activation

Supported environment

The IBM System Storage Enterprise Disk Systems can be connected across a broad range of server environment such as:

- ▶ System z
- ▶ System i
- ▶ System p
- ▶ System x™
- ▶ Servers from SUN Microsystems
- ▶ Servers from Hewlett-Packard
- ▶ Other non-IBM based server platforms

New versions of operating systems, servers, file systems, host bus adapters, clustering products and SAN components are constantly announced in the market. Information about the supported environments changes frequently. Therefore you are strongly advised always to refer to the online resources.

The Interoperability Matrix always provides the latest information about supported platforms, operating systems, host adapters and SAN infrastructure solutions. It contains detailed specifications about models and versions. It also lists special support items, such as boot support, and exceptions.

- ▶ The DS8000 Interoperability Matrix:
<http://www-03.ibm.com/servers/storage/disk/ds8000/pdf/ds8000-matrix.pdf>
- ▶ The DS6000 Interoperability Matrix:
<http://www-03.ibm.com/servers/storage/disk/ds6000/pdf/ds6000-interop.pdf>



IBM System Storage N series

This chapter describes the features and functionalities of the IBM System Storage N series.

It covers the following topics:

- ▶ N series hardware overview
- ▶ System Storage N3700
- ▶ System Storage N5000 series
- ▶ N series expansion units
- ▶ Data ONTAP software overview
- ▶ N series software solutions

9.1 N series hardware overview

The IBM System Storage N series offers additional choices to organizations facing the challenges of enterprise data management. The IBM System Storage N series is designed to deliver high-end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy. Figure 9-1 summarizes the current products available.

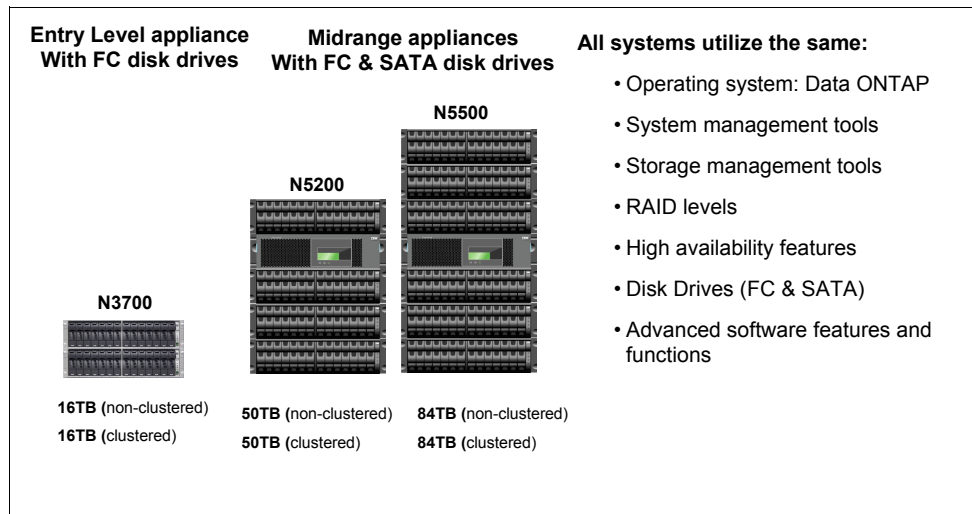


Figure 9-1 N series hardware overview

The N series products provide a wide-range of network attachment capabilities to a broad range of host and client systems using multiple network access protocols including file system NAS protocols (CIFS, NFS) and block I/O protocols including iSCSI and FCP all from a single hardware platform, simultaneously.

The N series products are very flexible, since they can be populated with both Fibre Channel and SATA disk drives.

An N series using Fibre Channel disk drives may be suitable for mission-critical high-performance data transaction environments.

An N series using SATA disk drives provides an economical platform for disk to disk backup, disaster recovery, data archive, or data like home directories which do not require high-performance transactional environments.

All N series systems utilize a single operating system (Data ONTAP) across the entire platform and offer a combination of multiple advanced function software features for comprehensive system management, storage management, onboard and outboard copy services, virtualization technologies, and disaster recovery and backup solutions.

Optional WORM data protection software provides additional data security in regulatory environments where data must be stored in non-erasable and non-rewritable formats to meet the industry's newest and strict regulatory requirements for retaining company data assets.

The N series portfolio of products offer ease-of-use tools that help manage database environments like Microsoft Exchange, Microsoft SQL, and Oracle. Patented RAID-DP (Raid Double Parity) helps ensure the highest availability and data loss prevention while using inexpensive SATA disk drive technology.

9.2 System Storage N3700

IBM System Storage N3700, shown in Figure 9-2, is a 3U rack-mountable storage system for storage in NAS, Internet Small Computer System Interface (iSCSI), Fibre Channel (FCP) or combined environments. The basic N3700 offering is a single-node model A10, which can be upgraded to the dual-node model A20 in the same rack space. The dual-node, clustered A20 supports failover and failback functions to maximize reliability. The N3700 can support 14 internal hot-plug disk drives, and can scale in capacity by attaching up to three 3U EXN2000 expansion units, each with a maximum of 14 drives. This gives the N3700 a maximum physical storage capacity of 16.8 TB physical storage capacity.

The N3700 also can connect to a Fibre Channel tape device for backup. A list of supported tape devices is at:

<http://www.ibm.com/servers/storage/nas/interophome.html>

Refer to the System Storage N series interoperability Matrix.



Figure 9-2 System Storage N3700

9.2.1 N3700 hardware features

There are two models available:

- ▶ N3700 A10 (Single Controller)
- ▶ N3700 A20 (Dual Controller)

The IBM System Storage N3700 has the following features:

- ▶ 3U rack-mountable integrated filer and disk storage enclosure
- ▶ Redundant hot plug power supplies and cooling
- ▶ Two integrated full duplex 10/100/1000 Ethernet ports per NAS controller
- ▶ Two integrated FC ports per NAS controller
- ▶ All N3700 devices shipped with root volume using RAID-DP (double-parity). There are 3 disks in the root volume (1 data and 2 parity).

N3700 Series Minimum/Maximum RAID configurations:

- ▶ RAID 4 Minimum: 1 data + 1 parity / Maximum: 13 data + 1 parity
- ▶ RAID-DP Minimum: 1 data + 2 parity / Maximum: 26 data + 2 parity (FC)

Disk storage expansion units supported: (note you can only attach one type of disk expansion unit to any particular N3700 - EXN2000 and EXN1000 cannot coexist on the same N3700). The base chassis does not support SATA disk drives.

- ▶ EXN2000, (formerly EXP600) - FC Disk Storage Expansion Unit
- ▶ EXN1000 - SATA Disk Storage Expansion Unit

Disk drive capacities supported (Note, you can only attach a maximum of two distinct drive types - rotational speed and capacity to any N3700 system):

- ▶ EXN2000 - 10K rpm FC disk drives (72 GB, 144 GB, 300 GB), 15K RPM FC disk drives (72 GB, 144 GB)
- ▶ EXN1000 - 7200 rpm SATA disk drives (250 GB, 320 GB, 500 GB)

9.2.2 N3700 specifications

Table 9-1 summarizes the specifications of the N3700.

Table 9-1 N3700 specifications

File specifications	N3700 A10	N3700 A20
IBM machine type/model	2863-A10	2863-A20
Max. raw capacity	16.8 TB	16.8 TB
Max. number of disk drives	56	56
Volumes and RAID groups (min./max. of 1 RAID group per volume)	1/7	1/7
Max. volume size	8 TB	8 TB
ECC memory	1 GB	2 GB
Nonvolatile memory	128 MB	256 MB
Ethernet 10/100/1000 copper	2	4
Copper FC adapter	1	2
Optical FC adapter (host-attach SAN/Tape SAN)	1	2
Clustered failover-capable	No	Yes
Rack Mount (in IBM 2101 Storage Solutions Rack Model 200 or other industry-standard 19-inch rack)	Yes	Yes

9.3 System Storage N5000 series

The IBM System Storage N5000 series, shown in Figure 9-3, supports NAS, iSCSI, FC, or combined environments.



Figure 9-3 System Storage N5000

The IBM System Storage N5000 series comes in two flavors, an appliance model, and a gateway model. Both models come without storage in the base chassis. The appliance model can attach one to twenty-four Fibre Channel EXN2000 and SATA EXN1000 disk expansion units. The gateway model supports the attachment of external storage from IBM, or other vendors.

For both product lines there are two versions of N5000 available, the N5200, and the N5500. The N5200 scales up to 50 TB (Gateway) or 84 TB (Appliance), and the N5500 to 84 TB (Gateway) or 168 TB (Appliance) physical storage capacity. Both systems are available as a single-node model (A10 or G10 for appliance/gateway respectively), which is upgradeable to an dual-node model (A20/G20).

Data ONTAP, with additional functions such as business continuance, is available through optional licensed functions.

9.3.1 System Storage N5000 series highlights

Reliable — Addresses the needs of business- and mission-critical applications through high data availability and system-level redundancy features

Versatile — Single, integrated architecture designed to support concurrent block I/O and file serving over Ethernet and FC SAN infrastructures

Fast — Supports high throughput and fast response times for database, e-mail, and technical applications

Flexible — FC and SATA disk drive capabilities allow for deployment in multiple solution environments including data compliant retention, nearline storage, disk-to-disk backup scenarios and high-performance, mission-critical I/O intensive operations

9.3.2 Differences between appliance and gateway models

Both the N series gateway and appliance models use the same storage controllers, operating system (Data ONTAP), system management tools, storage management tools, network access protocols, and high availability features. However, unlike the appliance, the gateway itself has no RAID functionality. Both models also utilize the same advanced software features and functions, with the exception of the features and functions which are mentioned in Table 9-2 on page 160.

Table 9-2 Differences between appliance and gateway models

	Appliance (A10 and A20)	Gateway (G10 and G20)
Filer Head (base chassis)	Filer heads are the same for both appliances and gateways	
Storage	Filer uses EXN2000 and SATA EXN1000 disk expansion units	Storage is provided by external RAID subsystem
Operating System	Data ONTAP	Data ONTAP slightly modified to remove RAID. RAID is dependent on the attached storage subsystem.
Software	35+ features and functions	Same as appliance models except for: Disk Sanitization, SnapLock, Compliance, LockVault Compliance, Nearline™ Feature, RAID-DP, RAID4
Performance	Dependent on filer head and storage system or disk expansion units	

Why use an N series gateway?

A typical heterogeneous environment can have a combination of storage arrays from IBM and other vendors, all separately managed (Figure 9-4).

With the N series gateway, you can extend N series unique management capabilities and ease of use to heterogeneous systems in both NAS and SAN environments. N series gateways provide a unified storage architecture virtualizing backend arrays, a single point of management, and multiprotocol access.

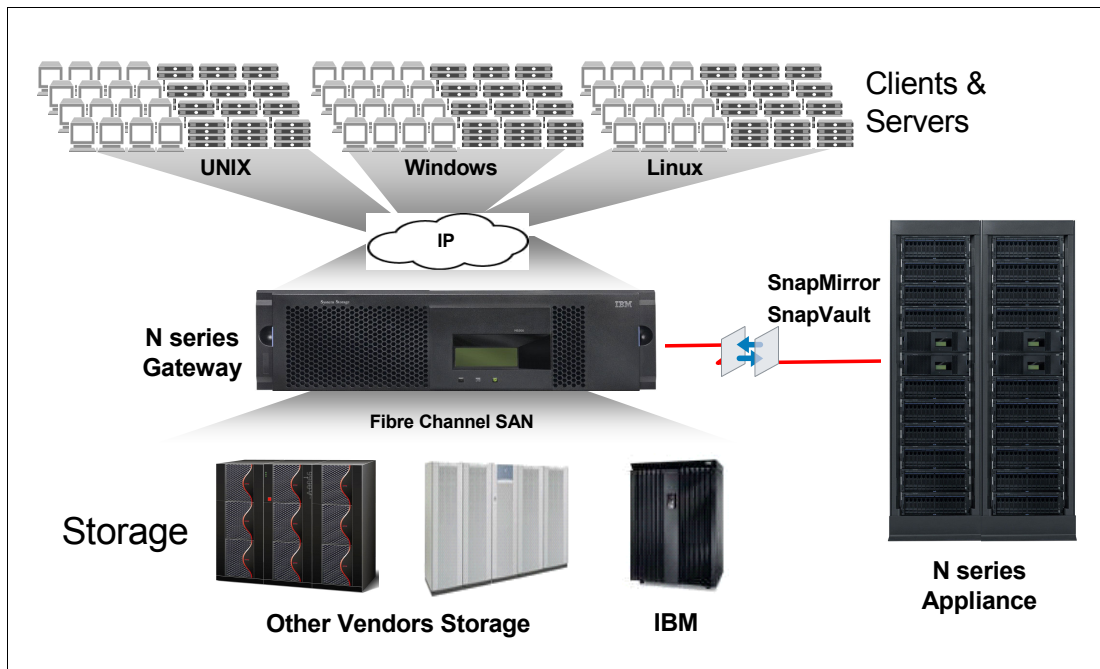


Figure 9-4 Gateway provides functionality of N series technology to other vendor storage subsystems

Finally, using the gateway allows the use N series replication technologies across all the different storage systems. This provides, in one solution, a heterogeneous backup and business continuance plan, a disaster recovery solution, and in the face of rising concerns regarding data compliance, a way to secure data.

9.3.3 N5000 hardware features

There are four Appliance models and four Gateway models available:

- ▶ N5200 A10 (Single Appliance)
- ▶ N5200 A20 (Clustered Appliance)
- ▶ N5500 A10 (Single Appliance)
- ▶ N5500 A20 (Clustered Appliance)
- ▶ N5200 G10 (Single Gateway)
- ▶ N5200 G20 (Clustered Gateway)
- ▶ N5500 G10 (Single Gateway)
- ▶ N5500 G20 (Clustered Gateway)

Each IBM System Storage N5000 has the following standard features:

- ▶ 19" rack-mount enclosure
- ▶ Dual redundant hot-plug integrated cooling fans and auto-ranging power supplies
- ▶ Four full-duplex 10/100/1000 Base-T Ethernet ports onboard
- ▶ Four 2 Gbps Fibre Channel ports onboard
- ▶ Built-in LVD SCSI Port

N5000 appliance series Default/Minimum/Maximum RAID group sizes:

- ▶ RAID-4 (single-parity) - FC-Data/Parity: 7+1 (13+1); SATA-Data/Parity: 7+1 (7+1)
- ▶ RAID-DP (double-parity) - FC-Data/Parity: 14+2 (26+2); SATA-Data/Parity: 12+2 (14+2)

Disk storage expansion units supported on N5000 appliances:

- ▶ EXN2000, (formerly EXP600) - FC Disk Storage Expansion Unit
- ▶ EXN1000 - SATA Disk Storage Expansion Unit

Disk drive capacities supported on N5000 appliances:

- ▶ EXN2000 - 10K RPM FC disk drives (72 GB, 144 GB, 300 GB), 15K RPM FC disk drives (72 GB, 144 GB)
- ▶ EXN1000 - 7200 RPM SATA disk drives (250 GB, 320 GB, 500 GB)

9.3.4 N5000 specifications

Table 9-3 summarizes the specifications of the N5000 models, Table 9-4 on page 162 shows the maximum disk capacities for various options on the appliance models, and Table 9-5 on page 162 shows the characteristics of various options on the gateway models, and Figure 9-6 shows the I/O expandability options.

Table 9-3 N5000 specifications

File specifications	N5200	N5200	N5500	N5500
IBM machine types - models	2864-A10 2864-G10	2864-A20 2864-G20	2865-A10 2865-G10	2865-A20 2865-G20
Storage configuration	Single storage controller	Dual (active/active) storage controllers	Single storage controller	Dual (active/active) storage controllers
ECC memory	2 GB	4 GB	4 GB	8 GB
Nonvolatile memory	512 MB	1 GB	512 MB	1 GB

File specifications	N5200	N5200	N5500	N5500
Onboard 10/100/1000 Ethernet ports	4	8	4	8
Onboard 2 Gbps Fibre Channel ports (configurable as storage-attached initiator or host-attached target)	4	8	4	8

Table 9-4 N5000 appliance specifications

Appliance disk specifications	N5200	N5200	N5500	N5500
IBM machine types - models	2864-A10	2864-A20	2865-A10	2865-A20
Max. number of Fibre Channel loops	4	8	4	8
Max. raw storage capacity EXN2000 with all				
72 GB FC disks	12 TB	12 TB	24 TB	24 TB
144 GB FC disks	24 TB	24 TB	48 TB	48 TB
300 GB FC disks	50 TB	50 TB	100 TB	100 TB
Max. raw storage capacity EXN1000 with all				
250 GB SATA disk	42 TB	42 TB	84 TB	84 TB
320 GB SATA disk	53 TB	53 TB	107 TB	107 TB
500 GB SATA disk	84 TB	84 TB	168 TB	168 TB
Max. storage expansion units EXN2000	12	12	24	24
Max. storage expansion units EXN1000	12	12	24	24
Max. number of disk drives or LUNs	168	168	336	336
Max. number of drives per Fibre Channel loop	84	84	84	84
Max. volume size	17.6 TB	17.6 TB	17.6 TB	17.6 TB

Table 9-5 N5000 gateway specifications

Appliance disk specifications	N5200	N5200	N5500	N5500
IBM machine types - models	2864-G10	2864-G20	2865-G10	2865-G20

Appliance disk specifications	N5200	N5200	N5500	N5500
Max. raw storage capacity	50 TB	50 TB	84 TB	84 TB
Max. number of LUNs on back-end disk storage array	168	168	336	336
Maximum LUN size on back-end disk storage array	500 GB	500 GB	500 GB	500 GB
Max. volume size: 1 TB=1,048,576,000,000 bytes.	16 TB	16 TB	16 TB	16 TB

Table 9-6 N5000 I/O expandability

I/O expandability	N5200-A10 Single storage controller	N5200-A20 Dual storage controller active/active configuration	N5500-A10 Single storage controller	N5500-A20 Dual storage controller active/active configuration
PCI-X expansion slots	3	6	3	6
Optional network connectivity - Max Dual-port Gigabit Ethernet (GbE) adapters (Fibre)	3	6	3	6
Max dual-port 2 Gbps FC disk adapters	3	6	3	6

9.4 N series expansion units

An IBM System Storage N5000 series application system requires at least one storage expansion unit per node, either an EXN1000 or an EXN2000. These expansion units are also available as options on the N3700.

EXN1000 SATA storage expansion units and EXN2000 FC storage expansion units cannot share Fibre Channel loops. A maximum of six storage expansion units (EXN1000 or EXN2000) are supported on a single Fibre Channel loop (although the N3700 only supports a maximum of three expansion units attached).

Within a single EXN1000 or EXN2000 expansion unit, all disk drives must be of a particular type (rotational speed/capacity).

9.4.1 EXN1000 expansion unit

The EXN1000 storage expansion unit provides a 3U, rack-mountable, disk enclosure containing up to 14 SATA disk drives of either 250 GB, 320 GB, or 500 GB. A minimum of five

drives must be installed in the EXN1000 enclosure. Any unused drive bays are covered with drive blank covers (FC 4099).

The EXN1000 is identical to the N3700 chassis except that the slot holding the CPU tray is replaced by an AT-FCX interface module.

9.4.2 EXN2000 expansion unit

The EXN2000 storage expansion unit provides a 3U rack-mountable disk enclosure containing up to 14 FC disk drives. A minimum of four drives must be installed in the EXN2000 enclosure. Any unused drive bays are covered with drive blank covers (FC 4099).

The EXN2000 supports the following FC disk drive speeds and capacities:

- ▶ 15,000 revolutions per minute (15K rpm) in 72 GB and 144 GB capacities
- ▶ 10,000 revolutions per minute (10K rpm) in 72 GB, 144 GB and 300 GB capacities

The EXN2000 is identical to the N3700 chassis except that the slot holding the CPU tray is replaced with an Electronically Switched Hub (ESH2). ESH2 provides a point to point connection to the drives rather than the traditional arbitrated loop. Switched Hub architecture has the benefit of additional availability, boosted performance in high I/O environments, and more powerful diagnostic abilities.

9.5 Data ONTAP software overview

The operating system for the IBM System Storage N series products is the Data ONTAP software. It is a highly optimized, scalable and flexible operating system that can handle heterogeneous environments. It integrates into UNIX, Windows, and Web environments.

Data ONTAP software includes the following standard base system features:

- ▶ Data ONTAP operating system software
- ▶ iSCSI SAN protocol support
- ▶ FlexVol
- ▶ Double parity RAID (RAID-DP)
- ▶ RAID4
- ▶ FTP file access protocol support
- ▶ SnapShot
- ▶ FilerView
- ▶ SecureAdmin
- ▶ Disk Sanitization
- ▶ iSCSI Host Attach Kit for AIX
- ▶ iSCSI Host Attach Kit for Windows
- ▶ iSCSI Host Attach Kit for Linux

The following protocols for the N series are available as extra charge features:

- ▶ CIFS - Provides File System access for Windows environments over an IP network.
- ▶ NFS - Provides File System access for Unix and Linux environments over an IP network.
- ▶ HTTP - allows a user to transfer displayable Web pages and related files.
- ▶ FCP - allows transfer of data between storage and servers in block I/O formats utilizing FCP protocols across a Fibre Channel SAN.

The following software products are available as extra charge features. These are described further in 9.6, "N series software solutions" on page 165:

- ▶ Cluster Failover

- ▶ FlexClone
- ▶ MultiStore
- ▶ SnapLock Compliance
- ▶ SnapLock Enterprise
- ▶ SnapMirror
- ▶ SnapMover
- ▶ SnapRestore
- ▶ SnapVault
- ▶ Open Systems SnapVault
- ▶ LockVault
- ▶ SnapDrive for Windows
- ▶ SnapDrive for Unix: AIX, Solaris, HP-UX, Linux
- ▶ SnapValidator
- ▶ SyncMirror
- ▶ SnapDrive
- ▶ SnapManager for SQL
- ▶ SnapManager for Exchange
- ▶ Single Mailbox Recovery for Exchange
- ▶ Data Fabric Manager

9.6 N series software solutions

Data ONTAP and the N series provides a number of advanced features in software.

9.6.1 Data availability

Traditional single-parity RAID technology offers protection from a single failed disk drive. So long as no other disk fails or uncorrectable bit errors occurs while the original failed disk is being reconstructed, all is well. If either event occurs during reconstruction, then some or all data contained in the RAID array or volume could be lost. With modern larger disk media, the likelihood of an uncorrectable bit error is fairly high, since disk capacities have increased but bit error rates have stayed the same. This makes traditional single parity RAID no longer reliable enough in mission critical environments.

N series gives you the choice of implementing either single parity (RAID4), giving protection against any one disk drive failure, or a double parity RAID-DP, giving protection against any two disk drive failures. The two RAID types are shown in Figure 9-5.

RAID-DP technology is especially designed to harden SATA drive technology.

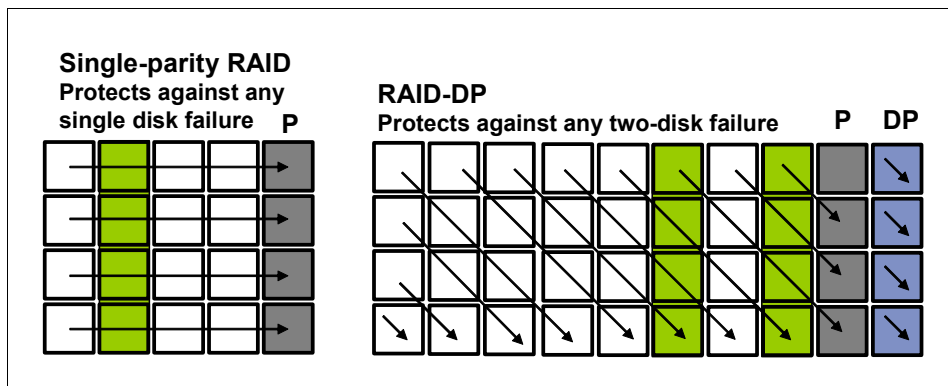


Figure 9-5 RAID4 and RAID-DP

9.6.2 Flexible data management and instant data replication

In today's world the needs of the dynamic enterprise inevitably change, often quickly and even dramatically. Traditional static storage systems, shown in Figure 9-6, require reprovisioning, (data must be manually copied from one volume to another) which causes disruption and increases management overhead and risk. To address this, IT organizations typically overprovision (allocating more storage resources than may actually be needed).

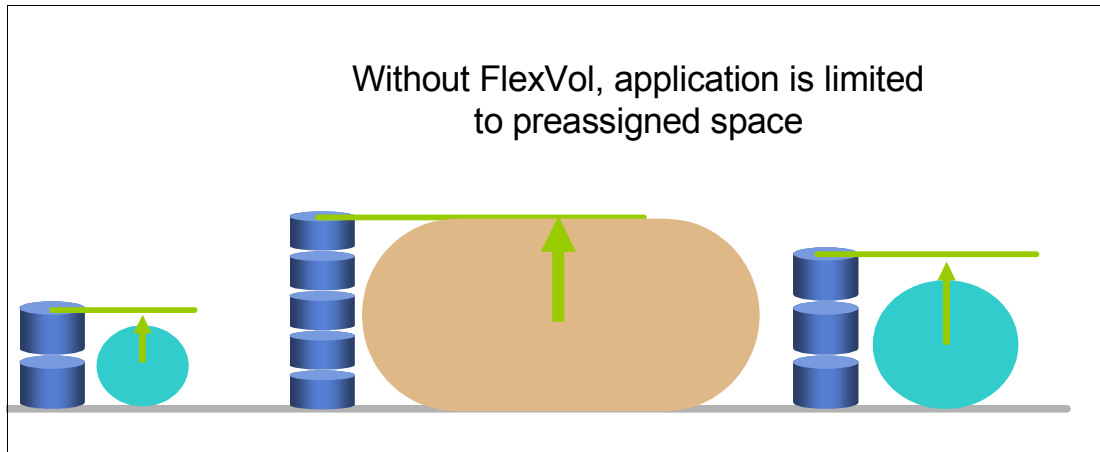


Figure 9-6 Traditional volume provisioning

In comparison, with the N series FlexVol function, a system administrator can create multiple flexible volumes across a large pool of disks. These flexible volumes can be dynamically increased or shrunk by adding/removing capacity, shown in Figure 9-7. FlexVol also enables “thin” provisioning, which makes it possible to oversubscribe free space and adapt rapidly to the changing needs of the enterprise.

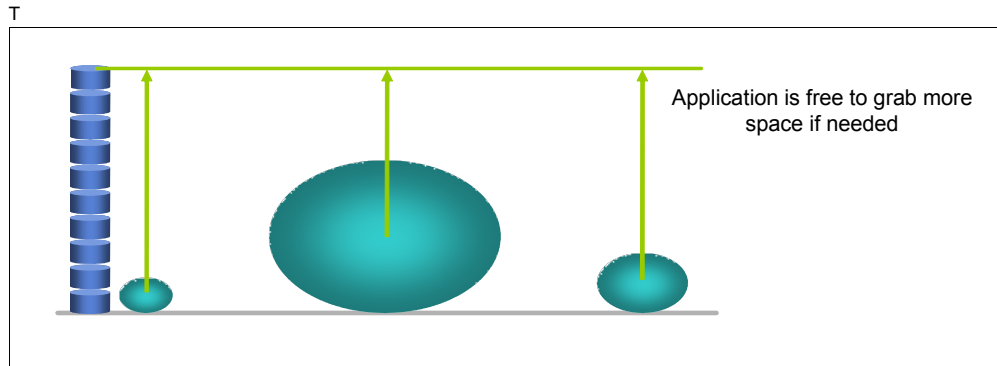


Figure 9-7 N series FlexVol technology

FlexClone technology (Figure 9-8) enables cloning, nearly instantaneous replication of data volumes and data sets using a minimal amount of additional storage space (8 KB) when created. Each cloned volume is a transparent, virtual copy with pointers to the original data.

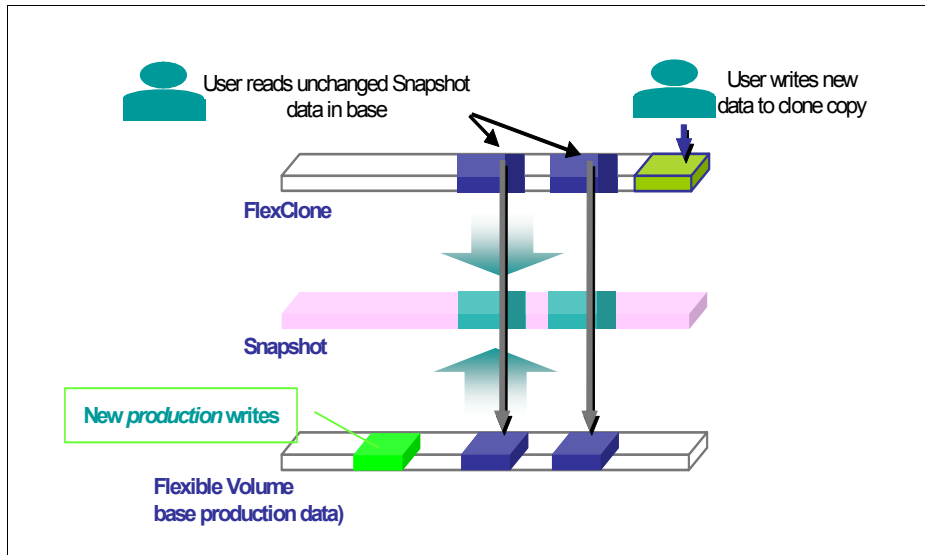


Figure 9-8 N series FlexClone technology

As the data on the original copy changes, those changes are reflected in the FlexClone copy. The FlexClone copy can then be used for essential enterprise operations, such as testing and bug fixing, or platform and upgrade checks with no affect on the original data.

9.6.3 Mixed environment support

Today's IT environments typically must provide a wide-range of network attach capabilities to a broad range of host and client systems using multiple network access protocols.

The N series is a multiprotocol system (Figure 9-9) - it provides NAS protocols (CIFS, NFS), and block I/O protocols including iSCSI and FCP all from a single hardware platform, simultaneously.

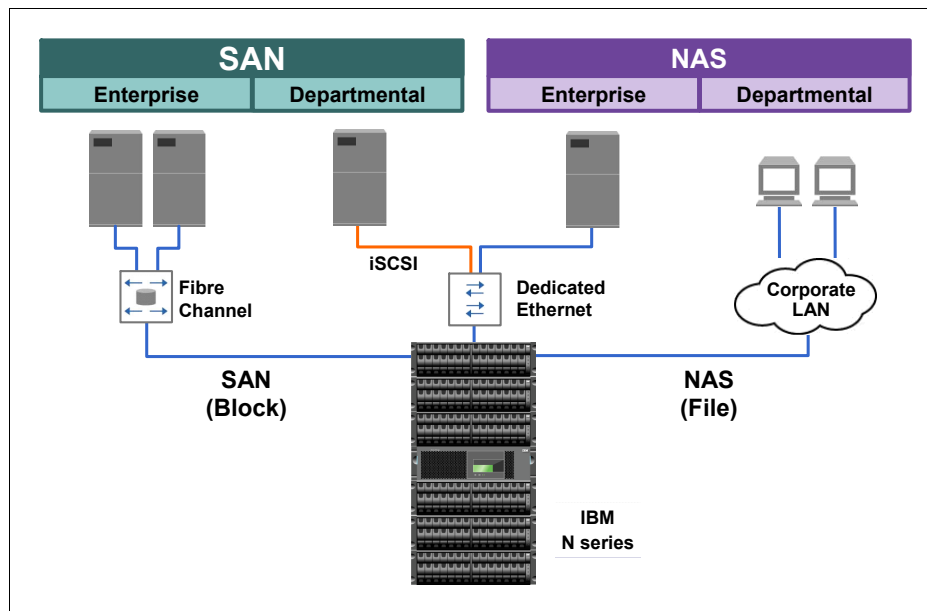


Figure 9-9 N series multiprotocol support

9.6.4 Server and storage consolidation with multistore

The MultiStore functionality of IBM N series is designed lets you quickly and easily create separate, private logical partitions in filer network and storage resources, shown in Figure 9-10. Each virtual storage partition can maintain separation from every other storage partition to prevent different enterprise departments that share the same storage resources from accessing or finding other partitions.

MultiStore helps prevent information on any virtual partition from being viewed, used or downloaded by an unauthorized users.

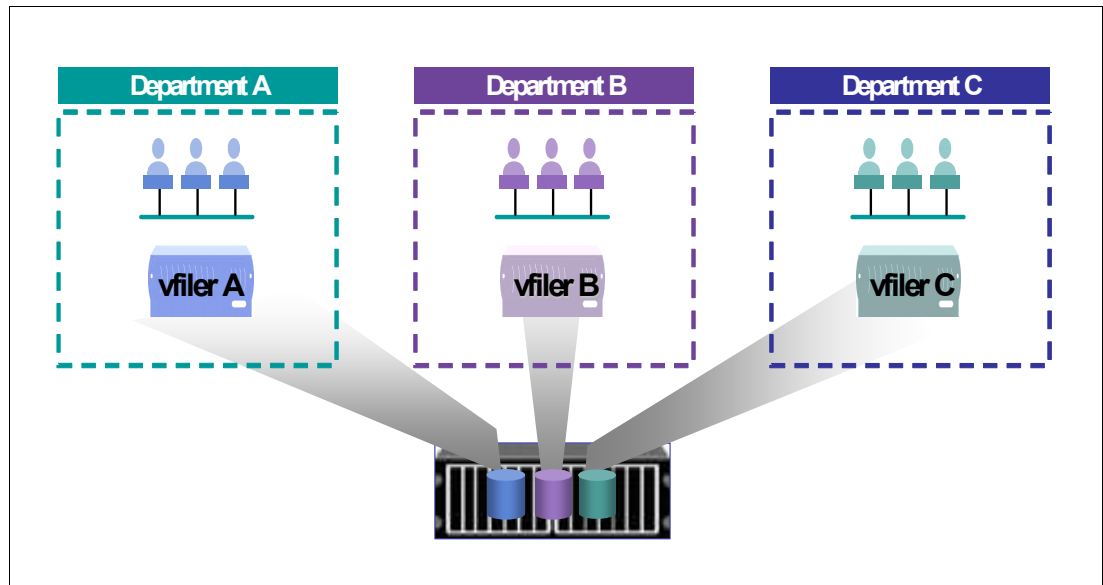


Figure 9-10 N series MultiStore

9.6.5 Tiered storage

With IBM N series, you can assign data to different types of storage (Figure 9-11) in order to reduce the total cost of ownership. The primary data (Tier 1) might be stored on high performance FC disks, the intermediate type of data (Tier 2), e.g. reference data, might be stored on less expensive nearline storage like SATA disks, and rarely used data might be stored on offline storage (Tier 3), e.g. optical or tape.

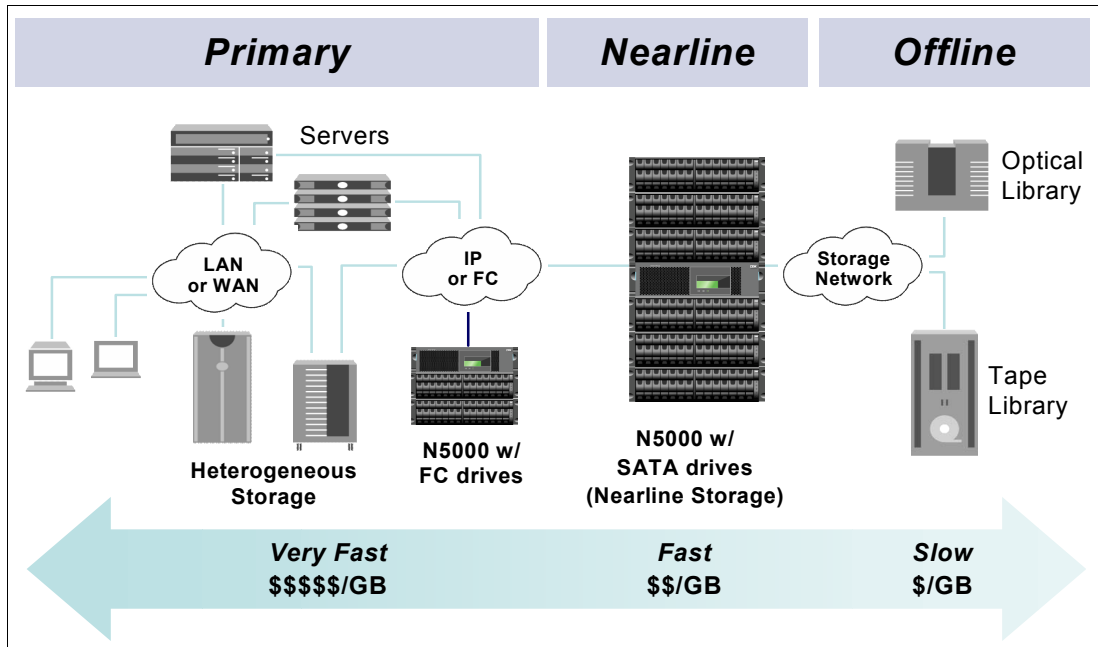


Figure 9-11 Tiered storage

9.6.6 Advanced data protection with SnapShot technology

IBM System Storage N series systems with SnapShot technology help deliver data stability, scalability, recoverability and performance. IBM System Storage N series systems can leverage the SnapShot technology as a foundation for developing a range of data protection solutions. These solutions can incorporate and extend the advantages of the SnapShot technology to support advanced enterprise data protection.

A SnapShot copy is a locally retained point-in-time image of data. SnapShot technology is a feature of the WAFL (Write Anywhere File Layout) storage virtualization technology that is a part of Data ONTAP. A SnapShot is a “frozen,” read-only view of a WAFL volume that provides easy access to old versions of files, directory hierarchies, and/or LUNs.

A SnapShot (Figure 9-12) can take only a few seconds to create, regardless of the size of the volume or the level of activity on the N series. After a SnapShot copy has been created, changes to data objects are reflected in updates to the current version of the objects, as though SnapShot copies did not exist. Meanwhile, the SnapShot version of the data remains completely unchanged. A SnapShot copy incurs little performance overhead; depending on available space, users can store up to 255 SnapShot copies per WAFL volume, all of which are accessible as read-only, online versions of the data.

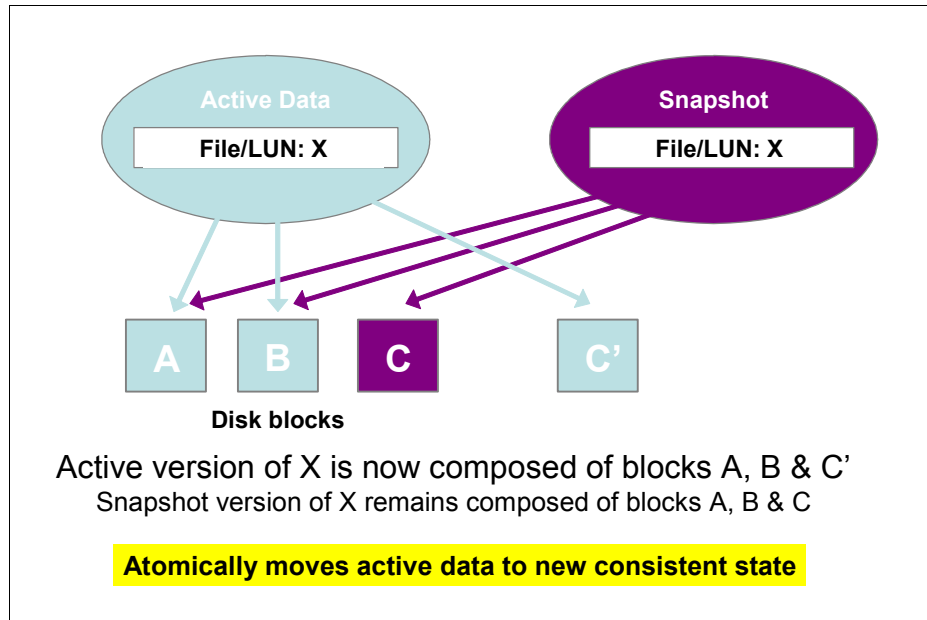


Figure 9-12 N series SnapShot - How it works

A SnapShot copy can be used to provide frequent, low-impact, user-recoverable backups of files, directory hierarchies, LUNs, and/or application data. A SnapShot copy can significantly improve the frequency and reliability of backups, since it is designed to avoid performance overhead and can be created on a running system.

A SnapShot supports near-instantaneous, user-managed restores. Users can directly access SnapShot copies to recover from accidental deletions, corruptions, or modifications of their data.

SnapRestore

IBM System Storage N series systems with SnapRestore capability helps recover data quickly when disaster strikes. SnapRestore technology can quickly recover large individual files or volumes through instant volume recovery. Volumes can be restored with a single command versus the file level restores that SnapShot offers.

SnapShot technology uses storage efficiently, as it stores only block-level changes between each successive SnapShot copy. Since the SnapShot process is automatic and incremental, backups are significantly faster and simpler. SnapRestore technology uses SnapShot copies to perform near-instantaneous data restoration. In contrast, non-point-in-time storage solutions may copy all of the data and require much more time and disk storage for the backup and restore operations.

9.6.7 Local and remote mirroring solutions

This section discusses mirroring capabilities of N series.

Synchronous local mirroring

The SyncMirror functionality, shown in Figure 9-13, of IBM N series provides synchronous local mirroring from one volume to another volume attached to the same filer. It maintains a strict physical separation between the two copies of your mirrored data. In case of an error in one copy, the data is still accessible without any manual intervention.

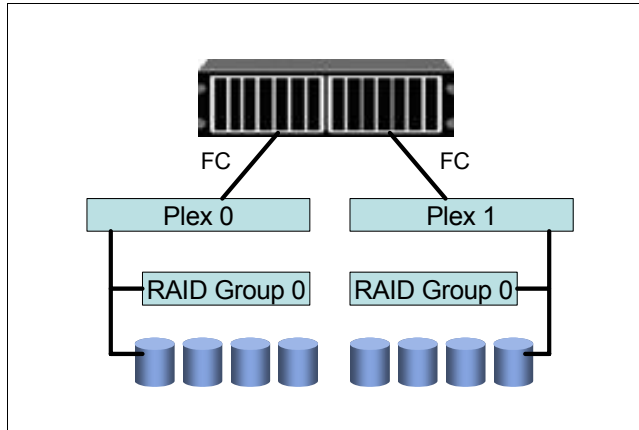


Figure 9-13 N series SyncMirror

With SyncMirror, filers can tolerate multiple simultaneous disk failures across the RAID groups within the WAFL file system. This redundancy goes beyond typical mirrored (RAID-1) implementations. Because each SyncMirror RAID group is also RAID-4 or RAID-DP protected, a complete mirror could be lost and an additional single drive loss within each RAID group could occur without data loss.

Remote mirroring

The SnapMirror function provides data set replication between network-connected N series systems for backup or disaster recovery purposes. After an initial baseline transfer of the entire data set, subsequent updates only transfer new and changed data blocks from the source to the destination, so that SnapMirror uses the network bandwidth efficiently. The destination file system is available for read-only access, or the mirror can be “broken” to enable writes to occur on the destination. After breaking the mirror, it can be reestablished by synchronizing the changes made to the destination back onto the source file system.

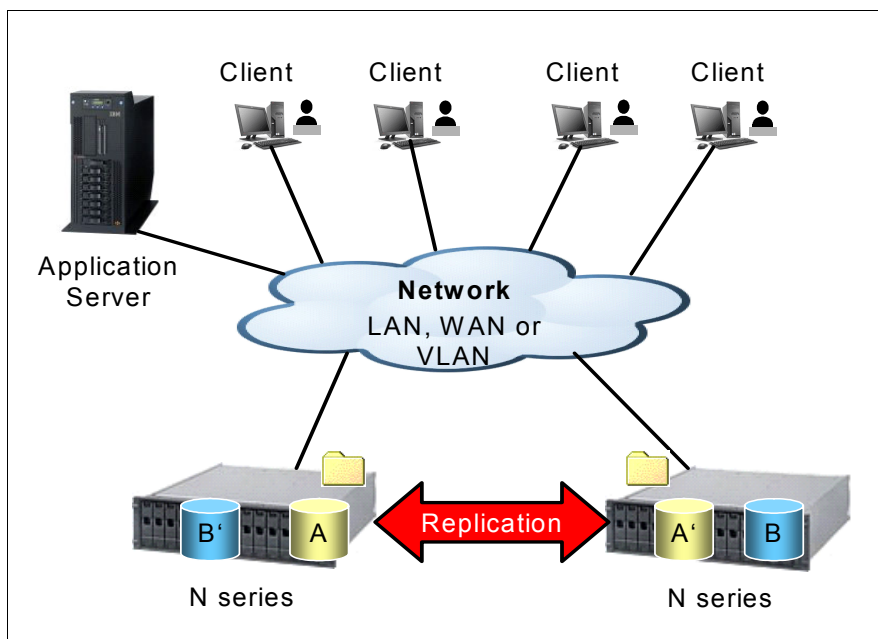


Figure 9-14 N series SnapMirror

SnapMirror can be used in three different modes: asynchronous, synchronous, and semi-synchronous.

In asynchronous mode, updates of new and changed data from the source to the destination occur on a schedule defined by the storage administrator. The update frequency can range from once per minute to once per week, depending on user needs.

Synchronous mode sends updates from the source to the destination as they occur, rather than on a schedule. This ensures that data written on the source system is protected on the destination even if the entire source system fails due to natural or human-caused disaster.

The semi-synchronous is a middle ground which keeps the source and destination file systems more closely synchronized than asynchronous mode, but with less performance impact. It is configured identically to synchronous mode, except that an option specifies how many writes can be unacknowledged by the destination before the source system delays acknowledging client writes.

N series SnapMirror can be use one to many (multiple copies), or many to one (consolidation) configurations.

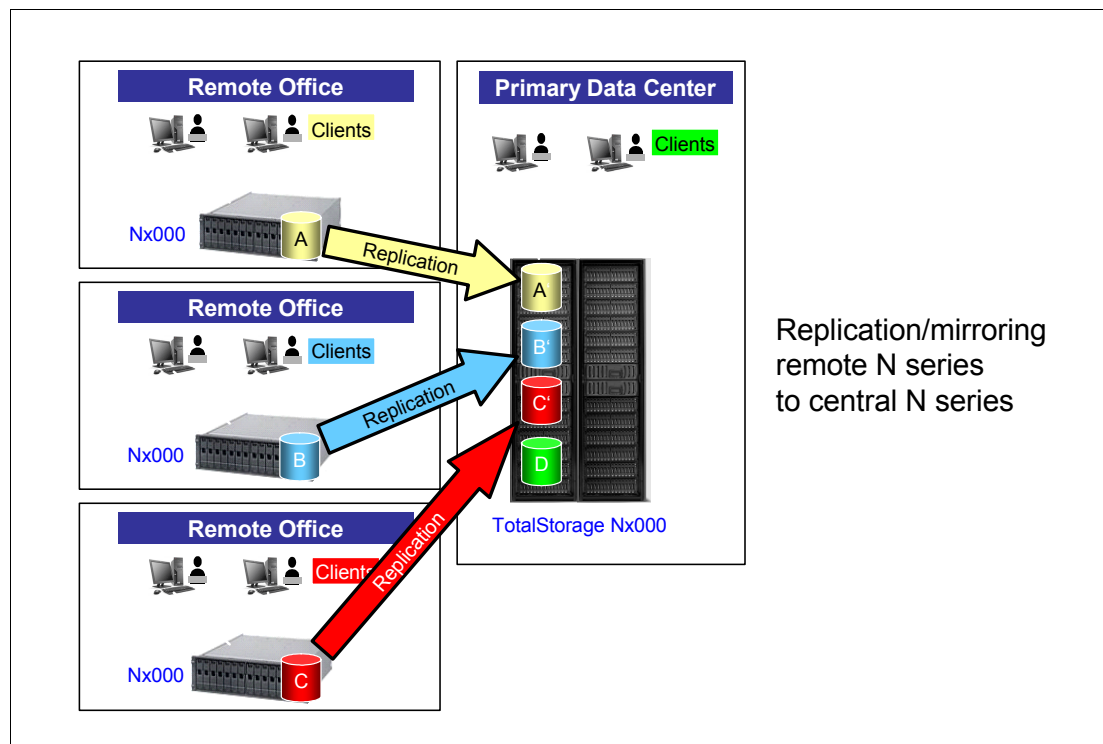


Figure 9-15 SnapMirror consolidation

Difference between SyncMirror and SnapMirror

The major difference between SyncMirror and Synchronous SnapMirror is in who owns the second copy of the data.

With SyncMirror, one host owns both copies. It simply writes to both copies from NVRAM. It also provides for instant failover if one copy fails for any reason.

With Synchronous SnapMirror, the destination filer owns the second copy. It moves blocks similar to how SnapMirror does, where blocks move over IP to the other filer rather than direct

FC writes to local disks. There is no automatic failover since the source host does not own the second copy of the volume and SnapMirror does not have this functionality.

9.6.8 Improving data availability with clustered failover

The System Storage N series with clustered failover is designed to provide a robust and highly available data service for business-critical environments. Installed on a pair of N series filers, clustered failover can help improve data availability by transferring the data service of an unavailable filer to the other filer in the cluster.

Clustered failover constantly monitors the health of clustered filers. If a system detects a catastrophic hardware failure on a filer, it can automatically initiate a failover operation to transfer the data service to its partner filer. The entire failover process is automatic, with no manual intervention typically required at any point.

This functionality assumes that all resources are accessible by each node - i.e. both nodes must have access to all disks physically (cabling) and logically (cluster software), shown in Figure 9-16.

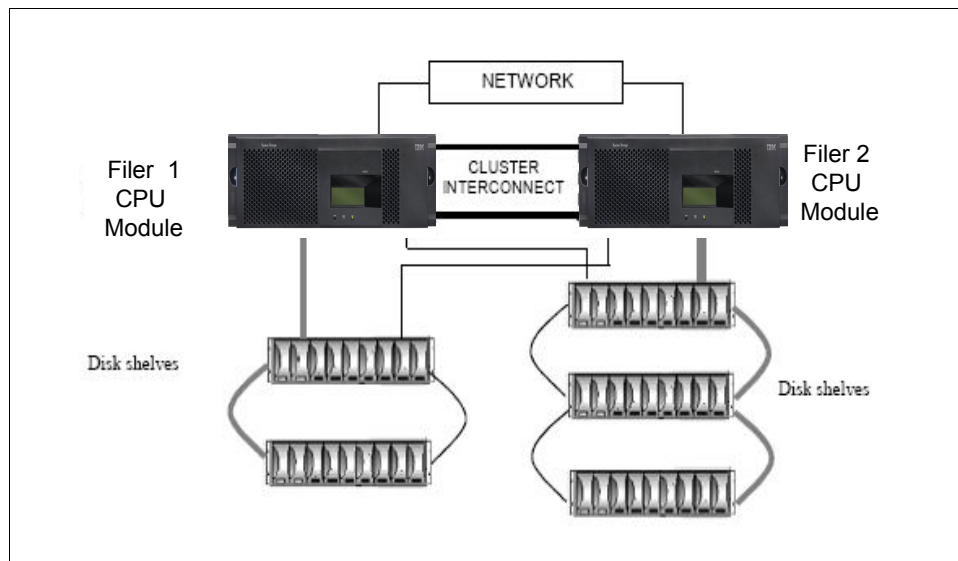


Figure 9-16 N series cluster failover

To achieve a higher level of protection, cluster failover can also be combined with SyncMirror, as shown in Figure 9-17. The physical separation of the cluster nodes and copies protects against a cluster node failure, and data loss in the case of a double-disk error or loss of disk connectivity.

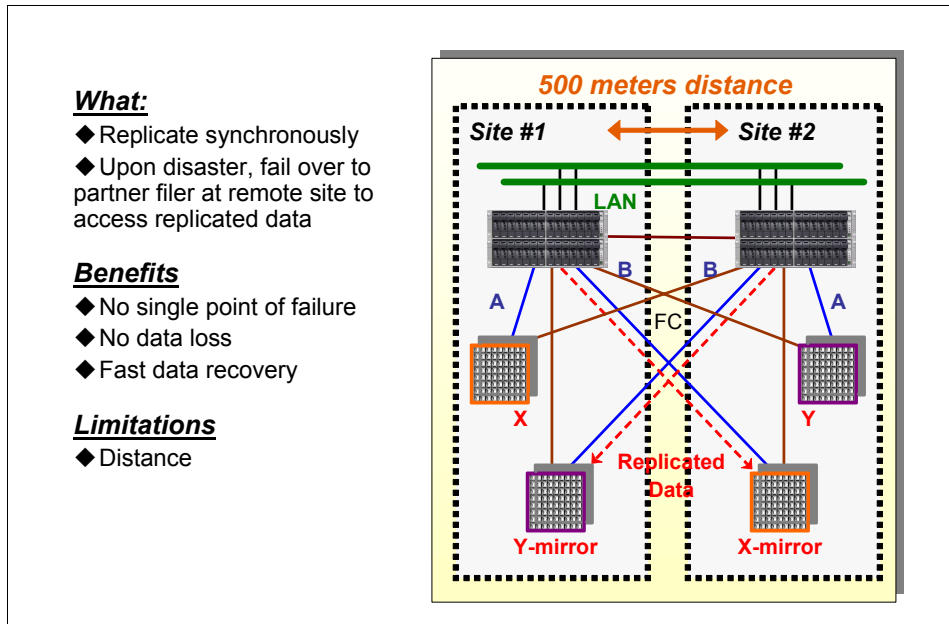


Figure 9-17 Clustered SyncMirror

9.6.9 Optimizing workloads with SnapMover

As companies grow, their storage requirements and policies change in response to new business initiatives and challenges. Data may need to migrate to scale the infrastructure to accommodate growing data sets and new clients, improve performance and response time for users, reduce the processing burden on an overloaded storage system, or meet a service-level agreement with users performing mission-critical work.

The challenge is to migrate workload to provide for improved performance and resource utilization, and to do it quickly, effectively, and economically. Most existing workload migration solutions, however, require IT to copy data from one location to another, which is tedious, time-consuming, and error-prone. In addition, users cannot access data during the migration process, which causes user dissatisfaction and impacts productivity.

SnapMover, shown in Figure 9-18, is designed as a fast and simple solution for migrating data among IBM System Storage N series systems while avoiding impact on data availability and user disruption. It supports the migration of data from one N series device to another without the need to copy and store it on other devices.

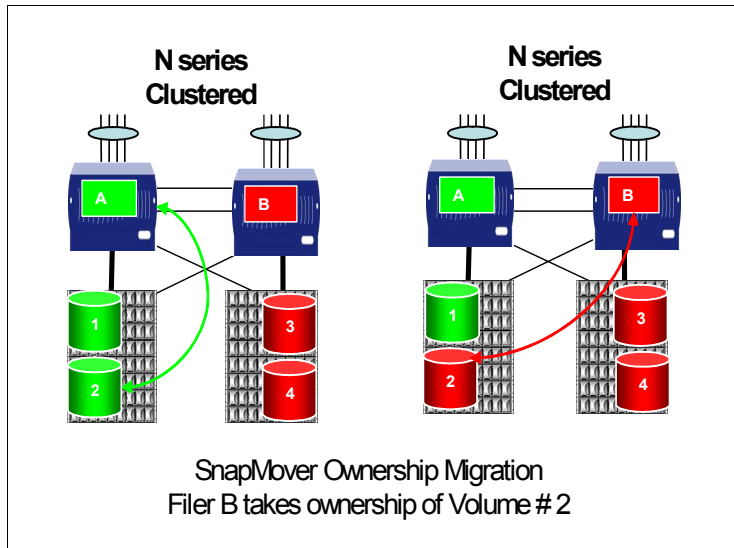


Figure 9-18 N series SnapMover

Note: MultiStore and Clustering are prerequisites for SnapMover.

9.6.10 Automated data backup with SnapVault

Today, while the volume of data generated in the average organization is increasing, backup windows are shrinking. Backup and restore operations can disrupt access to important information, leading to lost productivity.

IBM System Storage N series with SnapVault can help efficiently backup and restore critical data. By supporting the frequent backup of N series data, SnapVault provides a centralized, disk-based backup solution for multiple N series filers. Storing backup data in multiple SnapShot copies on the SnapVault secondary system allows enterprises to keep multiple backups made over time online for faster restoration.

By backing up multiple SnapVault primary storage systems to a single SnapShot copy on a large SnapVault secondary storage system, backup and restore operations can be centralized in a single location, as shown in Figure 9-19. SnapVault primary storage systems can be LAN-connected N series or located in remote offices connected over a WAN. The SnapVault secondary storage system can also be located at a remote disaster recovery site to help protect against production site failures. Tape backup operations can be performed from the SnapVault secondary storage system, consolidating and centralizing all data protection resources.

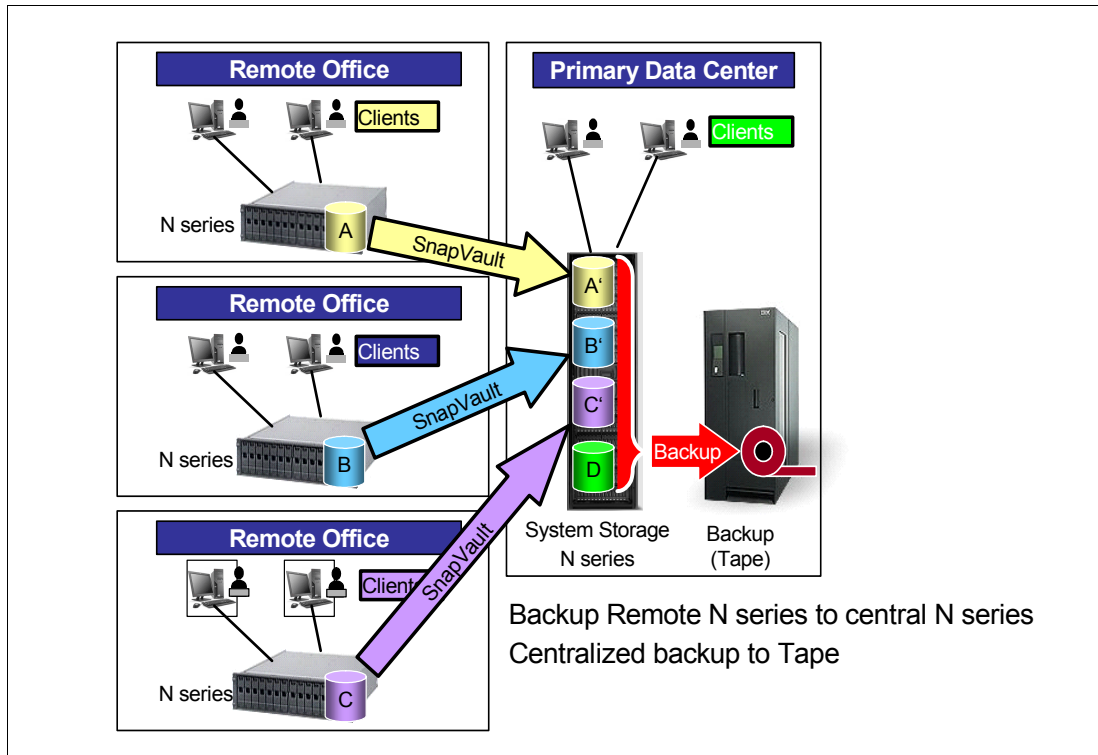


Figure 9-19 N series SnapVault

In contrast to SnapMirror, which is a disaster recovery or replication solution, SnapVault provides an hourly disk-based online backup and restore by periodically backing up snapshot copies to another system.

9.6.11 Disk-based backups for Open Systems

Open Systems SnapVault (OSSV) is a replication based backup for open systems hosts which provides online copies of data sets by storing them in snapshots on remote N series systems. The OSSV software serves as an agent on the open systems host and is available for the following server platforms:

- ▶ AIX
- ▶ HP-UX
- ▶ IRIX
- ▶ Red Hat Linux
- ▶ Windows
- ▶ Solaris

For the first backup, the N series requests an initial baseline transfer of the identified file systems from the open systems host. Subsequent transfers can either be whole files or only the changed blocks, depending on the OSSV configuration.

For each set of data transfers, OSSV creates a set of incremental snapshots that capture the changes to the N series and assigns each snapshot in the set a version number (0 for most current, 1 for second most recent, and so on).

If directory or file data needs to be restored to the open systems host, OSSV retrieves the data from one of the retained snapshots and transfers the data back to the open systems host that requests it.

Open Systems SnapVault can be implemented in combination with SnapVault. An example of a mixed OSSV/SnapVault environment is shown in Figure 9-20.

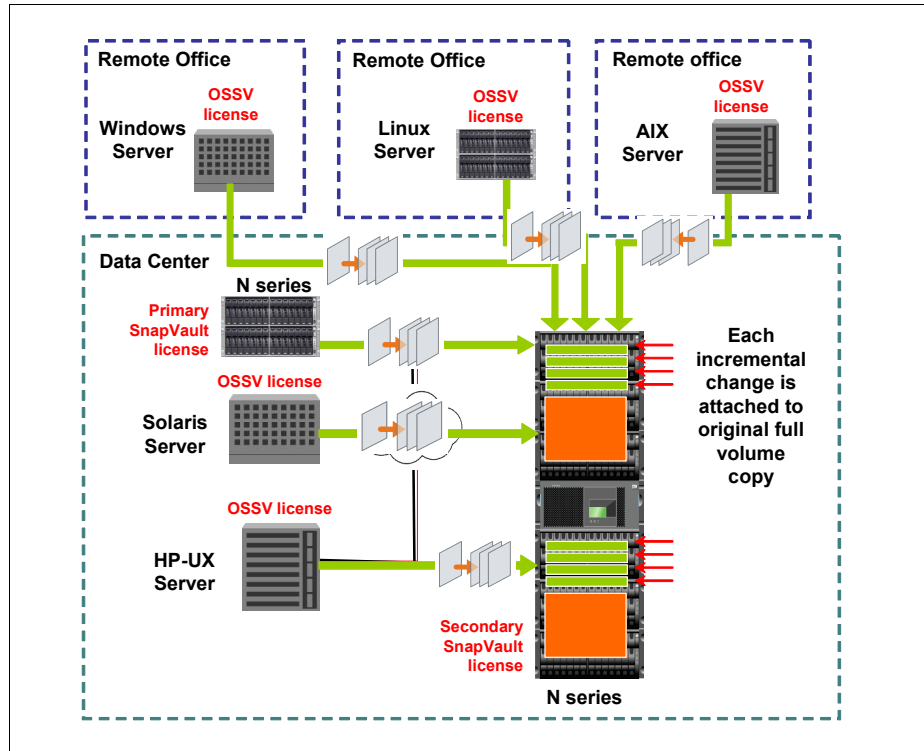


Figure 9-20 N series mixed SnapVault and Open Systems SnapVault environment

9.6.12 Backing up N series via NDMP and Tivoli Storage Manager

Tivoli Storage Manager Extended Edition includes support for the use of NDMP to back up and recover NAS file servers, like IBM System Storage N series.

Feature highlights include:

- ▶ Tivoli Storage Manager supports full image backup of NAS file systems images via NDMP
- ▶ Full backup can be followed by subsequent differential backups
- ▶ Tape resource sharing
- ▶ Individual file restore is possible
- ▶ Table of Contents (TOC) available for viewing/browsing
- ▶ SnapShots integrate with virtual file mapping feature
- ▶ Restore can be via Web Client or Tivoli Storage Manager command line interface

Within an NDMP backup the N series writes the backup data directly to tape and sends its metadata via TCP/IP to the Tivoli Storage Manager server. This metadata is called the Table of Contents (TOC). Unlike a LAN-free backup, the metadata is not stored within the Tivoli Storage Manager database but in a designated special Tivoli Storage Manager disk pool, that is only required for NDMP file level restores. The backup and restore is initiated and controlled by the Tivoli Storage Manager server.

The tape library can be controlled in two different ways:

- ▶ Tivoli Storage Manager server controls the SCSI library through a direct, physical connection to the library robotics control port
- ▶ The library robotics and the drives are physically connected directly to the N series system, and paths must be defined from the N series data mover to the library and drives. The Tivoli Storage Manager server controls library robotics by sending library commands across the network to the NAS file server. The NAS file server passes the commands to the tape library. Any responses generated by the library are sent to the NAS file server, and passed back across the network to the Tivoli Storage Manager server.

In Figure 9-21 shows the control and data flow for the second possibility.

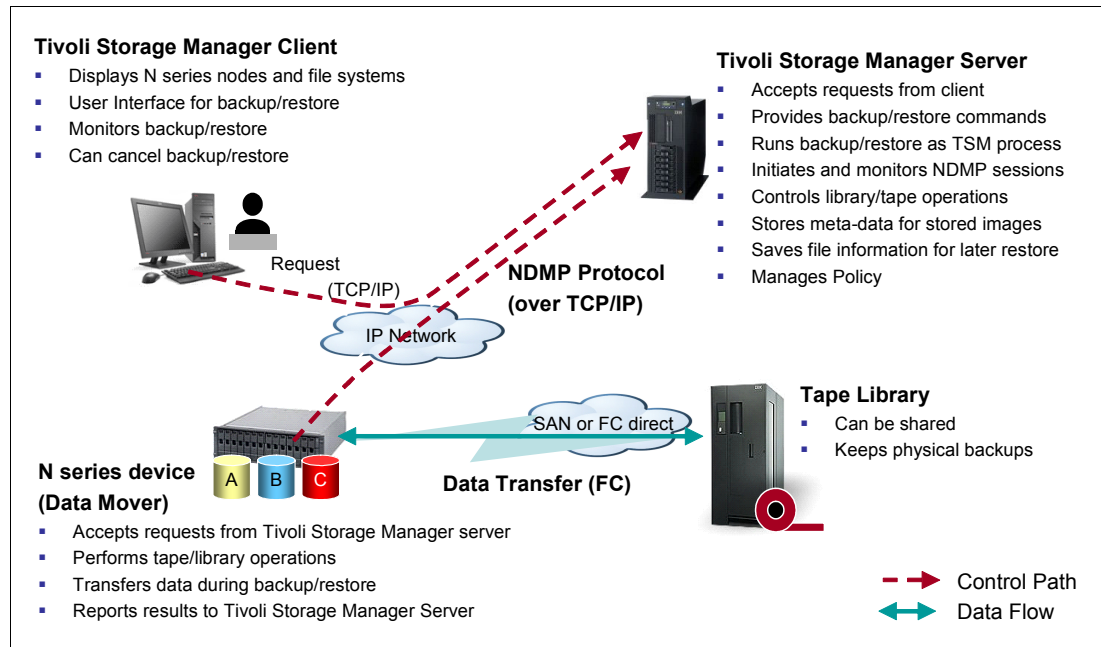


Figure 9-21 NDMP backup with Tivoli Storage Manager Extended Edition

For more information about Tivoli Storage Manager, see Chapter 24, “IBM Tivoli Storage Manager” on page 437.

9.6.13 Storage virtualization and simplified management with SnapDrive

IBM System Storage N series with SnapDrive software can help reduce the complexity of managing storage. SnapDrive software offers capabilities designed to virtualize and enhance storage management for Windows, Linux, AIX, HP-UX and Solaris environments.

It is tightly integrated with the native file system and provides a layer of abstraction between application data and physical storage associated with that data. Storage managed by SnapDrive software logically appears to come from a locally attached storage subsystem. In reality, the capacity comes from a centrally managed pool of networked storage equipped with enhanced attributes.

SnapDrive is designed to allow administrators to easily create virtual disks from pools of storage that can be distributed among several storage appliances. With SnapDrive you can add, delete, map, unmap, and mirror virtual disks online. SnapDrive provides the capability to expand capacity on-the-fly while avoiding negative impact to application or system performance.

SnapDrive integrates SnapShot technology (Figure 9-22) to offer near-instantaneous point-in-time images of application and user data. It enables access to SnapShot copies by mounting them as virtual disks. You can use these virtual disks for routine administrative tasks such as online backup, testing of new applications, or population of data marts while avoiding downtime to business-critical information. Combined with SnapRestore software, SnapDrive is designed to enable restoration of data in seconds.

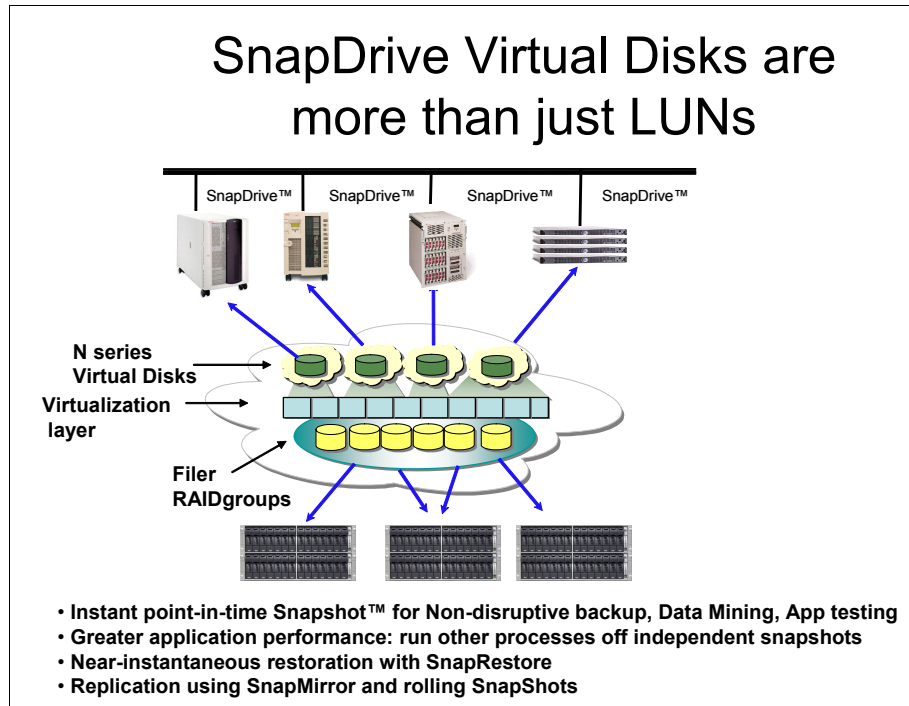


Figure 9-22 N series SnapDrive

SnapDrive, when configured with SnapMirror software, also supports online replication. The entire storage environment, including virtual disks and SnapShot copies can be mirrored to one or more N series.

SnapDrive can be seamlessly integrated into Microsoft Windows Cluster environments. It helps simplify management of cluster resources, virtual disks, and SnapShot copies. SnapDrive also supports the use of a virtual disk for a quorum device.

In the Windows environment, SnapDrive allows administration via the Microsoft Management Console (MMC) or a command line, making management simple and intuitive. Interactive wizards and easy-to-use interfaces guide administrators through management tasks and create automatic schedules of operations. In UNIX environments, SnapDrive comes with an intuitive command line for script-based automation.

9.6.14 Rapid backups and restores with SnapManager

N series SnapManager support rapid backup and restore of application environments.

SnapManager for Microsoft Exchange

SnapManager software can provide near-instantaneous hot backups and rapid restores for Exchange environments. It can schedule and automate Exchange database backups, use policy-based backup retention management, and simplify the migration of existing databases to N series filers. SnapManager software also offers high availability, with features that allow

expansion Exchange databases online. It supports tight integration with Microsoft Cluster Server (MSCS) and Multi Path I/O (MPIO). It also integrates with the N series Clustered Failover option and SnapMirror software to help simplify disaster recovery implementation.

SnapManager utilizes the Microsoft VSS Infrastructure, and in the Windows 2000 environment, it integrates with the Exchange Backup APIs and ESEutil to provide consistent online backups.

Depending on the restore requirements, there is a wide range of restore options: available: full Exchange Server content recovery, individual Exchange storage group recovery, individual Exchange database recovery, and virtual disk recovery. Even individual mailbox recovery may be performed with the Single Mailbox Recovery software.

Note: SnapManager for Microsoft Exchange requires:

- ▶ SnapDrive. software
- ▶ SnapRestore software
- ▶ FCP or iSCSI protocol
- ▶ Microsoft Windows 2000 or 2003 Server
- ▶ Microsoft Exchange Server 2000 or 2003

SnapManager for Microsoft SQL Server

SnapManager supports rapid SQL Server backup times, from hours to as little as seconds, and makes each backup a complete and consistent copy of the original. Backups are based on SnapShot copies, which require minimal disk space for each additional full backup. It allows back up or restore of several databases simultaneously, as well as volume expansion.

SnapManager software can integrate with the SQL Server application to automate the process of validating the consistency of data backup and checking that the data is available for restore.

The integration with SnapMirror technology supports performing remote replication of SQL Server data, thus helping to speed data recovery in the event of a disaster.

Note: SnapManager for SQL Server requires:

- ▶ SnapDrive software
- ▶ SnapRestore software
- ▶ Windows 2000 Server/Advanced Server SP2, SP3, SP4
- ▶ FCP or iSCSI protocol
- ▶ Windows 2003 Server/Advanced Server
- ▶ Microsoft SQL Server 2000
- ▶ Enterprise Server SP2 and SP3

9.6.15 Single mailbox recovery for Exchange

One of the most time-intensive (and frequent) tasks for Microsoft Exchange administrators is recovering single mailboxes or single messages. To recover single mail items from Exchange quickly, administrators must perform complex, time-consuming backups that involve backing up each mailbox separately (referred to as bricklevel backups). The alternative is a painful process of setting up a recovery server, loading the last full backup from tape, and then recovering a single mailbox.

The combination of IBM System Storage N series, SnapManager for Exchange, and Single Mailbox Recovery functionality supports the fast, accurate backup and recovery of Microsoft Exchange data. By directly reading the contents of SnapShot copies without the assistance of

the Exchange server, N series storage with Single Mailbox Recovery functionality can restore individual mail items from a recent (hourly, daily, weekly) SnapShot. It can restore individual mailboxes, folders, messages, attachments, calendar notes, contacts, and task items directly to the production Exchange server or to a new or existing offline Outlook® Personal Store (PST) file.

9.6.16 End-to-end data validation for Oracle databases

IBM System Storage N series with SnapValidator software (Figure 9-23) provides a high level of protection for Oracle data, helping detect potential data corruption before it occurs. By adding intelligence and database awareness to modular storage systems—across iSCSI SAN, FC SAN and NAS protocols—the software can help extend the advantages of checksum functionality to a greater variety of organizations.

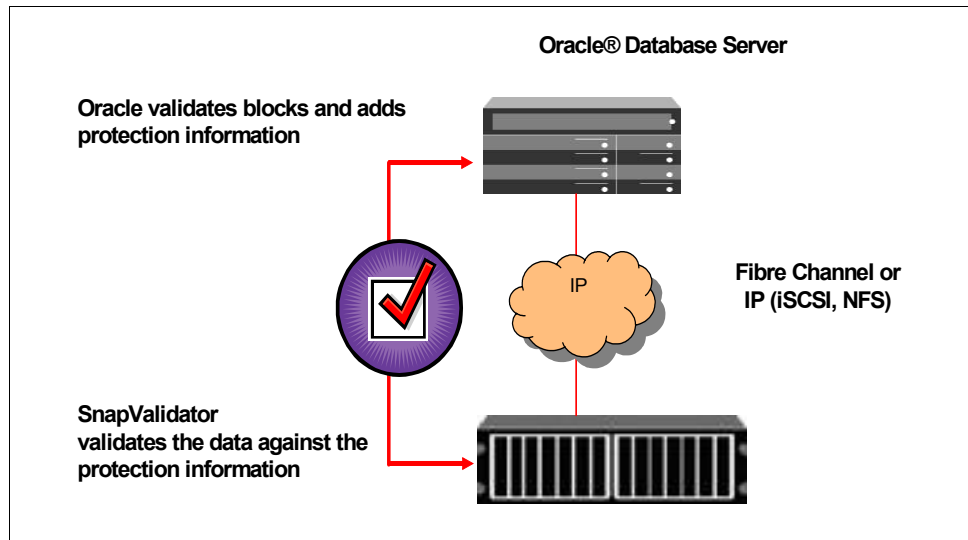


Figure 9-23 N series SnapValidator

SnapValidator is tightly integrated with the Oracle database architecture. It complies with the Oracle HARD initiative, which is intended to help companies achieve higher levels of data integrity. The HARD initiative incorporates several data validation technologies that can be embedded in storage devices to help prevent data corruption.

When initiating data transfers, the Oracle system assigns a unique value to each write request based on the sum of the data. It then sends that data set down the I/O path to the N series storage system. Upon receipt, SnapValidator calculates the sum of the data and compares it to the Oracle value. If these values match, the system writes the information to disk and reports a successful write. If the identifiers do not match, the system does not transfer the data, but instead produces an alert that offers options for resolving the problem immediately to prevent further problems.

SnapValidator helps prevent unplanned downtime resulting from data corruption.

9.6.17 Data permanence solutions

N series provides solutions for data permanence.

SnapLock Compliance and SnapLock Enterprise Software

Organizations today face increasingly strict regulations relative to records retention that require the archiving of e-mails, documents, patient records, design files, audit information, and other data for years, often on WORM (write once, read many) media. In today's world, government and internal records requests frequently combine highly specific search criteria with tight response times. Current WORM technologies based on optical disk and tape do not provide sufficiently rapid access.

The SnapLock function delivers high performance and high-security data function to diskbased nearline and primary N series storage, providing disk-based WORM capability. SnapLock is available in two versions, SnapLock Compliance and SnapLock Enterprise.

SnapLock Compliance helps organizations address records retention regulations such as SEC Rule 17a-4 (broker-dealers), HIPAA (healthcare), Sarbanes-Oxley (public companies), 21CFR Part 11 (life sciences), and DOD 5015.2 (government).

SnapLock Enterprise supports adherence to rigorous organizational best practices through functionality similar to that of SnapLock Compliance, but allows administrators to delete entire SnapLock Enterprise volumes.

A comparison between SnapLock Compliance and Enterprise is shown in Table 9-7.

Table 9-7 SnapLock Compliance and Enterprise compared

SnapLock Compliance	SnapLock Enterprise
"Strict" SnapLock ▶ Trust nobody	Flexible" SnapLock ▶ Trust administrator
Permanently nonerasable, nonrewritable disk storage (WORM) ▶ Until file expiration ▶ Safe from any keyboard attack	Revision-safe, long-term storage solution ▶ Virus and application bugproof ▶ Enables best practices business records retention
Complies w/ SEC regulations ▶ Meets SEC 17a-4 requirements ▶ Easy WORM-to-WORM replication	Partial storage admin control ▶ Admin can destroy volume to reclaim space ▶ Cannot modify/delete individual records

SnapLock also allows WORM data to be replicated securely and automatically between multiple N series filers using SnapMirror software, or to be backed up to tape for an additional level of data protection. It supports open, industry standard protocols such as NFS and CIFS, so it offers easy data access and application integration. Figure 9-24 shows how regulated and reference data can be easily archived with a partner archival application onto a SnapLock WORM volume.

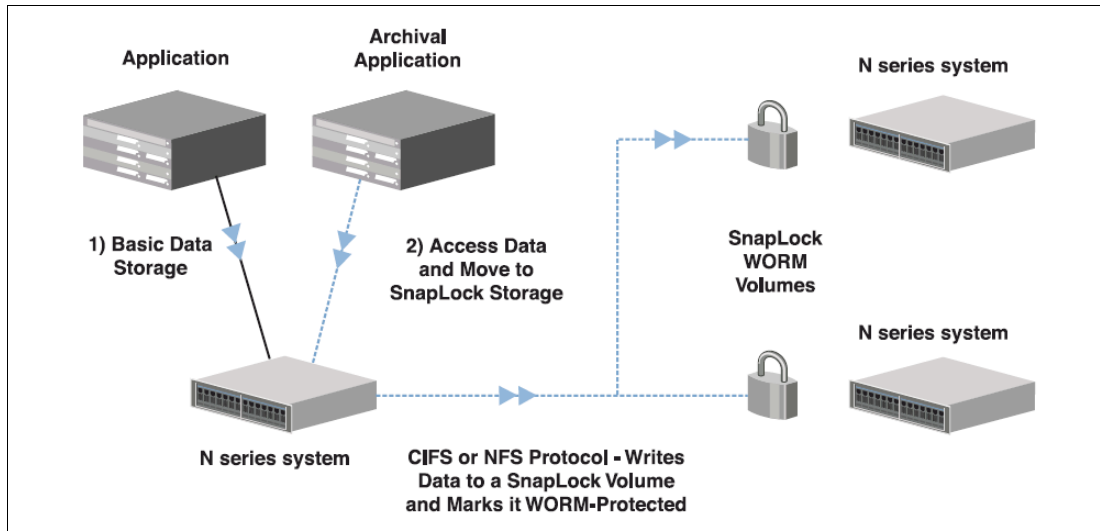


Figure 9-24 System Storage N series with SnapLock

LockVault compliance software

LockVault software integrates SnapLock and SnapVault technologies to create a solution specifically designed to help businesses address regulatory compliance requirements for unstructured data. WORM backups of unstructured data can be made by simply copying and storing only the unique blocks that have been written since the most recent incremental backup. Every block variation is safeguarded to help avoid data loss. LockVault can automatically provide online backups of unstructured data at multiple points in time. Each backup is immutable until its specified expiration date.

In addition, SnapMirror is designed to fully support WORM-to-WORM remote replication to provide the required duplicate copy of backup data. An N series LockVault solution can make regular WORM copies of file servers as well as mirrored copies of data as required.

A comparison between SnapLock and LockVault is shown in Table 9-8.

Table 9-8 SnapLock and LockVault compared

	SnapLock	LockVault
Best fit for	Structured and semistructured data	Unstructured data
Typical Mode of operation	Driven by archival application	Self-contained solution
Commit	Explicit commit required	Automatic commit and date assignment
Retention dates	Assigned to files	Assigned to Snapshot copies
Logging	Audit Logs	ComplianceJournal, Audit Logs
Version handling	Each version is a different file	Full original, then only changed blocks

9.6.18 N series management

The IBM System Storage N series systems can be managed by using the FilerView Web interface, the Command line interface (CLI), or the DataFabric Manager.

FilerView

FilerView is a Web-based administration tool, shown in Figure 9-25, that allows full remote N series administration and management. FilerView is a simple and intuitive Web-based single-appliance administration interface.

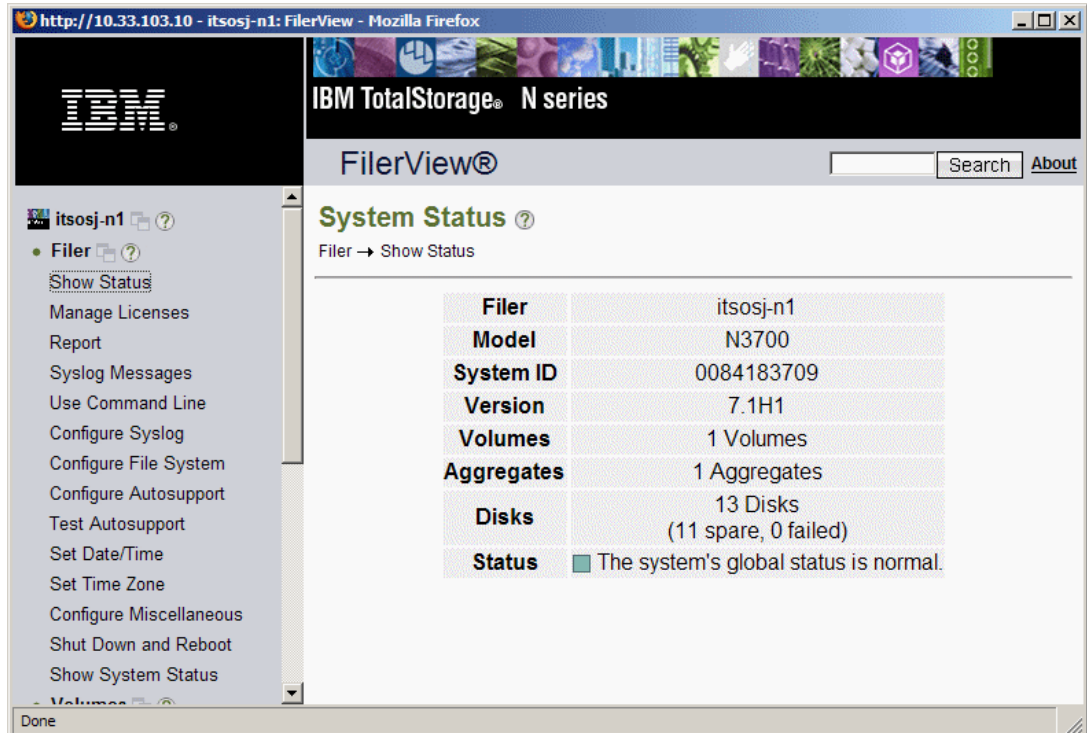


Figure 9-25 N series FilerView

Command line interface (CLI)

The CLI provides a fast and easy to use interface to N series in several ways:

- ▶ telnet client
- ▶ Secure shell interface (ssh) or remote shell (rsh)
- ▶ Serial console
- ▶ FilerView command line feature

The command line interface is accessed by clicking **Filer** in the FilerView navigation field (left part of FilerView window) then choosing **Use Command Line**.

DataFabric Manager

DataFabric Manager Software is an optional software feature which provides monitoring and management features for N series. DataFabric Manager allows an organization to rapidly deploy, provision, and manage a complete enterprise storage and content delivery network. It delivers a central point of control for alerts, reports and the configuration tool. DataFabric Manager capabilities are:

- ▶ Discovery
 - Filers
 - Aggregates, volumes, qtrees, and LUNs
- ▶ Monitoring
 - Status and health

- Alerts via e-mail and pager
- ▶ Performance and capacity analysis
 - Storage and content caching devices
 - Aggregates, volumes, qtrees, LUNs, disks, CPUs, and network links
- ▶ Configuration
 - FilerView launch
 - SnapMirror and SnapVault

Figure 9-26 shows a sample DataFabric Manager window.

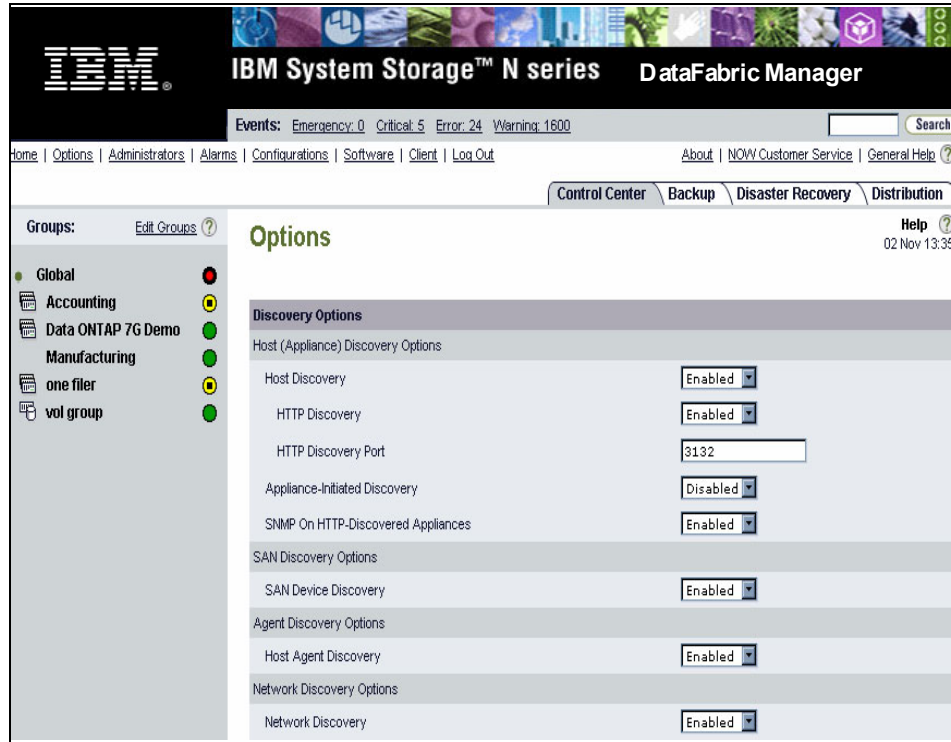


Figure 9-26 N series Data Fabric Manager

Note: DataFabric Manager runs on an external server and needs an additional license.

Encryption and authentication for administrative sessions

SecureAdmin is a Data ONTAP module which enables authenticated, command-based administrative sessions between an administrative user and Data ONTAP over an intranet or the Internet.

For environments with many administrators, internal security concerns, or the need to implement specific security policies, SecureAdmin can authenticate both the administrative user and the filer, creating a secured, direct communication link to the filer. It helps protect administrative logins, passwords, and session commands from “cleartext” snooping by replacing rsh and telnet with the strongly encrypted SSH 1.0 protocol.

9.6.19 Obliterating data with disk sanitization

Disk sanitization is the process of physically obliterating data by overwriting disks with specified byte patterns or random data to help prevent recovery of the deleted data by any known recovery methods, shown in Figure 9-27.

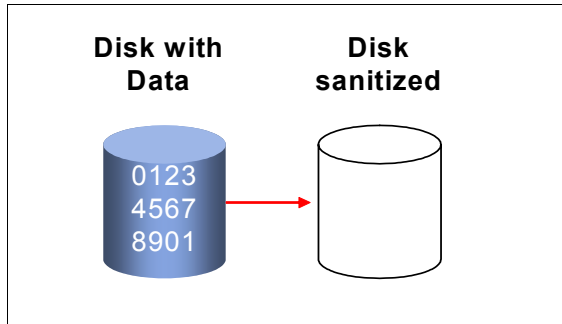


Figure 9-27 N series disk sanitization

This feature provides disk sanitization using three successive byte overwrite patterns per cycle and a default six cycles per operation.

9.7 More information

For more information about the IBM System Storage N series, see the following documents and Web sites:

- ▶ <http://www-03.ibm.com/servers/storage/nas/>
- ▶ *The IBM System Storage N Series*, SG24-7129
- ▶ *Using the IBM System Storage N Series with IBM Tivoli Storage Manager*, SG24-7243



IBM System Storage DR550

This chapter discusses the IBM System Storage DR550 (DR550) and IBM System Storage DR550 Express which can be used to address the needs of data retention and other business or legal compliance requirements.

The DR550 has been designed to meet the requirements of efficient and effective long-term data retention and protection at a significantly lower total cost of ownership. It can be configured with up to 89.6 TB of disk storage for the IBM System Storage DR550. What differentiates the DR550 from the competition is its ability to incorporate tiered storage by utilizing many other kinds of media seamlessly into its storage pools.

Policies can be established to store an object on the DR550's disk storage for a period of time, say three months, and then retained for the rest of its lifecycle on tape. At the end of its lifecycle, the DR550 can automatically delete the object.

The IBM System Storage DR550 is built with the latest IBM server technology (POWER5 processors), storage technology (IBM System Storage DS4100 subsystems), and data protection technology (IBM System Storage Archive Manager).

10.1 Data retention and compliance

There is a rapidly growing class of data, which is best described by the way in which it is managed rather than the arrangement of its bits. The most important attribute of this kind of data is its retention period, hence it is called *retention managed data*, and it is typically kept in an archive or a repository. In the past it has been variously known as *archive data*, fixed content data, reference data, unstructured data or other terms implying its read-only nature. It is often measured in terabytes and is kept for long periods of time, sometimes forever.

Not only are there numerous state and governmental regulations that must be met for data storage, but there are also industry-specific and company-specific ones as well. And of course these regulations are constantly being updated and amended. Organizations need to develop a strategy that must ensure that the correct records are kept for the correct period of time. Needless to say these records must be readily accessible when they need to be retrieved at the request of regulators or auditors. It is easy to envisage the exponential growth in data storage that will result from these regulations and the concomitant requirement for a means of managing this data. Overall, the management and control of retention managed data is a significant challenge for the IT industry when taking into account factors such as cost, latency, bandwidth, integration, security, and privacy.

10.1.1 Characteristics of retention managed data

When considering the safekeeping of retention managed data, companies also need to consider storage and data characteristics that differentiate it from transactional data.

Storage characteristics of retention managed data are:

- ▶ **Variable data retention period:** Usually a minimum of a few months, up to forever.
- ▶ **Variable data volume:** Many customers start with 5 to 10 TB of storage in an enterprise. It also usually consists of a large number of small files.
- ▶ **Data access frequency:** Write Once Read Rarely or even Write Once Read Never.
- ▶ **Data read/write performance:** Write: handle volume; Read: varies by industry and application.
- ▶ **Data protection:** Requirement for non-erasability, non-rewritability, and deletion when the retention policy expires.

Data characteristics of retention managed data are:

- ▶ **Data lifecycle:** Typically the data is frequently accessed near its creation, then the access frequency diminishes exponentially, near zero. Some industries have peaks that require access – like check images in tax season.
- ▶ **Data rendering:** Ability to view or use data in a very old data store. There is a risk that the data format used to store the data may no longer be supported in 25 years.
- ▶ **Data mining:** With all this data being saved there is intrinsic value in the content of the archive that could be exploited using data mining. But this implies indexing and categorization of the data when it is initially written.

10.1.2 IBM strategy and positioning

Regulations and other business imperatives, as briefly outlined above, stress the need for an Information Lifecycle Management process and tools to be in place. IBM unique experiences with the broad range of ILM technologies, and its broad portfolio of offerings and solutions,

can help businesses address this particular need and provide them with the best solutions to manage their information throughout its lifecycle.

IBM provides a comprehensive, open, set of solutions to help. IBM has products that provide content management, data retention management, and sophisticated storage management, along with the storage systems to house the data. To specifically help companies with their risk and compliance efforts, the IBM Risk and Compliance framework is another tool designed to illustrate the infrastructure capabilities needed to help address the myriad of compliance requirements. Using the framework, organizations can standardize on the use of common technologies to design and deploy a compliance architecture that may help them deal more effectively with compliance initiatives. Key products to the IBM data retention and compliance solutions are IBM System Storage Archive Manager (formerly IBM Tivoli Storage Manager), IBM DB2 Content Manager (formerly IBM Tivoli Storage Manager for Data Retention), IBM DB2 Content Manager, along with any needed disk and tape-based storage.

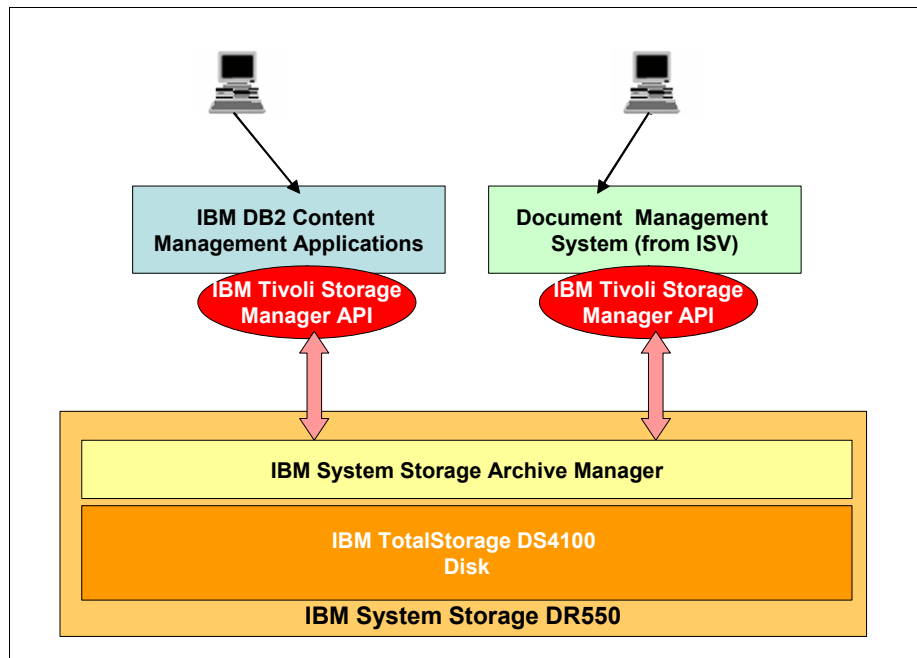


Figure 10-1 Representation of components of the DR550

The main focus of the IBM System Storage DR550 is to provide for a secure storage system, where deletion or modification of data is completely disallowed except through a well-defined retention and expiry policy.

The IBM System Storage DR550 is the repository for regulated business information. It does not create the information. A complete solution includes applications that gather information such as e-mail, instant messaging, and content management applications such as IBM DB2 Content Manager - or other Document Management Systems from Independent Software Vendors-, which can index, archive, and can later present this information to compliance reviewers and finally storage systems that retain the data in accordance with regulations.

10.2 Positioning of the DR550

The DR550 is made of off-the shelf hardware and software components that can be used in many different contexts, it is important to stress the intended use of the offering.

What it is

The DR550 is designed and built primarily to address the issue of compliance with government and industry regulations, in particular, the SEC and the NASD stipulations as it pertains to retaining critical data. It is also an ideal offering for banking and insurance companies for example where data retention and implementing, and managing data retention with appropriate retention and expiration policies is a critical business need.

The main focus is to provide for a secure storage system where deletion or modification of data is completely disallowed except under a well-defined retention and expiry policy.

What it is not

The DR550 is *not* meant to and cannot be used as a standard backup/restore offering even though many of the components of the offering have been deployed as general purpose backup and restore offerings.

DS4100 with EXP100, which is the storage used in this offering, by itself, is a standard external enterprise storage system with wide range of applications such as workgroup and departmental storage including on demand storage access. It is a highly scalable storage system that supports multiple host systems. But the implementation used in this offering uses the Serial Advanced Technology Attachment (SATA) technology, which unlike the Fibre Channel technology is not well suited for many enterprise applications that are mission-critical. The bundling of the offering also does not allow for adding multi-platform hosts. The only access to the storage is through the two servers provided with the offering via the IBM System Storage Archive Manager API.

10.3 IBM System Storage DR550 overview

The IBM System Storage DR550 brings together off-the-shelf IBM hardware and software products.

10.3.1 Hardware

Note: The hardware comes pre-mounted in a secure rack; the software is pre-installed and to a large extent pre-configured. This makes it easy to deploy.

Figure 10-2 on page 191 shows the IBM System Storage DR550.



Figure 10-2 DR550 with console kit

The IBM System Storage DR550 is available in several different configurations: There are two single node and five dual node configurations that vary depending on installed disk storage capacity.

The base single node configuration consists of one IBM @server p5 520 (dual-processor), one IBM System Storage DS4100 Storage Server with 5.6 TB of raw disk capacity, one IBM TotalStorage SAN Switch B16 that connects the p5 520 server to the storage, and a convenient Hardware Management Console, everything installed in a 7014 rack.

The second single node configuration comes with an additional EXP100 expansion enclosure, increasing the total disk storage capacity to 11.2 TB.

Dual node configurations are designed for high availability, and consist of two IBM @server p5 520 (dual-processor) in an HACMP active-passive cluster, two Fibre Channel switches 2005-B16 and one or two DS4100 Storage Servers. Depending on the number of installed EXP100 expansion enclosures, the dual node configurations are available in configurations of 5.6, 11.2, 22.4, 44.8, and 89.6 TB total disk storage capacity. The 89.6 TB configuration features a second 7014 rack to accommodate the second DS4100 Storage Server and the additional seven EXP100 expansion units.

All of these configurations, except for the 89.6 TB dual-node setup, are available with an Enhanced Remote Mirroring (ERM) option. Also included are the ERM activation keys for the DS4100s. Please note that this option requires two DR550s, preferably at different sites. The ERM option uses DS4100 Global Mirror function.

Each configuration is so-called *tape-ready*, and includes additional hardware required to attach tape devices. Tape devices are not included in the configuration - you need to order this separately. You can attach any tape drive/library which is supported by the pSeries, AIX, and IBM System Storage Archive Manager, for example, the IBM System Storage TS1120 Tape Drive. WORM media is almost also a requirement, so you can consider IBM TotalStorage 3592 WORM media, in conjunction with the TS1120, or LTO 3 drive/libraries with WORM capability. Using WORM media allows you to use the DR550 capabilities for non-erasable and non-rewritable data in a tape storage pool.

10.3.2 Software

The software bundle includes the IBM AIX V5.3 operating system, HACMP cluster management software (dual node only), IBM System Storage Archive Manager, and the IBM System Storage DS4000 Storage Manager, customized for additional protection, all running on the p520 server(s). The only way to communicate with the IBM Tivoli Storage Manager server (provided by IBM System Storage Archive Manager) on the DR550 is via the IBM System Storage Archive Manager API. As such, document management applications need to be capable of communicating with this API in order to archive data on the DR550.

Tape device attachment

While the DR550 does not come with a tape drive in its standard configuration, the IBM System Storage Archive Manager in the DR550 can use tape media for the following purposes:

Migration of the data

Migrating data off the primary disk storage pool after a certain percentage of the pool capacity has been reached onto a tape storage pool can tremendously extend, at a reasonable cost, the storage capacity of the DR550.

Backup of the data

Keeping a backup of the archived data in a remote tape storage pool protects from disasters that could happen to the disk storage pool or even the entire site where the DR550 is located. With data being an essential asset for many enterprises, having the possibility to recover from a disaster by restoring from the tape backup pool is invaluable.

Backup the System Storage Manager database

Like the actual data, the System Storage Archive Manager database needs protection from the very same scenarios described above.

The redbook *Understanding the IBM System Storage DR550*, SG24-7091 shows some examples of how to attach and use tape in the context of the DR550 archiving solution, including the IBM TotalStorage 3494 Tape Library and IBM TotalStorage 3592 Tape Drive, including WORM capability.

10.4 IBM System Storage DR550 Express

The IBM System Storage DR550 Express has many of the features and benefits of the DR550, but at a much lower cost making it ideal for small to medium businesses.

The DR550 Express is shipped with approximately 1 TB of storage capacity, using the internal SCSI disk of the p5 520 server (eight 146 GB Ultra™ SCSI 3 disk drives). The DR550 Express is tape ready. Storage can be expanded to provide a capacity of either 3.9 TB or 6.7 TB using a DS4100 Storage Server populated with up to fourteen 400 GB SATA disks.

The DR550 Express can be integrated into a 25U 49 inch rack.

Figure 10-3 shows the basic pSeries server used in the IBM System Storage DR550 Express.



Figure 10-3 IBM System Storage DR550 Express

The IBM System Storage DR550 Express consists of one IBM eServer p5 520 (dual-processor), one IBM System Storage DS4100 Storage Server (optional, for expanded storage capacity), one IBM TotalStorage SAN Switch B16 that connects the p5 520 server to the storage, and a convenient Hardware Management Console, everything installed in a 7014-S25 rack.

The software bundle includes the IBM AIX V5.3 operating system, IBM System Storage Archive Manager, and the IBM System Storage DS4000 Storage Manager, customized for additional protection, all running on the p520 server. The DS4000 Storage Manager is used only if the DS4100 disk expansion is attached. The only way to communicate with the IBM Tivoli Storage Manager server (provided by IBM System Storage Archive Manager) on the DR550 is via the IBM System Storage Archive Manager API. As such, document management applications need to be capable of communicating with this API in order to archive data on the DR550.

The DR550 Express has similar optional tape attachment capabilities to the DR550.

10.5 Data retention features

The key software component of the DR550 and DR550 Express is IBM System Storage Archive Manager. IBM System Storage Archive Manager is what gives the DR550 its unique data retention features.

Note: The IBM System Storage DR550 is an archive solution focused on providing a data retention repository for applications that utilize the provided API to archive data to the IBM System Storage Archive Manager server.

It is beyond the scope of this document to explain IBM System Storage Archive Manager in detail. For a detailed overview of IBM Tivoli Storage Manager and its complementary products, refer to *IBM Tivoli Storage Manager Administrator's Guide*, GC32-0768, *IBM Tivoli Storage Management Concepts*, SG24-4877, and *IBM Tivoli Storage Manager Implementation Guide*, SG24-5416, and Chapter 24, "IBM Tivoli Storage Manager" on page 437.

In this section we discuss the IBM System Storage Archive Manager data retention features. We concentrate on explaining added archive retention features such as event based archive retention, deletion hold, and data retention protection.

In order to use the archive function of IBM System Storage Archive Manager, you must define valid policies that preclude defining a policy domain, policy set, management class or classes, and an archive copy group, as well as setting archive retention parameters in the

archive copy group, and associating your application clients (applications using the API) with the IBM System Storage Archive Manager policies.

The archive copy group parameters that govern retention are RETVER, RETINIT, and RETMIN. The RETINIT and RETMIN parameters were introduced in IBM Tivoli Storage Manager V5.2.2 to make it possible for applications utilizing the API to further control the retention period (RETVER) for archive objects.

10.5.1 Methods of archive retention

There are two methods of archive retention, which are defined by the parameters of the archive Copy Group:

- ▶ Chronological archive retention
- ▶ Event based archive retention

We now look at the parameters of the archive copy group, and their possible values for the two archive retention methods.

Existing archive retention parameter

- ▶ RETVER (retain version)

Possible value: `retver=0` to 30,000 days or NOLIMIT

The retain version parameter (RETVER) within the archive copy group specifies the number of days to retain each archive object. Possible values are 0 to 30,000 days or NOLIMIT, which means that an archive copy is maintained indefinitely - that is, you can never delete it.

New archive retention parameters

- ▶ RETINIT (retention initiation)

`retinit= creation or event`

The retention initiation (RETINIT) parameter specifies when the time specified by the retain version (`RETVER=n days`) attribute is initiated. The possible values for this parameter are `creation` or `event`. The default value is `creation`. We explain both values below.

Retinit=creation (Chronological archive retention)

By setting this parameter to `creation` (`RETINIT=creation`) in the archive copy group you specify that the retention time specified by the RETVER attribute (`retver=n days`) is initiated right at the time an archive copy is stored on the server. This is referred to as *chronological archive retention*.

Retinit=event (Event based archive retention)

By setting this parameter to `event` (`RETINIT=event`) in the archive copy group, you specify that the retention time (`RETVER=n days`) for the archived data is initiated by an application that utilizes API function calls. If the application never initiates the retention, the data is retained indefinitely. This method of archive retention is referred to as *event based archive retention*.

- ▶ RETMIN (retain minimum)

`retmin= 0` to 30,000 days

The retain minimum (RETMIN) parameter applies only to event based archive retention policy, and specifies the minimum number of days to retain an archived object regardless of the value of RETVER. The default value is 365. Possible values are 0 to 30,000 days (approx. 82 years).

Chronological archive retention

Figure 10-4 shows a simplified view of a chronological retention policy. With `RETINIT=creation` and `RETVAR=365` days, a file that is archived on day 0 is retained 365 days and becomes eligible for expiration. In this case, after 365 days from the time the data was created, all references to that data are deleted from the database, making the data irretrievable from IBM System Storage Archive Manager storage volumes. This kind of archive retention is called *chronological retention*. By default the `RETINIT` value is set to creation.

Tip: Use chronological archive retention when the application that is doing the archiving is not able to send retention events such as activate, hold, and release. Chronological archive retention is also to be used when you are archiving to a regular IBM Tivoli Storage Manager Server (as opposed to IBM System Storage Archive Manager) through the normal backup/archive client (not applicable for DR550).

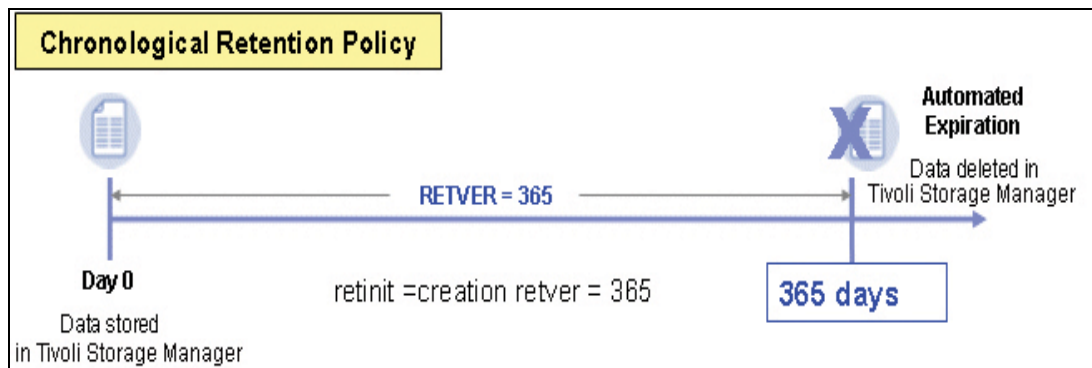


Figure 10-4 Chronological retention policy

Archive copy groups using chronological retention policy satisfy many archive retention requirements.

Event based retention policy

In certain situations, data retention periods cannot be easily defined, or they depend on events taking place long after the data is archived. Event based archive retention is designed to meet these requirements. Event based retention policy is designed for applications that utilize the API function calls to trigger events also known as *retention events*.

Figure 10-5 on page 196 shows a timeline depicting event based policy. In this example, an application utilizing the API archives data using the retention values shown. The archived data is retained for a minimum of 2555 days (`RETMIN=2555`). If the retention time (`RETVAR`) is activated through an API retention event, IBM System Storage Archive Manager assigns an expiration date for this object. The expiration date that IBM System Storage Archive Manager assigns is whichever comes later, either:

- ▶ The date the object was archived plus the number of days specified in the `RETMIN` parameter.
- Or,
- ▶ The date the event was signaled, plus the number of days specified in the `RETVAR` parameter.

After reaching this expiration date, the data is eligible for expiration. When the time for expiration occurs, all references to that data are deleted from the IBM System Storage

Archive Manager database making the data irretrievable from IBM System Storage Archive Manager storage volumes. This kind of archive retention is referred to as event based retention.

Tip: Use event based archive retention, if the archive application you are using (such as IBM DB2 Content Manager together with Record Manager) utilizes the API function calls to activate the retention period of the archived data objects.

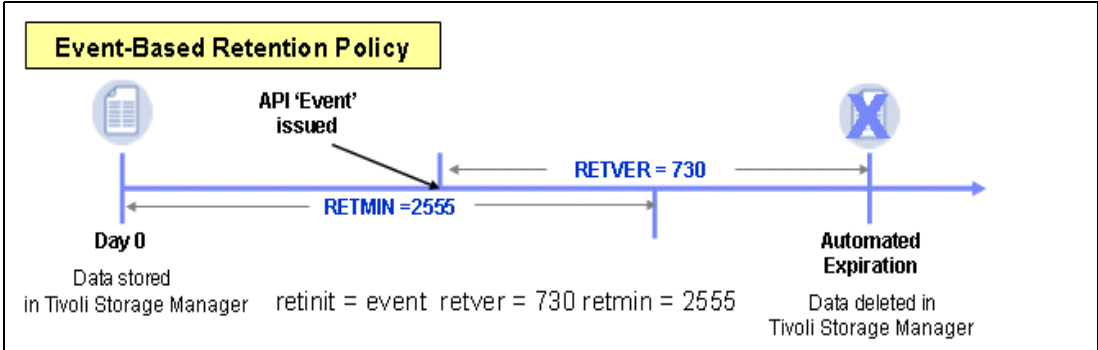


Figure 10-5 Event based retention policy

Deletion hold and release

Some regulations require that the data be retained longer than the minimum retention period in certain cases. This might be any litigation, legally or a company required audit, or criminal investigation requiring the data as evidence. The API supports new function calls used to place a *deletion hold* on an archive object. These functions are also called retention events. A deletion hold can be applied at any point in time during the retention period for an archived object. The object will then be retained until a *deletion release* is applied. If a deletion release is not applied, the object is retained indefinitely. Although deletion hold and release are events, they can be applied to objects archived not only using the event based policies, but also the chronological, creation based policies.

Figure 10-6 shows a timeline depicting deletion hold and release.

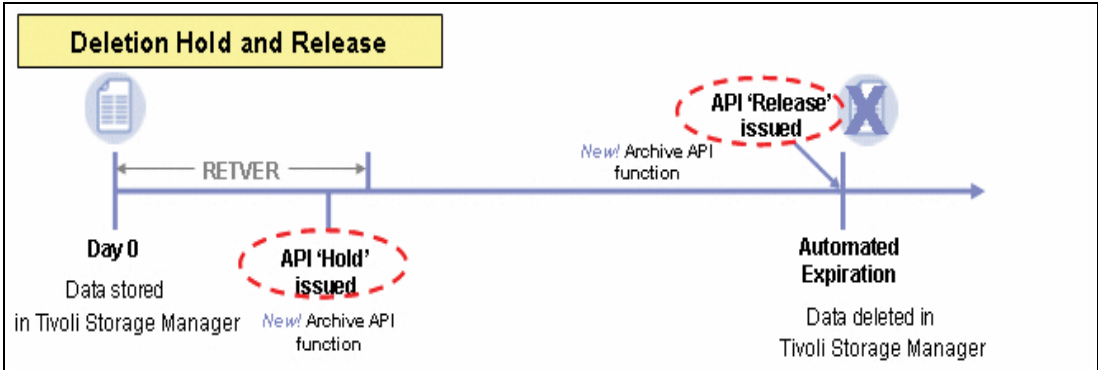


Figure 10-6 Deletion hold and release

10.6 Document Management System

Any Document Management System (DMS) that can use the archive and retention functions offered through the IBM System Storage Archive Manager API should be able to

communicate and use the DR550. Indeed, *only* applications or middleware using the API can send data to, or retrieve data from DR550. Information regarding the IBM System Storage Archive Manager API client may be found at:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v1r1/index.jsp?toc=/com.ibm.itsmfdt.doc/toc.xml>

In addition to the IBM DMS solution, DB2 Content Manager, several vendors have qualified (or have committed to qualify) one or more of their applications with DR550 and DR550 Express. These vendors include:

- ▶ AXS-One, Inc.
- ▶ BrainTribe (formerly Comprendium)
- ▶ Caminosoft Corporation
- ▶ Ceyoniq
- ▶ d.velop AG
- ▶ Easy Software AG
- ▶ FileNet corporation
- ▶ Heilig and Schubert (H&S) Software AG
- ▶ Hummingbird
- ▶ Hyland Software (OnBase)
- ▶ Hyperwave AG
- ▶ IRIS Luxembourg Software (Documentum Connector)
- ▶ MBS Technologies (iSeries Connector for IBM CM V5)
- ▶ OpenText (formerly IXOS)
- ▶ Princeton Softech Optim 6.2 (for Pe
- ▶ Saperion
- ▶ SER Solutions GmbH
- ▶ Triade Information Systems GmbH (NFS/CIFS/FTP Gateway)
- ▶ Symantec Enterprise Vault (formerly Veritas/KVS)
- ▶ Windream GmbH
- ▶ Zantaz

Additional information about qualified ISVs may be found at:

<http://www.ibm.com/servers/storage/disk/dr/interop.html>

You can also find information about ISV support for the IBM System Storage Archive Manager API client on the Tivoli Web pages:

<http://www.ibm.com/software/tivoli/partners/public.jsp?tab=connect&content=rft-validation-tsm>

Additional information about IBM DB2 Content Manager can be found in *Understanding the IBM System Storage DR550*, SG24-7091.

10.7 More information

For a more detailed understanding of the DR550 refer to these sources:

- ▶ *Understanding the IBM System Storage DR550*, SG24-7091
- ▶ *IBM System Storage DR550 Installation, Setup, Operations, and Problem Determination Guide*:

<http://www-1.ibm.com/servers/storage/disk/dr/pdf/DR550-Users-Guide.pdf>

- ▶ IBM System Storage DR550 Web page:

<http://www-1.ibm.com/servers/storage/disk/dr/index.html>



Part 2

Tape systems

In Part 2 we review IBM System Storage and TotalStorage offerings for tape drives, tape libraries, and virtual tape solutions.



IBM TotalStorage and System Storage tape drives

This chapter provides information about IBM tape drives. It includes the Linear Tape-Open (LTO) technology and related Ultrium specifications, as well as the IBM TotalStorage 3590 Tape Drive and the IBM System Storage TS1120 Tape Drive.

We first provide some general information about LTO, 3592, and 3590 tape technology, and then cover the following specific products. Note that some models have recently been renamed. The former names are also given here.

- ▶ **IBM TotalStorage 3580 Tape Drive:**
 - IBM TotalStorage 3580 Tape Drive Model L33
 - IBM TotalStorage 3580 Tape Drive Model L3H
- ▶ **IBM System Storage TS1030 Tape Drive**
 - IBM System Storage TS1030 Tape Drive Model F3B (TS1030 tape drive, LTO Gen 3 tape drive; formerly known as 3588 F3B)
- ▶ **IBM TotalStorage 3590 Tape Drive**
 - Model E11 and E1A (3590 E11 and E1A)
 - Model H11 and H1A (3590 H11 and H1A)
- ▶ IBM TotalStorage 3592 Tape Drive Model J1A (TS1120 tape drive, formerly known as 3592 J1A)
- ▶ IBM System Storage TS1120 Tape Drive (TS1120 tape drive; formerly known as 3592 E05)
- ▶ IBM System Storage TS1120 Tape Controller Model C06 (formerly known as 3592 C06)
- ▶ IBM TotalStorage 3592 Tape Controller Model J70 (3592 J70)

We provide information about each model covering their characteristics in terms of capacity, performance, scalability, and reliability.

11.1 Technology overview

IBM tape drive products support various platforms including IBM System i, System p, System z, System x, other UNIX platforms (HP-UX, Sun Solaris), and Windows.

11.1.1 LTO technology

Linear Tape-Open (LTO) is an open format technology. This means that users can have multiple sources of products and media. The LTO technology establishes a new open format specification for high-capacity, high-performance storage products and addresses a growing customer need for improved data interchange across platforms.

LTO technology was developed jointly by IBM, Hewlett Packard (HP) and Seagate in 1997 to provide a clear and viable choice in an increasing complex array of tape storage options. The consortium created two specifications, Accelis and Ultrium.

The current technology provider companies are IBM, HP, and Certance LLC (owned by Quantum). The Accelis technology has not been pursued by manufacturers as it is apparent that Ultrium meets the market needs.

Some technical standards defined for LTO technology are:

- ▶ Cartridge dimension (approximate) is 4.1 x 4.0 x 0.8 inches (105 x 102 x 21 mm)
- ▶ The single-hub design allows for the cartridge to be optimally packed with media. High capacity is further enhanced by the use of an LTO technology data compression algorithm with two control modes to maximize compression efficiency.
- ▶ There are three currently available Ultrium formats:
 - Ultrium format generation 1: 200 GB capacity (2:1 compression) per cartridge. Ultrium 1 drives are no longer being actively sold, although the media can still be read by Ultrium 3 drives, and read/written by Ultrium 2 drives.
 - Ultrium format generation 2: 400 GB capacity (2:1 compression) per cartridge.
 - Ultrium format generation 3: 800 GB capacity (2:1 compression) per cartridge.

The next generation, Ultrium 4, is on the published roadmap, and is intended to support 800 GB native/1.6 TB compressed per cartridge.

LTO technology delivers the following benefits:

- ▶ Easily integrates into current operating environments
- ▶ Four generation roadmap protects investment today and in the future as each new generation reads previous generations formats.
- ▶ Migration paths focused on increasing maximum transfer rates to reduce backup window. In addition, capacity will double in each successive generation.
- ▶ Simplified product planning means faster cycle time for new features.
- ▶ Compliance testing ensures that LTO Ultrium drives and media cartridges conform to the specification to deliver data interchange among multiple vendors' products.
- ▶ Instant access to data; this is made possible via LTO CM (Cartridge Memory), which is a passive, contactless silicon storage device that is physically a part of the cartridge.
- ▶ Dedicated Dual Servo; the servo bands are pre-written on the tape during the tape cartridge manufacture process. If one servo element becomes defective, the head will continue to track as a result of the second "redundant" servo system.

For details about the LTO or Ultrium, refer to:

<http://www.lto-technology.com>
<http://www.ultrium.com>

Interleaved recording

The LTO drive uses an interleaved, serpentine, longitudinal recording format. The first set of 8 or 16 data tracks is written from near the physical beginning of the tape to near the physical end of the tape. The head then repositions to the next set of tracks for the return. This process continues until all tracks are written and the tape is full.

The format of the recording of the data and servo tracks is defined as part of the LTO specification in order to meet the requirement for interchange between different manufacturers' implementations.

Servo tracks

Servo tracks (also called servo bands) enable accurate positioning of the tape drive head over the data track, ensuring that the head does not stray onto an adjacent track. They are necessary to support high-data densities on the tape where the tracks are very close together. The servo bands are written at time of cartridge manufacture, before the cartridge is usable for data storage and retrieval. If the servo bands should be erased, the tape will be unusable. Servo tracks are like lane markings on a multi-lane highway. Imagine how difficult it would be to drive on the highway without any lane markings. Lane markings help by positioning you on the lane, just as servo tracks support the drive recording head to position on the data tracks.

Track following

Each pair of servo bursts is at an angle to each other, and the servo heads move such that they keep a constant value for the distance between the bursts. In this way the servo is able to follow a straight line within the servo band; any small deviation away from the correct path causes a variation (plus or minus) in the gap between the bursts (see Figure 11-1). Provided that the servo head element follows a straight line along the servo band, then the distance "x" shown in the figure remains constant. IBM LTO drives use two servo bands simultaneously during write operations to provide two sources of servo information, and therefore increased accuracy.

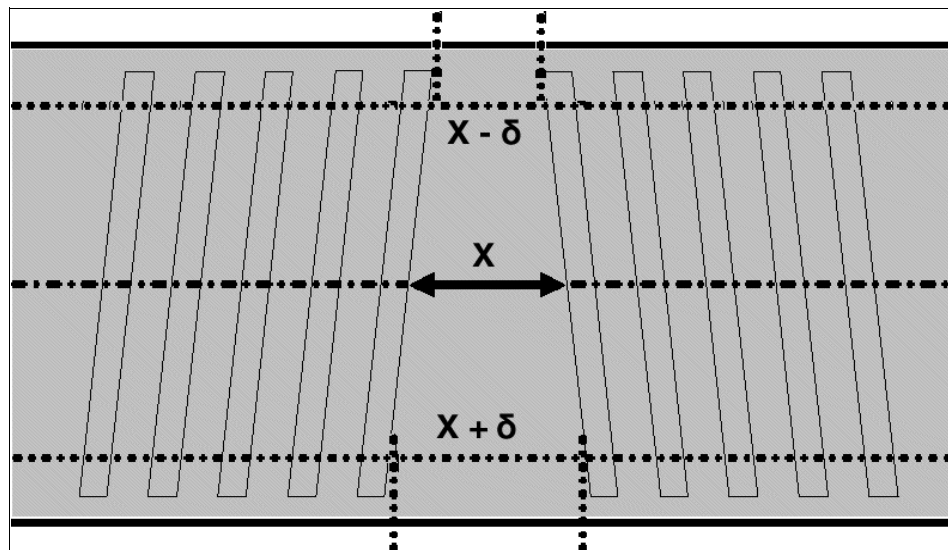


Figure 11-1 Magnified servo band showing a pair of servo bursts

Longitudinal positioning

The LTO servo band is designed not only for track following, but also for recording the longitudinal position (LPOS). The absolute location down the length of the tape and the manufacturer data are recorded in LPOS *words* approximately every quarter-inch (.7 cm) along the tape. The LPOS word consists of symbols constructed from bit sequences (ones and zeros); these bits are encoded within the servo frames. Each servo frame encodes one bit using the first pair of servo bursts. When servo stripes 2 and 3 (out of the five) are shifted inward (see Figure 11-2), this encodes a zero; when servo stripes 2 and 3 are shifted outward, this encodes a one.

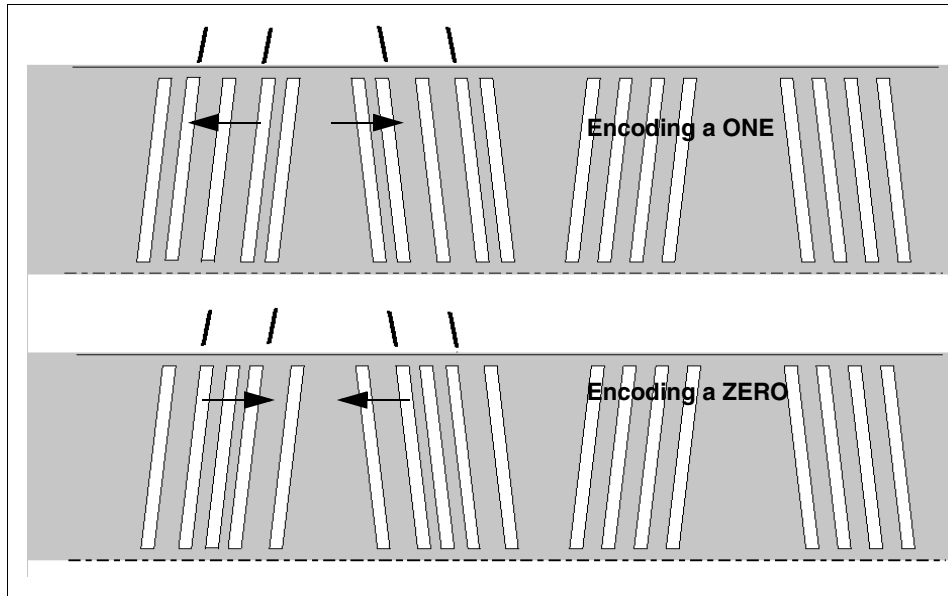


Figure 11-2 Encoding bits using the servo stripes within the servo bursts

The LPOS word contains 36 bits and therefore has a length of 36 servo frames. Each of the 5 servo bands on the tape may be uniquely identified by the relative positions of the frames down the tape, in adjacent servo bands. The offset of the frames between servo band n and servo band $n+1$ are specific to each servo band (0 and 1, 1 and 2, 2 and 3, or 3 and 4). Therefore, the drive can move the head directly from the physical beginning of the tape to a specific logical position for reading or writing.

Data tracks

The area between adjacent servo bands is a data band. There are four data bands numbered 2, 0, 1, and 3, where data band number 2 is nearest the reference edge of the tape and data band 3 is farthest away, as in Figure 11-3 on page 205. The data bands are written in sequence beginning with 0 (in the center of the tape) and ending with 3.

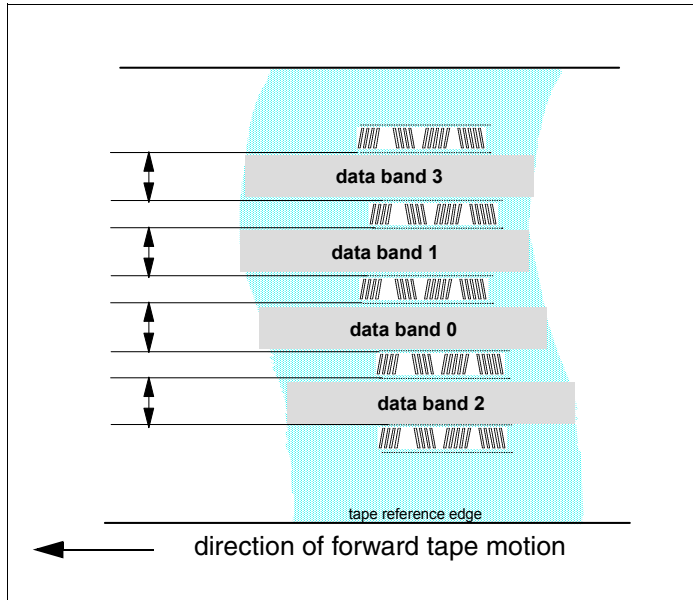


Figure 11-3 Four data bands written between the servo tracks

Backwards compatibility

The LTO standard specifies a backwards compatibility of writing one generation backwards and reading two generations backwards. To make this possible, LTO uses a technique called *shingling*. When using shingling, a write track may overlap the bottom of a previously written track. LTO Generation 2 uses shingling when writing data to an LTO Generation 2 Cartridge. The very first two passes write to the tape in the normal way. The following passes can partially overwrite previously written data tracks. The IBM LTO Generation 2 write head width is that of the LTO Generation 1. Therefore the LTO Generation 2 drives can write an LTO 1 cartridge in full track width, and when writing to an LTO 2 cartridge it uses the shingling write function. To read the residual LTO generation 2 data tracks, the read head must of course be narrower than the LTO generation 1 read head (Figure 11-4 on page 206).

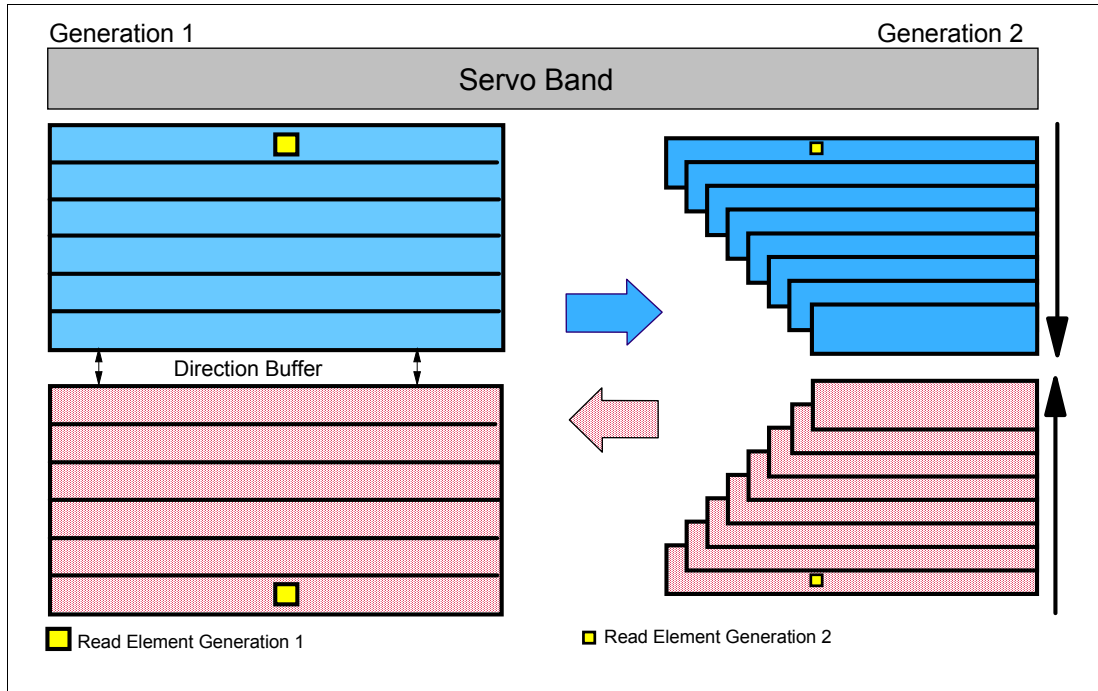


Figure 11-4 Shingling - writing to tape generation 1 versus generation 2

Again, LTO Generation 3 is conceptually the same, but technically different. The Ultrium 3 data is written 16 tracks at a time utilizing the shingling technique. To achieve compatibility, the write heads are equal to the residual track width of the Generation 2 format, and the spacing between alternate (every second) Generation 3 write heads is the same as between the Generation 2 write heads. Thus Generation 3 data is written using the shingling method, but Generation 2 data is written using only every second write head and writing the full Generation 2 data width. Generation 1 data can similarly be read by every other read head.

Linear density

The linear density for LTO 1 is 4880 bits per mm. The linear density was improved for LTO 2 to 7398 bits per mm, and for LTO 3 to 9638 bits per mm. Both LTO 1 and LTO 2 cartridges are of the same length, 610 m (2000 feet). To achieve the required doubling of capacity, the LTO 3 tape is slightly longer, 680 m (2231 feet).

11.1.2 The IBM LTO Ultrium 3 technology



IBM LTO Ultrium 3 tape drive is the latest in IBM LTO Ultrium Tape Drive technology available.

Description

The LTO Ultrium 3 tape drive incorporates third-generation IBM LTO Ultrium technology. It offers the following significant improvements over the Ultrium 2 Tape Drive:

- ▶ Maximum tape drive throughput **data rate performance is more than doubled**, up to 80 MBps native data transfer rate. Data tracks are now written 16 at a time. IBM Ultrium 3 Tape Drives can read and write, at eight data tracks at a time, LTO Ultrium 2 Data Cartridges at Ultrium 2 rates, and read LTO Ultrium 1 Data Cartridges at Ultrium 1 rates.

Note: Although the LTO Ultrium 3 Tape Drive provides the capability for excellent tape performance, other components of the system may limit the actual performance achieved. Also, although the compression technology used in the tape drive can typically double the amount of data that can be stored on the media, the actual degree of compression achieved is highly sensitive to the characteristics of the data being compressed.

- ▶ The **tape cartridge capacity is doubled** over the Ultrium 2 Data Cartridge up to 400 GB native physical capacity (800 GB with 2:1 compression), with the use of the IBM System Storage LTO Ultrium 400 GB Data Cartridge. This is achieved by increasing the linear density, the number of tape tracks, and the media length. The media length in this case has been increased by 231ft (70 m). The tape itself is an advanced metal particle tape developed to help provide durability and capacity.
- ▶ **Ultrium 2 cartridge compatibility** — The Ultrium 3 Tape Drive can read and write on Ultrium 2 cartridges.
- ▶ **4 Gbps and 2 Gbps Fibre Channel attachment** — The 3588 Model F3A, for older models of the IBM TotalStorage 3584 Tape Library) supporting the LTO3 technology, comes with a 2 Gbps Fibre Channel interface. The IBM System Storage TS1030 Tape Drive, used in the IBM System Storage TS3310 Tape Library and IBM System Storage TS3500 Tape Library, comes with a 4 Gbps Fibre Channel interface. Both drives support a large variety of operating systems.
- ▶ **New dual stage 16 channel head actuator** — designed to provide precision head alignment to help support higher track density and improved data integrity.
- ▶ **New independent tape loader and threader motors and positive pin retention** — designed to help improve the reliability of loading and unloading a cartridge, and to retain the pin even if tension is dropped. With an independent loader motor coupled with the positive pin retention, the tape threads with a higher level of reliability
- ▶ **Larger internal data buffer** — There is a 128 MB internal data buffer in the Ultrium 3 Tape Drive as compared to a 64 MB internal data buffer in the Ultrium 2 Tape Drive
- ▶ **Highly integrated electronics using IBM-engineered copper technology** — designed to reduce the total number of components in the drive, lower chip temperatures, and reduce power requirements, helping to provide for a more reliable drive. The generation 3 drive electronics are also designed to provide “on-the-fly” error correction capability for soft errors in the memory arrays in data and control paths.
- ▶ **Graceful dynamic braking** — In the event of power failure, reel motors are designed to maintain tension and gradually decelerate instead of stopping abruptly, helping reduce tape breakage, stretching, or loose tape wraps during a sudden power-down.

Enhanced features

Enhanced features in the IBM LTO Ultrium 3 Tape Drive include:

WORM tape format

Beginning with LTO Ultrium format generation 3, Write Once Read Many (WORM) functionality provides for non-erasable, non-rewritable operation with tape media and is designed for long-term, tamper-resistant record retention.

The LTO Ultrium format generation 3 specification for WORM includes low-level encoding in the Cartridge Memory (CM) and also mastered into the servo pattern as part of the manufacturing process. This encoding is designed to prevent tampering. Data can be appended at the end of a WORM cartridge to which data was previously written allowing the full use of the high-capacity tape media. LTO Ultrium format generation 3 non-WORM and WORM drives can coexist.

Servo and track layout technology

There are 704 data tracks in Ultrium 3 versus 512 data tracks in Ultrium 2. High bandwidth servo system features a low-mass servo to help more effectively track servo bands and improve data throughput with damaged media in less-than-optimal shock and vibrate environments.

Surface Control Guiding Mechanism

IBM patented Surface Control Guiding Mechanism is designed to guide the tape along the tape path in the Ultrium 3 drive. This method uses the surface of the tape, rather than the edges, to control tape motion. This helps reduce tape damage (especially to the edges of the tape) and tape debris, which comes from the damaged edges and can accumulate in the head area.

Magneto Resistive (MR) head design

Use of flat lap head technology in MR heads for Ultrium 3 helps minimize contact, debris accumulation, and wear on the tape as it moves over the read/write heads.

Digital speed matching

The Ultrium 3 Tape Drive performs dynamic speed matching (at one of five speeds, 40, 50, 60, 70, 80 MBps) to adjust the drive's native data rate as closely as possible to the net host data rate (after data compressibility has been factored out). This helps reduce the number of backhitch repositions and improve throughput performance. Speed matching on Ultrium 3 ranges from 40 to 80 MBps versus 17.5 to 35 MBps on Ultrium 2.

Robust drive components optimized for automation environments

Drive designed using some of the most robust components available, such as: (1) all metal clutch, (2) steel ball bearings in loader, (3) robust leader block design, (4) single circuit card, to help enhance reliability and prolong the life of the drive.

Power management

The Ultrium 3 Tape Drive power management function is designed to control the drive electronics to be either completely turned off or to be in a low-power mode when the circuit functions are not needed for drive operation.

Adaptive read equalization

Designed to automatically compensate for dynamic changes in readback signal response.

Dynamic amplitude asymmetry compensation

Designed to dynamically optimize readback signals for linear readback response from magneto resistive read head transducers.

Separate writing of multiple filemarks

Separate writing of multiple filemarks is designed to cause any write command of two or more filemarks to cause a separate data set to be written containing all filemarks after the first. This feature has two advantages, first it helps improve performance if a subsequent append overwrites somewhere after the first filemark. Second, write of multiple filemarks typically indicates a point where an append operation might occur after the first of these filemarks.

This change helps prevent having to rewrite datasets containing customer data and the first filemark in cases if such an append occurs.

LTO Data Compression (LTO-DC)

The Ultrium 3 uses LTO-DC which is an implementation of a Lempel-Ziv class 1 (LZ-1) data compression algorithm. LTO-DC is an extension of Adaptive Lossless Data Compression (ALDC) and an improvement over previous IBM lossless compression algorithms. IBM patented “Scheme-Swapping” compression is designed to look ahead at incoming data, and determine the most efficient storage method (either ALDC or pass-thru mode) to help provide optimal data compression and increased data throughput.

LTO Cartridge Memory (LTO-CM)

Contained within the LTO Ultrium data cartridge is the LTO-CM, which is a passive, contactless silicon storage device that is physically a part of the cartridge. The LTO-CM is used to hold information about that specific cartridge, the media in the cartridge, and the data on the media. The storage capacity of the LTO-CM is 4,096 bytes. Communication between the drive and the LTO-CM is via a low-level RF field transmitted by the drive to the cartridge.

Statistical Analysis and Reporting System (SARS)

The Ultrium 3 Tape Drive uses SARS to help isolate failures between media and hardware. The SARS use the cartridge performance history saved in the CM module and the drive performance history kept in the drive flash EEPROM to help determine the more likely cause of failure. SARS can cause the drive to request a cleaner tape, to mark the media as degraded, and to indicate that the hardware has degraded.

Reliability

The IBM LTO Ultrium tape format differs from earlier IBM products. Reliability and availability features include:

- ▶ Data integrity: The drive performs a read after write, for verification.
- ▶ Power loss: No recorded data is lost as a result of normal or abnormal power loss while the drive is reading or writing data.
- ▶ Error correction: Two levels of error correction that can provide recovery from longitudinal media scratches.
- ▶ Integrated head cleaner: During the load process, a brush integrated into the drive mechanism cleans the head before it is used with the tape.
- ▶ Surface control guiding: Guides the tape along the tape path using the surface of the tape rather than the edges to control tape motion.
- ▶ Flat lap head: Improves contact between the read and write recording elements and the tape, giving higher quality recording and readback of data.
- ▶ Statistical Analysis and Reporting System (SARS): Only IBM LTO drives provide this level of preventive diagnostic reporting to assist in isolating failures between media and hardware.

Ultrium 1, 2 and 3 compatibility

The IBM Ultrium 3 tape drives support read and write to Ultrium 2 cartridges and read only for the Ultrium 1 cartridges. IBM Ultrium 2 tape drives (both standalone and in IBM Ultrium libraries) support both Ultrium 1 and Ultrium 2 cartridges. An Ultrium 1 cartridge in an Ultrium 2 drive will be written at the same 100 GB native capacity, but with improved performance (20 MBps). Ultrium 1 drives cannot read or write an Ultrium 2 or 3 cartridge. If you put an Ultrium 2/3 cartridge in an Ultrium 1 drive, then you will get an “Unsupported Cartridge Format” failure. Figure 11-5 on page 210 shows the compatibility.

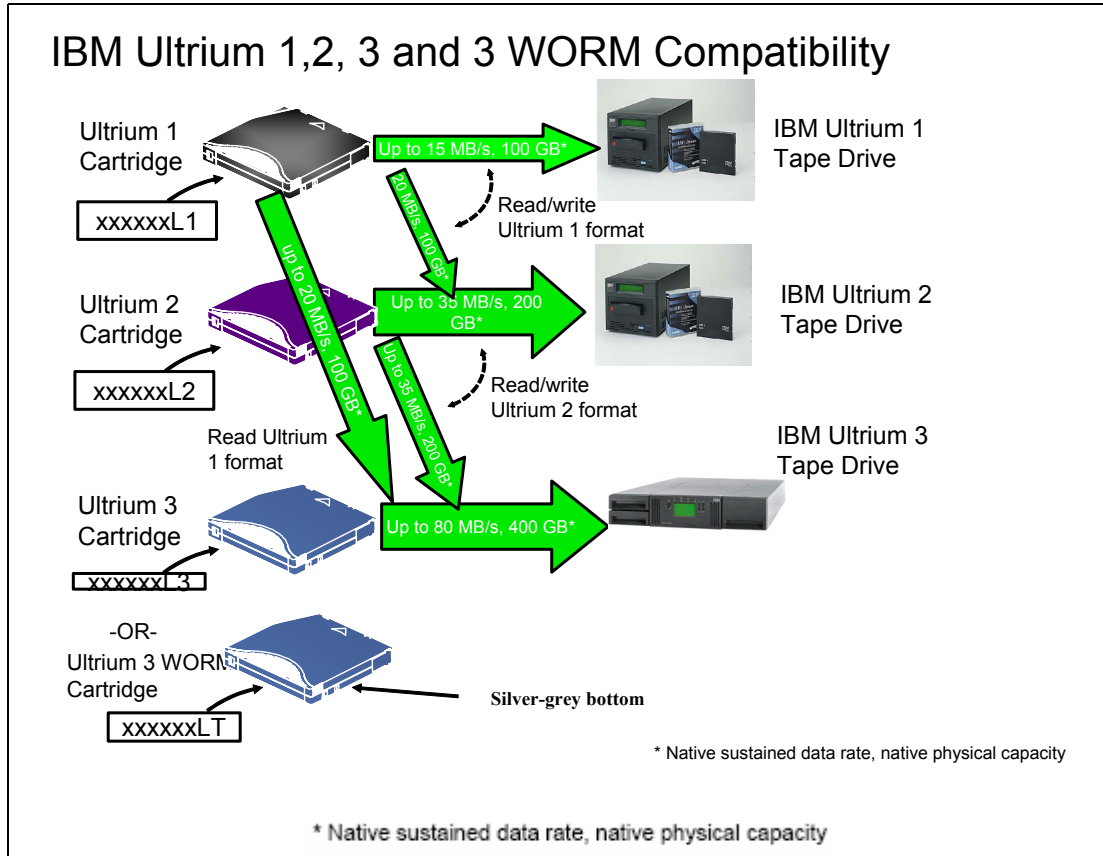


Figure 11-5 Ultrium 1, 2 and 3 compatibility

For more information, refer to this Web site:

<http://www.ibm.com/servers/storage/tape/lto/>

11.1.3 3592 technology

The 3592 models (IBM TotalStorage 3592 Tape Drive Model J1A and IBM System Storage TS1120 Tape Drive - formerly known as 3592 Model E05) exploit a number of technologies to provide improved speed, capacity, and reliability. For simplicity, we will refer to them collectively in this section as the 3592 models.

Interleaved recording

The 3592 drive and recording technology has many similarities with the technology introduced in 11.1.1 “LTO technology” on page 202.

The 3592 models use the same interleaved recording technique as also used in “Interleaved recording” on page 203. Other LTO technology like “Servo tracks” on page 203, “Track following” on page 203, and “Longitudinal positioning” on page 204 also apply.

Feeds/speeds/speed matching

The IBM System Storage TS1120 Tape Drive has a maximum native data rate of up to 104 MBps.

The IBM TotalStorage 3592 Tape Drive has a maximum native data rate of up to 40 MBps.

Both drives use the proven Digital Speed Matching technology shipped in IBM LTO tape drives to monitor the effective (after data compression) host data rate and switch to lower the read/write speed to keep pace with the host. Digital Speed Matching has two important effects. First, it helps dramatically reduce the number of back hitches required to read or write a tape when the drive is attached to a system that does not have an effective data rate high enough to stream the drive at its maximum native data rate. Digital speed matching helps significantly reduce the stress on the tape that can result from constant backhitching. Second, it helps reduce the time required to reposition, or backhitch, the tape.

Cartridge Memory (CM)

This is identical to what was described in “LTO Cartridge Memory (LTO-CM)” on page 209. Contained within the cartridge is the Cartridge Memory (CM), a passive, contactless silicon storage device (4,096 bytes) that is physically a part of the cartridge. The CM is used to hold information about that specific cartridge, the media in the cartridge, and the data on the media. It is designed to support the High Resolution Tape directory feature. The CM is located in the left-rear corner of the cartridge and is mounted at an angle similar to LTO to allow the possibility of interfacing to the CM from the rear of the cartridge by a picker or other device. Communication between the drive and the CM is performed via a non-contact passive radio frequency interface that eliminates the need for physical connections to the cartridge for power or signals.

Virtual backhitch (non-volatile caching)

As other modern tape drives do, the 3592 models stage write data through an intermediate DRAM buffer on its way to tape. This buffer is volatile in that it will not retain what is stored in it if power is lost. For streamed writes (or reads) this buffer yields considerably improved performance. When a pre-3592 drive is performing a streamed write to tape and the buffer empties, or if a synchronizing command is received which forces the buffer to be written to tape, then the streamed writing will cease for want of data. Any non-immediate write-type command (for example, how file marks are typically written by most mainframe applications) is considered a synchronizing command. Non-immediate write-type commands require the drive to store data to tape before returning Command Complete (with good status) in response to that command. This by definition forces all the data in the volatile buffer to be written to tape.

When streaming writes cease, a pre-3592 tape drive halts the tape and repositions it upstream of where writing ended. This allows subsequently received data to be written immediately following the previously written data, so as to eliminate the waste of the considerable length of tape from the point at which good status is returned to the host to the point at which the host has subsequently sent enough data to resume writing. For example, if tape is streaming at 4.74 m/s when the buffered data falls below the threshold, an entire meter of tape can pass unwritten in about 210 milliseconds. Substantial lengths of unwritten tape can significantly reduce capacity. Heretofore, a backhitch has been used by typical tape drives to eliminate this capacity loss following a synchronizing write to tape.

Non-volatile caching, or NVC, is a 3592 feature which can help greatly improve write performance through backhitch reduction. This system temporarily reserves portions of physical tape for cache areas. Data received from the host is written to the volatile buffer as usual, and also to nonvolatile tape cache areas—with the exception that no backhitch is typically necessary when writing temporary copies to cache areas of tape. This temporary capacity loss is easily recouped. The data is written to temporary cache areas and is not released in the volatile buffer, but instead it accumulates. This accumulation typically continues until the buffer is nearly full. At this time the accumulated data in the buffer is rewritten via a streamed write to the standard area of tape. When the rewrite is complete, the temporary cache areas of tape are released so that they may be overwritten. Writing temporary copies to the cache areas of tape without backhitching until the buffer is nearly full,

and then streaming a rewrite of the data to the standard area of tape, can help significantly improve the average write throughput to tape.

Capacity scaling and segmentation

The 3592 models support scaling and segmentation modes on the cartridge to enable customers to trade off capacity for improved access times. The IBM System Storage TS1120 Tape Drive supports capacity scaling of JA tape cartridges (native format 500 GB in Generation 2) to 100 GB. Capacity scaling allows the utilized length of tape to be logically shortened, allowing improved data access times in trade-off for reduced capacity. The tapes can subsequently be scaled back to full capacity as needed. Multiple scale settings are supported on the E05 drive including a 100 GB, 20% scaled JA cartridge.

Capacity scaling of tape medium is the action of modifying internal formatting indicators in the medium (and in the cartridge memory (CM) chip) so that the normal serpentine track format is altered in such a way as to limit the recorded portion of the tape to a specified fraction of the linear dimension of the medium. This action is normally accomplished by the drive responding to a command, in this case a Mode Select command. It pertains only to the currently loaded cartridge; it is not persistent.

The consequences of capacity scaling a tape to a percentage value, say 20% for example, is that the maximum number of recordable gigabytes are reduced to 20% of the normal value, and the average time to locate a random record on a full tape starting from load point is (very roughly) 20% of the time to locate a random record from load point for a full, unscaled tape. To compare: the average time to locate a random record on an unscaled (serpentine) tape that has only been filled to 20% capacity is nearly the same as the average time to locate a random record on an unscaled tape that has been filled to 100%. Scaling cuts the access time proportionately, and it also introduces normal end-of-tape programmatic warnings when approaching the scaled capacity limit in the same sense that those indicators are returned at end of tape when unscaled.

Channel calibration

The channel calibration feature is designed to allow for customization of each read/write data channel for optimum performance. The customization can enable compensation for variations in the recording channel transfer function, media characteristics, and read/write head characteristics. The 3592 models automatically perform recalibration in the field if it detects degraded performance.

11.1.4 3590 technology

The IBM TotalStorage 3590 Tape Drive (previously known as the IBM Magstar® 3590), was first introduced in July 1995. The original cartridge maintained the external form factor of the 3490 (Figure 11-6 on page 213), had a capacity of 10 GB uncompact (30 GB compressed), and the data rate was 9 MBps. Later drive models and new media increased these figures. See 11.3.3 “IBM TotalStorage 3590 Tape Drive” on page 219 for the latest specifications. The data format is not compatible with the 3490.

We include this information for historical reasons, since this technology is being replaced by improved 3592 technology.

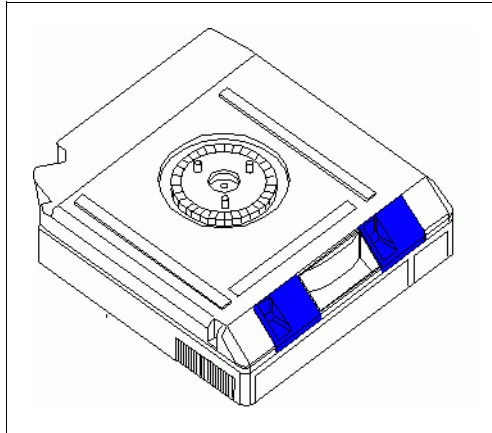


Figure 11-6 3590 tape cartridge

Magneto Resistive (MR) head

The IBM 3590 provides high capacity, performance, reliability, and a wide range of host connectivity. This technology exploits a fourth generation magneto resistive (MR) head, a 16 MB buffer, predictive failure analysis, and state-of-the-art electronic packaging.

Interleaved recording

While reading or writing 16 tracks at a time, the 3590 models use serpentine, interleaved, longitudinal recording technology for a total of four, eight, or twelve round trips from the physical beginning to the physical end of the tape and back again. The tape read/write head indexes, or moves vertically, when it completes each round trip so that the recorded tracks are interleaved across the width of the tape.

Figure 11-7 shows the recording element of the IBM TotalStorage 3590 Tape Drive. It also shows the way in which the read/write heads are moved over the width of the tape medium.

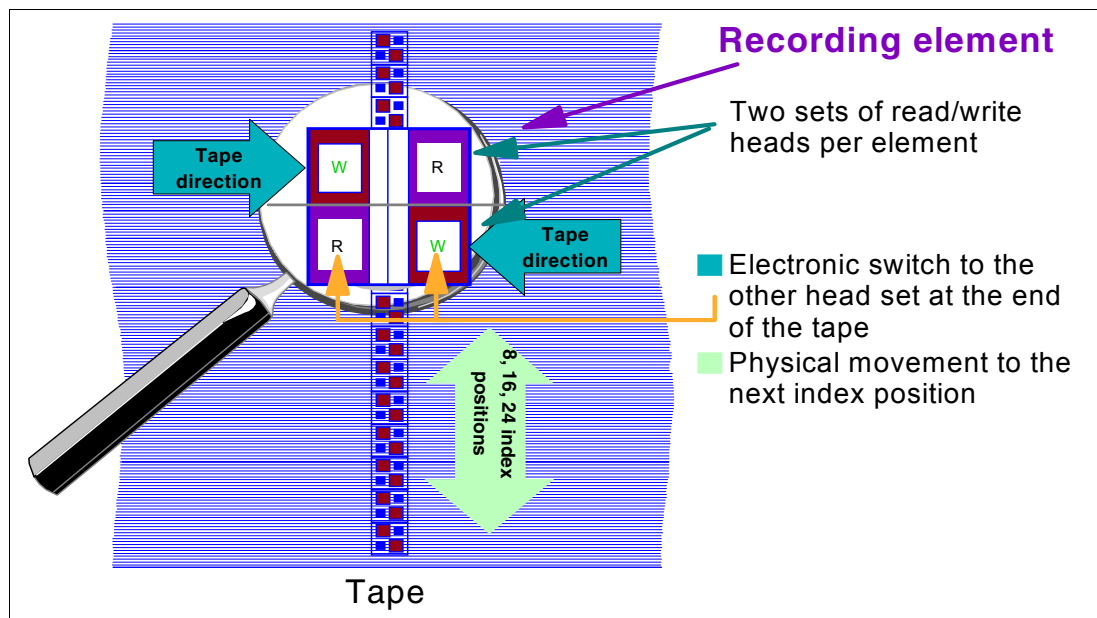


Figure 11-7 IBM 3590 recording

High-speed data access

The 3590 longitudinal serpentine recording technique allows for high performance read operations. The 3590 Model E tape drive makes eight round trip passes over the tape, where the Model H tape drive makes 12 passes. If a required block of data is, logically, one-eighth or one-twelfth of the way along the length of the tape, the read head need only to index upward by one position to access this data. No tape movement is required.

The IBM 3590 tape subsystem was the first to use a volume control region, inaccessible to user applications, to locate data on the tape. When a read request takes place, the tape drive uses a high-speed block search to position the tape directly at the required data block. It indexes the head to the correct set of tracks to allow for high-speed access to blocks or files. The tape drive performs this work while it is logically disconnected from the channel. This enables the channel to do other work.

The IBM 3590 uses IBMLZ1 hardware compression to give a greater compression ratio than the IDRC compression technique used in 3490 drives. This facility allows for access speeds that are significantly faster than previous generations of tape technology. The 3590 can search for data at up to 332 MBps.

Data integrity with the 3590

The IBM TotalStorage 3590 Tape Drive stripes data horizontally and vertically. If a media error occurs, even one that covers several tracks, the error correction code can reconstruct the data for the application. The 3590 E and H tape drives have enhanced error correction code for improved data reliability compared to the Model B drives.

The way that the 3590 tape drive writes data onto the tape allows it to achieve RAID-like tape storage:

- ▶ Multiple write elements are used for improved performance.
- ▶ Data is spread across multiple tracks to achieve improved availability.
- ▶ Recovery bits are written for improved error recovery.

Multi-system attached

The 3590 attaches via SCSI or Fibre Channel interface to a wide variety of platforms. The tape drives can also attach to (ESCON and FICON channels on S/390® or zSeries systems through an IBM TotalStorage Enterprise Tape Controller 3590 model A60.

PFA

Predictive failure analysis (PFA) microcode gives early warning of drive and media problems to maximize availability. The *Statistical Analysis and Reporting System* (SARS) analyzes and reports on tape drive and tape cartridge performance to help you determine whether the tape cartridge or the hardware in the tape drive is causing errors, determine if the tape media is degrading over time, and determine if the tape drive hardware is degrading over time.

Block ID

The drives automatically write a block ID after each block of data. This is used to locate data rapidly; in high-speed search or locate mode, the drive positions rapidly to a requested block. The drive is aware which group of 16 tracks holds the block and does not read every intervening block. Both models read and write 16 tracks of data in parallel. When they reach the end of a tape, the heads index and the tape reverses and a further set of 16 tracks is written in the opposite direction. This continues until all the tracks have been written. This form of recording is described as serpentine longitudinal. An additional benefit is that a full tape does not need to be rewound; it is already at load point.

Metal particle media

A chromium dioxide medium was used in the IBM 3480 and 3490 cartridges. The IBM 3590 cartridge uses a metal particle medium, which has a significantly increased coercivity and therefore permits a much higher data recording density in comparison with chromium dioxide media. The linear density is proportional to the medium's coercivity, and therefore the linear density of the IBM 3590 tape is approximately three times that of the IBM 3480 and 3490. The track density is also improved approximately fourfold. Advances in the metal particle coatings and media binders afford reliability and magnetic stability equal or superior to chrome media.

The 3590 tape drives use a metal particle medium in the tape cartridge that can store 10, 20, 30, 40, or 60 GB of uncompact data, depending on the cartridge type and the drive model. The integrated control unit uses a compaction algorithm that can increase the storage capacity of these cartridges. Assuming a compression ratio of three to one (3:1), the cartridge capacity increases to 60 GB on E models and to 90 GB on H models.

11.2 IBM LTO tape drives

The IBM LTO tape drive product family provides includes:

- ▶ IBM System Storage TS1030 Tape Drive
- ▶ IBM TotalStorage 3580 Tape Drive

The IBM System Storage TS1030 Tape Drive is a newer 4 Gbps Fibre Channel attached model (2 Gbps Fibre Channel LTO models are also available), while the IBM TotalStorage 3580 Tape Drive has SCSI attachment.

11.2.1 IBM System Storage TS1030 Tape Drive

The IBM System Storage TS1030 Tape Drive Model F3B (formerly the 3588 as shown in Figure 11-8) can be installed in an IBM System Storage TS3500 Tape Library (formerly machine type 3584) and IBM System Storage TS3310 Tape Library. It offers high capacity, performance, and technology designed for the midrange open systems environment. It is an IBM Ultrium 3 drive, with a native data transfer of up to 80 MBps, and native Ultrium 3 cartridge capacity of 400 GB (800 GB with 2:1 compression). The TS1030 can read and write LTO Ultrium 2 Data Cartridges and read LTO Ultrium 1 Data Cartridges. The TS1030 Model F3B has a 4 Gbps Fibre Channel interface for attachment to IBM System p, IBM System i, or IBM System z products, and System x, HP, Sun, UNIX, Linux, or Windows servers.



Figure 11-8 IBM System Storage TS1030 Tape Drive

The TS1030 Model F3B can be installed in new or installed IBM System Storage TS3500 Tape Library Models L53 and D53, IBM System Storage TS3310 Tape Library, and also in the 3584 older models L52, L32, D52, and D32.

For more information

See the Web site:

http://www-306.ibm.com/common/ssi/0IX.wss?DocURL=http://d03xhttpc1001g.boulder.ibm.com/common/ssi/rep_ca/0/897/ENUS106-340/index.html&InfoType=AN&InfoSubType=CA&InfoDesc=Announcement+Letters&panelurl=&paneltext=

11.2.2 IBM TotalStorage 3580 Tape Drive

The IBM TotalStorage 3580 Tape Drive (Figure 11-9) is an external, 3rd generation LTO drive. This is an external stand-alone or rack-mountable unit, similar to previous models of the 3580 and is the entry point for the family of IBM Ultrium tape products. The 3580 model L33 can read and write LTO Ultrium 2 and LTO Ultrium 3 cartridges and read LTO Ultrium 1 cartridges.

The IBM TotalStorage 3580 Tape Drive features a SCSI Ultra160 Low Voltage interface for connection to IBM System i, IBM System p, IBM System x, HP, Sun, UNIX, and Intel-PC servers.



Figure 11-9 The IBM TotalStorage 3580 Tape Drive

Note: There is another model available, 3580 model L3H, also known as the IBM TotalStorage 3580 Tape Drive Express. It has a new 19-inch rack mount shelf option which accommodates a single 3580 drive.

The L33 and L3H models use the Ultrium 3 LTO cartridge, which has a capacity of 400 GB (800 GB with 2:1 compression) with a sustained data rate of up to 80 MBps (uncompressed). Both models can read and write to Ultrium 2 LTO cartridges and read Ultrium 1 cartridges.

The IBM TotalStorage 3580 Tape Drive has an LCD display and indicators for write protection, tape in use, drive cleaning, and data compression. It is a very cost-effective solution for save-and-restore and archiving functions.

More information

For additional product details, refer to:

<http://www-.ibm.com/servers/storage/tape/3580/>

For additional product support, refer to;

<http://www.ibm.com/servers/storage/support/lto/3580/>

11.3 TotalStorage 359x tape drive

We discuss three types of 359x tape drives:

- ▶ IBM TotalStorage 3592 Tape Drive - first generation 3592-type drive

- ▶ IBM System Storage TS1120 Tape Drive - second generation 3592-type drive
- ▶ IBM TotalStorage 3590 Tape Drive - original Magstar-type drive

We also discuss the enterprise controllers:

- ▶ IBM TotalStorage 3592 Tape Controller Model J70
- ▶ IBM System Storage TS1120 Tape Controller Model C06

11.3.1 IBM TotalStorage 3592 Tape Drive (3592 J1A)

The IBM TotalStorage 3592 Tape Drive Model J1A provides high capacity and performance for storing mission critical data. By offering significant advancements in capacity and data transfer rates over the 3590 technology, the 3592 tape drive helps address storage requirements that are often filled by two types of drives - those that provide fast access for data access and those that provide high capacity for backups. The 3592 tape drive handles both types of use, helping simplify your tape infrastructure. Additionally, the 3592 tape drive offers WORM functionality that is designed to help support data retention needs and applications requiring an audit trail.

The 3592 drive also helps protect your existing investment in tape automation by offering compatibility with the IBM TotalStorage 3494 Tape Library, in which it can coexist with 3490 and 3590 tape drives, the IBM TotalStorage 3584 Tape Library, the IBM System Storage TS3500 Tape Library, the IBM TotalStorage 3494 Virtual Tape Server, and selected StorageTek™ Tape Libraries (ACS). Figure 11-10 shows the IBM TotalStorage 3592 Tape Drive Model J1A.

The 3592 drive can attach via 2 Gbps Fibre Channel, ESCON, and FICON (including 2 Gbps). Its maximum native and compressed data transfer rates are 40/120 MB per second respectively.



Figure 11-10 IBM 3592 J1A Tape Drive

The 3592 tape drive is significantly lighter and more compact than its predecessor, the IBM TotalStorage 3590 Tape Drive. The 3592 Tape Drive and canister weighs only 5.7 kg as compared to IBM 3590's 40 kg. Figure 11-11 on page 218 shows the two products side by side, where the 3592 is on top.



Figure 11-11 The IBM 3590 and IBM 3592 Tape Drive Model J1A

Media support

- ▶ Supports WORM cartridges of 60 GB and 300 GB native capacity
- ▶ Supports rewritable cartridges of 60 GB and 300 GB native capacity

The 3592 tape drive can initialize short length JJ cartridges to 60 GB and initialize (or reinitialize) standard JA length cartridges to either 60 GB (to support fast time to data) or 300 GB (to support high capacity).

More information

- ▶ For additional product details, refer to:
<http://www.ibm.com/servers/storage/tape/3592/>
- ▶ For additional product support, refer to:
<http://www.ibm.com/servers/storage/support/tape/3592/>

11.3.2 IBM System Storage TS1120 Tape Drive

The IBM System Storage TS1120 Tape Drive (machine type 3592-E05) is the second generation 3592-type drive. It is supported in IBM tape libraries, IBM frames that support stand-alone installation, and in an IBM TotalStorage 3592 Model C20 (3592 C20 frame) attached to a Sun StorageTek™ 9310 library.

The TS1120 tape drive uses the existing 3592 media, which is available in rewritable or WORM format to store 100 GB or 500 GB native capacity depending on cartridge type. The 3592 JA/JW media helps reduce resources to lower total costs, whereas the 3592 JJ/JR media is designed to support applications that require rapid access to data.

TS1120 tape drives can be shared among supported open system hosts on a SAN, or between FICON and ESCON mainframe hosts when attached to an IBM System Storage TS1120 Tape Controller Model C06 (TS1120 tape controller). Sharing drives optimizes drive utilization and helps reduce infrastructure requirements. Figure 11-12 on page 219 shows the TS1120 tape drive.

The TS1120 features dual-port 4 Gbps Fibre Channel interfaces.

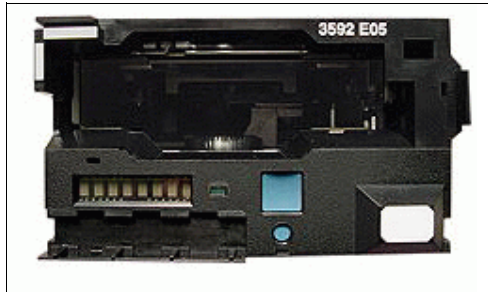


Figure 11-12 IBM TS1120 tape drive

Highlights

- ▶ Supports IBM Systems and selected open system platforms
- ▶ Supported on existing IBM and Sun StorageTek automation
- ▶ Offers native data transfer rate of up to 104 MBps - over 2.5 times as fast as the IBM TotalStorage 3592 Tape Drive and over seven times as fast as the 3590.
- ▶ Up to 500 GB native cartridge capacity, a 1.6 times increase over the IBM TotalStorage 3592 Tape Drive and a five times as much as the 3590 H Models
- ▶ Supports up to 1.5 TB (uncompressed in a System z9environment) on a 3592 JA/JW cartridge
- ▶ Supports fast access with a 3592 JJ/JR cartridge

Media support

The TS1120 tape drive can initialize short length JJ cartridges to 60 or 100 GB and initialize (or re-initialize) standard length JA cartridges to 60, 100, 300 or 500 GB to support fast access to data or to help address data growth and facilitate interchange. At typical compression ratios, the 3592 JA cartridge can provide usable capacity of up to 1 TB in an open system environment, and up to 1.5 TB in an IBM System z9™ environment when used with a TS1120 tape drive

More information

- ▶ For additional product details, refer to:
<http://www.ibm.com/servers/storage/tape/ts1120/>
- ▶ For additional product support, refer to:
<http://www.ibm.com/servers/storage/support/tape/ts1120/>

11.3.3 IBM TotalStorage 3590 Tape Drive

The IBM TotalStorage 3590 Tape Drive (Figure 11-13 on page 220) family includes four models (the E11, E1A, H11 and H1A). The H models with the Extended Length Cartridges made available in 2002, have a capacity and data rate of 60 GB (180 GB assuming 3:1 compression) and 14 MBps native respectively, while maintaining backward compatibility for reading with the base E models.

Note: IBM has announced it will withdraw all models of the IBM TotalStorage 3590 Tape Drive on September 29, 2006. It is replaced by the IBM TotalStorage 3592 Tape Drive Model J1A and IBM System Storage TS1120 Tape Drive.



Figure 11-13 IBM TotalStorage Enterprise Tape Drive 3590 in a rack

The IBM TotalStorage 3590 Tape Drive E11 and H11 is frame-mounted or rack-mounted with a 10-cartridge Automated Cartridge Facility (ACF) autoloader for unattended operation and up to 1.2 TB capacity for Model H11.

The IBM TotalStorage 3590 Tape Drive 3590 E1A and H1A is supported in an IBM TotalStorage 3494 Tape Library, IBM TotalStorage 3494 Virtual Tape Server, and IBM Enterprise Tape 3590 Silo Compatible Subsystem. This model does not have an ACF.

All 3590 model H tape drives

The 3590 H models may be ordered with either UltraSCSI or Fibre Channel interfaces. Each drive has two interfaces for availability reasons.

More information

- ▶ For additional product details, go to:
<http://www.ibm.com/servers/storage/tape/3590/>
- ▶ For additional product support, go to;
<http://www.ibm.com/servers/storage/support/tape/3590/>

11.3.4 IBM TotalStorage 3592 Tape Controller Model J70

The IBM TotalStorage 3592 Tape Controller Model J70 as seen in Figure 11-14 on page 221 is the follow-on to the original IBM TotalStorage Tape Controller 3590 Model A60 and is designed for environments with ESCON or FICON attached servers with IBM 3592 drives.

The IBM TotalStorage 3592 Tape Controller Model J70 (model 3592-J70) provides ESCON and FICON attachment of either 3592 J1A or 3590 tape drives in an IBM 3494 library, Silo Compatible Frame or standalone rack or frame.

The 3592 J70 (shown in Figure 11-14 on page 221) controller is built from “best-of-breed” components that include IBM eServer pSeries server technology, the AIX operating system and the PCI-X bus architecture. Redundant hot-swappable components are designed to failover automatically to provide high availability include power supplies and cooling fans. The

3592 J70 controller provides nondisruptive addition of 3592 J1A drives which helps enhance configuration flexibility and availability.

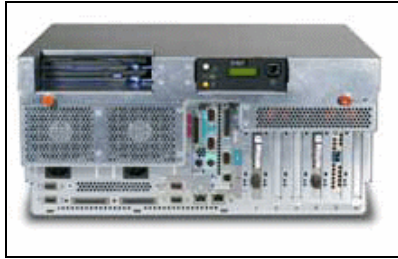


Figure 11-14 IBM 3592 model J70 tape controller

More information

- For additional product details, refer to:

http://www.ibm.com/servers/storage/tape/3592/datasheet_j70.html

11.3.5 IBM System Storage TS1120 Tape Controller Model C06

The IBM System Storage TS1120 Tape Controller Model C06 (model type 3592-C06) provides performance and reliability for IBM System z customers. The IBM System Storage TS1120 Tape Controller Model C06, has up to four 4 Gbps FICON attachments. The TS1120 Tape Controller also has up to eight ESCON attachments, or an intermix of ESCON and FICON attachments. Up to 16 IBM System Storage TS1120 Tape Drives or IBM TotalStorage 3592 Tape Drives can be attached to a single TS1120 Tape Controller. The controller can be installed in an IBM System Storage 3952 Tape Frame Model F05 , an IBM 3953 Tape Frame Model F05, or in a stand-alone rack, supporting 3592 Tape Drives installed in IBM 3494 frames, IBM 3584 frames, IBM 3592 Model C20 frames, and stand-alone racks. TS1120 tape drives can be shared among supported open system hosts on a Storage Area Network (SAN), or between FICON and ESCON mainframe hosts when attached to an IBM System Storage TS1120 Tape Controller (TS1120 tape controller). Sharing drives optimizes drive utilization and helps reduce infrastructure requirements.

The TS1120 Tape Controller provides up to 1.7 times the throughput of the IBM TotalStorage 3592 Tape Controller Model J70, with 4 Gbps FICON attachment using the 3592 Model J1A Tape Drive

Figure 11-15 shows the IBM TS1120 Tape Controller.



Figure 11-15 IBM TS1120 Tape Controller

11.4 More information

- ▶ *IBM TotalStorage Enterprise Tape: A Practical Guide, SG24-4632*
- ▶ *Implementing IBM Tape in UNIX Systems, SG24-6502*
- ▶ *Implementing IBM Tape in Linux and Windows, SG24-6268*
- ▶ *IBM System Storage Tape Library Guide for Open Systems, SG24-5946*

11.5 Tape media

There is separate tape media for each tape drive type - LTO, 3592, and 3590. Each drive type only supports its own media - e.g. you cannot use an LTO cartridge in a 3592 tape drive.

11.5.1 LTO cartridge

The Ultrium cartridge is a single-reel cartridge. This means that the whole tape is wrapped around a single reel when the cartridge is not loaded in a drive. During the loading process, the threader of the drive catches the leader pin of the tape and threads it through the drive and the machine reel. During the read/write process the tape is stored on the machine reel and the cartridge.

Two views of the tape cartridge are shown in Figure 11-16 and Figure 11-17 on page 223.

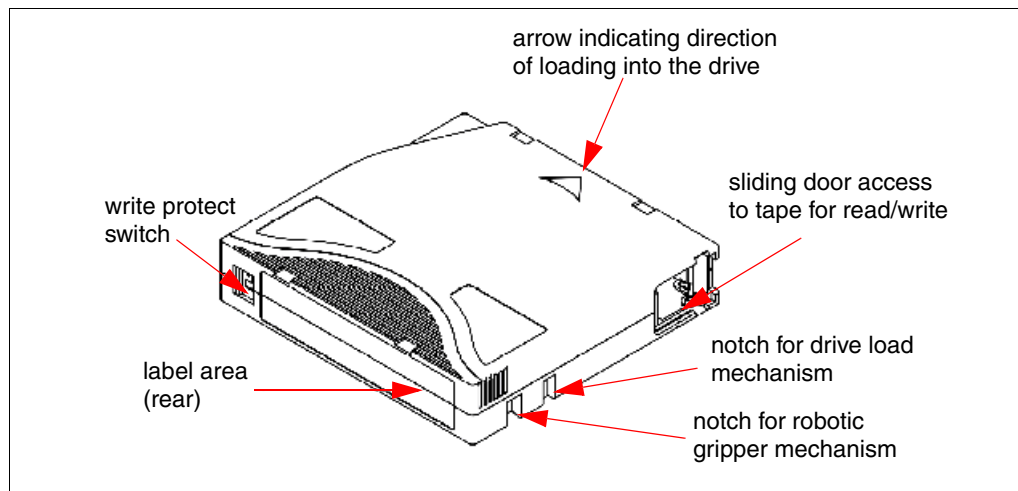


Figure 11-16 Ultrium cartridge view from top and rear

The cartridge is approximately 10.2 cm long, 10.5 cm wide, and 2.2 cm high (approximately 4 x 4.16 x 0.87 inches). The cartridge contains 1/2-in (12.6 mm), metal-particle tape with a high density recording area. With Ultrium 3 there is both a regular rewritable tape, and a write-once-read-many (WORM) tape. To achieve the 400 GB native capacity, this tape is slightly longer than previous LTO media, at 680 m (2231 feet).

Figure 11-17 on page 223 shows some of the components of the cartridge.

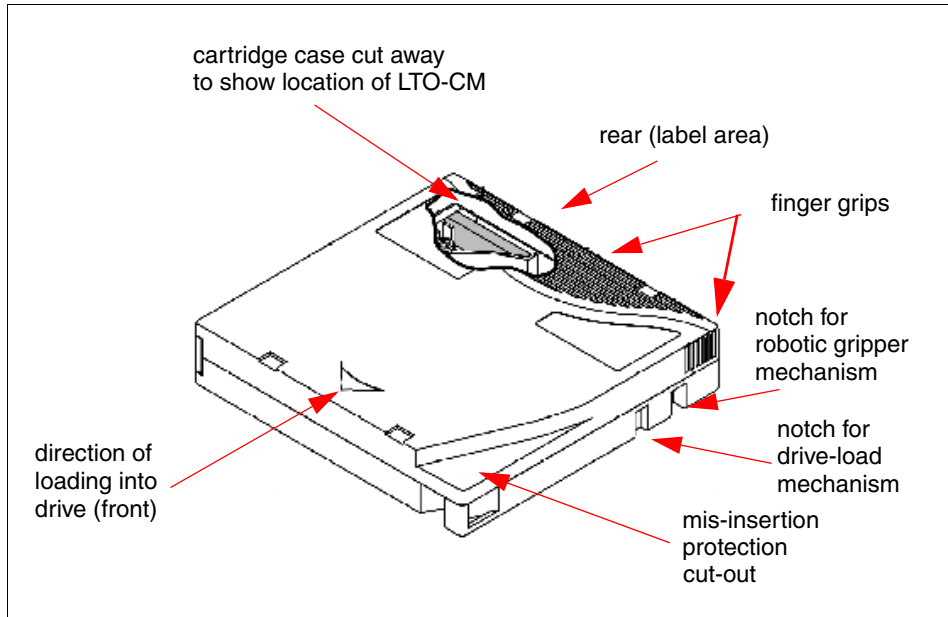


Figure 11-17 Ultrium cartridge view from top and front

11.5.2 IBM 3592 cartridge

3592 cartridges have both a rewritable and WORM version.

Rewritable cartridge

Rewritable 3592 cartridges are available in short length (JJ) or standard (JA) length. The short length is intended to provide a more economically priced cartridge for storing smaller quantities of data.

The IBM TotalStorage 3592 Tape Drive can initialize JJ cartridges to 60 GB and initialize (or re-initialize) JA cartridges to either 60 GB (to support fast time to data) or 300 GB (to support high capacity).

The IBM System Storage TS1120 Tape Drive can initialize short length JJ cartridges to 60 or 100 GB and initialize (or re-initialize) standard length JA cartridges to 60, 100, 300 or 500 GB to support fast access to data or to help address data growth and facilitate interchange. At typical compression ratios, the 3592 JA cartridge can provide usable capacity of up to 1 TB in an open system environment, and up to 1.5 TB in an IBM System z9™ environment when used with a TS1120 tape drive. T

WORM cartridge

The IBM TotalStorage 3592 Tape Drive and IBM System Storage TS1120 Tape Drive support WORM JR and JW cartridges to store data in a non-erasable, non-rewritable format. This is intended to help support the long term retention of reference data and meet the requirements of regulatory bodies worldwide. The short length JR and standard length JW cartridges have advanced security features that are designed to prevent the alteration or deletion of stored data while allowing data to be appended to existing cartridges or files.

Note: 3592 tape cartridges are not compatible with 3590 tape drives and, likewise, 3590 tapes cannot be used in the 3592 drives. However 3592 cartridges can coexist in the same library with LTO and 3590 cartridges, where both drive types are supported.

Advanced media

The IBM 3592 tape cartridge contains an advanced fourth-generation metal particle formulation in a dual layer coating on a half-inch-wide tape. The IBM tape uses an advanced magnetic coating and process designed to provide a high output and signal quality to support the current TS1120 and 3592 J1A tape drives. The tape features an ultra-smooth and uniform magnetic layer less than 0.2 microns thick and a specially refined coating formulation designed to help improve media reliability and performance as well as minimize wear of the tape heads and components. A precision timing-based servo with enhanced features helps enable high track densities, high data rates and data access performance as well as high reliability and stop-start performance. In addition, modifications to the cartridge design and construction are designed to help improve pin retention, hub and clutch engagement, spool alignment and tape stacking within the cartridge. These enhancements are designed to help improve reliability and durability of not only the media, but of the tape drive as well.

Cartridge memory

A cartridge memory chip that stores access history and media performance information is built into every cartridge. Records are written to the memory chip every time the cartridge is unloaded from a TS1120 or 3592 J1A tape drive. These records can be used by the IBM Statistical Analysis and Reporting System (SARS) programs to analyze and report on tape drive and cartridge usage and help diagnose and isolate tape errors. SARS can also be used to proactively determine if the tape media or tape drives are degrading over time.

11.5.3 IBM 3590 cartridge

The 3590 cartridge use metal-particle half-inch tape in an enclosed cartridge shell.

The Extended High Performance Cartridge Tape (EHPCT) providing 20/40/60 GB before compression. It is identifiable by the green tabs on the cartridge case.

The highest of the three capacities is achieved in 384-track mode on a 3590H drive in an uncompressed format. The 3590 drives may require the Extended Media Support feature to be able to use EHPCT cartridges. Check with your IBM customer representative for verification as to drive capabilities. All 3590 model H drives are shipped with this support as standard.



IBM Tape Libraries

This chapter discusses IBM Enterprise Tape Products, which provide the highest levels of performance and reliability of any IBM tape subsystem.

This chapter covers the following products:

► **IBM System Storage TS3100 Tape Library**

- IBM System Storage TS3100 Tape Library Express Model L2U (TS3100 tape library; 3573 L2U)
- IBM System Storage TS3100 Tape Library Express Model L3S (TS3100 tape library; 3573 L3S)
- IBM System Storage TS3100 Tape Library Express Model F3S (TS3100 tape library; 3573 F3S)

► **IBM System Storage TS3200 Tape Library**

- IBM System Storage TS3200 Tape Library Express Model L4U (TS3200 tape library; 3573 L4U)
- IBM System Storage TS3200 Tape Library Express Model L3H (TS3200 tape library; 3573 L2H)
- IBM System Storage TS3200 Tape Library Express Model F3H (TS3200 tape library; 3573 F3S)

► **IBM System Storage TS3310 Tape Library (TS3310 tape library)**

- IBM System Storage TS3310 Tape Library Module Model E5B (TS3310 tape library; 3576 L5B)
- IBM System Storage TS3310 Tape Library Module Model E9U (TS3310 module; 3576 E9U)

► **IBM System Storage TS3500 Tape Library** - larger modular library with the potential to house a maximum of 192 tape drives in as many as 16 frames. Formerly known as IBM TotalStorage Tape Library 3584

- IBM System Storage TS3500 Tape Library Frame Model L23 (3584 L23; L23)
- IBM System Storage TS3500 Tape Library Frame Model L53 (3584 L53; L53)
- IBM System Storage TS3500 Tape Library Frame Model D23 (3584 D23; D22)

- IBM System Storage TS3500 Tape Library Frame Model D53 (3584 D53; D52)
- IBM System Storage TS3500 High Availability Option Model HA1 (3584 HA1; HA1)
- ▶ **IBM TotalStorage 3953 Tape System for zSeries**
 - IBM TotalStorage 3953 Tape System Model F05 (3953 F05)
 - IBM TotalStorage 3953 Library Manager Model L05 (3953 L05)
- ▶ **IBM TotalStorage 3494 Tape Library**
 - IBM TotalStorage 3494 Tape Library Frame Model L12 (3494 L12; L12)
 - IBM TotalStorage 3494 Tape Library Frame Model L14 (3494 L14; L14)
 - IBM TotalStorage 3494 Tape Library Frame Model L22 (3494 L22; L22)
 - IBM TotalStorage 3494 Tape Library Frame Model D12 (3494 D12; D12)
 - IBM TotalStorage 3494 Tape Library Frame Model D14 (3494 D14; D14)
 - IBM TotalStorage 3494 Tape Library Frame Model D22 (3494 D22; D22)
 - IBM TotalStorage 3494 Tape Library Frame Model D24 (3494 D24; D24)
 - IBM TotalStorage 3494 High Availability Frame Model HA1 (3494 HA1; HA1)
 - IBM TotalStorage 3494 Tape Library Frame Model S10 (3494 S10; S10)
- ▶ **Enterprise Tape System 359x family**
 - Enterprise Tape Controller 3590 Model A60 (3590 A60)
 - Silo Compatible Frame 3590 Model C12 (3590 C12)
 - Silo Compatible Frame 3590 Model C10 (3590 C10)
- ▶ **Other models**
 - **IBM TotalStorage 3581 Tape Autoloader:** A single-drive unit, but it has a seven-cartridge autoloader within the device. It is a stand-alone desktop unit; however, optional additional hardware is available for installation in a standard 19-inch rack.
 - **IBM TotalStorage 3582 Tape Library:** A small robotic library, accommodating one to two drives and providing space for up to 23 Ultrium cartridges. The IBM 3582 may be stand-alone or, with an optional feature, housed in a rack.
 - **IBM TotalStorage 3583 Tape Library:** A small robotic library, accommodating from one to six drives and providing space for up to 72 Ultrium cartridges. The IBM 3583 may be standalone or, with an optional feature, housed in a rack.

12.1 Overview

IBM System Storage provides a complete spectrum of tape libraries highlighting high performance and capacity for entry, midrange and enterprise system scenery. The IBM libraries handle backups, save and restore, and archival data storage needs.

IBM System Storage TS3100 Tape Library

The IBM System Storage TS3100 Tape Library is well-suited for handling backup, save and restore, and archival data-storage needs for small to medium-size environments. It has a single Ultrium 3 tape drive and 22 tape-cartridge capacity.

IBM System Storage TS3200 Tape Library

The IBM System Storage TS3200 Tape Library is designed for backup, save and restore, and archival data-storage needs for small to medium-size environments. The TS3200 is an

external 4U standalone or rack-mountable unit that incorporates up to two Ultrium 3 tape drives and 44 tape cartridges, as well as 3 media mail slots and 1 dedicated cleaning cartridge slot.

IBM System Storage TS3310 Tape Library

IBM System Storage TS3310 Tape Library offers simple, rapid expansion as processing needs grow. Its entry level configuration is a single 5 EIA rack unit high library. Over time, as the need for tape backup expands, you can add an additional 9U expansion module, with space for additional cartridges, tape drives and a redundant power supply.

IBM System Storage TS3500 Tape Library

The IBM System Storage TS3500 Tape Library combines IBM automation and drive technology to provide a highly scalable, automated tape library for System z and open systems backup and archive in midrange to enterprise environments.

IBM TotalStorage 3494 Tape Library

The IBM TotalStorage 3494 Tape Library is an excellent solution for today's large storage requirements. Modular, flexible, and reliable, the 3494 uses 3590 tape drives, is cost-effective, and is backed by the service expertise of the IBM service organization. The 3494 offers a variety of models and features to fit all your needs and will grow as your business grows.

12.2 IBM System Storage TS3100 Tape Library Express Model

IBM System Storage TS3100 Tape Library provides compact, high-capacity, low-cost solutions for simple, unattended data backup. The TS3100 houses up to 22 tape cartridges and a 1-slot I/O Station in a compact 2U form factor with easy access to tape cartridges via two removable magazines. The library comes with one Ultrium 3 LVD SCSI or Ultrium 3 4 Gb Fibre Channel drive, to connect to a wide variety of open systems servers. TS3100 media capacity is up to 8.8 TB (17.6 TB with 2:1 compression) data storage per unit. Figure 12-1 shows the IBM System Storage TS3100 Tape Library.



Figure 12-1 IBM System Storage TS3100 Tape Library

The IBM System Storage TS3100 Tape Library is available in 3 models:

- ▶ IBM System Storage TS3100 Tape Library Express Model L3S (3573-L3S)
Comes with an Ultrium 3 LVD SCSI drive
- ▶ IBM System Storage TS3100 Tape Library Express Model F3S (3573-F3S)
Comes with an Ultrium 3 4 Gbps Fibre drive
- ▶ IBM System Storage TS3100 Tape Library Express Model L2U (3573-L2U)
Can be ordered with either an Ultrium 3 LVD SCSI drive (feature code 8043) or an Ultrium 3 4 Gbps Fibre drive (feature code 8044)

12.2.1 Features

These are the features for the IBM System Storage TS3100 Tape Library Express Model:

- ▶ IBM LTO Ultrium 3 tape drive.
- ▶ Supports 4 Gbps Fibre Channel and LVD SCSI attachment
- ▶ 2U form factor with 22 data-cartridge slots. The cartridges are housed in two removable magazines, left and right.
- ▶ A single mail slot on the left cartridge magazine provides import/export facilities for cartridges, without interrupting library operation
- ▶ Sequential or random access mode with a standard bar-code reader
- ▶ Native data rate of 80 MBps
- ▶ Ultrium 3 native data physical capacities up to 8.8 TB (up to 17.6 TB using 2:1 compression)
- ▶ Operator front control panel contains power button, front panel LEDs, control keys, and the operator control panel display.
- ▶ Remote library management through a Web interface

12.2.2 Specifications

Table 12-1 summarizes the capabilities of the IBM System Storage TS3100 Tape Library.

Table 12-1 IBM System Storage TS3100 Tape Library Express specifications

Model	3573 Model L2U, L3S, F3S
Drive options	Ultrium 3: LVD SCSI—FC (8043) 4 Gbps Fibre Channel—FC(8044)
Optional feature codes	Rack mount 7002 Additional right-side magazine set 8106
Tape drive type	IBM LTO Ultrium 3
Number of drives	1
Number of tape cartridges	22
Number of mail slots	1
Physical capacity	Up to 800 GB per cartridge compressed; 400 GB native Up to 17.6 TB per tape library compressed; 8.8 TB native
Data transfer rate	Up to 80 MBps native with LTO Ultrium 3
Attachment and systems support	LVD Ultra160 and 4 Gbps Fibre Channel interfaces, attaching to IBM System p (AIX and Linux), IBM System x (Windows 2000/2003 and Linux), System i (i5/OS and OS/400), System z (Linux and Fibre Channel only), HP, SUN, and other Intel servers

For the most up to date and detailed operating system and attachment requirements, see:

http://www.ibm.com/servers/storage/tape/compatibility/pdf/ts3100_ts3200_interop.pdf

12.3 IBM System Storage TS3200 Tape Library

IBM System Storage TS3200 Tape Library is a dual or single drive tape library, offering high capacity and performance technology for the midrange environments. The TS3200 is an external 4U standalone or rack-mountable unit that incorporates up to two LTO 3 tape drives. Its native data rate is up to 80 Mbps per drive. The TS3200 can be ordered with up to two LVD SCSI or two 4 GB Fibre Channel LTO 3 drives, for connection to a wide variety of open systems servers. The TS3200 has four removable cartridge magazines, providing 44 data cartridge slots, one three port I/O station, and one dedicated cleaning cartridge slot. Remote management and a bar code reader are standard in the library, and it can run in sequential or random access mode. Optional features available are rack mount kit, additional power supply, and Path Failover.

The TS3200 has IBM patented Multi-Path Architecture for sharing the library robotics. This allows a library with two drives to be partitioned into two logical libraries, for sharing between servers and/or applications.

Figure 12-2 shows the IBM System Storage TS3200 Tape Library.



Figure 12-2 IBM System Storage TS3200 Tape Library

The IBM System Storage TS3200 Tape Library is available in 3 models:

- ▶ IBM System Storage TS3200 Tape Library Express Model L3H (3573-L3H)
One Ultrium 3 LVD SCSI drive. Can order one additional SCSI drive, making two in all.
- ▶ IBM System Storage TS3200 Tape Library Express Model F3H (3573-F3H)
One Ultrium 3 4 Gb FC drive. Can order one additional FC drive, making two in all.
- ▶ IBM System Storage TS3200 Tape Library Express Model L4U (3573-L2U)
Can be ordered with one or two Ultrium 3 LVD SCSI drives (feature code 8043) or Ultrium 3 4 Gbps FC drives (feature code 8044). Drive types can be mixed on this model, that is, it can be ordered with two SCSI drives, two FC drives, or one of each.

12.3.1 Features

These are the features for the IBM System Storage TS3200 Tape Library:

- ▶ IBM LTO Ultrium 3 tape drive, for high capacity and performance.
- ▶ 4 Gbps Fibre Channel and LVD SCSI attachment
- ▶ 4U form factor with 44 data-cartridge
- ▶ Native data physical capacities up to 17.6 TB (35.2 TB with 2:1 compression)
- ▶ Sequential or random access mode with a standard bar-code reader

- ▶ Native drive data rate of 80 MBps
- ▶ The front panel integrates the power button, front panel LEDs, control keys, and the operator control Panel display, for direct library management.
- ▶ Remote library management through a Web interface
- ▶ Figure 12-3 shows the internal view of TS3200.

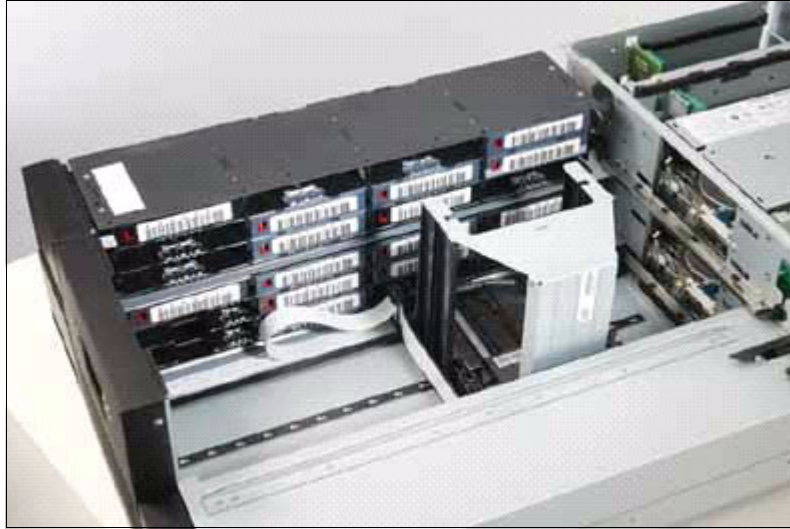


Figure 12-3 TS3200 internal view

Figure 12-4 shows the I/O station which is in the left magazine.



Figure 12-4 TS3200 Library station I/O in the left magazine

12.3.2 Specifications

Table 12-2 shows the specifications of the TS3200.

Table 12-2 IBM System Storage TS3200 Tape Library specifications

Model	Type 3573 Models L4U, L3H, F3H
Drive options	Ultrium 3: LVD SCSI—FC 8043, part number 23R7260 4 Gbps Fibre Channel—FC 8044, part number 23R7261
Optional feature codes	Rack mount 7002 Right-side magazine set 8106 Path Failover 1682 Additional Power Supply 1901

Tape drive type	IBM LTO Ultrium 3
Number of drives	1-2
Number of tape cartridges	22
Number of mail slots	44
Number of mail slots	3
Physical capacity	Up to 800 GB per cartridge compressed; 400 GB native Up to 17.6 TB per tape library compressed; 8.8 TB native
Data transfer rate	Up to 80 MBps native with LTO Ultrium 3
Attachment and systems support	LVD Ultra160 and 4-Gbps Fibre Channel interfaces, attaching to IBM System p (AIX and Linux), IBM System x (Windows 2000/2003 and Linux), System i (i5/OS and OS/400), System z (Linux and Fibre Channel only), HP, SUN, and other Intel servers

For the most up to date and detailed operating system and attachment requirements, see:

http://www.ibm.com/servers/storage/tape/compatibility/pdf/ts3100_ts3200_interop.pdf

12.4 IBM System Storage TS3310 Tape Library

The IBM System Storage TS3310 Tape Library delivers high performance, capacity, and technology designed for reliability and the heavy demands of tape storage. The TS3310 is highly modular and scalable - with a base 5U control module library and additional 9U expansion units. The TS3310 uses high-performance LTO Ultrium 3 tape drives (either LVD Ultra 160 SCSI or 4 Gbps FC), with 80 MBps native data transfer rate.

The IBM System Storage TS3310 Tape Library (3576-L5B) basic configuration contains the library control module, fixed tape cartridge storage of 30 slots, I/O station of 6 slots, a touch screen display, cartridge handling robotics, and up to two LTO Ultrium 3 tape drives. It also includes a license key which enables one half of the total storage slots of a Model E9U Expansion Module for future expansion of the TS3310.

The TS3310 Expansion module (3576-E9U) use 9U of space. Each 9U expansion module can accommodate up to 4 LTO Ultrium 3 tape drives, up to 80 tape cartridge slots, and 12 configurable I/O station slots (the configurable slots can be either I/O slots or storage slots). Up to two 9U expansion modules can be attached, giving a total maximum library capacity of 85 TB native capacity in 212 slots, and a maximum of 10 LTO drives.

Figure 12-5 on page 232 shows the IBM TS3310 base L5B unit with one expansion module.



Figure 12-5 IBM System Storage TS3310 Tape Library Tape Library - Base and Expansion

12.4.1 Features

These are the features for the IBM System Storage TS3310 Tape Library:

- ▶ TS3310 is a modular library, 12 TB up to 85 TB native capacity (30 to 212 LTO storage slots)
 - From 1 to 10 LTO Gen 3 LVD SCSI or 4 Gbps hot-swappable tape drives
 - Supports Logical Partitioning using Multipath Architecture - up to 10 logical libraries can be defined for attachment to different servers, operating systems, and applications.
 - Optional features; Capacity on demand, Rack Mounting, Path Failover (Data and Control), Redundant power
- ▶ TS3310 Base library module supports, 12 TB native capacity (30 slots), up to two LTO 3 drives (FC/SCSI), 6 slot cartridge I/O station, robotics and control logic, barcode reader and remote management capability are standard, native SMI-S support, and license key to support initial E9U expansion. Optional base library features are:
 - Path failover
 - Redundant power supply
- ▶ Figure 12-6 on page 233 shows the TS3310 Base Library front and rear view. In the rear view are two LTO drives, and one standard power supply installed. A second power supply.



Figure 12-6 TS3310 Base unit rear and back view

- ▶ One or two expansion modules can be installed. Each Expansion Module supports:
 - Up to 37 TB native capacity (92 storage slots)
 - Up to 4 LTO 3 drives (FC/SCSI)
 - Option of zero or 12 cartridge I/O slots
 - Optional redundant power supply
- ▶ Combined maximum configuration (model L5B + E9U +E9U) supports
 - 85 TB native capacity (212 storage slots*)
 - Up to 10 LTO 3 drives (FC/SCSI)
 - 6, 12, 18, 24 or 30 cartridge I/O slots
- ▶ Capacity On Demand: The Capacity Expansion feature (#1640), ordered against the TS3310 Model L5B base library, enables the unused storage slots within an expansion library via a firmware license key. Capacity Expansion license keys enable the remaining one half of the storage slots in the E9U Expansion Library and enable one half the storage slots in the next Model E9U Expansion Library. Without this feature, only half the slots in the expansion module are available. This feature helps to reduce the initial hardware cost.

Note: For latest information about the TS3310, see:

<http://www.ibm.com/servers/storage/tape/ts3310/index.html>

12.4.2 Specifications

The specifications for IBM System Storage TS3310 Tape Library are listed in Table 12-3 on page 234.

Table 12-3 TS3310 specifications

Type and Model	3576 Model L5B	3576 Model L5B and 1 E9U	3576 Model L5B and 2 E9U
Configuration	Base library	Base library and 1 expansion unit	Base library and 2 expansion units
LTO storage slots (max)	30	122	212
LTO Input/Output slots (min/max)	0/6	0/18	0/30
Maximum tape drives	2	6	10
Total Physical Capacity (fixed + I/O slots) - uncompressed	12 TB (with 30 storage slots)	49 TB (zero I/O slots)	85 TB (zero I/O slots)
Capacity on Demand increments	n/a	46 cartridges	46 cartridges
Maximum logical libraries	2	6	10
Hot-swap components	Power supplies, tape drives SMI-S enabled		
Systems management			
Operating systems/servers supported	Attaches to IBM System p (AIX and Linux), IBM System x (Windows 2000/2003 and Linux), System i (i5/OS and OS/400), System z (Linux and Fibre Channel only), HP, SUN, and other Intel servers		

For the complete detailed operating system and server attachment support, see:

http://www.ibm.com/servers/storage/tape/compatibility/pdf/ts3310_interop.pdf

12.5 IBM System Storage TS3500 Tape Library

The IBM System Storage TS3500 Tape Library is a highly scalable, automated tape library for mainframe and open systems backup and archive in midrange to enterprise environments. It is formerly known as IBM TotalStorage Tape Library 3584.

The IBM System Storage TS3500 Tape Library supports System z via the IBM TotalStorage 3953 Tape System. The 3953 Tape System enables System z hosts to access the TS3500 cartridge inventory and allows connection to 3592 J1A and TS1120 tape drives. More information about these drives is in 11.3.1, “IBM TotalStorage 3592 Tape Drive (3592 J1A)” on page 217 and 11.3.2, “IBM System Storage TS1120 Tape Drive” on page 218.

The TS3500, when attached to open systems, uses the IBM System Storage TS1030 Tape Drive, (see 11.2.1, “IBM System Storage TS1030 Tape Drive” on page 215).

The TS3500 can form a VTS configuration - see 13.2, “IBM TotalStorage 3494 Virtual Tape Server” on page 269.

The TS3500 Model HA1 allows two robotic accessors to operate simultaneously in two to 16 frame configurations.

Figure 12-7 on page 235 shows a maximum 16 frame TS3500 configuration.



Figure 12-7 IBM System Storage TS3500 Tape Library

The configuration highlights are:

- ▶ 1 - 16 frames (plus two service frames)
- ▶ 1 - 192 tape drives
 - Up to 192 System z attached TS1120 tape drives
 - Up to 192 open system attached tape drives
 - TS1030 LTO or TS1120/3592 J1A tape drives are supported
 - Tape drives types and associated media require unique frames
- ▶ Up to 9.39 TB maximum compression
 - Up to 2.755 TB native/5.51 TB compressed capacity with IBM Ultrium 3 cartridges
 - Up to 3.130 TB native/9.39 TB compressed capacity with 3592 cartridges and 3:1 compression
- ▶ Designed for concurrent maintenance
 - Hot-swappable drives
 - Hot-swappable drive and accessor power supplies
 - Redundant drive and accessor power
 - Redundant AC power
- ▶ Dual Ethernet interface for TSSC (TS3500 System Console) connection
 - High Availability Library Management / Robotics
 - Adds left and right service frames to the library
 - Adds second active accessor and library management node
 - Each accessor operates independently in flexible zones
 - Support dynamic workload balancing across zones
- ▶ IBM System z and z/OS options
 - Supports redundant ESCON/FICON attachment to supported servers
 - Redundant 3953 F05 frame power and library control path interfaces
 - Redundant 3953 F05 VTS switches features (optional for TS1120 controller)
 - Optional redundant 3953 Library Manager

The TS3500 can attach to IBM System p (AIX and Linux), IBM System x (Windows 2000/2003 and Linux), System i (i5/OS and OS/400), System z (Linux and Fibre Channel only), HP, SUN, and other Intel servers, NAS Gateway 500.

For the latest detailed requirements, see:

http://www.ibm.com/servers/storage/tape/compatibility/pdf/ts3500_interop.pdf

12.5.1 Structure of the IBM System Storage TS3500 Tape Library

This section defines the base and expansion frames, and describes the models of the IBM System Storage TS3500 Tape Library.

The basic IBM System Storage TS3500 Tape Library is a single storage unit known as the base frame (Model L23 for 3592 and TS1120 tape drives, or Models L53 for TS1020 drives). The library's capacity is increased by adding up to fifteen additional storage units, called expansion frames. The frames join end to end, with the base frame on the left and the expansion frame on the right. The additional expansion frames are supported by a common cartridge accessor that requires no pass-through mechanism. Each frame in the library may contain up to twelve drives.

For information about drives that are compatible in frames, drives that are compatible in a logical library, and the compatibility of drives and cartridges, see the *IBM System Storage TS3500 Tape Library Introduction and Planning Guide*, GA32-0559.

The IBM System Storage TS3500 Tape Library features an optional second cartridge accessor. If you order dual accessors, two service bay frames are required. Service bay A is known as Model HA1 and service bay B is a Model D23 or Model D53 frame.

TS3500 L23 / L53 frames common features

- ▶ Operator control panel and access door
- ▶ Cartridge accessor with two grippers
- ▶ Enhanced Frame Controller Assembly
- ▶ Redundant hot-swap power supplies
- ▶ Second Ethernet port for TS3000 System Console (TSSC)
- ▶ Dual AC power line cords
- ▶ Remote Support Facility (call home)
- ▶ 16 / 32 cartridge I/O station

TS3500 L23 frame

- ▶ Supports up to 12 TS1120 or 3592 J1A drives, System z or open system attached
- ▶ Stores 237 - 260 cartridges with 16 cartridge I/O and 0 - 12 drives
- ▶ Stores 199 - 222 cartridges with 32 cartridge I/O and 0 - 12 drives

TS3500 L53 Frame

- ▶ Supports up to 12 TS1030 drives
- ▶ Open system attach only
- ▶ Stores 261 - 287 cartridges with 16 cartridge I/O and 0 - 12 drives
- ▶ Stores 219 - 245 cartridges with 32 cartridge I/O and 0 - 12 drives

Figure 12-8 on page 237 shows the internal parts of TS3500 Frame L23/L53.

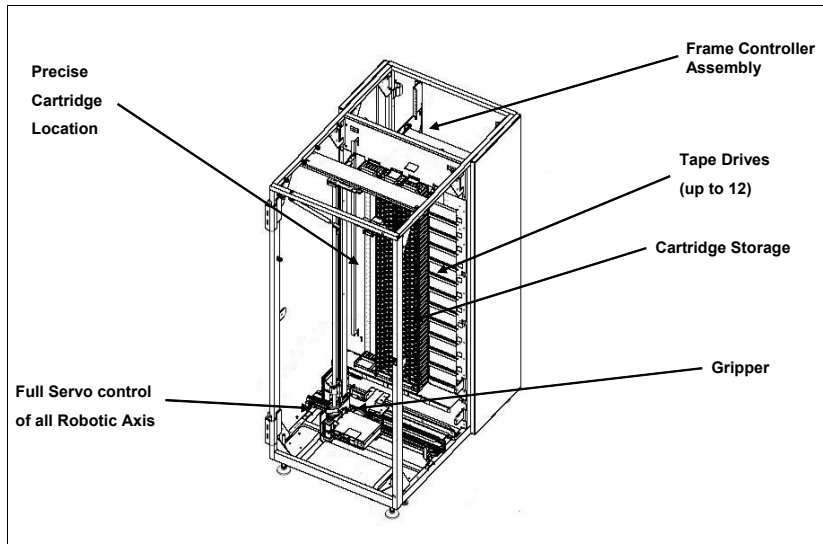


Figure 12-8 TS3500 Frame L23/53

TS3500 D23 / D53 Frames common features

- ▶ Support for up to 12 tape drives
- ▶ Enhanced Frame Controller Assembly (optional)
- ▶ Redundant power
- ▶ Dual AC power line cords

TS3500 D23 frame

- ▶ Supports System z and open system attached TS1120 and 3592 J1A tape drives
- ▶ Stores 360 - 400 cartridges with 0 – 12 drives

TS3500 D53 frame

- ▶ Supports open system attached LTO Tape drives
- ▶ Stores 396 - 440 LTO cartridges with 0 - 12 drives

Figure 12-9 on page 238 shows a Frame D23/D53 without the accessor.

Model HA1

The Model HA1 provides for the installation and operation of a second library accessor in a TS3500 to help improve availability and performance. It is required when the optional second accessor is ordered. The Model HA1 itself provides only a frame, that is always configured as Service Bay A for the original accessor that came with the base frame. It does not contain any tape drives or storage cells for customer data cartridges. An additional Model Dxx Frame is required as a Service Bay B for the second accessor.

The Model HA1 is approximately 307 mm (12 in.) shorter in depth than Models Lxx and Dxx. More information is in 12.5.2, “Dual accessors and service bays in the TS3500” on page 241.

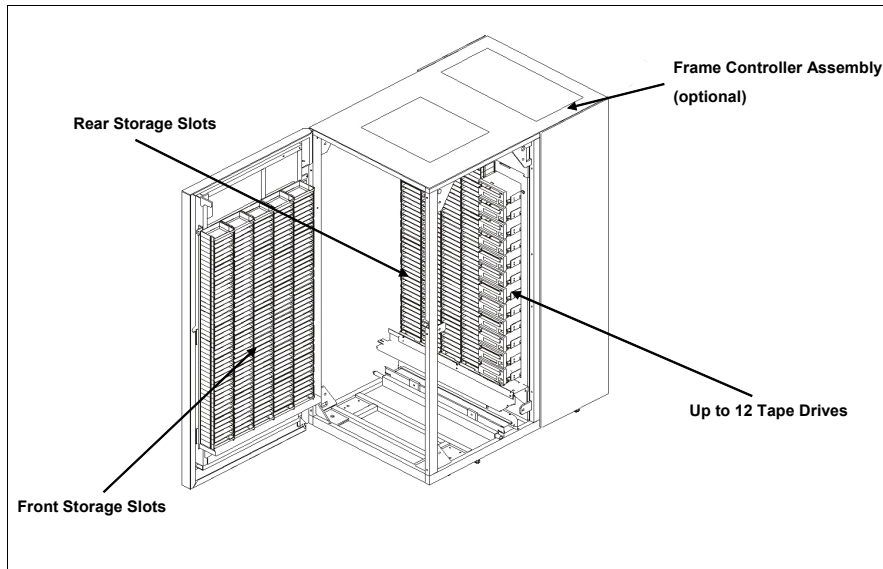


Figure 12-9 TS3500 Frame D23 or D53

Previous models

The models of the IBM System Storage TS3500 Tape Library vary, depending on the type of drives that they contain and whether the frame is a service bay, base frame, or expansion frame. The following is a description of each frame. Some of these models are no longer being sold; however they are still widely seen in customer environments.

Model L22 (3584-L22)

A base frame that uses up to twelve 3592 Tape Drives and up to 260 IBM TotalStorage 3592 Enterprise Tape Cartridges. The Model L22 is approximately 307 mm (12 in.) shorter in depth than Model L32.

Model D22 (3584-D22)

An expansion frame that uses up to twelve 3592 tape drives and up to 400 IBM TotalStorage 3592 Enterprise Tape Cartridges. This frame can optionally be configured as service bay B. The Model D22 is approximately 307 mm (12 in.) less deep than the Model D32.

Model L32 (3584-L32)

A base frame that uses up to twelve Ultrium tape drives and up to 281 IBM LTO Ultrium Tape Cartridges. The Model L32 is approximately 307 mm (12 in.) deeper than the Model L52. The D32 was withdrawn from marketing on 1 Oct, 2004.

Model D32 (3584-D32)

An expansion frame that uses up to twelve Ultrium Tape Drives and up to 440 IBM LTO Ultrium Tape Cartridges. The Model D32 is approximately 307 mm (12 in.) deeper than the Model D52. The D32 was withdrawn from marketing on 1 Oct, 2004.

Model L52 (3584-L52)

A base frame that uses up to twelve Ultrium Tape Drives and up to 287 IBM LTO Ultrium Tape Cartridges. The Model L52 is approximately 307 mm (12 in.) less deep than the Model L32.

Model D52 (3584-D52)

An expansion frame that uses up to twelve Ultrium Tape Drives and up to 440 IBM LTO Ultrium Tape Cartridges. This frame can optionally be configured as service bay B. The Model D52 is approximately 307 mm (12 in.) less deep than the Model D32.

Robotic cartridge accessor

The cartridge accessor is the assembly that moves tape cartridges between storage slots, tape drives, and the I/O station. The accessor assembly moves horizontally through the library frames using a rail system; it uses both top and bottom rails. The accessor assembly consists of a dual gripper mounted on a vertical pole. The gripper can move up and down vertically, and also rotates to access cartridge slots on both the back walls and front doors of the library frames. A barcode reader is mounted on the accessor and can scan the cartridges in one frame in less than a minute.

Barcode reader

The barcode reader reads the barcode on a label that is attached to a cartridge or at the rear of every storage slot (which indicates an empty storage slot). The barcode reader is mounted on the pivot assembly, and is used during inventories, audits, insertions, and inventory updates. The inventory is updated whenever the door is opened, and determines whether cartridges have been added to, removed from, or moved within the library). Because all storage slots have empty storage cell labels, the library can easily and quickly recognize if there is a labeled cartridge or an empty storage slot in every location. This eliminates the need to reread or manually intervene in storage cells if no label is readable. Without this approach the library cannot differentiate between a slot that is unlabeled, badly labeled, or empty.

Calibration sensor

This provides a way to locate certain positions within the library very precisely during the calibration operation. The calibration sensor is mounted on the underside of Gripper 1. All positions are calculated from these locating positions.

Control path failover

Alternate path support, available on certain IBM libraries and attached hosts, configures multiple physical control paths to the same logical library within the device driver and provides automatic failover to an alternate control path when a permanent error occurs on one path. This is transparent to the running application. Figure 12-10 on page 240 shows a simple multi-path architecture connection consisting of two HBAs in an AIX, Linux, Solaris, Windows or HP-UX host that are connected to a library with two or more drives. Two drives have the control ports enabled. The two HBAs are connected to the first and second control port drives, respectively.

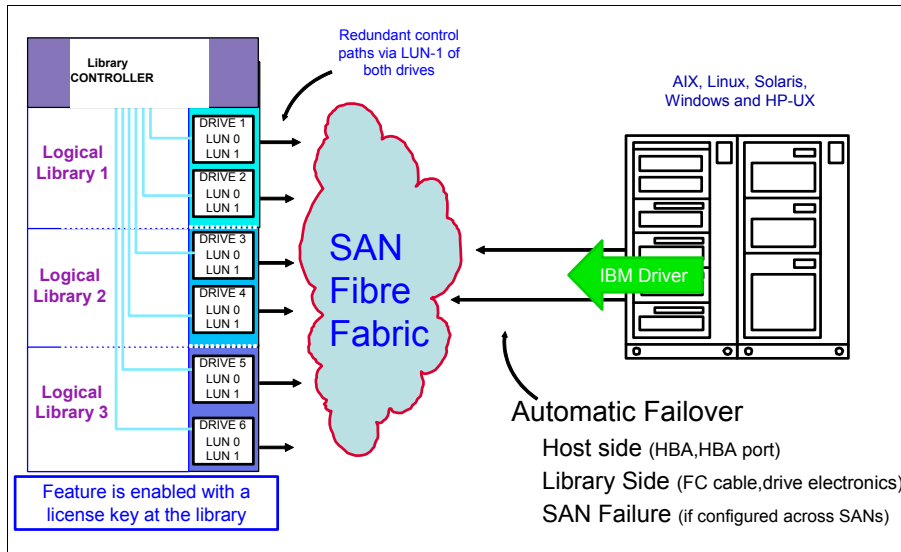


Figure 12-10 IBM Multipath host with two HBAs to the Library System

Figure 12-11 shows how two physical control paths are provided to the library for redundancy if one path from an HBA to the library fails.

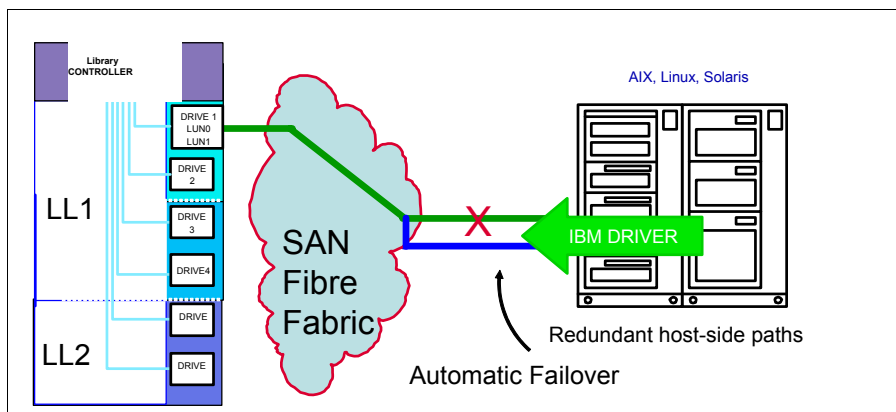


Figure 12-11 IBM Multipath fail over from Host Side

Reliability

The TS3500 is designed for high availability and reliability. Most of its essential components are redundant, so there is no single point of failure. Here is a summary of its high-availability features:

- ▶ Redundant grippers. A failure of one gripper will cause the library to switch to a second gripper.
- ▶ Redundant library power. Each drive frame provides one additional power supply. A single frame library contains one redundant power supply. Optional redundant frame power available. The library can operate on a single power supply. The library automatically monitors and controls redundant power distribution.
- ▶ Redundant drive power. Each drive bay power module supplies redundant power to another drive bay.
- ▶ Redundant control paths. Any LTO drive can be used as a library control path. Automatic control path failover available on certain platforms.

- ▶ Redundant copies of Vital Library Data. Includes configuration data, calibration data, setup data, and so on. One processor card contains the primary copy and another processor card contains a backup copy.
- ▶ Backup/restore process is completely automated.
- ▶ Redundant copies of library firmware. Each processor card contains the firmware for every other processor card. Each processor card contains two copies of operational firmware protecting the library from potentially harmful firmware update disruptions.
- ▶ Component replacement is simplified.
- ▶ Helps reduce the risk of memory failures.
- ▶ Closed loop servo systems. Each servo system uses feedback.
- ▶ Includes horizontal motion, vertical motion, pivot motion, gripper extend, and retract motion.
- ▶ Velocity and position are monitored allowing higher performance, greater control, reliable closed loop and avoiding collisions and gripper damage.
- ▶ Dual Accessors - this will increase reliability by operating without disruption when any component of the working accessor fails.

12.5.2 Dual accessors and service bays in the TS3500

When an optional second accessor is installed, the TS3500 becomes more available, because it can use that accessor to operate without disruption when any component of the working accessor fails.

Note: To use dual accessors and service bays, the TS3500 must be installed with the Advanced Library Management System (ALMS).

A library with the optional second accessor installed can also optimize cartridge mount performance. A mount occurs when the accessor removes a cartridge from a drive, returns it to its storage slot, collects another cartridge from a random storage slot, moves it to and loads it into the drive. The second accessor is part of feature code 1440 (Service Bay B Configuration), which also includes a D23 or D53 frame as service bay B. If you order a second accessor you must also order a TS3500 high availability (HA1) frame, which is also known as service bay A.

When dual accessors are installed and an attached host issues a command for cartridge movement, the library automatically determines which accessor can perform the mount in the most timely manner. If the library's primary accessor fails, the second accessor automatically assumes control. Although the library uses defaults to specify the zones (areas) in which the accessors operate, you can specify particular zones by using the Tape Library Specialist Web interface. This process is called setting the preferred zone. Figure 12-12 on page 242 shows how using a dual accessor increases performance in a TS3500 from 4 frames to 16 frames.

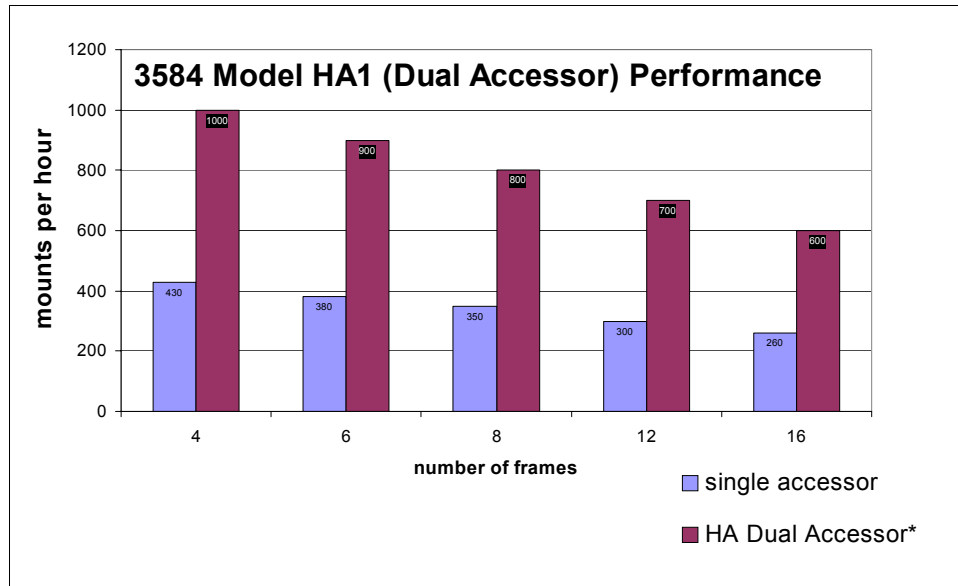


Figure 12-12 TS3500 Performance single accessor versus Dual accessor¹

12.5.3 Advanced Library Management System (ALMS)

The Advanced Library Management System (ALMS), virtualizes the locations of cartridges in the TS3500. Logical libraries can then consist of unique drives and ranges of volume serial numbers instead of fixed locations.

ALMS is an extension of IBM patented Multi-Path Architecture. With ALMS, the TS3500 is the industry's first standards-based tape library to virtualize the locations of cartridges (called SCSI element addresses) while maintaining native SAN attachment for the tape drives. ALMS enables logical libraries to consist of unique drives and ranges of volume serial (VOLSER) numbers, instead of fixed locations.

When you enable ALMS with its license key, you can assign tape drives to any logical library by using the Tape Library Specialist Web interface. Logical libraries can also be added, deleted, or easily changed without disruption. Storage capacity can be changed without impact to host applications.

ALMS offers dynamic management of cartridges, cartridge storage slots, tape drives, and logical libraries. It enables the TS3500 to achieve unprecedented levels of integration for functionality through dynamic partitioning, storage slot pooling, and flexible drive assignment. ALMS eliminates downtime when you add Capacity on Demand (CoD) storage, add or remove logical libraries, or change logical library storage allocation. ALMS also reduces downtime when you add expansion frames, add or remove tape drives, or change logical drive allocation.

The capabilities of ALMS include:

- ▶ Dynamic partitioning (storage slot pooling and flexible drive assignment)
- ▶ The transparent ability to add or remove storage capacity to any host application
- ▶ The ability to configure drives or to configure Model L22, L23, L32, L53 and L52 storage capacity without taking the library offline

¹ *Assumes the library is partitioned into at least 2 logical libraries such that each of the cartridges located in the 2 zones will be mounted in drives within each zone resulting in no accessor path overlap. This requires four columns of empty storage slots at the zone boundary.

- ▶ Cartridge assignment policy

Virtual I/O Slots

ALMS also supports virtual I/O slots. Virtual I/O slots increase the quantity of available I/O slots by allowing storage slots to appear to attached hosts as I/O slots. The goal of virtual I/O slots is to reduce the dependencies between the system administrator and library operator so that each performs their import and export tasks without needing the other to perform any actions. With virtual I/O slots, the library automatically moves cartridges from the I/O stations to physical storage slots and from physical storage slots to the I/O stations. Each logical library (which can scale from 1 to 192) supports up to 255 dedicated virtual I/O slots. The Virtual I/O slot feature is designed to support continuous library operations during insert/eject operations that involve 100's of cartridges.

Virtual I/O Slots simplifies operation:

- ▶ Cartridges for multiple logical libraries can be inserted in any order and then logically inserted by the backup application
- ▶ Allows up to 255 insert operations per logical library without any host or application intervention
- ▶ Inserted cartridges are automatically moved to storage slots (“virtual I/O slots”) without application involvement
- ▶ Cartridges ejected by an application are “virtually ejected” and then queued for a physical move to the I/O Station

12.5.4 Tape drives supported by the IBM System Storage TS3500 Tape Library

This section gives information about the types of drives that can be installed in the IBM System Storage TS3500 Tape Library.

- ▶ IBM LTO and TS1120/3592 J1A tape drives
 - TS1030 tape drives are supported in L53 and D53 frames
 - D53 frames can be added to any TS3500 library
 - TS1120 and 3592 J1A tape drives are supported in TS3500 L23 / D23 frames
 - D23 frames can be added to any TS3500 library
- ▶ IBM 3490 virtual drives¹ (using B10 or B20 VTS models)
- ▶ TS1120 tape controllers¹ (FICON and/or ESCON attached)
 - Only TS1120 or 3592 J1A tape drives are supported
- ▶ B10 and B20 Virtual Tape Servers¹ (FICON and/or ESCON)
 - Up to eight VTSs can be attached to a single TS3500 tape library
 - Support PtP VTS ‘peer’ in separate 3494 or TS3500 tape library

12.6 IBM TotalStorage 3953 Tape System for System z

The 3953 F05 Frame and L05 Library Manager allow attachment of IBM System z host systems to the IBM System Storage TS3500 Tape Library. The System z hosts attach via either Enterprise System Connection ESCON or Fibre Connection FICON to the following components: IBM System Storage TS1120 Tape Controller (3592 Model C06), IBM TotalStorage 3592 Tape Controller Model J70, or IBM TotalStorage 3494 Virtual Tape Server Models B10 and B20. Any number of subsystems between one and sixteen can attach to the 3953 L05 Library Manager. If two or more subsystems are attached, a maximum of two can be VTSs. Figure 12-13 on page 244 shows a diagram connection of the 3953.

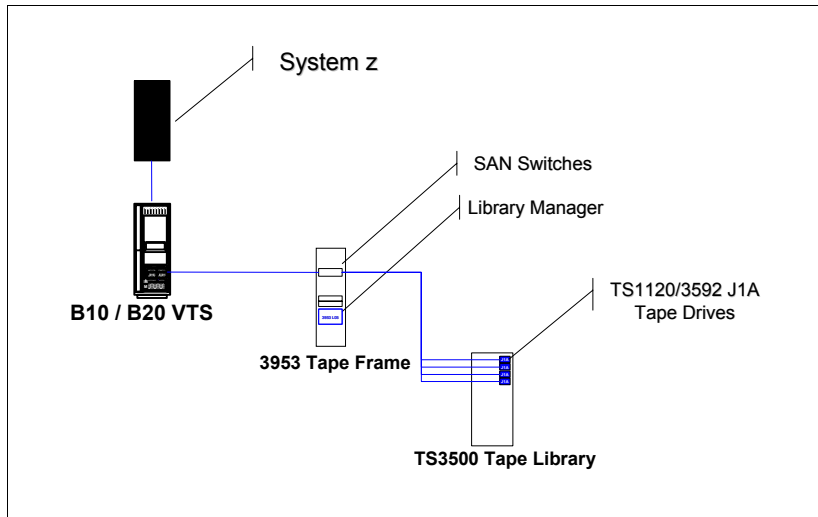


Figure 12-13 IBM TotalStorage 3953 Tape System for zSeries diagram

12.6.1 Features

These are the features for IBM TotalStorage 3953 Tape System for System z:

- ▶ Builds on recent tape library enhancements
 - Dual robotics and nondisruptive frame addition
 - Advanced Library Managements System
 - TS1120 and 3592 J1A tape drive support
- ▶ Requires two separate components
 - One 3953 F05 Tape Frame
 - One 3953 L05 Library Manager
- ▶ Supports VTS B10 / B20 and TS1120 controller
 - Attached to TS1120 and/or 3592 J1A tape drives
- ▶ Supports System z operating systems
 - zLinux is supported natively
- ▶ Supports open systems
 - Coexists with open systems in their own logical library
 - Securely isolated from System z environment
- ▶ 3953 L05 Library Manager
 - Functionally equivalent to 3494 Library Manager
 - Employs identical microcode and server
- ▶ The 3953 L05 is mounted in the first 3953 F05 frame
 - Requires dual internal LAN features
 - Requires TotalStorage Master Console (TSMC)
- ▶ Optional 3953 L05 provides redundancy
 - Mounted in first 3953 F05 frame
 - Facilitates microcode updates
- ▶ Issues TS3500 mount requests
 - Accepts PLF commands from VTS or TS1120 controller

- Issues SCSI media manager 'mount request

Figure 12-14 shows a 3593 configuration.

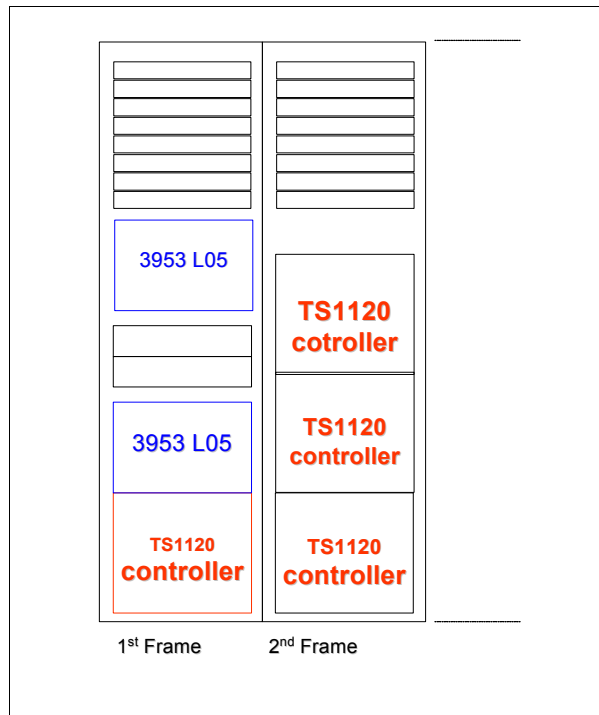


Figure 12-14 IBM TotalStorage 3953 Tape System for System z

12.7 IBM TotalStorage 3494 Tape Library

The IBM TotalStorage 3494 Tape Library (Figure 12-15 on page 246) is an excellent solution for large storage requirements. The IBM TotalStorage 3494 Tape Library consists of individual frame units for modular expansion that provides a wide range of configurations.



Figure 12-15 IBM TotalStorage 3494 Tape Library

The library manager enables the attachment of multiple host platforms, maintains a database of the cartridges held in the library, holds status information about library components, and provides an easy-to-use GUI for operations staff. The library manager controls all of the operations in the 3494. The TotalStorage Expert Enterprise Tape Library Specialist runs in the library manager to provide Web-based monitoring of library functions.

The main components of the IBM TotalStorage 3494 Tape Library are:

- ▶ Tape Library Base Frame
- ▶ Tape Drive Expansion Frame
- ▶ Tape Storage Frame
- ▶ High Availability Frames
- ▶ Virtual Tape Server

12.7.1 Consolidation and multiplatform attachment

The 3494 can be used for data consolidation and reduced requirements for tape drives and cartridges, environmental controls and personnel. The 3494 supports multiplatform attachment to selected System z, System p, System i and System x servers running IBM operating systems and Linux as well as selected open system servers from Sun and HP. The 3494 library also supports Intel processor-based systems running Microsoft Windows 2000, Windows 2003 and Linux operating systems.

12.7.2 Performance

The average mount-access time in a single-frame 3494 is only seven seconds. The 3494 can perform up to 265 cartridge exchanges per hour with a single gripper, up to 305 exchanges per hour with the optional dual gripper and up to 610 exchanges per hour with a dual gripper

and dual active accessors. Due to increased accessor travel, the exchange capability of the accessor can decrease as the number of frames increases. The time required to inventory tape cartridges in any single library frame (control unit, drive unit or storage unit) is approximately four minutes. In a 3494 HA1 environment, both accessors are used to reduce the time to perform an inventory.

12.7.3 Enterprise tape drive support

The 3494 supports up to 128 open system attached IBM System Storage TS1120 Tape Drives, up to 128 IBM TotalStorage 3592 Tape Drive Model J1A, or up to 92 IBM TotalStorage 3590 Tape Drives Model E1A or H1A. The 3494 library can support up to 132 3592 tape drives, or up to 76 3590 tape drives in a System z environment when attached to IBM System Storage TS1120 Tape Controller. For more information about these drives see Chapter 11, “IBM TotalStorage and System Storage tape drives” on page 201.

12.7.4 Storage capacity

Total capacity varies according to the tape technology installed and the application’s ability to utilize the capacity of the cartridges. The 3494 supports a maximum of 6,240 cartridges which provide a native physical capacity of up to 3.12PBs.

12.7.5 Tape Library Base Frame

The initial building block of the 3494 is the Lxx control unit frame, which contains a library manager, a cartridge accessor, up to two tape drives, and slots for the storage of tape cartridges. To the Lxx frame, you can add drive frames, storage unit frames, the IBM TotalStorage 3494 Virtual Tape Server, and a high-availability model to create a maximum configuration of 16 frames and two service bays.

The tape library base frame is a central and important component of the 3494. Each base frame contains:

- ▶ Tape subsystem (3592 tape drives (TS1120), 3590 tape drives, or 3490 tape drives)
- ▶ Library manager - manages the robotic tape accessor and the library tape cartridge inventory
- ▶ Cartridge accessor with robotic accessor and gripper which controls tape cartridge movement within the library
- ▶ The optional dual gripper allows up to 610 mounts or demounts per hour with the Dual Active Accessor Feature (DAA)
- ▶ Optional convenience I/O station - two separate optional features that allow up to either 10 or 30 cartridges to be added or removed from the library without interrupting library operation
- ▶ Cartridge storage cells

There are four control unit models:

- ▶ L10 - contains up to two IBM 3490E tape drives (SCSI, ESCON, or Parallel attached)
- ▶ L12 - contains up to two 3590 Tape Drives (SCSI or FC attached)
- ▶ L14 - contains up to two 3590 Tape Drives (ESCON attached)
- ▶ L22 - contains up to four 3592 Tape Drives (FC attached)

Note: The L10 and L14 frames are withdrawn from marketing but are still widely available in customer environments.

One control unit frame is required in each 3494.

12.7.6 Tape Drive Expansion Frame

The 3494 Tape Drive Expansion Frame provides additional tape drives and tape cartridge storage.

Each expansion frame contains:

- ▶ Zero to twelve drives per frame depending on the model
- ▶ Up to 400 cartridge storage cells
- ▶ Up to fifteen drive units can be attached to a single Lxx control unit

The five expansion frame models D10, D12, D14, D22 and D24 differ in cartridge storage cell capacity, tape subsystem supported, and host platform attachment capability:

- ▶ D10 can attach from up to two IBM 3490E tape drives and store from 270 to 400 tape cartridges
- ▶ D12 can attach up to six native Fibre Channel SCSI attached 3590 tape drives and store from 250 up to 400 tape cartridges
- ▶ D14 can attach up to four ESCON/FICON attached 3590 tape drives and store from 305 to 400 tape cartridges
- ▶ D22 can attach up to twelve native Fibre Channel SCSI attached 3592 tape drives and store 250 to 400 tape cartridges
- ▶ D24 can attach up to eight ESCON/FICON attached 3592 Tape Drives and store 305 to 400 tape cartridges

A 3590 Model A60 Controller installed in a D14 drive unit frame can provide ESCON/FICON support for 3590 Tape Drives in an adjacent L12, L14, or D12 library frame. A 3592 Model J70 Controller installed in a D24 drive unit frame can provide ESCON/FICON support for 3590 and 3592 Tape Drives in an adjacent L12, L14, L22, D12 or D22 library frame.

The 3494 supports the following tape drives within the same library, but in separate 3494-D12 3494-D14, 3494-D22 or 3494-D24 drive frames:

- ▶ IBM TotalStorage 3590 Tape Drive Model E1A
- ▶ IBM TotalStorage 3590 Tape Drive Model H1A

Note: The D10 frame is withdrawn from marketing but is still widely available in customer environments.

12.7.7 Tape Storage Frame

The 3494-S10 Tape Storage Frame provides additional tape cartridge storage capacity. Each storage unit has a capacity of up to 400 additional tape cartridges. The cartridges can be a mix of 3490E, 3590 or 3592 (TS1120) tape cartridges.

Up to fifteen storage frames can attach to any Lxx control unit frame.

12.7.8 High Availability Frame

The 3494-HA1 High Availability Frame provides a second library manager (LM) and a second cartridge accessor. In the unlikely event of a failure of either of these components, they can be switched over to replace the failed unit. In addition, there is a dual active accessor feature that will allow both cartridge accessors to operate simultaneously.

The 3494-HA1 consists of two service bay units, one placed at each end of the library, plus a second library manager and accessor. The HA1 will operate in hot-standby mode if there is no dual active accessor option.

In the event of a cartridge accessor or library manager failure, the standby accessor and/or library manager will take control of all operations in the library. The model HA1 is designed to eliminate the library manager and the tape cartridge accessor as single points-of-failure.

An optional feature for the High Availability Frame is the Dual Active Accessor. This feature activates the library's second cartridge accessor enabling both accessors to operate simultaneously so that this option not only improves availability, but also adds to the throughput of the library by increasing the maximum mount rate.

IBM TotalStorage 3494 Tape Library enhancements include:

- ▶ Increased open system hosts- to increase the number of open system host TCP/IP connections from 32 to 64 for increased connectivity.
- ▶ Increased open systems tape drives- to increase the number of allowed open system tape drive attachments from 32 to 128 for expanded open system support.
- ▶ LAN PCI Library Manager- to support the above attachments.
- ▶ Dual power on the Models D22 and D24 - which helps eliminate the current power distribution unit and customer power source as a single point of failure when a 3494 Model HA1 is installed, and allows you to take advantage of the n+1 power characteristics of the 3592 Model J1A Tape Drive.

12.7.9 IBM 3494 Library Manager

The Library Manager, the operational focal point of the 3494, provides the support to set up, maintain, configure, and operate the library. It consists of a controller, a personal computer (PC), a display, a keyboard, and the library manager application.

The Library Manager GUI is provided by the library manager application and runs on the library manager PC.

The Library Manager provides three user interfaces:

- ▶ **Library Manager Console:** Installed on the rear side of the IBM 3494 Model L12 frame. Provides monitoring of the 3494 and appropriate actions.
- ▶ **IBM TotalStorage Enterprise Tape Library (ETL) Specialist:** A Web browser interface for monitoring the 3494 and for take specific actions for an installed VTS or Peer-to- Peer (PtP) VTS.
- ▶ **IBM TotalStorage Master Console:** Integrates service monitoring of the 3494.

Remote Library Manager Console: Provides remote access to the Library Manager to issue commands or take actions. In addition, host software commands allow you to retrieve information from the Library Manager, or to send commands to the library.

12.7.10 More information

You can find more information about IBM Enterprise tape products from these Web sites:

- ▶ http://www.storage.ibm.com/tape/libraries-autoloaders/3494_vts/index.html
- ▶ <http://www.storage.ibm.com/tape/libraries-autoloaders/3494/index.html>
- ▶ <http://www.storage.ibm.com/tape/drives/3590/index.html>
- ▶ <http://www.storage.ibm.com/tape/drives/3592/index.html>

Refer to this redbook for in-depth information about IBM Enterprise tape products:

- ▶ *IBM TotalStorage Enterprise Tape: A Practical Guide, SG24-4632*

12.8 Enterprise Tape System 359x family

The TotalStorage 359x Enterprise Tape System family includes tape drives and controllers that can be configured in a broad range of solutions, including stand-alone frames, racks, or automated library solutions. This variety of solutions enables the 359x family to meet a diverse set of customer requirements.

The Enterprise Tape System 359x family includes a variety of members of 359x tape drives and autoloader, 359x tape cartridges, tape controller 3590 model A60/C10/C12 and tape controller 3592 model J70. More information about this Tape system is in 11.3, "TotalStorage 359x tape drive" on page 216.

12.9 Other tape libraries

The following tape libraries are also available. Some have been withdrawn from marketing but are still found in customer environments.

12.9.1 IBM TotalStorage 3581 Tape Autoloader

The IBM TotalStorage 3581 Tape Autoloader (Figure 12-16) offers a single Ultrium 3 tape drive and storage for up to eight tape cartridges. Models L38 and F38 use SCSI LVD and 2 Gbps Fibre channel interfaces to the LTO drive respectively.



Figure 12-16 IBM TotalStorage 3581 Tape Autoloader

The IBM LTO Ultrium 3 tape drives support up to 80 MBps native data transfer rates. With the IBM System Storage LTO Ultrium 400 GB Data Cartridge, an IBM 3581 has capacity of up to 6.4 TB with 2:1 compression or 3.2 TB native.

The IBM 3581 features Low Voltage Differential (LVD) Ultra160 SCSI (Model L38 for Ultrium 3 or Model L28 for Ultrium 2) or native switched 2 Gbps Fibre Channel (Mode F38 for Ultrium 3 or Model F28 for Ultrium 2) drive interfaces and can attach to a wide variety of open system servers.

Product description

The IBM TotalStorage 3581 Tape Autoloader is available as four separate model types, depending on the capacity and required SCSI2 or FC interface. All models have 8 cartridge slots.

- ▶ L28 model has a 1.6 TB native (3.2 TB compressed) data capacity, using LTO2 LVD Ultra160 SCSI (also High-Voltage Differential (HVD) with converter)
- ▶ F28 model has a 1.6 TB native (3.2 TB compressed) data capacity using LTO2 native switched 2 Gbps Fibre Channel
- ▶ L38 model has a 3.2 TB native (6.4 TB compressed) data capacity using LTO3 LVD Ultra160 SCSI which can connect to a wide variety of open system servers.
- ▶ F38 model has a 3.2 TB native (6.4 TB compressed) data capacity using LTO3 native switched 2 Gbps Fibre Channel.

The L28 and F28 models both use the same Ultrium 2 LTO cartridge, which has a capacity of 200 GB (400 GB with 2:1 compression), and are capable of sustaining a data rate of up to 35 MBps native (or 70 MBps compressed).

The L38 and F38 models both use the same Ultrium 3 LTO cartridge, which has a capacity of 400 GB (800 GB with 2:1 compression), and are capable of sustaining a data rate of up to 80 MBps native (or 160 MBps compressed).

The IBM 3581 can be operated in sequential or random-access modes. In sequential mode, it loads cartridges one after another, controlled by the hardware when it receives the unload command from the host server. In random-access mode, it relies on application software for cartridge management.

Although it has only a single tape drive, it appears as two SCSI devices on the SCSI bus. In other words, the autoloader and the drive have separate SCSI addresses.

The IBM 3581 only operates on a single path, and control of the autoloader is handled by a single server. Connection to multiple servers can be achieved when using SAN technology and appropriate application software for cartridge management, such as IBM Tivoli Storage Manager.

Two additional optional features are available:

- ▶ Rack Mount Option so that it can be installed as a stand-alone single unit, or two units can be mounted side-by-side in a standard 19-inch rack, requiring 5 EIA units of rack space. This chargeable feature provides the necessary hardware to mount the autoloader in the rack.
- ▶ Barcode Reader enables the autoloader to read cartridge information contained in a barcode label on the IBM Ultrium cartridges. User-installed application software provides the inventory management functions enabled by the barcode reader feature.

For more information, refer to this Web site:

<http://www.storage.ibm.com/tape/lto/index.html>

For specific information about the IBM TotalStorage 3581 Tape Autoloader, refer to:

<http://www.storage.ibm.com/tape/lto/3581/index.html>

12.9.2 IBM TotalStorage 3582 Tape Library

The IBM TotalStorage 3582 Tape Library is shown in Figure 12-17 on page 252. It is a high-performance, reliable, scalable tape subsystem. Designed for tape automation, it can be attached to System i, System p, System x, Intel (running Windows or Linux), Sun SPARC, Hewlett-Packard, and other open systems using SCSI or Fibre Channel attachment.



Figure 12-17 IBM TotalStorage 3582 Tape Library

The IBM TotalStorage 3582 Tape Library Model L23 can contain one or two IBM Ultrium 3 and/or Ultrium 2 Tape Drives with up to 80 MBps native data transfer rates per drive (or 35 MBps with the Ultrium 2 Tape Drives). The generation 3 drive provides physical tape cartridge capacity up to 400 GB native capacity (800 GB with 2:1 compression) when using the new IBM System Storage LTO Ultrium 400 GB Data Cartridge.

Product description

The compact 4U IBM 3582 comes standard with Multipath architecture and the ability to partition the library into two logical libraries, two removable cartridge magazines, and a bar code reader for quick media verification and inventory. The library can be configured as a stand-alone unit or can be mounted in an industry-standard 19-inch rack.

Additional optional features include: Control Path Failover, Data Path Failover, and a Remote Management Unit/Specialist for remote library management. An LCD display on the front panel supports convenient local review of library status and provides control over local functions.

The IBM 3582 comes standard with a one cartridge I/O station and 23 data cartridge slots giving a library physical capacity of 9.6 TB native (19.2 TB with 2:1 compression) when using generation 3 LTO media and drives. The IBM Ultrium 3 and Ultrium 2 Tape drives come in 2 Gbps fibre Channel and Low Voltage (LVD) Ultra160 SCSI varieties to attach to a wide spectrum of open system servers.

Cartridge storage

As well as the installed tape drives, the library enclosure contains cartridge storage slots, arranged in two rows. The row toward the front of the library is made of two removable magazines of seven slots each. The row toward the rear of the library contains nine slots. The magazines are designed so that tape cartridges can only be inserted in the proper orientation. Once inserted, the tape cartridges will be retained in the magazine so that they remain in place even when the magazine is inverted and shaken lightly. The magazines can only be inserted one way into the mounting columns in the library.

Barcode scanner

A barcode scanner is provided as standard with the L23 model, and it does not affect the slot capacity of the libraries. The barcode scanner is used during the inventory process to locate

all cartridges inserted into the library. This action is repeated every time the front door is opened to ensure that the inventory is updated if a cartridge has been manually added, moved, or removed while the door was open.

I/O station

This facility enables the insertion and ejection of cartridges without interrupting the normal operation of the library. There is a single-slot I/O station where a cartridge can be inserted or ejected by opening the I/O station door.

Robotic system

In conjunction with the library control microcode, the robotic system identifies and moves cartridges between the storage slots, tape drives, and the I/O station. It has a number of components:

- ▶ A cartridge picker for placing cartridges in storage slots, tape drives, or the I/O station
- ▶ A bar code scanner used to set up the library initially when it identifies the types of storage arrays and tape drives installed in the library, and in normal operation for reading the external labels on the cartridges, when it locates and categorizes all cartridges installed in the library
- ▶ X-axis and Z-axis drive motors for moving the picker assembly inside the library enclosure

Library control and operation

The library control unit contains the electronics and logic for autochanger operations. It controls all operations in the library, including the interaction between the library and operators.

For more information, refer to this Web site:

<http://www.storage.ibm.com/tape/lto/index.html>

For specific information about the IBM TotalStorage 3582 Tape Library, refer to:

<http://www.storage.ibm.com/tape/lto/3582/index.html>

12.9.3 IBM TotalStorage 3583 Tape Library

Note: All models of the 3583 were withdrawn from marketing on Jan 27, 2006.

The IBM TotalStorage 3583 Tape Library, shown in Figure 12-18 on page 254 can be attached to System i, System p, System x, AS/400®, RS/6000, Netfinity®, and non-IBM servers, workstations, and personal computers that support SCSI, HVD, SCSI LVD, and Fibre Channel interfaces. It uses IBM Ultrium (1, 2, and 3 generation) tape drives.



Figure 12-18 IBM TotalStorage 3583 Tape Library

Product description

The IBM 3583 features the patented multipath architecture which is designed to allow simultaneous attachment of heterogeneous servers and applications to LTO logical library partitions, including mixed Ultrium drives and media. The IBM 3583 comes standard with a barcode reader for quick media verification and inventory. The IBM 3583 features storage capacities of 18, 36, and 72 data cartridge slots, providing total Ultrium 3 capacities of 7.2 TB,

The IBM 3583 comes in three different models - the major difference between the models is the number of storage cells shipped with the initial order:

- ▶ Model L18 has space for 18 cartridges and 1 to 6 Ultrium drives
- ▶ Model L36 has space for 36 cartridges and 1 to 6 Ultrium drives
- ▶ Model L72 has space for 72 cartridges and 1 to 6 Ultrium drives

The model numbering is an indication of how many cartridge slots the model was originally shipped with. However, all of its models are the same physical size, and the two smaller models, L18 and L36, can be field-upgraded to hold up to 72 cartridges by the addition of upgrade features. Table 12-4 shows the capacities, assuming Ultrium 3 drives are used.

Table 12-4 IBM TotalStorage 3583 Tape Library model summary

Models	Cartridge Slots	Data Capacity (Native)	Data Capacity (Comp.)	IBM Ultrium Tape Drives
L18	18	3.6 TB	7.2 TB	1-6
L36	36	7.2 TB	14.4 TB	1-6
L72	72	14.4 TB	28.8 TB	1-6

Upgrades and optional features

You can add features to each model type in the IBM 3583 in order to add capacity in terms of drives and cartridge cells, and to add the 12-cartridge I/O station. This effectively takes the place of model upgrades. Model L18 cannot be upgraded to Model L36 or L72. Similarly, Model L36 cannot be upgraded to Model L72. Any model can be extended to the maximum capacity of the largest model (L72) with the addition of MES features.



IBM Tape Virtualization products

This chapter presents the IBM tape virtualization products, solutions, and service offerings:

- ▶ IBM Virtualization Engine TS7510 (TS7510 Virtualization Engine™; 3954 CV5)
 - IBM Virtualization Engine TS7510 Model SV5 (TS7510 Cache Controller; 3955 SV5)
 - IBM Virtualization Engine TS7510 Model SX5 (TS7510 Cache Module; 3955 SX5)
 - IBM Virtualization Engine TS7510 Software Version 1 Release 1 (TS7510 V1R1 Software)
- ▶ IBM TotalStorage 3494 Virtual Tape Server Model B10 and B20 (B10 / B20 VTS)
 - IBM TotalStorage Virtual Tape Frame 3494 Model CX1 (CX1)
 - IBM TotalStorage 3494 Peer-to-Peer Virtual Tape Server (PtP VTS)
- ▶ Enterprise Removable Media Manager (eRMM)

13.1 IBM Virtualization Engine TS7510

This section covers the IBM open tape virtualization product, the IBM Virtualization Engine TS7510.

13.1.1 Introduction to open tape virtualization

One of the biggest challenges with backup planning today is that the amount of data being backed up is growing, but the time allotted for a backup (the backup window) is shrinking or remaining static. Applications need to be up and operational nearly 24 hours a day. To manage the need for increased data capacity and data protection, customers must find ways to shrink their backup windows and recover as quickly as possible.

In traditional backup environments, shown in Figure 13-1, the client data is backed up in two different ways:

- ▶ LAN clients write their backup data via the backup server to tape
- ▶ LAN-free clients utilize the SAN for direct backup to the tape devices

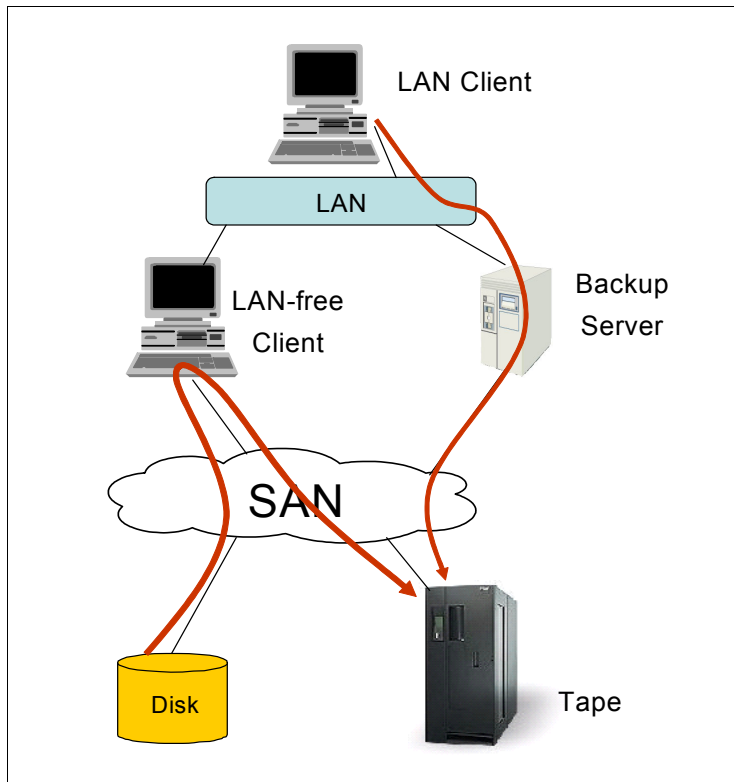


Figure 13-1 Traditional backup

The problems with these traditional backups are:

- ▶ LAN backups cannot benefit from fast tape drives, because the network transfer rate is the limiting factor; data cannot be delivered fast enough to keep the tape drives streaming.
- ▶ LAN-free clients require dedicated tape drives for every backup session. So, many tape drives are needed during the backup window.

A solution for problems is to use disk to disk (D2D) or disk to disk to tape (D2D2T) backups.

For LAN backups this can be done by using a disk buffer, as shown in Figure 13-2. The client writes the backup data via LAN first to a disk buffer on the backup server. From there it can be recalled quite rapidly if needed. After the backup on disk is completed, the backup server can migrate the data from the disk buffer to a physical tape with no impact on the client.

Furthermore, the disk buffer works as a cache. The most recently backed up data resides in the fast disk buffer until its overwritten by newer data.

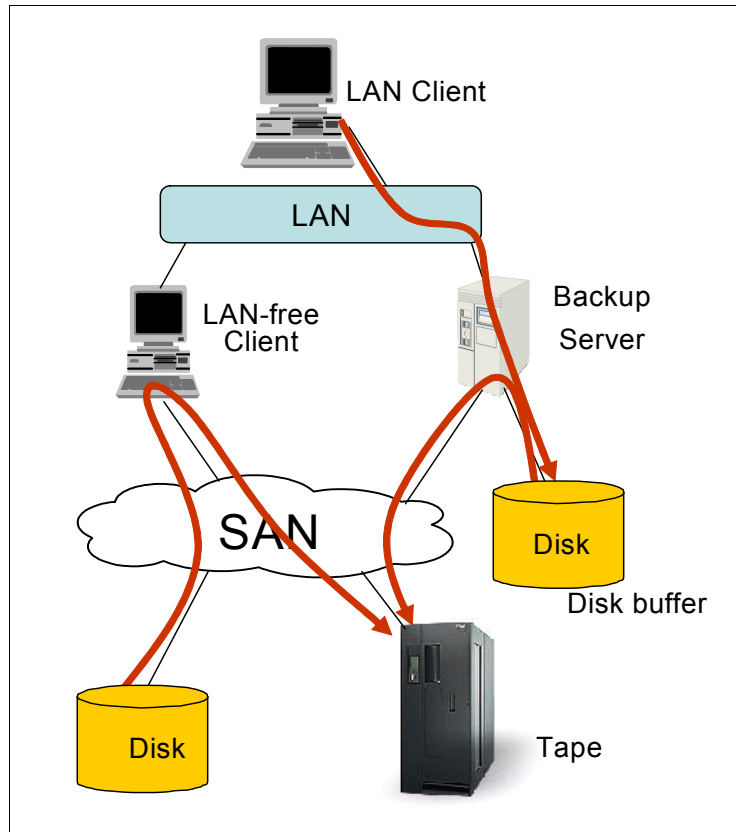


Figure 13-2 Backup to disk buffer

This is a good solution if the backup software supports it. IBM Tivoli Storage Manager supports a hierarchy of storage pools, and can automatically migrate from disk to tape, but other backup applications may not have this capability. For LAN-free clients with Tivoli Storage Manager, SANergy® is also required to enable them to write to a disk buffer.

An alternative solution to optimizing backup, is to introduce a virtualization and emulation layer. This layer appears to both the backup server and the LAN-free clients as a tape drive; however internally, the backup data is stored as *logical volumes* on disk space. Since the logical or virtual volumes are not physically mounted (as they are not actual tapes!), the emulation layer also acts as a library robotic - or medium changer.

The combination of hardware (servers, cache controllers, and cache modules), and software (tape drive and robotic emulation) is called a *Virtual Tape Library (VTL)*. In Figure 13-3 you can see a backup environment with the IBM Virtualization Engine TS7510.

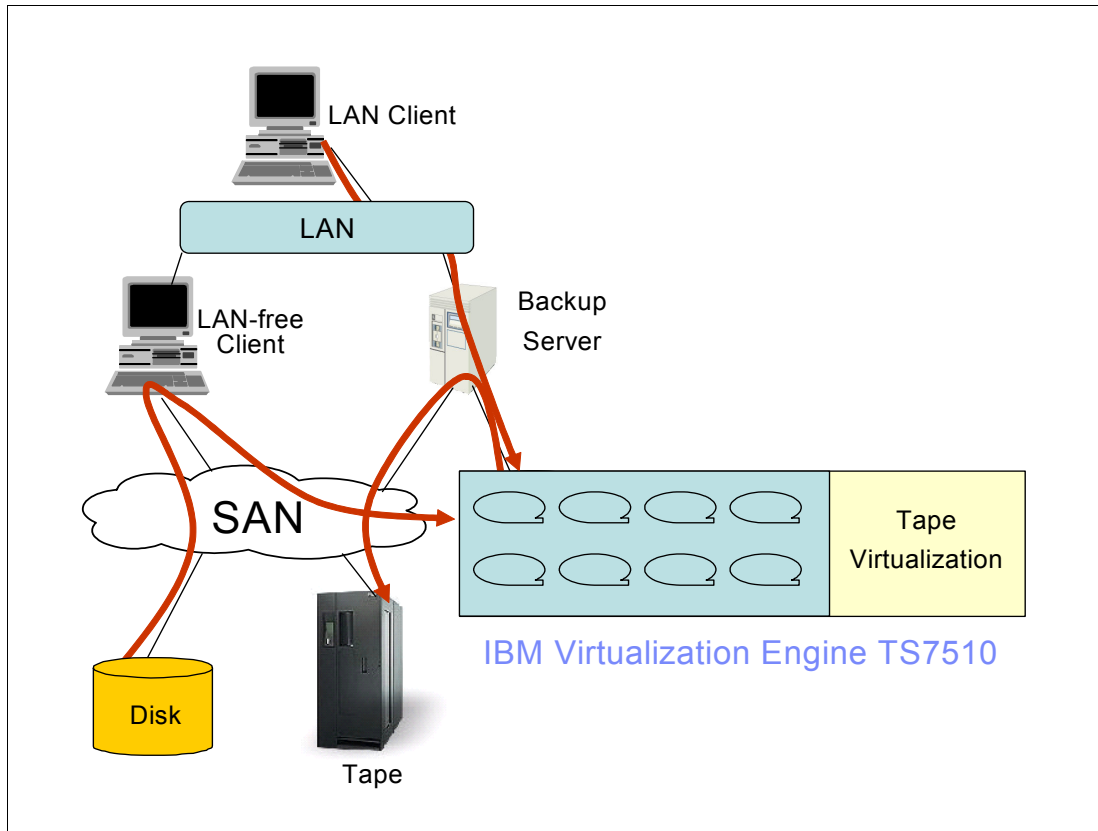


Figure 13-3 Tape virtualization with IBM Virtualization Engine TS7510

13.1.2 IBM Virtualization Engine TS7510 overview

The IBM Virtualization Engine TS7510 (TS7510 Virtualization Engine 3954-CV5) is the first member of the IBM Virtualization Engine TS7000 Series of virtual tape libraries. It combines IBM server and disk technology into an integrated solution which virtualizes, or emulates tape libraries, tape drives, and tape media. Real tape resources can also be attached to the TS7510 to help address information lifecycle management and business continuance.

The TS7510 can help to achieve the following goals:

- ▶ Reduce backup window
- ▶ Improve restore process
- ▶ Facilitate data sharing and resource virtualization
- ▶ Lower total cost of ownership (TCO)
- ▶ Allow multiple disparate backup applications to share the same physical resources

The TS7510 consists of three hardware machine types and a Virtualization Engine for the tape software product. The IBM System Storage 3952 Frame Model F05 (3592 F05 frame) is an independent frame used to contain the components of the TS7510.

The 3952 F05 frame can contain:

- ▶ One or two IBM Virtualization Engine TS7510 Model CV5 servers
- ▶ Two IBM Virtualization Engine TS7510 Model SV5 cache controllers
- ▶ Up to six IBM Virtualization Engine TS7510 Model SX5 cache modules

An adjacent 3952 frame can contain:

- ▶ Up to eight additional TS7510 Model SX5 cache modules

Figure 13-4 shows a two frame configuration of the IBM Virtualization Engine TS7510:

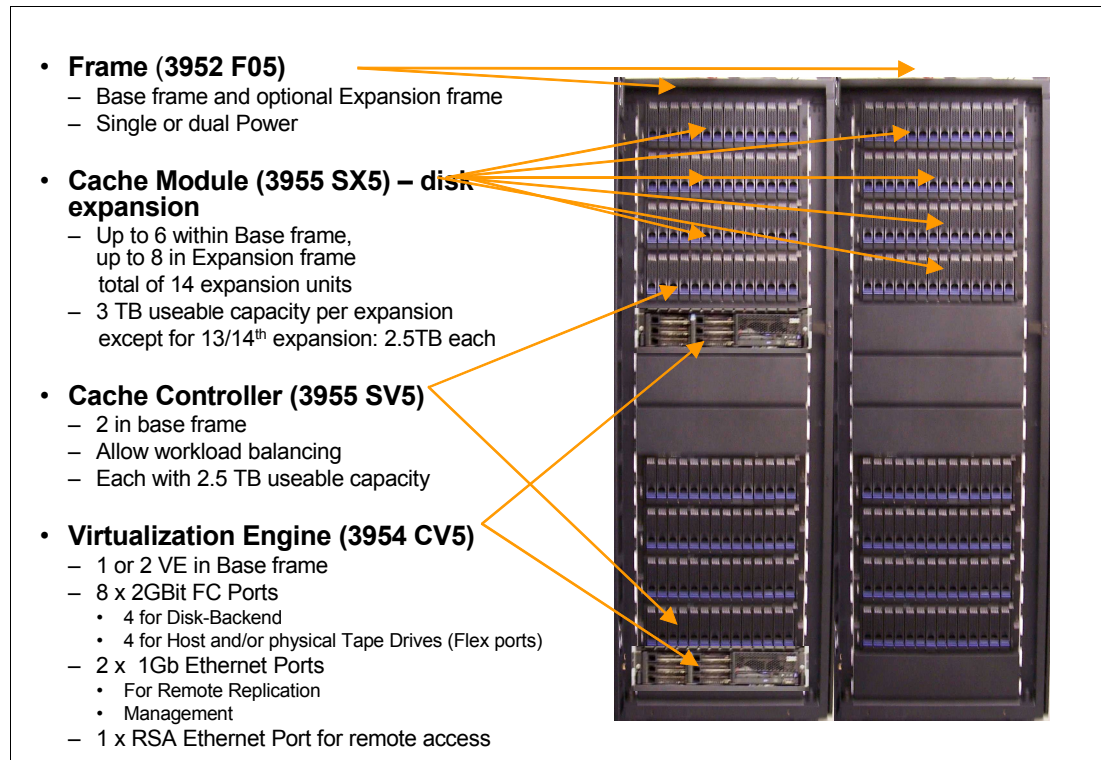


Figure 13-4 IBM Virtualization Engine TS7510 hardware overview

The TS7510 can provide up to 56 TB of raw cache capacity (46 TB of usable cache) to satisfy data backup and recovery requirements. This is made up of 2.5 TB usable capacity in each of the standard 2 cache controllers (3955 SV5) plus the maximum expansion, using the Cache Modules (3955 SX5). Each TS7510 SX5 cache module drawer offers 3.5 TB of raw cache - cache drawers can be added singly or in multiple increments. Each cache module drawer actually provides 3 TB usable capacity, except for the last two drawers - which have 2.5 TB of usable cache. Therefore, in a fully populated TS7510 there is 2 x 2.5 TB usable with just the 2 base 3955 SV5, plus 12 * 3 + 2 * 2.5 with the 14 expansion drawers. This comes to 46 TB total usable cache. All disk is configured as RAID 5.

A TS7510 with a single server provides up to 512 virtual tape drives, up to 64 virtual tape libraries and up to 4096 virtual volumes. In a high availability (HA) configuration with dual TS7510 Model CV5 servers, twice as many virtual resources are available.

The TS7510 supports the following hardware and software at the time of writing:

Servers

- ▶ IBM System p and IBM RS/6000 servers
- ▶ IBM eServer System z servers (zSeries Linux)
- ▶ IBM System x and Intel servers running the operating systems below
- ▶ Selected HP-UX servers running the operating systems listed below
- ▶ Selected SUN servers running the operating systems listed below

Tape storage systems

- ▶ IBM System Storage TS3500 Tape Library (3584) and IBM TotalStorage 3494 Tape Library
- ▶ IBM TotalStorage Ultrium LTO2/LTO3 tape drive
- ▶ IBM TotalStorage 3592 Tape Drive Model J1A

These tape storage systems are all physically attached

Operating systems

- ▶ AIX 5L V5.1, or later
- ▶ Sun Solaris 8 or 9
- ▶ Microsoft Windows 2000
- ▶ Microsoft Windows 2003
- ▶ HP-UX 11.00, 11i (64 bit), 11.23i, and 11.23pi
- ▶ Linux distributions: Red Hat Enterprise Linux 3 (RHEL 3) and SUSE LINUX Enterprise Server 8 and 9 (SLES 8 and 9)

Please check this URL for the most up to date list of supported configurations for the TS7510:

<http://www.ibm.com/servers/storage/tape/ts7510/index.html>

13.1.3 IBM Virtualization Engine TS7510 (3954 CV5)

The IBM Virtualization Engine TS7510 (3954 CV5) is a 2-way processor engine. Every TS7510 configuration must contain at least one Virtualization Engine Server. With an optional second Virtualization Engine Server the TS7510 can operate in a dual node high availability configuration. A TS7510 dual node high availability configuration can have more virtual cartridges, virtual volumes, and interface ports and must be ordered with a failover/failback option for redundancy. Each Virtualization Engine Server comes with eight Fibre Channel connections. Four of the Fibre Channel ports are connected to the TS7510 Cache Controllers (SV5), and four are available for host server or tape attachment. In a dual node system there are eight Fibre Channel ports available for host server or tape attachment.

Tip: *IBM Virtualization Engine TS7510 Introduction and Planning Guide, GC26-7767*, is a good place to find information about all of the physical site planning for the TS7510.

Table 13-1 IBM Virtualization Engine TS7510 at a glance

Form factor/height Frame dimensions are:	43.4" x 25.4" x 71" (WxDxH) One or two frames per system
Processor (max)	2 server nodes
Native cache (max with all cache module expansion)	46 TB
Connectivity	4 FC ports per node 8 FC ports per system (two nodes)
Network interface	3 Ethernet ports, 1 for service, 2 for IP replication
Hot-swap components	Server Nodes, disk drives, HBA, power supplies
RAID support	RAID 5
Systems management	SNMP traps for integration into ISV software for error reporting

Form factor/height Frame dimensions are:	43.4" x 25.4" x 71" (WxDxH) One or two frames per system
Limited warranty	1 year

Table 13-2 IBM Virtualization Engine TS7510

Specification	Single Node	Dual Node
Virtual Drives	512	1024
Min / Inc / Max Virtual Libraries	64	128
Min / Inc / Max Virtual Volumes	4096	8192
Performance (~ max) MBps	500	600
Native capacity (TB)	23	46
Max Number of Disk Drawers	8	16
Total FC ports	8	16
Total Ports to disk Cache	4	8
Total Host/Tape ports	4	8
Ethernet Replication Ports	2	4
Tape Drives Support via emulation	LTO 2, LTO 3, 3592 J1A	
Tape Library Support	3494 and TS3500	

13.1.4 IBM System Storage 3952 Tape Frame Model F05

The IBM System Storage 3952 Tape Frame Model F05 is the 19-inch rack in which all of the components reside and is required for the configuration. The rack comes standard with one Power Distribution Unit (PDU), although we recommend that you purchase a second PDU for the initial installation.

13.1.5 The TS7510 Cache Controller (3955 SV5)

The TS7510 Cache Controller (3955 SV5) provides disk cache capacity for the TS7510. It is a modified IBM System Storage DS4100 (see 7.3, "IBM System Storage DS4100" on page 82) with dual controllers, and provides fourteen internal 250 GB SATA disk drives. There are four 2 Gbps FC ports to attach to the CV5 (Virtualization Engine) on dual controllers, and two 2 Gbps FC ports to attach to the SX5 (Cache Modules) on dual controllers. It also provides dual AC power. Two TS7510 Cache Controllers must be installed in the base unit 3952 Tape Frame for the TS7510.

The Cache Controller disk is configured as RAID-5 with *five data*, plus *one parity*, plus *one spare* (5+P+S), giving a usable capacity of 2.5 TB per Cache Controller, or 5 TB in a basic TS7510 without any Cache Modules. It is managed in-band from the Virtualization Engine CV5 itself - the Ethernet and serial ports are not used.

13.1.6 The TS7510 Cache Module (3955 Model SX5)

The TS7510 Cache Module (3955 Model SX5) provides additional disk cache capacity for the TS7510. It is an IBM System Storage DS4000 EXP100 (see 7.9, "IBM System Storage

DS4000 EXP100 Expansion” on page 96) expansion unit with dual enclosure service modules, and provides fourteen 250 GB SATA disk drive. There are two 2 Gbps FC ports to attach to the SV5 Cache Controller or other SX5 Cache Modules on dual enclosure service modules. It also provides dual AC power.

The Cache Module disk is configured as RAID-5 (six data, plus one parity), which leads to a usable capacity of 3 TB per SX5. This is true except for the 13th and 14th Cache Modules installed, which are configured as RAID-5 (5+P+S), giving a usable capacity for these two drawers of 2.5 TB each.

13.1.7 IBM Virtualization Engine TS7510 Management Console

The IBM Virtualization Engine TS7510 Management Console system hosts the IBM Virtualization Engine TS7510 Software. The Management Console is used to configure, manage, and service support the TS7510. This console is required by the TS7510 and can be supplied in advance, or optionally ordered from IBM. It must be installed before installing the IBM Virtualization Engine TS7510.

Note: The Management Console requires an Intel server running Windows 2000 or 2003. For detailed requirements, check the *IBM Virtualization Engine TS7510 Introduction and Planning Guide, GC26-7767*.

Figure 13-5 shows the Java based GUI interface of the TS7510 Management Console:

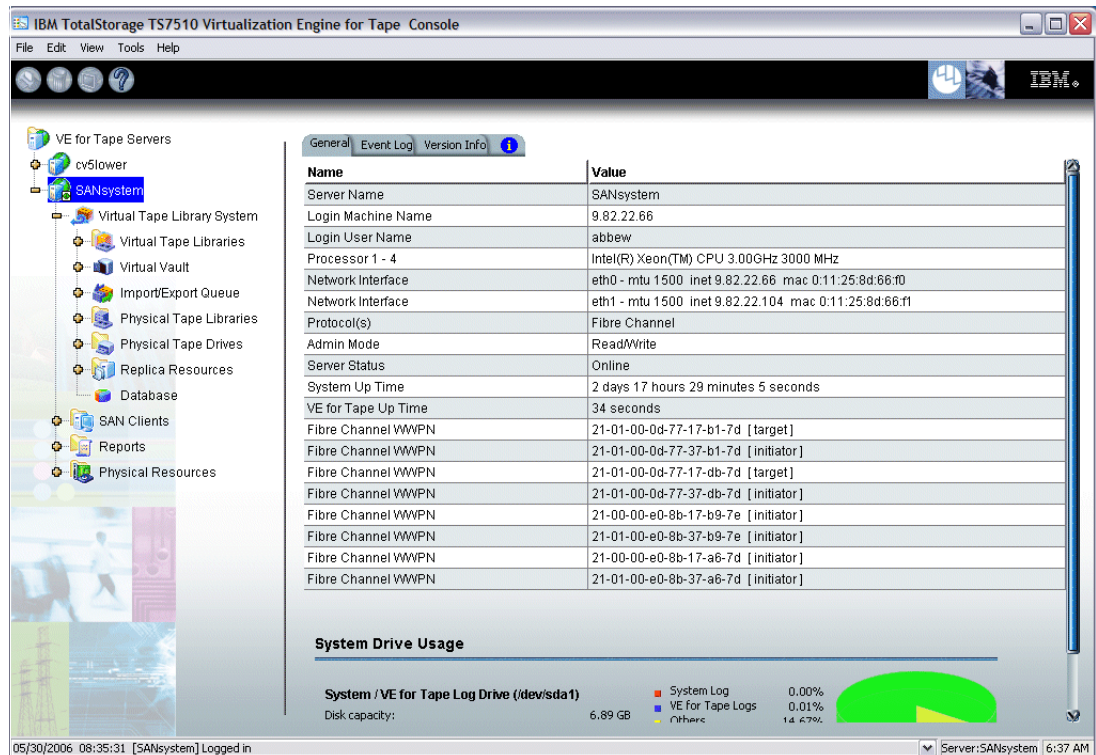


Figure 13-5 TS7510 Management Console

13.1.8 TS7510 Data management and advanced features

The IBM Virtualization Engine TS7510 application software provides tape library and tape drive emulation including virtual volumes. It also includes backup and network compression, Import and Export, Network Replication, Copy functions.

Tape library emulation

The IBM Virtualization Engine TS7510 has been designed to create a virtual copy of the IBM System Storage TS3500 Tape Library (3584). You can configure the TS7510 to act exactly like the TS3500 with any slot configuration you need for the virtual unit. The TS7510 will look like a TS3500 and act like a TS3500 to all the backup software applications that attach to it. Each backup client can attach to the tape library across the SAN the same way it did to an actual library - no extra client software is required.

Tape drive emulation

You can configure the IBM Virtualization Engine TS7510 to look like three different types of tape drives:

- ▶ IBM LTO 2 tape drives with 200 GB cartridges
- ▶ IBM LTO 3 tape drives with 400 GB cartridges
- ▶ IBM TotalStorage 3592 Tape Drive Model J1A tape drives with 300 GB cartridges

Each virtual library can have only one type of virtual tape drive within it.

Virtual tape volumes

Virtual tape volumes are associated with the drive type that is chosen. If you choose LTO 3 drives, the virtual tape volumes have LTO 3 capacity. The default size for all media types is 5 GB. This means that at least 5 GB of space will be required for all virtual tape volumes once they have been mounted into a virtual drive. While writing to a virtual volume, the volume expands its size in increments defined by the volume type.

The increment sizes are as follows:

- ▶ LTO 2 starting size is 5 GB with an incremental growth of 5 GB.
- ▶ IBM 3592 starting size is 5 GB with an incremental growth of 5 GB.
- ▶ LTO 3 starting size is 5 GB with an incremental growth of 7 GB.

Figure 13-6 shows how the virtual tape volume sizing works.

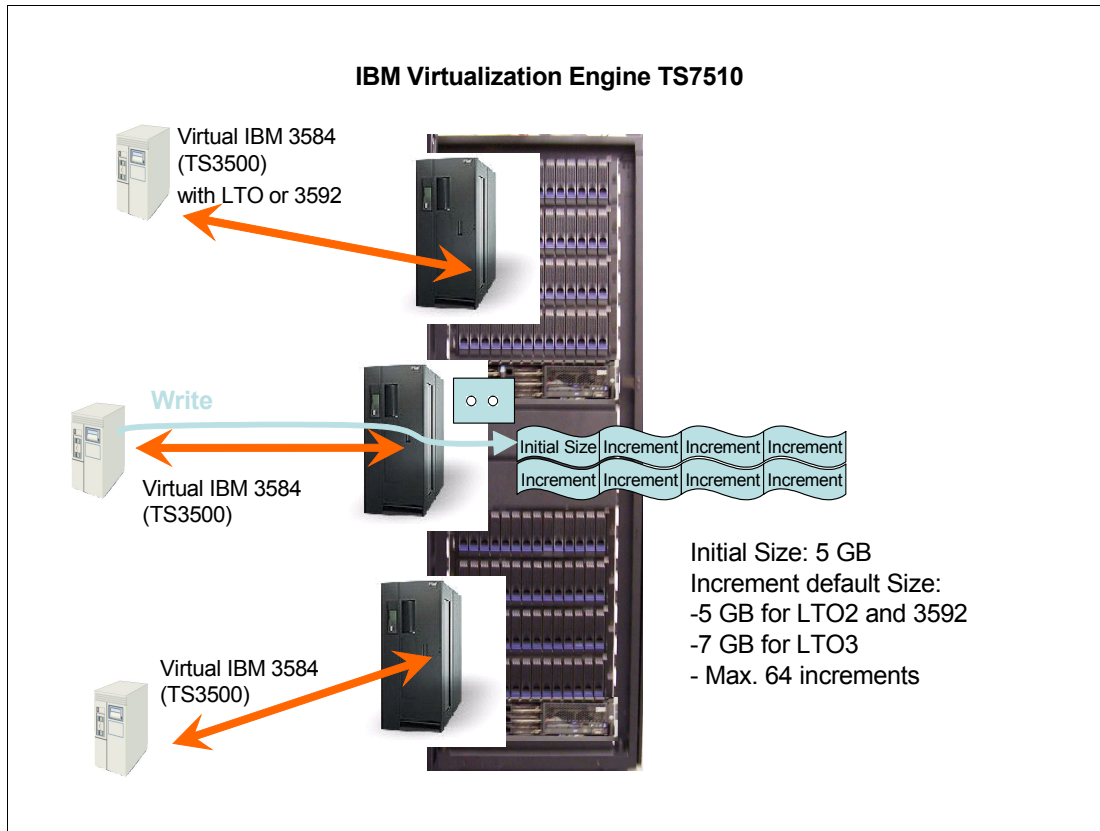


Figure 13-6 TS7510 OnDemand Virtual Tape Space Allocation

Note: When a virtual tape volume is reused by the backup application, the IBM Virtualization Engine TS7510 reduces the volume to the initialized size of 5 GB.

Backup compression

Backup compression is an option for data backup that you can turn on or off at any time from the Virtualization Engine console. Backup compression decreases the performance speed of the Virtualization Engine. If performance is not an issue, compression can significantly increase the amount of data placed into each virtual tape volume.

Import/export

Data from physical tapes can be imported into the TS7510. If it is ever required to recover files from physical tapes, the TS7510 can be used to access those tapes for immediate recovery.

The import function allows you to:

- ▶ Copy the contents of a physical tape to a virtual tape
- ▶ Directly access a physical tape without copying the entire tape
- ▶ Recycle a physical tape

For additional data protection, the data on virtual tapes can be exported to physical tapes for long-term data archiving. Data can also be copied to physical tapes using your backup application's copy function. When you export, the TS7510 copies/moves the contents of the virtual tape to a physical tape. Because some backup applications alter what they write to the tape depending on the type of cartridge used, the TS7510 only exports tapes to like media.

You cannot export to a dissimilar physical tape. Exporting tape to like media enables the backup application to accept the tape as valid; from the backup application's point of view, there is no difference between the virtual and physical tape.

Figure 13-7 illustrates the overall process of the Import and Export commands.

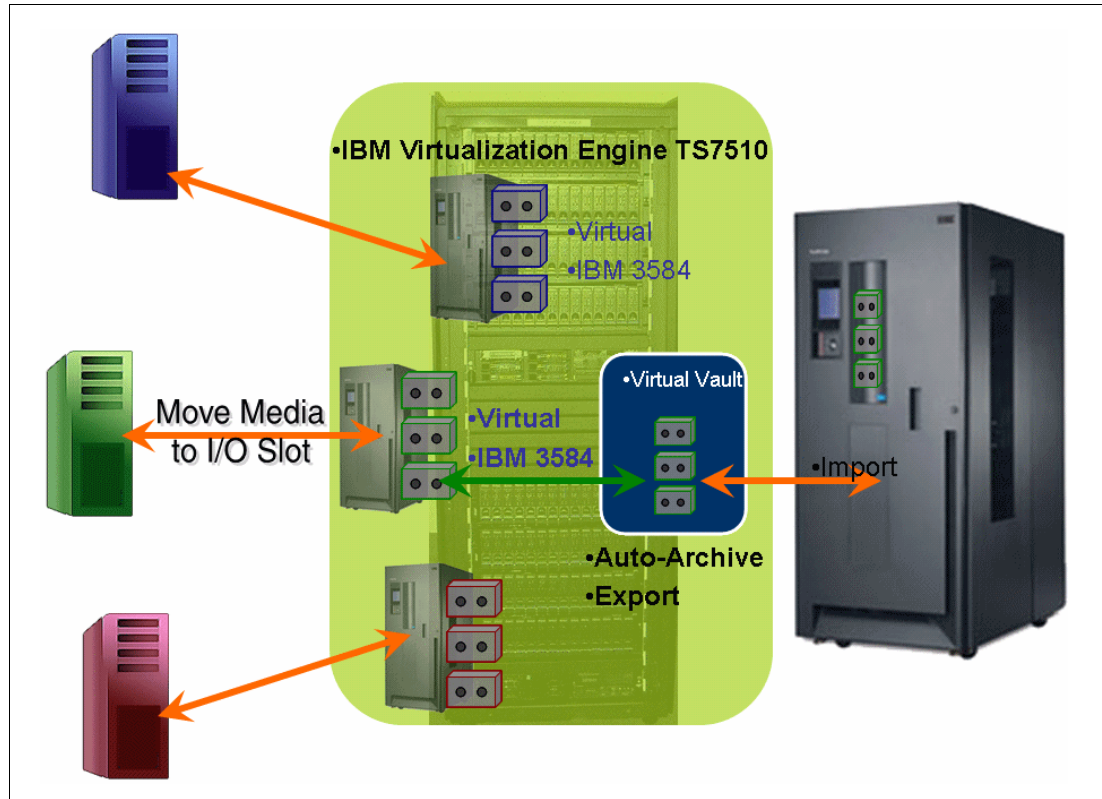


Figure 13-7 TS7510 Import/Export function

Note: If you will be importing data from physical tapes into your virtual tape library or exporting virtual tapes to physical tapes, you must assign your physical tape libraries/drives to the TS7510.

Auto archive

Auto Archive (also called “Export Physical Copy” or “Physical Copy on Export”) exports data to physical tapes whenever a virtual tape is exported from a virtual library (such as from a backup application or other utility). In order to use this feature, the physical tape library must support barcodes. When the TS7510 attempts to export to a physical tape it must find a matching barcode in a physical library (you do not need to specify which physical library). If you use Auto Archive, determine if you want the virtual tape copied (retained) or moved (removed) after the data is transferred. If you select to move the virtual tape, indicate how long to wait before deleting it. Also, indicate if you want to export your physical tapes after archiving.

Network replication

Replication protects the data on a virtual tape by maintaining a remote copy of the virtual tape on another TS7510 (called the target server) over the IP network. At prescribed intervals, when the tape is not in use, changed data from the primary virtual tape is transmitted to the replica resource on the target server so that they are synchronized. The target server is

usually located at a remote location. Under normal operation, backup clients do not have access to the replica resource on the target server. If a disaster occurs and the replica is needed, the administrator can promote the replica to become the primary virtual tape so that clients can access it.

Figure 13-8 illustrates replication movement.

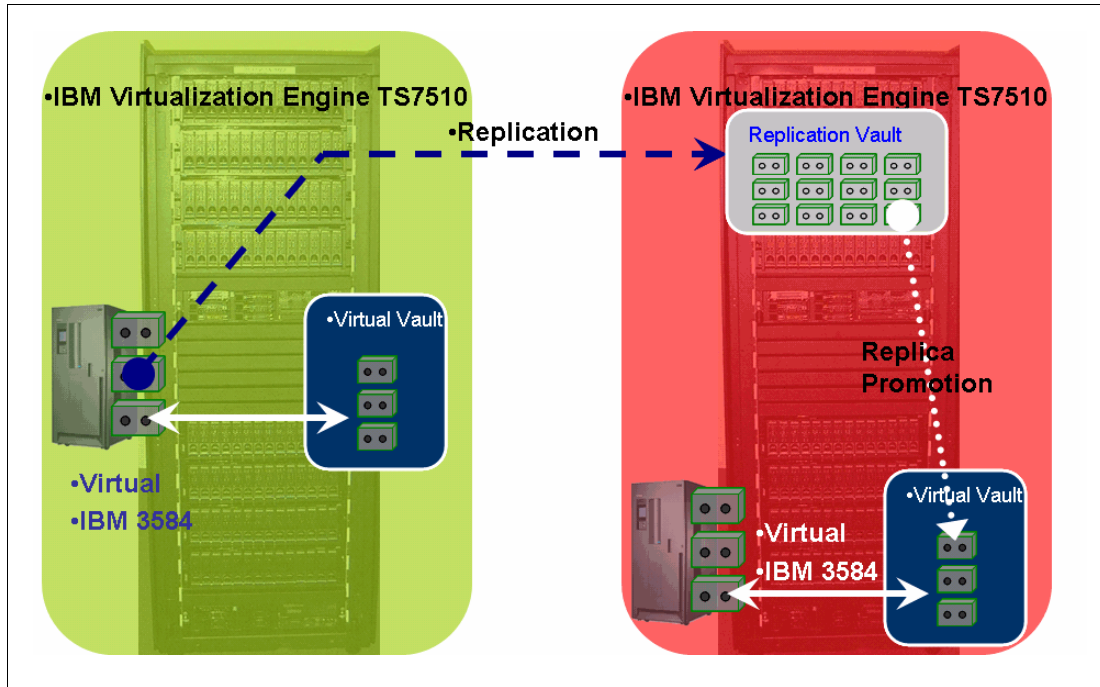


Figure 13-8 TS7510 Network Replication function

Auto replication

Auto Replication (also called “Export Network Copy” or “Network Copy on Export”) replicates data to another TS7510 whenever a virtual tape is exported from a virtual library (such as from a backup application or other utility). If you use Auto Replication, determine whether you want the virtual tape copied (retained) or moved (removed) after the data is replicated.

Figure 13-9 illustrates the Auto Replication process. The left side shows the primary engine, and the right side shows the backup engine. The primary initiates the Auto Replication function.

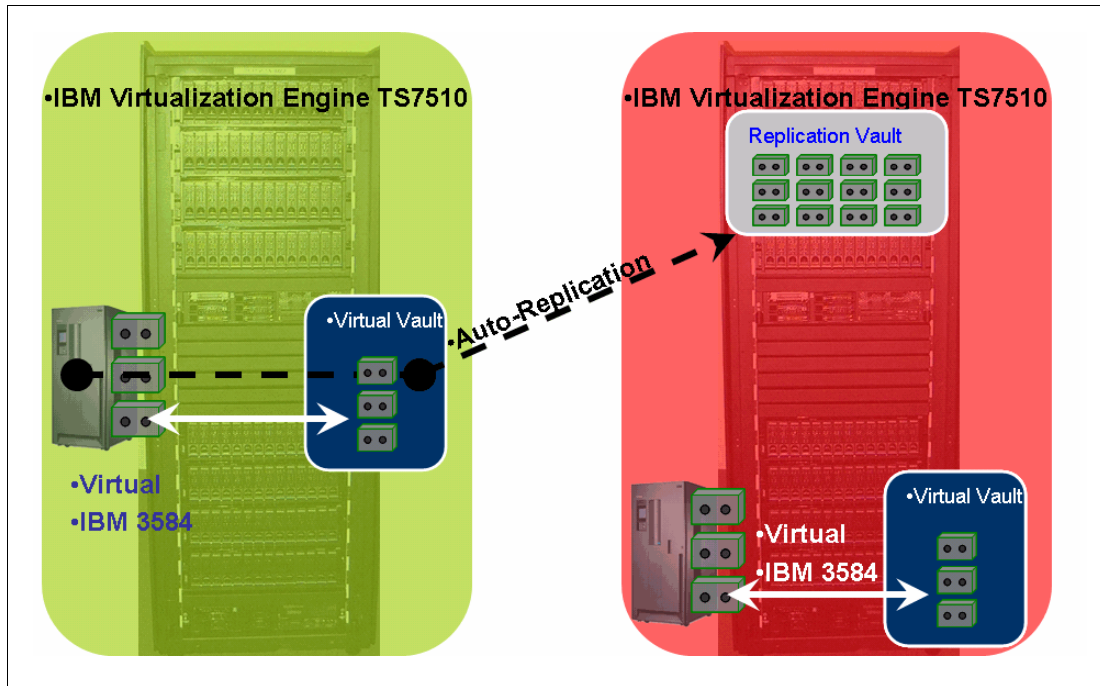


Figure 13-9 TS7510 Auto Replication function

Remote Copy

You can copy the contents of a single tape to a remote target server. Because the Remote Copy feature replicates the full tape rather than appending to an existing virtual tape, you can only copy a tape if there is no virtual tape on the remote target server with the same barcode. Therefore, if you have copied this tape before, you must delete the copy from the remote target server before continuing.

Figure 13-10 illustrates the Remote Copy movement. The primary Virtualization Engine is on the left, and the remote backup is on the right.

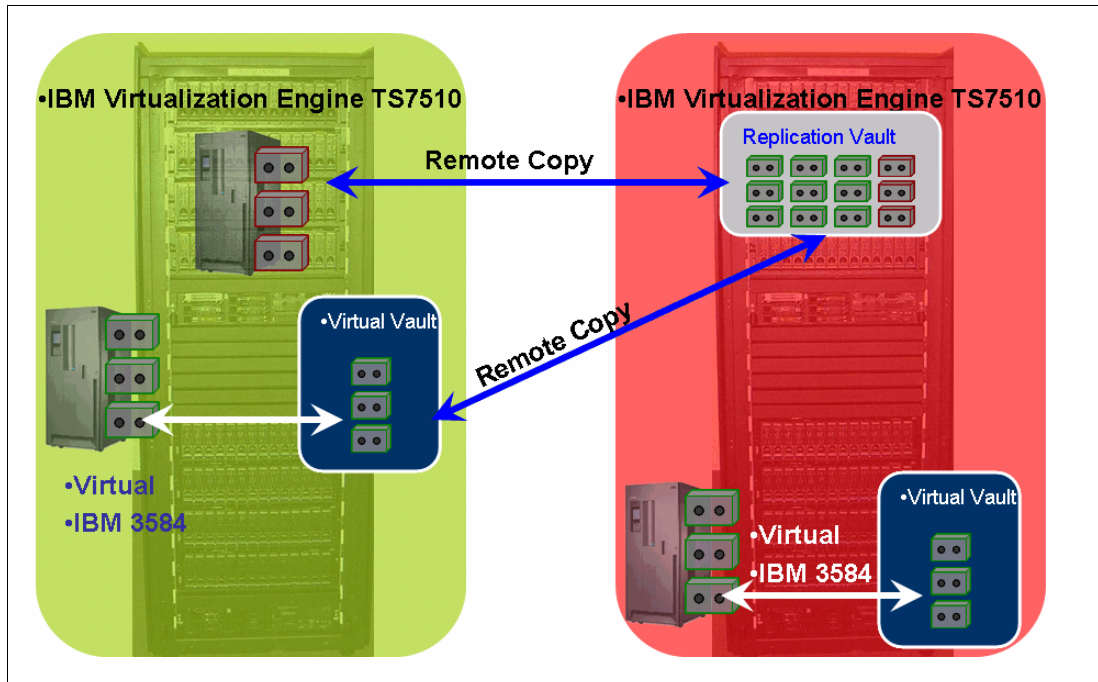


Figure 13-10 TS7510 Remote Copy function

Failover/failback

The TS7510 failover/failback option provides high availability for TS7510 operations by eliminating the down time that can occur should a TS7510 Virtualization Engine Server (software or hardware) fail. In the TS7510 failover design, a TS7510 Virtualization Engine Server is configured to monitor another TS7510 Virtualization Engine Server. In the event that the server being monitored fails to fulfill its responsibilities to the SAN Clients it is serving, the monitoring server will take over its resources.

13.1.9 IBM Virtualization Engine TS7510 solution areas

The TS7510 can form part of the storage solutions areas.

TS7510 and Information Life Cycle Management (ILM)

The TS7510 introduces a two tiered storage system consistent with the Information Life Cycle Management (ILM) tiered storage model. The disk cache can provide rapid access to more frequently used backup data. Less frequently accessed data is migrated to tape, providing cost-effective storage.

TS7510 and Business Continuity (BC)

The TS7510 offers improved restore times for volumes resident in the cache. In combination with the optional IP replication and failover features, and IBM tape libraries it provides a very high level of business continuity. The IP replication feature is particularly attractive for IT organizations looking to reduce the need for the physical transportation of tapes, while the failover/ failback feature helps improve the availability of the virtualization systems.

TS7510 and Infrastructure simplification (IS)

The TS7510 assists with Infrastructure simplification by allowing multiple, disparate backup servers and applications to share the same virtual tape library. This can help reduce the need

for multiple, disparate tape libraries and improve the overall utilization of the tape storage systems.

13.2 IBM TotalStorage 3494 Virtual Tape Server

This section covers IBM enterprise tape virtualization product, the IBM TotalStorage 3494 Virtual Tape Server.

13.2.1 Introduction to IBM TotalStorage 3494 Virtual Tape Server

The demand for storage and 24x7 access to data increases every day. Meeting this demand is critical, especially given the heightened importance of data backup and the role of storage technology in disaster recovery. Cost control is a key issue as enterprises try to operate more efficiently. To address these requirements, many corporations have turned to tape storage as a strategic tool.

The IBM TotalStorage 3494 Virtual Tape Server (VTS), shown in Figure 13-11, is an enterprise virtual tape system. It combines the benefits of high-speed disk access with low-cost tape storage by virtualizing data in a RAID-5 disk cache buffer, which helps eliminate many of the physical delays associated with tape read/write activities. This approach allows the VTS to help boost system performance and reduce batch processing times and backup windows.



Figure 13-11 IBM TotalStorage 3494 Virtual Tape Server

Because data written to tape is often reused shortly after it is stored, the VTS creates virtual volumes in a disk cache for later read access. Based on a Least Recently Used (LRU) algorithm, these volumes are retained on disk even after the VTS has written a copy to tape. If the data requested from a host is not in the cache, the full virtual volume is retrieved from the cartridge, uploaded into the virtual cache and then transferred to the host from the cache. In practice, the VTS satisfies most mount requests by accessing the virtual volumes in the disk cache.

In enterprise applications, most cartridges stored in conventional tape libraries contain one volume per tape, a poor use of space with the data often occupying less than half a physical tape. The VTS can fully utilize cartridge capacity, shown in Figure 13-12. This can

dramatically reduce the number of tapes in the library, the amount of floor space needed to store tapes, and the work associated with it.

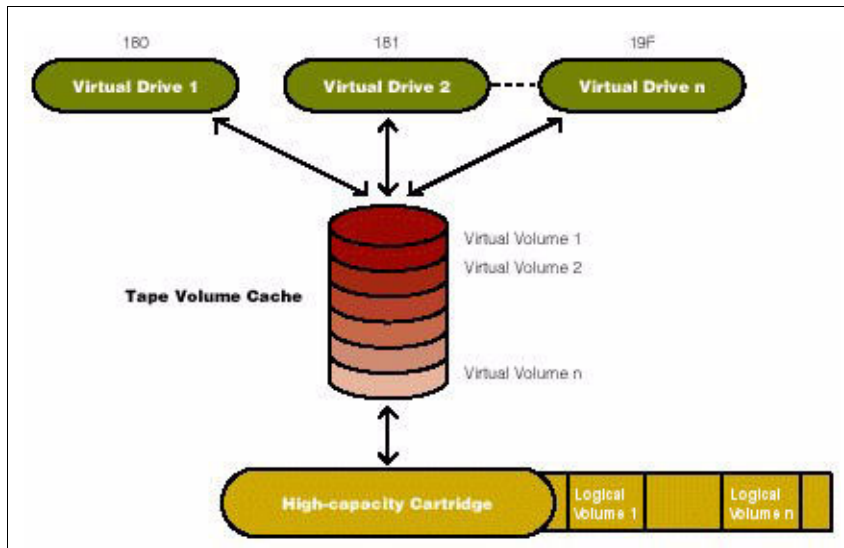


Figure 13-12 VTS virtual drives and hard disk caching

The VTS runs automated storage management software that optimizes the use of tape and the disk cache. Over time, some of the stacked volumes residing on the physical tapes become invalid, while other data must be retained for years. The VTS automatically copies the valid logical volumes to a new stacked tape and reclaims the free tape space for new data. The VTS also automates the migration of data to higher capacity volumes and the export of data to other VTS systems for disaster backup, workload balancing and archiving purposes.

Furthermore, the VTS provides operational flexibility by eliminating one bottleneck of a conventional tape environment: the number of physical tape drives available. The VTS can offer hundreds of logical tape devices that can be easily allocated to any host application.

The VTS can be attached to IBM and non-IBM servers using FICON, ESCON and SCSI connections. A VTS can be supported at distances of up to 100km using channel directors or switches. Even greater distances are supported with WAN or ATM connections that use supported channel extension products.

The VTS can also extend the current level of security and disaster tolerance. By installing the system at a remote site, servers can be backed up across the network, eliminating the need to transport and manage cartridges off-site. Alternatively, a configuration of two VTS systems coupled by peer-to-peer controllers can automatically duplicate tape data and can further protect against data loss in case of failures. These coupled systems can also be installed at two geographically distant sites to minimize potential server downtime in case of a disaster. Because the VTS systems reside in two locations, operations can typically continue even if one of the facilities becomes nonoperational during planned maintenance, service or system upgrade.

A Peer-to-Peer (PtP) VTS can also be integrated into an IBM Geographically Dispersed Parallel Sysplex (GDPS) environment (Tier 6 for disaster recovery). GDPS uses automation software to switch sites quickly, managing the complete switch of all resources from one site to another as required during a planned or unplanned failover.

For ESCON attached servers, the PtP VTS can support distances up to 50 km using the supported communication extension solutions. With FICON attachment, supported distance is up to 100 km.

13.2.2 Virtual Tape Server logical components and functions

In this part of the chapter we describe the key logical components and functions of a Virtual Tape Server.

Emulation of 3490E tape drives

From a System host or server perspective, the VTS emulates 3490E subsystems, each with 16 tape drives. This provides a total of 64, 128, or 256 virtual tape depending on the configuration, shown in Figure 13-13.

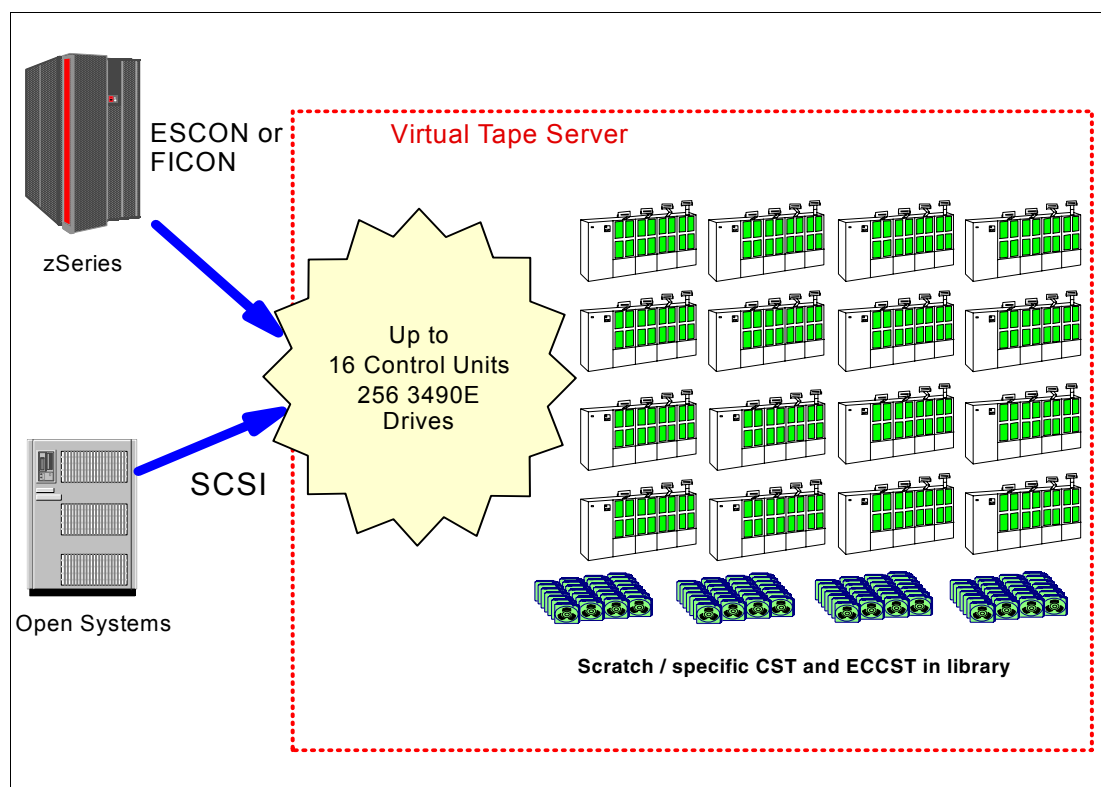


Figure 13-13 VTS Virtual Drives emulation

Each emulated drive is called a *virtual tape drive*. The subsystem handles all 3490E tape commands. There is no direct relationship between a virtual tape drive and the real tape drive. Data is written and read as though it is stored on a real media cartridge. The amount of data stored on a *virtual volume* is variable up to a maximum as determined by the media type selected.

Virtual Volumes

A virtual volume is created in the Tape Volume Cache (TVC) when the host writes data to the VTS subsystem. All host interaction with tape data in a VTS is through virtual volumes and virtual tape drives. Each virtual volume has the same characteristics as a real volume.

Tape Volume Cache (TVC)

After an application closes a virtual volume, if it was modified, a copy of it is made by the VTS storage management software onto a physical tape. The virtual volume remains available on the disk storage until the space it occupies is needed to satisfy another mount request. The disk storage thus acts as a cache for the tape volumes, providing for fast access.

Logical Volumes

When a virtual volume is copied from the TVC to a physical tape cartridge, it becomes a logical volume. When a logical volume is moved from a physical cartridge to the TVC, the process is called *recall* and the volume becomes a virtual volume again.

Figure 13-14 illustrates the relationship between virtual and logical volumes.

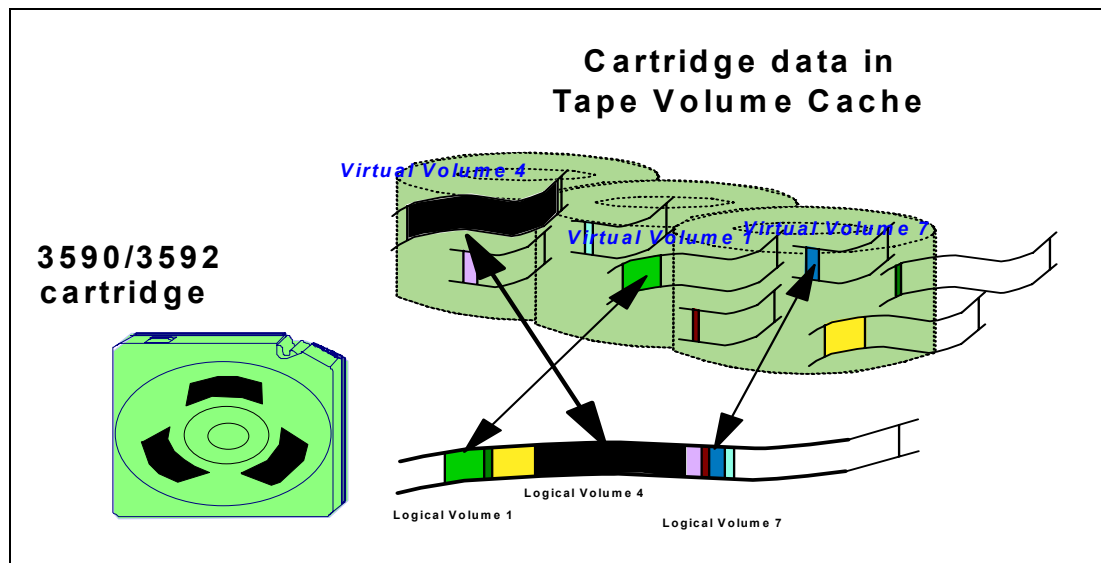


Figure 13-14 Data in the TVC and on physical tape

As virtual volumes are copied from the TVC to a physical cartridge, they are stacked on the cartridge end to end, taking up only the space written by the host application. This arrangement maximizes utilization of a cartridge's storage capacity. The storage management software within the VTS manages the location of the logical volumes on the physical cartridges. The only control the user has over the location of the data is where volume pooling is used as part of Advanced Policy Management. Without Advanced Policy Management there is no control.

Advanced Policy Management

Advanced Policy Management (APM) is an optional feature which provides:

- ▶ Physical volume pooling allows grouping of logical volumes onto a specified range of physical volumes, with:
 - Various reclamation policies
 - Secure data erase
- ▶ Selective Dual Copy written to separate physical volumes
- ▶ Peer-to-Peer Copy mode control
- ▶ Tape Volume Cache Management - allow volumes to remain in the cache for extended periods of time
- ▶ Extended logical volume sizes - up to 4 GB uncompressed capacity

- ▶ Bulk Volume Information Retrieval (BVIR)
- ▶ Export/Import

This gives you control over the handling of logical and physical volumes. It allows use of DFSMS Automated Class Selection (ACS) routines (see 26.1, “Managing Enterprise data” on page 484).

Stacked volumes

The physical cartridges used by the VTS to store logical volumes are completely under the control of the VTS and are not known to the hosts. The physical volumes are called *stacked volumes*. The stacked volumes must have unique machine-readable volumes and external labels like any other cartridges in a tape library.

Through the Library Manager Console, you define which physical cartridges are to be used by the VTS. Logical volumes stored on those cartridges are mapped by the internal storage management software. With Advanced Policy Management installed, your stacked volumes can be assigned to individual pools. Logical volumes can then be assigned to the stacked volume pools.

13.2.3 IBM TotalStorage 3494 Virtual Tape Server Model B10 and B20

Currently, the VTS models B10 and B20 are available. The first and second generation VTS models B16 and B18 are no longer available, but existing models can still be upgraded to the latest models.

The third generation of Virtual Tape Server products provides an entry model, the Model B10 VTS, and an enhanced model, the Model B20 VTS. These new models integrate the advanced performance, capacity, and data integrity design of the 3592 and 3590 tape drives with high-performance disk and a new advanced RISC controller to form a storage hierarchy managed by robust storage management firmware with extensive self management capability. They provide a variety of host attachments, disk storage capacities, and tape drive attachments. They include also functions such as advanced policy management to control physical volume pooling, cache management, secure data erase, import/export, dual copy, peer-to-peer dual copy, copy mode control, and larger logical volumes.

The VTS is attachable to an IBM TotalStorage 3494 Tape Library with IBM 3590 or 3592 tape technology or to an IBM System Storage TS3500 Tape Library (IBM 3584 Tape Library) with the IBM 3953 Library Manager and with IBM 3592 tapes. Supported 3590-B1A, E1A, or H1A drives reside in a D12 Frame, 3592 drive models J1A and E05 reside in an L22 or D22 frame.

Model B10

The Model B10 VTS provides host connection of up to four FICON channels, four Extended Performance ESCON Channels, up to eight SCSI bus attachments, and up to 1.3 TB of tape volume cache capacity. Each Model B10 VTS provides a maximum of 64 virtual tape drives and up to 250,000 logical volumes each with a maximum capacity of 2.4 GB (assuming 3:1 compression) or 6 GB (assuming 3:1 compression and using the 2000 MB volume size and the larger logical volume size support). A Model B10 VTS with the Peer-to-Peer Copy features can be interconnected with another Model B10 or B18 VTS using four IBM TotalStorage Virtual Tape Controllers (VTCs) in an IBM TotalStorage Peer-to-Peer VTS (PtP VTS) configuration to provide both remote and local peer-to-peer copy capability between two VTSs using ESCON connections.

Model B20

The Model B20 VTS provides host connection of up to eight FICON channels, up to 16 Extended Performance ESCON Channels, up to eight SCSI bus attachments, and over 5 TB of tape volume cache capacity. Each Model B20 VTS provides a maximum of 256 virtual tape drives and up to 500,000 logical volumes each with a maximum capacity of 2.4 GB (assuming 3:1 compression) or 12 GB (assuming 3:1 compression and using the 4000 MB volume size and the larger logical volume size support). A Model B20 VTS with the Peer-to-Peer Copy features can compression). A Model B20 VTS with the Peer-to-Peer Copy features can be interconnected with another Model B20 or B18 VTS using four or eight IBM TotalStorage Virtual Tape Controllers (VTCs) in an IBM TotalStorage Peer-to-Peer VTS (PtP VTS) configuration to provide both remote and local peer-to-peer copy capability between two VTSs using ESCON or FICON connections.

The minimum and maximum configurations of the B10 and B20 VTS models are shown in Table 13-3.

Table 13-3 Virtual Tape Server Model B10 and B20 specifications

Specifications	B10 Minimum	B10 Maximum	B20 Minimum	B20 Maximum
Tape volume cache (GB)	648	1,296	2,592	5,184
Virtual drives	64	64	128	256
3590 tape drives	4	6	4	12
3592 tape drives	4	12	4	12
3590/3592 tape drives	N/A	N/A	4/4	6/12
Virtual volumes	250,000	250,000	250,000	500,000
FICON-only channels	2	4	4	8
ESCON-only channels	2	4	8	16
SCSI-only channels	8	8	N/A	N/A
FICON/ESCON channels	2/2	2/2	4/4	4/8
ESCON/SCSI interfaces	4/4	4/4	8/8	8/8

The VTS frames contain the VTS controller, its associated storage management software and the RAID disk arrays that make up the TVC. The Model B10 VTS or Model B20 VTS frame is a stand-alone unit that can be located a maximum of 14 m from the associated 3590 drive frame and up to 25 m from a 3592 drive frame, when installed in an IBM TotalStorage 3494 Tape Library. If the VTS is connected to an IBM System Storage TS3500 Tape Library/3953 configuration, it can be located up to 62 m (200 feet) away from the D22 frames of the TS3500 as shown in Figure 13-15.

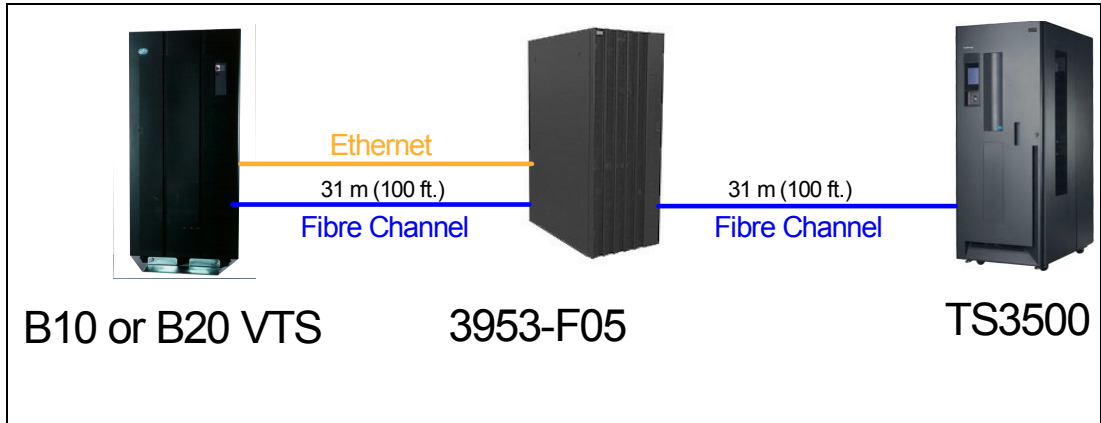


Figure 13-15 VTS distance to IBM 3584 and 3953

Several frames are needed in addition to the Bxx frame for the VTS: library frames for its drives and the IBM 3494 Library Manager hardware, which is installed in the IBM 3953 Library Manager frame or if attached to an IBM TotalStorage 3494 Tape Library in the Library Base Frame (Lxx Frame). Figure 13-16 shows a diagram of the physical hardware used for the VTS.

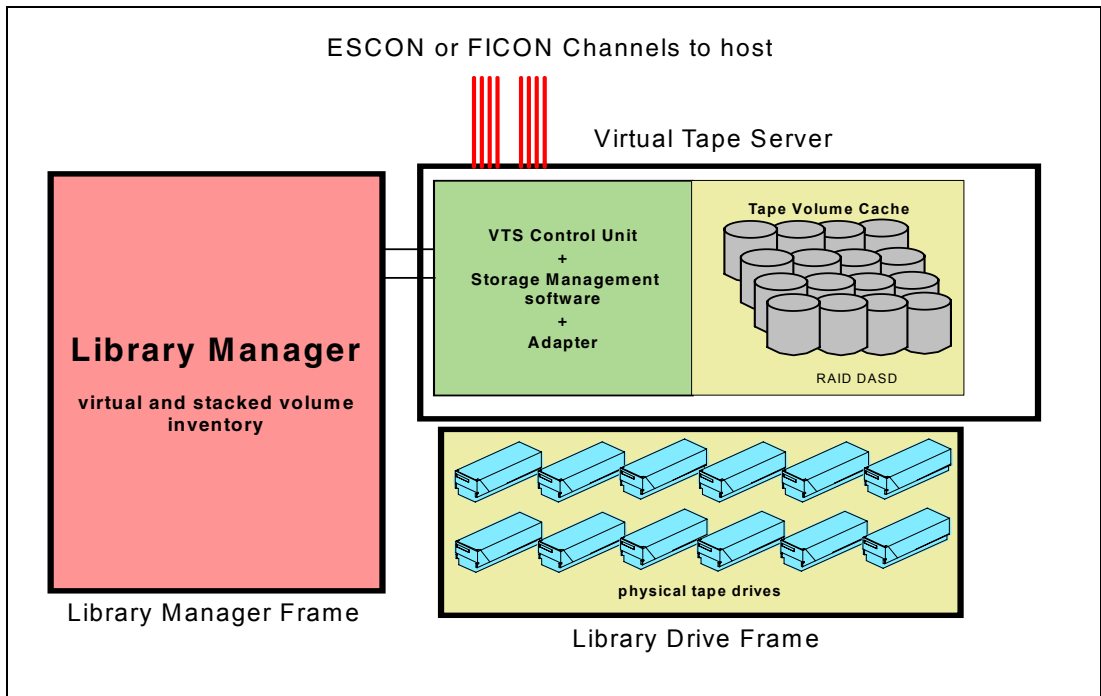


Figure 13-16 Hardware components of the VTS transparent to the host system

VTS Import/Export capability

Volumes required for disaster backup can be exported from a stand-alone VTS using the import/export facility of the Advanced Function feature. These volumes can be stored remotely and entered into a disaster recovery VTS using the import facility of the Advanced Function and Advanced Policy Management features.

13.2.4 IBM TotalStorage Virtual Tape Frame 3494 Model CX1

The IBM TotalStorage Virtual Tape Frame 3494 Model CX1 is one of the components of the IBM TotalStorage Peer-to-Peer Virtual Tape Server (PtP VTS). A PtP VTS uses robust distributed servers, Virtual Tape Controllers (VTCs) housed in the Virtual Tape Frame Model CX1, to interconnect two VTSs (Model B10, or B20) into a single PtP VTS system for FICON or ESCON attachment to a host system. The VTS Models B10 and B20 can connect to ESCON or FICON channels. The Model CX1 provides the housing and power for two or four VTC features that interconnect two VTSs. There are two power control compartments, each with its own power cord to allow connection to two power sources.

A component overview is shown in Figure 13-17.

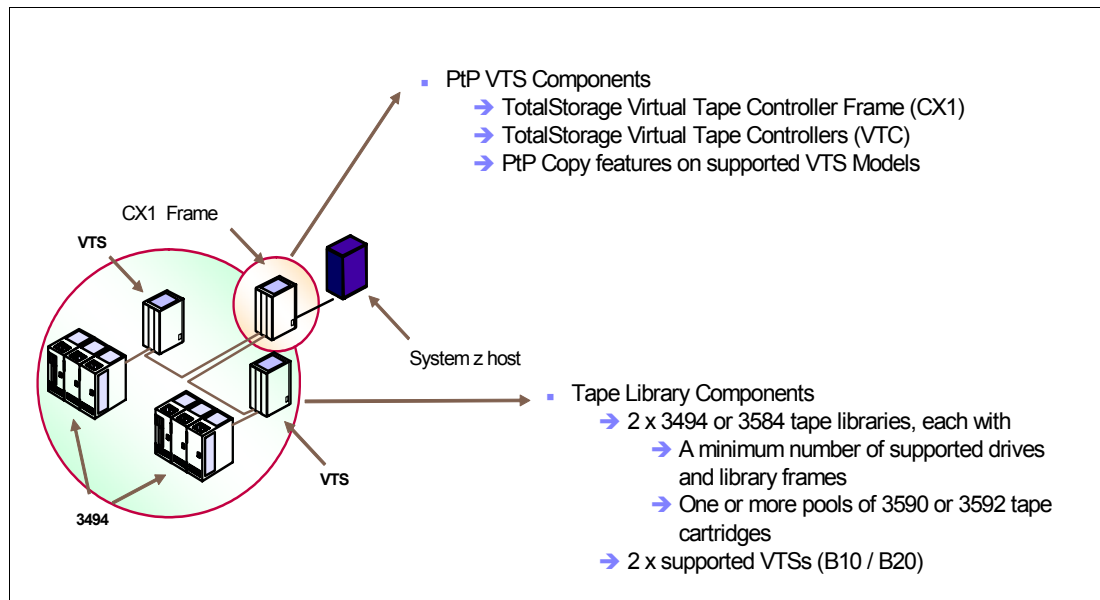


Figure 13-17

13.2.5 IBM TotalStorage 3494 Peer-to-Peer Virtual Tape Server

The VTS 3494 Models B10, B18, or B20 can include Peer-to-Peer Copy features. These features allow two VTSs to be interconnected into a single system to form a PtP VTS with no single point of failure that prevents access to data. The differences about the physical and the logical view are shown in Figure 13-18.

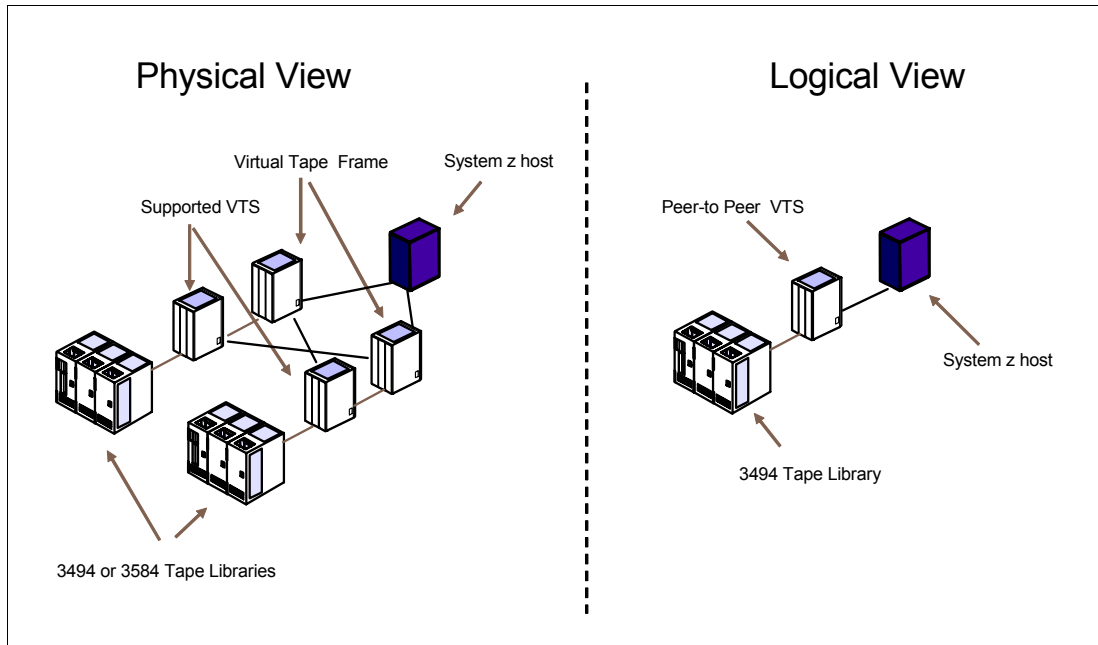


Figure 13-18 PtP VTS physical and logical view

Each of the VTSs in a PtP VTS continues to operate during scheduled or unscheduled service of the other VTS. The VTS Peer-to-Peer Copy features provide automatic copy of virtual volumes into both VTSs in a PtP VTS. When a volume is requested, the PtP VTS automatically retrieves one of the copies from either VTS. For even greater data protection, the VTSs in a PtP VTS can be in different locations to protect against disasters.

A VTS Model B10, B18, or B20 with the Peer-to-Peer Copy features is an independently operating distributed server node in a PtP VTS. It incorporates all the functional capability of the VTS Model B10, B18, or B20 without Peer-to-Peer Copy features, except for SCSI host attachment and the Import/Export capability of Advanced Function. The VTS incorporates new channel commands to support large block transfers of compressed logical volumes to and from the VTC. These commands reduce the time needed to copy logical volumes, especially across extended distances. Other functions allow the VTC to maintain copy synchronization of the two VTSs and provide expedited deletion of redundant logical volume copies from the tape volume cache.

The PtP VTS enhances those capabilities by maintaining a copy of all virtual volumes in both VTSs of the PtP VTS. The VTCs will automatically create the second copy and maintain the copy synchronization of the two VTSs. The accessibility of the copy will depend on the mode of operation selected:

- ▶ Immediate Copy - Creates a copy in the companion connected VTS upon receipt of a Rewind Unload command and provides the highest level of protection of your data in the VTS
- ▶ Deferred Copy - Schedules creation of a copy in the companion connected VTS after receipt of a Rewind Unload command as activity permits

All VTCs within a PtP VTS must use the same mode of operation.

The specifications of the B10/B10 and B20/B20 Peer-to-Peer VTS Server are shown in Table 13-4 on page 278.

Table 13-4 Peer-to-Peer Virtual Tape Server specifications

Specifications	B10/B10	B20/B20
Virtual Tape Controller Features	4	4/8
Tape Volume Cache (TB)	1.3/2.6*	5.2/10.4*
Virtual drives	64	128/256*
Physical tape drives (3590)	8/12*	8/24*
Physical tape drives (3592)	8/12*	8/24*
Physical tape drives (3590 and 3592 mix)	NA	8/12* (3590) and 8/24* (3592)
Virtual volumes	250,000	250,000
ESCON channels	8	8/16*
FICON channels	8	8/16*
FICON/ESCON channels	4/4	4/4 to 8/8*
Supported environments	z/OS, z/VM and VM/ESA® environments	
Note: * Indicates minimum/maximum configurations.		

It is highly recommended that both VTSs be configured with the same disk storage capacity. A PtP VTS will appear to attached host processors as a single automated tape library.

Remote Operation

The components of the PtP VTS can be physically separated to provide remote copy of virtual volumes, shown in Figure 13-19. One of the VTSs with its would be in one location and the other VTS with its in another location. The required four or eight VTCs can be flexibly placed between the two locations.

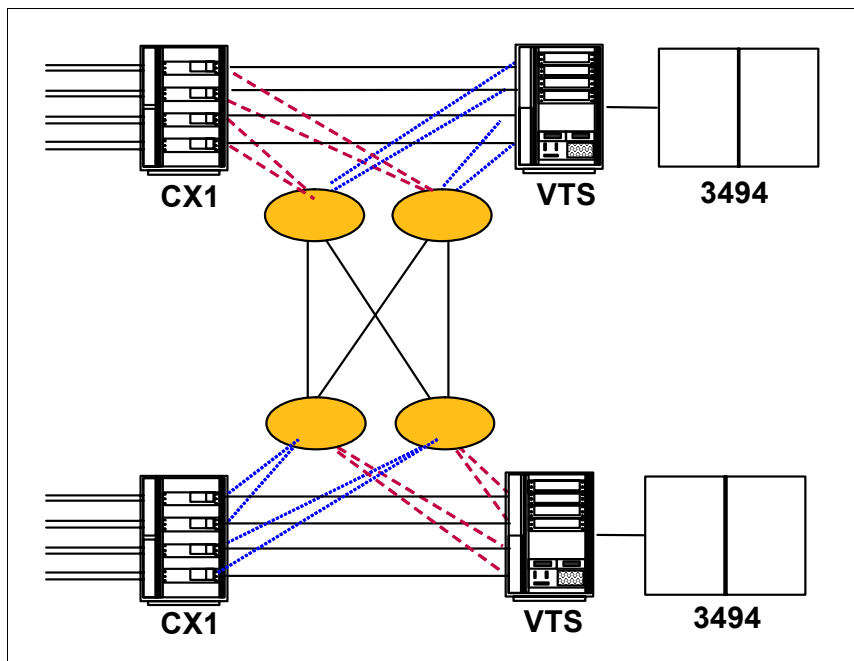


Figure 13-19 Remote PtP VTS implementation

ESCON directors, Fibre Channel Directors, or Dense Wave Division Multiplexors (DWDM) can be used to extend the maximum distance between components as shown in Table 13-5.

Table 13-5 Maximum distances for PtP VTS

	ESCON	ESCON with DWDM	FICON
Host Server to VTC	43 km	75 km	100 km
VTC to VTS	26 km	50 km	100 km

Note: The maximum **end-to-end distance** of the host system to VTC ESCON interface, including the DWDM pair length, is 75 km. The maximum **end-to-end distance** of the VTC to VTS ESCON interface, including the DWDM pair length, is 50 km.

The VTS FICON attachments support 2 Gbps link speed when attached to a System z server, or an appropriate FICON/FC switch, with 2 Gbps FICON features and the appropriate levels of software to utilize the 2 Gbps capability. Extended distance connections can be provided by channel extenders (virtually unlimited distance) certified for use by providers with a PtP VTS ESCON configuration.

Geographically Dispersed Parallel Sysplex (GDPS) Control

The PtP VTS configuration allows the GDPS service offering to include tape data in its management of data consistency and integrity across sites. An I/O VTS selection option is provided especially for use with GDPS such that all virtual volumes are processed from a primary VTS, and a copy is stored on the secondary VTS. Control capability has been added to allow GDPS to 'freeze' copy operations so that tape data consistency can be maintained across GDPS managed sites during a switch between the primary and secondary VTSs, and also provide synchronization of system data sets such as catalogs, the tape control database, and tape management databases with the PtP VTS after an emergency switchover.

Expanded DFSMS Tape Library Application Interface

PtP VTS operations can be user controlled through use of the expanded DFSMS Tape Library Application Interface. You can now dynamically control the copy mode of operation and I/O VTS selection mode. Additional functions include enable/disable copy operation, and query virtual drive and volume status. Users can tailor the control of their PtP VTS by writing their own application program that exploits the application interface.

Virtual Tape Controller

A Virtual Tape Controller (VTC) in the IBM TotalStorage Virtual Tape Frame 3494 Model CX1 provides interconnection between two VTSs with the Peer-to-Peer Copy features, and provides two host attachments for the PtP VTS. There must be four (for the Model B10 or B18) or eight (for the Model B20 or B18) VTCs in a PtP VTS configuration. Each VTC is an independently operating, distributed node within the PtP VTS and continues to operate during scheduled or unscheduled service of another VTC.

Each VTC also provides two ESCON or two FICON host server attachments, 16 or 32 (with a B20 only) virtual addresses, and transfers data between host channels and the VTSs. Each ESCON attachment supports 64 logical paths and each FICON attachment supports 128 logical paths. A VTC with host server FICON attachment will also have two FICON attachments for VTS connectivity. Similarly, a VTC with host server ESCON attachment will have two ESCON attachments for VTS connectivity. Both Long Wavelength and Short Wavelength FICON connection are provided for server and VTS attachment. Within a Peer-to-Peer VTS, ESCON and FICON VTC intermix is provided.

VTC Operation

For a virtual mount on any of its addresses, the VTC will select one of the two VTSs to support the requested volume processing activities.

Each VTC performs the following:

- ▶ Maintains synchronization of the copy of logical volumes
- ▶ Creates logical volume copies using large block transfers of compressed logical volumes
- ▶ Balances workload between the VTSs
- ▶ Directs specific volume mounts to the VTS with a cached copy of the requested virtual volume
- ▶ Provides status displays on a customer-selected Web site via the customer's intranet

When one of the VTSs is off-line for service, all activity will be directed to the remaining active VTS. When a VTS is returned to an active state, each of the Virtual Tape Controllers will resume copy operations to bring the VTSs into synchronization.

PtP Specialist

Each VTC provides an optional Web user interface to the IBM TotalStorage Peer-to-Peer Virtual Tape Server Specialist. The PtP Specialist provides a Web page that may be used to access PtP configuration, status, statistics, and online product documentation. This Web page may only be accessed through the Ethernet LAN connection provided on each VTC. The configuration and management of the Ethernet LAN connection must be provided by the user and its use is optional.

Standby Virtual Tape Controllers

An additional 3494 Model CX1 with four VTCs can be installed as a standby for four active VTCs in a PtP VTS configuration. The standby VTCs at the remote site are intended for use where the VTSs are physically separated to provide disaster backup. If all the VTCs at the primary site become unavailable, then, with intervention by a service representative, the remote VTS and the standby VTCs can be started in read/write disconnected mode of operation to provide read and write access to volumes stored in the VTS. A host system must be configured and available to the remote site.

“Call Home” Service Support

When activated, the “Call Home” function automatically opens a service alert when a problem occurs with the 3494 Models B10, B20, or CX1. The “Call Home” process can also be initiated by a user from the library manager operator panel.

13.3 Enterprise Removable Media Manager

This section covers Enterprise Removable Media Manager (eRMM), which is an IBM service offering for advanced tape management.

Enterprise Removable Media Manager provides features known from the mainframe's DFSMSrmm™ for open systems. It complements IBM Open Storage Software Family to provide storage virtualization and advanced storage management for removable media. eRMM automatically configures drives for IBM Tivoli Storage Manager and it gathers audit trails and statistical data for the complete cartridge lifecycle.

13.3.1 Introduction to Enterprise Removable Media Manager

Management for removable media is one of the biggest challenges for today's heterogeneous open systems tape environments.

The following issues are widespread with tapes and tape media:

- ▶ Tape resources are statically linked to applications
- ▶ Resource sharing in heterogeneous and even in large homogeneous configurations is very limited
- ▶ Adding or changing tape hardware requires changes to every application
- ▶ Cartridge management has to be done by each application
- ▶ Multitude of media changer and/or management interfaces (SCSI, IBM 3494, STK ACSLs)
- ▶ Lack of centralized management and monitoring

Figure 13-20 shows the static link between applications and tape library hardware in traditional tape environments.

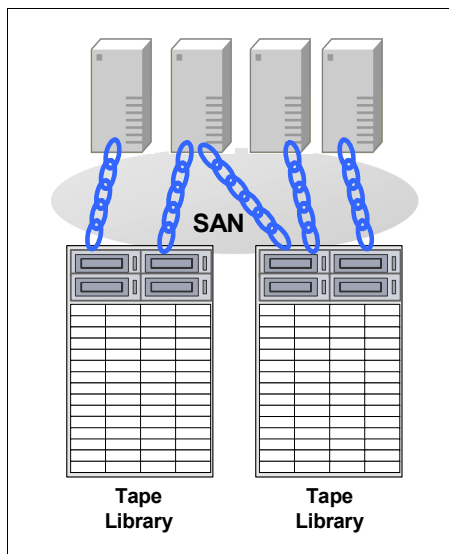


Figure 13-20 Customer issues with tape

eRMM provides a virtualization layer between applications like Tivoli Storage Manager and the tape library hardware. Essentially, eRMM decouples tape resources from applications which simplifies the sharing of tape resources even in large heterogeneous environments.

eRMM provides these benefits:

- ▶ Decouples tape resources and applications
- ▶ Simplifies the sharing of resources even in large heterogeneous configurations
- ▶ Allows to change the hardware without changing all applications
- ▶ Provides policy-based cartridge management
- ▶ Virtualizes the media changer interface (IEEE 1244 or IBM 3494)
- ▶ Provides centralized management and monitoring

Figure 13-21 shows the virtualization layer between applications and tape library hardware in an eRMM tape environments.

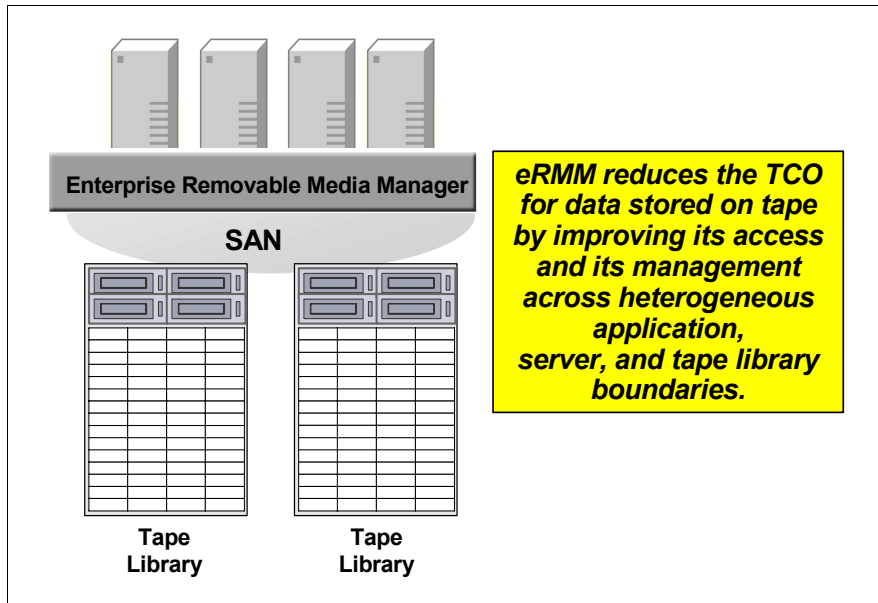


Figure 13-21 eRMM tape library hardware virtualization

Sample case study: A customer had to replace 26 IBM 3592 drives by 30 IBM TS1120 drives, in a Tivoli Storage Manager environment. eRMM automatically detected the changed hardware and provisioned the new drives to Tivoli Storage Manager without further changes to Tivoli Storage Manager. All cartridges had to be rewritten with a new Tivoli Storage Manager label to increase the capacity of the cartridges from 300 GB to 500 GB. A workflow in eRMM assured that each cartridge was relabelled exactly once.

13.3.2 eRMM central reporting and monitoring

eRMM enables central reporting and monitoring across heterogeneous application and tape library boundaries.

eRMM automatically detects which drive is available on which server and by which device handle (e.g., \\.\Tape1, /dev/rmt2). Thus eRMM can provision the device handles to applications eliminating the need to define all device handles by each application.

eRMM checks a device handle before it provisions it to a application. Thus eRMM detects and reports centrally broken paths from the application server to the tape drives. eRMM helps the administrator to answer questions like: Is there a single path from one server to one drive broken? Are there multiple paths between the same server and multiple drives broken? Are there multiple paths between multiple server and the same drive broken? Are there many paths between multiple server and multiple drives broken?

eRMM collects historical data on cartridge mounts including throughput and errors. This is helpful for the administrator for answering questions like 'What is the best window to schedule an additional backup task?' and 'Which cartridge was mounted in which drive and by which application?' The historical data enables proactive management, for example by identifying servers, HBAs, drives and cartridges with downgraded performance.

Figure 13-22 displays an example report for drive utilization. This report shows long-term trends for drive utilization helping to plan purchases of additional drives to satisfy an increasing tape workload.

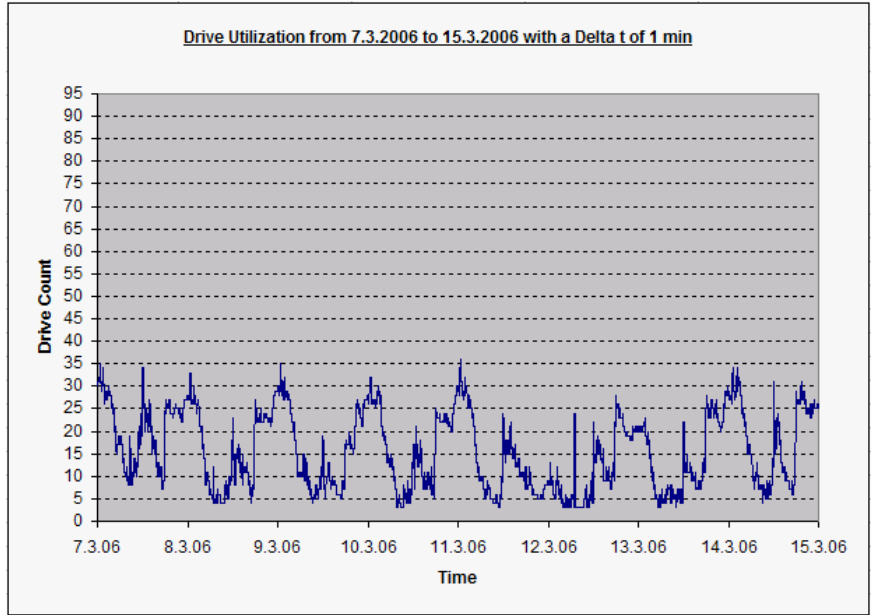


Figure 13-22 eRMM drive utilization

Figure 13-23 displays a drive histogram. This reports helps to determine whether additional backup jobs can be added to the existing tape infrastructure and it helps to identify time frames for scheduling these new backup jobs.

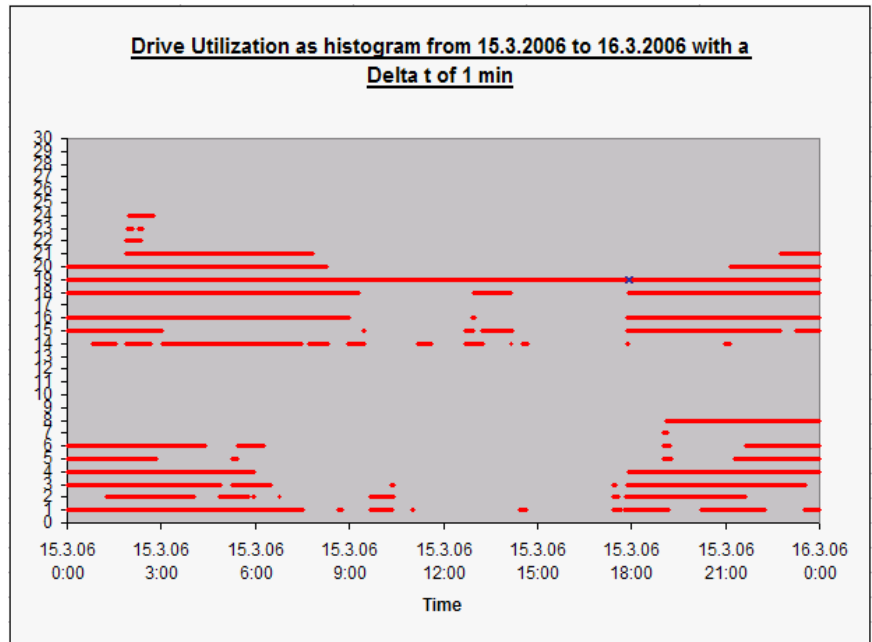


Figure 13-23 eRMM drive histogram

13.3.3 eRMM logical components and functions

This section describes the names and terms that are used in eRMM environments, as well as the key logical components and functions of eRMM.

DriveGroups and DriveGroupApplications

A DriveGroup object is a named group to which a drive can belong. DriveGroups are used to aggregate drives, and then to support both an access permissions model and a preferential usage policy. Each drive must belong to a DriveGroup.

DriveGroupApplication objects are used to allow applications to access a particular DriveGroup.

A DriveGroup may span across Libraries.

CartridgeGroups and CartridgeGroupApplications

CartridgeGroups are logical collections of cartridges. They are used to control an application's access to cartridges. A single CartridgeGroup can contain cartridges from more than one library.

CartridgeGroupApplication objects are used to allow applications to access particular CartridgeGroups.

ScratchPools

ScratchPools are a special kind of CartridgeGroup. They are searched first for empty Cartridges before any other CartridgeGroup.

If a Volume is allocated on a Cartridge which belongs to a ScratchPool, the Cartridge is moved to another ordinary CartridgeGroup to which the application issuing the command also has access. If the application doesn't have access to another ordinary CartridgeGroup the Cartridge will not be moved out of the ScratchPool, but it will be set to the "not allocatable" state which will prevent its usage by another application.

In order to see how many Cartridges are currently available it is recommended to create at least one ScratchPool and another ordinary CartridgeGroup so that the ScratchPool only contains empty cartridges. By default the eRMM installation creates a ScratchPool and an ordinary CartridgeGroup.

Figure 13-24 shows an example of DriveGroups, CartridgeGroups and ScratchPools.

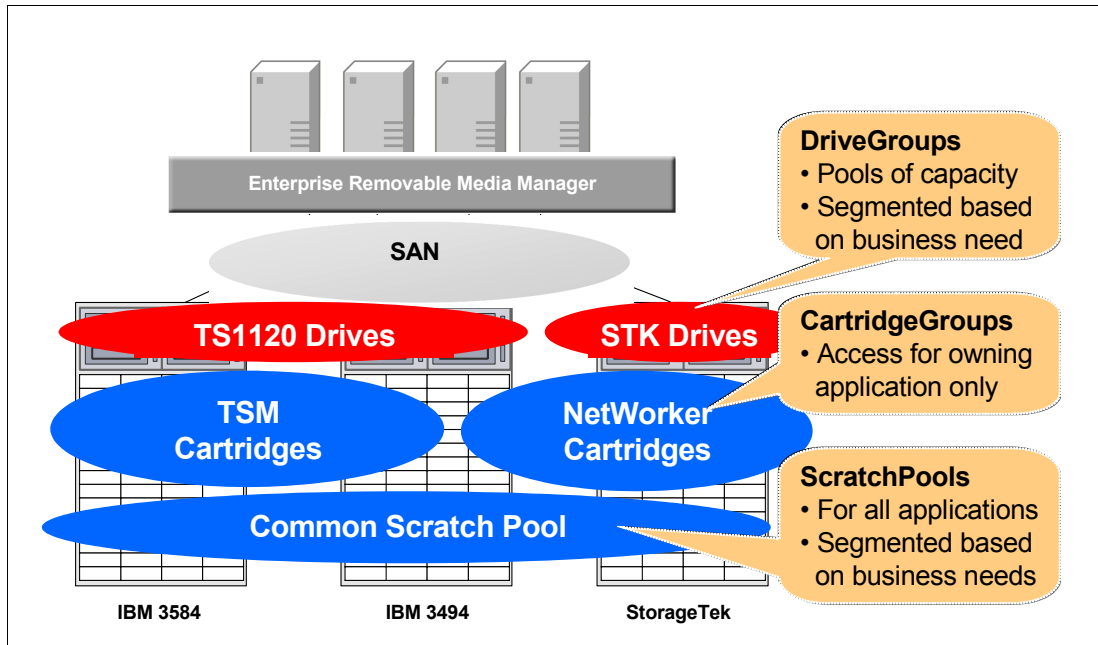


Figure 13-24 eRMM Groups and Pools

Media Manager (MM)

The Media Manager (MM) is the central “server” component which, among other tasks, coordinates access to drives and cartridges, handles volume allocation and deallocation requests, and stores a log of all activities. MM uses IBM DB2 for persistent storage.

Library Manager (LM)

The Library Manager (LM) provides MM access to library media changers. It reports all slots, tapes, and cartridges to the media manager, controls libraries on behalf of the media manager, and encapsulates (virtualizes) the library hardware. This allows to integrate new library hardware without any changes to an already installed eRMM Media Manager.

Host Drive Manager (HDM)

The Host Drive Manager (HDM) reports all local device handles to MM, handles mount and unmount commands, checks the path when a cartridge is loaded, and reports statistical data to MM when a cartridge is unloaded.

Admin Console

The Admin Console offers a Command Line Interface (CLI) and a Web Graphical User Interface (WebGUI) which enable configuration and administration of eRMM.

External Library Manager (ELM)

The External Library Manager (ELM) Adapter for IBM Tivoli Storage Manager enables Tivoli Storage Manager to utilize eRMM for media management purposes.

Figure 13-25 displays the logical architecture of eRMM with the components Media Manager, Library Manager, Host Drive Manager, and Admin Console.

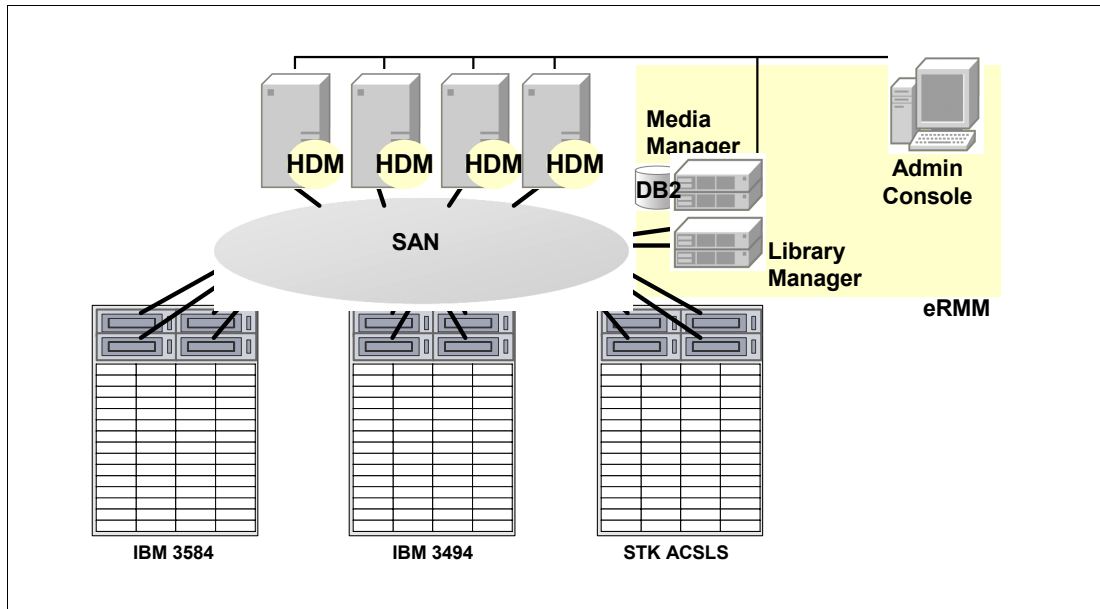


Figure 13-25 eRMM logical architecture

All eRMM logical components can run on the same or on different servers, therefore different options are available for scaling and high availability.

Application servers which need access to eRMM managed resources (libraries, drives, cartridges) must run the HDM. In addition, Tivoli Storage Manager servers need the ELM for eRMM.

13.3.4 eRMM control flow

The following three figures (Figure 13-26, Figure 13-27, and Figure 13-28) show the control flow for a tape mount in an eRMM environment. The workflow for Tivoli Storage Manager Storage Agents is similar to the workflow for Tivoli Storage Manager server.

Step 1: The Tivoli Storage Manager server wants to mount a new scratch volume of type 3592. The Tivoli Storage Manager server starts the ELM and sends it the mount request via the Tivoli Storage Manager External Library Manager Interface (see Tivoli Storage Manager Administrator's Guide, Appendix B 'Appendix B. External Media Management Interface Description' for further details).

Step 2: The eRMM ELM forwards the mount request via TCP/IP to the eRMM MM.

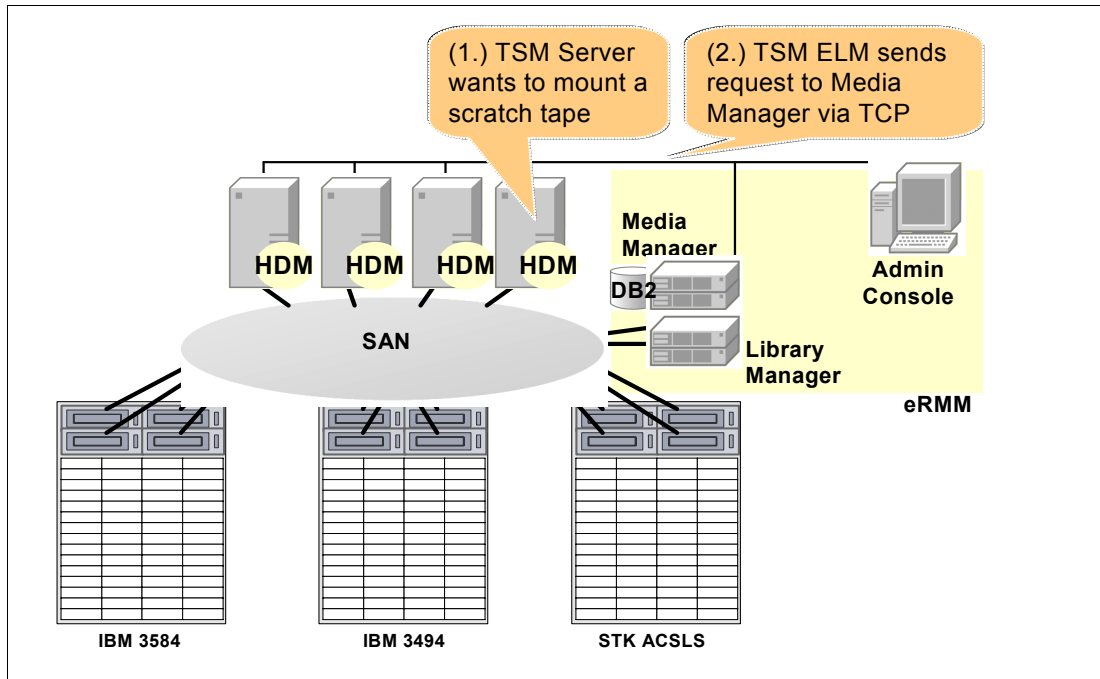


Figure 13-26 eRMM Control Flow - Steps 1 to 2

Step 3: The eRMM MM queries its database for idle drives and scratch cartridges and selects a drive and a cartridge according to the access rules. eRMM takes into account which drives are configured in the operating system of the requesting Tivoli Storage Manager server.

Step 4: The eRMM MM forwards the specific mount request via TCP/IP to the respective Host Drive Manager (running on the same server as the Tivoli Storage Manager server) and to the eRMM LM.

Step 5: The eRMM LM converts the mount request into a library specific command (SCSI for IBM 3584, Imcp for IBM 3494, ACSLS for STK) and loads the cartridge into the drive.

Step 6: The eRMM HDM queries the drive to ensure that the path to the drive is healthy.

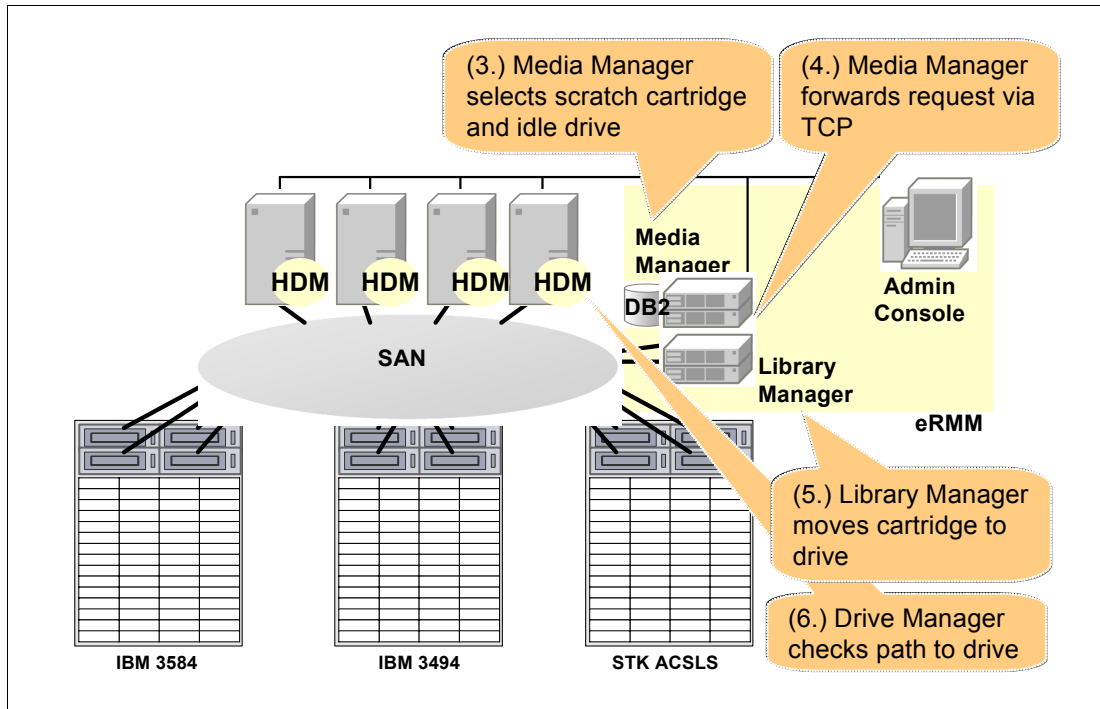


Figure 13-27 eRMM Control Flow - Steps 3 to 6

Step 7: The eRMM MM updates the status in the eRMM DB.

Step 8: The eRMM MM sends the device handle (e.g., /dev/rmt5, \\.\Tape3) and the cartridge VOLSER via TCP to the eRMM ELM. The eRMM ELM returns the Device Handle and the VOLSER to the Tivoli Storage Manager server via the Tivoli Storage Manager External Library Manager Interface.

Step 9: The Tivoli Storage Manager server updates its volume inventory and directly accesses the drive. eRMM is not involved for read and write operations. This design allows to put eRMM into or pull eRMM out of an existing Tivoli Storage Manager environment without recopying the data on tape.

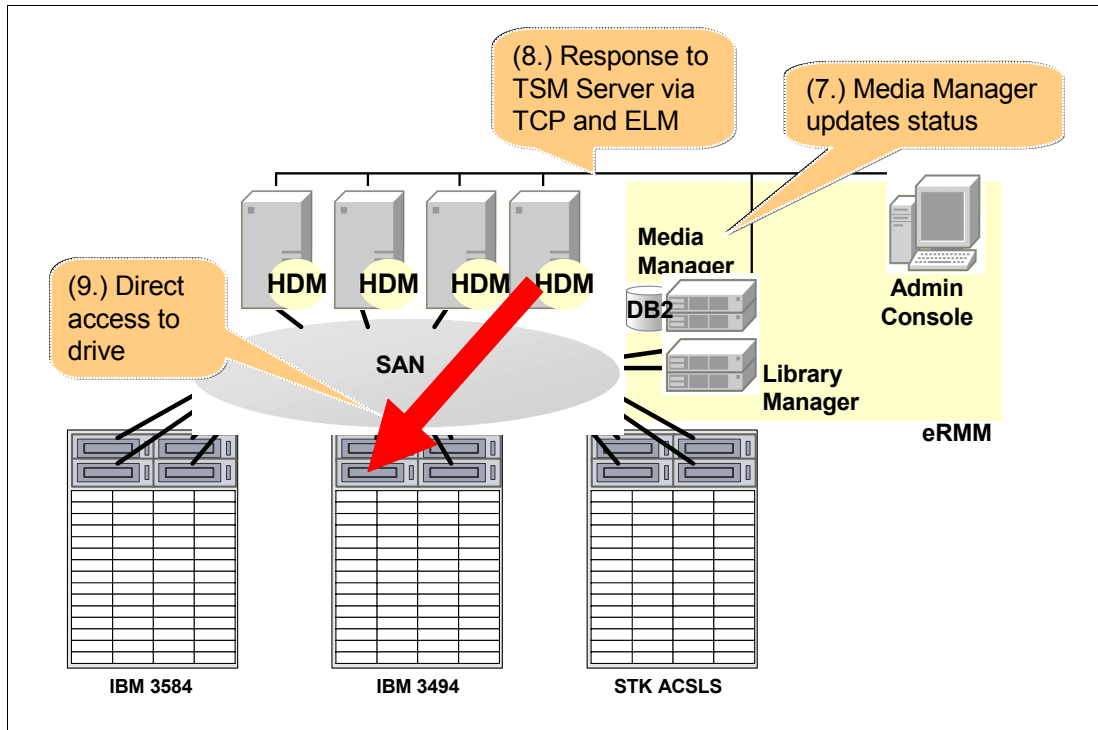


Figure 13-28 eRMM Control Flow - Steps 7 to 9

This example illustrates how Tivoli Storage Manager installations with Library sharing and LAN-free Backup can benefit from eRMM's provisioning of tape resources. It is no longer required to configure and maintain tape paths in Tivoli Storage Manager.

13.3.5 Supported features (eRMM 1.2.4)

The following features are supported by eRMM at the current (at the time of writing) release - eRMM v1.2.4:

- ▶ Automatic configuration of drives for Tivoli Storage Manager
- ▶ Centralized access control, administration, problem determination and reporting
- ▶ Policy-based drive and cartridge allocation
- ▶ Dynamic and heterogeneous drive sharing
- ▶ Dynamic and heterogeneous drive and cartridge pooling
- ▶ Mount request queuing
- ▶ Audit trails and statistical data for the complete cartridge lifecycle
- ▶ Dynamic drive sharing
- ▶ Dynamic drive and cartridge pooling including common scratch pool management
- ▶ Tape library virtualization

The following features are considered to advance eRMM's value proposition:

- ▶ Policy based cartridge and lifecycle management
- ▶ Offsite media management and tracking (vaulting)
- ▶ Advanced reporting and auditing

13.3.6 Supported platforms (eRMM 1.2.4)

The following platforms are supported by eRMM.

Applications

- ▶ Tivoli Storage Manager on AIX, Solaris, HP-UX, Linux (Intel and System z) and Windows
- ▶ Native OS commands like tar, dd, and mksysb
- ▶ EMC Legato NetWorker (on request)

Tape Libraries

- ▶ IBM System Storage TS3500 Tape Library Open Systems
- ▶ IBM TotalStorage 3494 Tape Library Open Systems
- ▶ ACSLS managed StorageTek libraries
- ▶ Other SCSI Libraries (e.g. IBM 3583, ADIC on request)

Support for additional applications and tape libraries are under consideration as enhancements to eRMM.

13.3.7 Strategic fit and positioning

How does eRMM fit into the strategic framework of the business solution areas Infrastructure Simplification, Business Continuity and Information Lifecycle Management?

eRMM and infrastructure simplification

- ▶ Automated detection and configuration of drives for Tivoli Storage Manager
- ▶ Provisioning and simplified sharing of drives and cartridges to Tivoli Storage Manager
- ▶ Policy based drive and cartridge utilization
- ▶ Policy based cartridge management
- ▶ Virtualization of the libraries' media changer interface

eRMM and business continuity

- ▶ High available library sharing for Tivoli Storage Manager on AIX, Solaris, HP-UX, Linux, and Windows
- ▶ High available provisioning of scratch cartridges to Tivoli Storage Manager

eRMM and ILM

- ▶ Policy based vaulting for tired storage management (considered for future releases of eRMM)
- ▶ Audit trails and statistical data of complete cartridge lifecycle for regulatory compliance and for tape quality management (considered for future releases of eRMM)

13.4 More Information

For more information about the IBM tape virtualization products and service offerings consult the following documents and Web sites:

- ▶ <http://www.ibm.com/servers/storage/tape/3494vts/index.html>
- ▶ <http://picard.mainz.de.ibm.com/A182/index.jsp?page=projects/ermm>
- ▶ <https://w3.webahead.ibm.com/w3ki/display/ermm>
- ▶ *IBM Virtualization Engine TS7510 Introduction and Planning Guide*, GC26-7767
- ▶ *IBM Virtualization Engine TS7510 User's Guide*, GC26-7769
- ▶ *IBM Virtualization Engine TS7510: Tape Virtualization for Open Systems Servers*, SG24-7189

- ▶ *IBM TotalStorage Enterprise Automated Tape Library (3494) Introduction and Planning Guide, GA32-0448*
- ▶ *IBM TotalStorage Virtual Tape Server: Planning, Implementing, and Monitoring, SG24-2229*
- ▶ *IBM TotalStorage Peer-to-Peer Virtual Tape Server Planning and Implementation Guide, SG24-6115*



Complementary storage products

This chapter discusses entry-level tape products that are complementary to the LTO series of drives and libraries discussed earlier. The entry-level tape products cover a wide range of formats including 4 mm, 8 mm, VXA-2, SDLT, and SLR/MLR QIC to suit varied business needs. Two additional products (DVD-RAM and disk) are also briefly mentioned.

14.1 Overview

Table 14-1 and Table 14-2 present a summary of the entry-level storage products that are available as part of the System Storage portfolio. Details on each of the drives are presented below.

Table 14-1 Tape storage products - part 1

Product	External Disk Drive	4 mm DDS Gen 5 (DAT72)	8 mm VXA-2 Tape Drive
Machine	7204	7206	7206
Model	573-73 GB 646-146 GB	336-DDS5	VX2 VX3 (VXA-320 media)
Product Strengths	External deskside disk storage	Cost-effective streaming tape drive	Low-cost, high-capacity VXA-2 and VXZ-3 technology
Technology	N/A	Helical Scan	Helical Scan
Number of Heads/Tracks	8 heads 4 disks	Rotating Drum	Rotating Drum
Number of Drives	1	1	1
Max Number of Cartridges	N/A	1	1
Cartridge Capacity Native/Compressed	N/A	36/72 GB	VX2: 80/160 GB VX3: 160/320 GB
Max System Capacity Compressed	573: 73 GB 646-146 GB	72 GB	VX2: 160 GB VX3: 320 GB
Max Drive Data Rate Native/Compressed	80 MBps	3/6 MBps	VX2: 6/12 MBps VX3: 12/24 MBps
Time to Data	70 ms	50 seconds	50 seconds
Interface	SCSI-3 ULTRA, LVD/SE, 160/320	SCSI-2 F/W SE, LVD/SE	SCSI-3 ULTRA, LVD/SE, 160/320
Supported Platforms	System p	System p	System p, System i
Warranty	IBM Onsite Repair 1 year, 24x7	IBM Onsite Repair 1 year, 24x7	IBM Onsite Repair 1 year, 24x7

Note: The 7205-550 was withdrawn from marketing on June 2, 2006.
The 7207-122 was withdrawn from marketing on March 27, 2006.
The 7212-312 was withdrawn from marketing on June 2, 2006.

Table 14-2 Tape storage products - part 2

Product	SDLT Drive Tape Drive	SLR (QIC) Compatible External Tape Drive	Storage Device Enclosure	VXA-2 Autoloader
Machine	7205	7207	7212	7212
Model	550	122 330	102 103	312
Product Strengths	Cost-effective save/restore/archive solution	Backward read/write compatible with iSeries internal	Rack-mountable 2-drive enclosure utilizes only 1U of space	Rack-mountable Only 1 EIA rack unit high
Technology	Longitudinal Serpentine	SLR (QIC format)	DDS, DVD, VXA plus SLR, LTO (on 103)	Helical Scan
Number of Heads/Tracks	8/536	1/1	DDS/VXA: Rotating Drum	Rotating Drum
Number of Drives	1	1	1-2	1
Max Number of Cartridges	1	1	2	10
Cartridge Capacity Native/Compressed (GB)	160/320	122: 4/8 330: 30/60	DDS-4: 20/40 DAT72: 36/72 VXA-2: 80/160 SLR60:37.5/75 SLR100:50/100 LTO2:200/400 VXA-320:160/320 DVD-RAM2:4.7/9.4	80 /160
Max System Capacity Compressed (GB)	320	122: 4/8 330: 30/60	320 with two VXA-2 drives 640 with two VXA-320 drives 144 with two DAT72 drives 80 with two DDS4 drives 150 with two SLR60 drives 200 with two SLR100 drives 800 with two LTO2 drives 18.8 with two DVD-RAM drives	1600

Product	SDLT Drive Tape Drive	SLR (QIC) Compatible External Tape Drive	Storage Device Enclosure	VXA-2 Autoloader
Max Drive Data Rate Native/Compressed (MBps)	16/32	122:.38/.76 330: 4/8	DDS/DAT72: 3/6 VXA: 6/12 VXA-320: 12/20 SLR60:4/8 SLR100: 5/10 LTO2:24/48 DVD-RAM:2.77	6/12 MBps
Time to Data	70 seconds	122: 85 seconds 330: 50 second	VXA-2: 40 seconds VXA: 40 seconds DAT72: 50 seconds	50 seconds
Interface	SCSI-2 F/W, Diff PCI, Ultra2 SCSI LVD	SCSI-2 SE, ULTRA, LVD/SE, 160/320	SCSI-3 ULTRA, LVS/SE, 160/320	SCSI-3 Ultra, LVD- SE, 160/320
Supported Platforms	System p	System p, System i	System p, System i	System p, System i
Warranty	IBM Onsite Repair 1 Year, 24x7	IBM Onsite Repair 1 Year, 24x7	IBM Onsite Repair 1 Year, 24x7	CRU, IBM Onsite Repair 1 year, 24x7

14.2 IBM 7204 external disk drive

The IBM 7204 External disk drives are stand-alone devices that provide a single Ultra3 SCSI disk drive with a formatted capacity of up to 146 GB. The maximum media data transfer rate is 100 MBps. The IBM 7204 External Disk Drives use a 68-pin LVD interface and attach to the System p servers with a SCSI cable to a variety of SCSI adapters. Obviously, multiple 7204 devices can be installed on a single system.



Figure 14-1 IBM 7204 external disk drive

There are two models which differ solely on the size of the disk drive that is installed in the enclosure (see Table 14-1 on page 294).

14.3 IBM 7205 SDLT tape drive

The IBM 7205 Model 550 is an external, stand-alone, SCSI-attached, DLT drive that attaches to System p workstations and servers via a SCSI Ultra2 LVD interface. The IBM 7205 Model

550 features a native transfer rate of up to 16 MBps, with a native cartridge capacity of 160 GB. The IBM 7205 Model 550 also supports 2:1 data compression, providing an effective capacity of up to 320 GB and an effective data transfer rate of up to 32 MBps.

To protect the data availability of tapes written by older generations of DLT drives, the 7205 Model 550 is designed to be capable of reading tapes written on SDLT320, DLT7000 or DLT8000 tape - enabling cost-effective backup and restore functions. As DLT8000 tape technology nears the end of its evolution, the new SDLT320 standard, through its backward compatibility with older DLT drives, can provide a migration path to greater tape storage capacity.

The IBM 7205 Model 550 includes its own device driver that is used by save/restore utilities provided with the AIX4.3.3 (or later) operating system.



Figure 14-2 IBM 7205 SDLT tape drive

14.4 IBM 7206 tape drive

Figure 14-3 shows the 7206 4 mm tape drive. There are several models available.



Figure 14-3 IBM 7206 4 mm tape drive

14.4.1 Model 336

The IBM 7206 Model 336 External DDS Gen 5 (DAT72) Tape Drive can be attached to System p servers via an Ultra2, Ultra3, or Ultra320 SCSI LVD interface.

The 7206 Model 336 tape drive can achieve a media capacity up to 72 GB with 2:1 data compression, nearly twice the capacity of the previous IBM 7206-220 DDS4 tape drive. The 7206 Model 336 offers a sustained data transfer rate of up to 6 MB per second (compressed).

14.4.2 Model VX2

The IBM 7206 Model VX2 External VXA-2 Tape Drive is a higher capacity, cost-effective alternative to DDS tape technology offered in the Model 336. The 7206 Model VX2 Tape Drive supports three lengths of cartridges with up to 160 GB of compressed storage capacity and data rates of 12 MB per second (compressed) which is double that achieved by DDS. The interface to System p or System i servers is via a standard Ultra2 SCSI LVD connector.

14.4.3 Model VX3

The IBM 7206 Model VX3 External VXA-320 Tape Drive supports cartridges of up to 320 GB of compressed storage capacity and data rates of 24 MB per second (compressed). The interface to System p or System i servers is via a standard Ultra2 SCSI LVD, Ultra3 or Ultra320 connector.

14.5 IBM 7207 SLR tape drive

The following models are available:

14.5.1 Model 122

The IBM 7207 Model 122 4 GB External SLR5 QIC is a tape drive aimed as a low-cost entry-level backup solution for the iSeries range of servers. It is backwardly compatible with System i internal QIC drives and can also be attached to System p via the SCSI-2 interface.

The 7207 Model 122 provides a native media capacity of 4 GB per cartridge and a data transfer rate of up to 0.38 MBps. With 2:1 compression, the 7207 Model 122 has a capacity of up to 8 GB and a data transfer rate of up to 0.76 MBps. In addition to reading and writing on 1/4-inch tape cartridges using the QIC 4 GB format, the 7207 Model 122 is read/write compatible with QIC 120/150, 525, 1000, and QIC 2 GB formats.



Figure 14-4 IBM 7207 SLR tape drive

14.5.2 Model 330

IBM TotalStorage 7207 Model 330 External SLR60 Tape Drive is designed to offer improved capacity and performance over model 122. To assist with migration from older QIC tape formats Model 330 is designed to be capable of reading and writing tapes written on other systems that use SLR100, MLR3, and MLR1 tape formats; it is read-only compatible with SLR5 and DC9250 media. There are two cartridge types that can be used in the tape drive

either 30 GB (native) or 5 GB (native). The data rate achieved by this drive is 4 MB per second.

14.6 IBM 7210 DVD-RAM drive

The IBM 7210 External DVD-RAM Drive is a self-powered stand-alone drive, designed for the open systems environment, which includes the IBM @server i5, iSeries, and AS/400, and p5, pSeries, and RS/6000 server. It comes in two models which are described in more detail below.



Figure 14-5 IBM 7210 DVD-RAM drive

14.6.1 Model 025

Note: The 7210-025 was withdrawn from marketing on Feb. 24, 2006.

The 7210 Model 025 DVD-RAM Drive is an external storage device. This self-powered stand-alone drive can attach to System i and System p servers, using a SCSI-2 interface. The 7210 DVD-RAM Drive operates at speeds dependent on the type of media used.

- ▶ CD-ROM media maximum speed is 3.6 MBps
- ▶ DVD-ROM media maximum speed is 2.7 MBps
- ▶ DVD-RAM media maximum speed is 1.3 MBps Write, 2.7 MBps Read

The media used in this device determines the operations possible. The 7210 DVD-RAM Drive can read and write to all DVD-RAM media (2.6 GB to 9.4 GB). The 7210 DVD-RAM Drive can only read media marked as CD-R, CD-R/W, and CD-ROM. The drive features the following:

- ▶ CD-ROM speeds of up to 3600 KB/second
- ▶ DVD-ROM speeds of up to 8115 KB/second
- ▶ DVD-RAM speeds of 1352 to 2705 KB/second (dependent on media type)
- ▶ Capacity of 2.6 GB, 4.7 GB, 5.2 GB, and 9.4 GB per DVD
- ▶ SCSI-2 that supports synchronous and asynchronous data transfer modes

14.6.2 Model 030

Model 030 is similar in most respects to the model 025, except that can only read and write to 4.7/9.4 GB DVD-RAM, and supports Ultra-3 SCSI (LVD) attachment.

14.7 IBM 7212 storage device enclosure

The IBM 7212 Storage Enclosure is a versatile product that provides efficient and convenient storage and storage automation. The 7212 is designed to accept two storage devices and be mounted in 1 EIA unit of a standard 19-inch rack using an optional rack hardware kit, or it can be configured for desktop mounting. The enclosure connects to System i and System p.

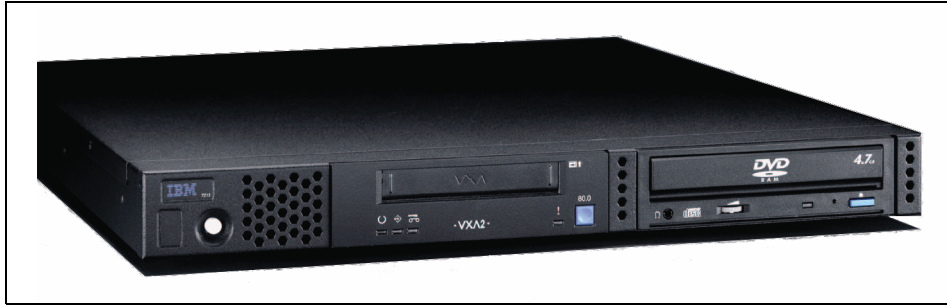


Figure 14-6 IBM 7212 storage device enclosure

It offers several storage device options including the following:

- ▶ Low cost VXA-2 technology, and VXA-320
- ▶ DDS Gen 5 (DAT72) and DDS4 4 mm
- ▶ SLR60 and SLR100 (QIC format) tape drives
- ▶ LTO 2 drives
- ▶ DVD-RAM and DVDROM optical drives

It provides a solution for environments in which cabling space and server storage bays are limited.

14.8 IBM 7212-312 VXA-2 autoloader

The IBM TotalStorage 7212 Model 312 Tape Autoloader features a high performance, low-cost VXA-2 tape drive. The 7212 Model 312 packaging is a low profile, rotating carousel design that stores up to 10 data cartridges that can be automatically positioned in the tape drive or at a cartridge access point for removal. These automated features are designed to reduce the attended operator time for data storage, archival, and retrieval that is often required in single or multi-server networks without automation features. The VXA-2 tape drive, with physical storage capacity of up to 80 GB and a data transfer rate of up to 6 MBps, is designed to offer excellent capacity and performance at a low price.

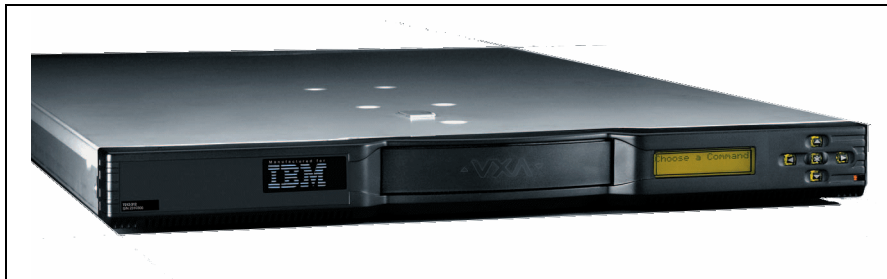


Figure 14-7 IBM 7212-312 VXA-2 autoloader

Part 3



Storage networking

Part 3 introduces some of the important different technologies and protocols that can be used to interconnect storage devices. These include SAN, NAS, and iSCSI.



Introduction to storage networking

Variations for storage networking seem to be materializing faster than they can be implemented. Storage networking offers significant capabilities and flexibility for accessing stored data that was not previously available. Understanding the technology basics is essential to making the best choices.

This chapter provides an introduction to different storage networking options that you can choose to build the infrastructure for accessing your stored data. We mainly cover the fundamentals of SAN and different SAN expansion technologies.

15.1 Overview

Why are there so many forms of storage networking? For one, new technologies emerge and evolve but don't replace the investment in previous technologies overnight. And no single storage networking approach solves all problems, or optimizes all variables. There are trade-offs in cost, ease-of-management, performance, distance, and maturity, to name but a few. For the foreseeable future, multiple storage networking alternatives will coexist—often within the same organization.

Three popular models for attaching storage for the network are Storage Area Network (SAN), Network Attached Storage (NAS), and iSCSI. These methods help to remove direct attachments between storage and server giving more flexibility in storage access.

The SAN can be viewed as an extension to the storage bus concept that enables storage devices and servers to be interconnected using similar elements as in LAN: routers, hubs, switches, directors and gateways. Storage resides on this dedicated network, providing an any-to-any connection for processors and storage on that network. The data traffic is based on a block level I/O (requests access devices directly). The most common media is Fibre Channel. Fibre Channel Protocol (FCP) is the I/O protocol for open systems, and Fibre Connectivity (FICON) replaces it for mainframe environments. Ethernet-based SANs are also emerging. Today, there are several popular protocols used to build or extend a SAN besides FCP like Internet SCSI (iSCSI), Fibre Channel over IP (FCIP) and Internet Fibre Channel Protocol (iFCP).

The NAS device is attached to a TCP/IP-based network (LAN or WAN), and accessed using CIFS, NFS or specialized I/O protocols for file access and file sharing. The data traffic is based on a file level I/O. It receives an NFS or CIFS request over a network and has an internal processor which translates that request to the block-I/O commands to access the appropriate device only visible to the NAS product itself.

15.2 Storage Area Network (SAN)

A SAN is a dedicated network for storage devices and the processors that access those devices. Figure 15-1 on page 305 shows a picture of a SAN. SANs today are usually built using Fibre Channel technology, but the concept of a SAN is independent of the underlying type of network. I/O requests to disk storage on a SAN are called "block I/Os" because, just as for direct-attached disk, the read and write I/O commands identify a specific device (disk drive or tape drive) and, in the case of disks, specific block (sector) locations on the disk.

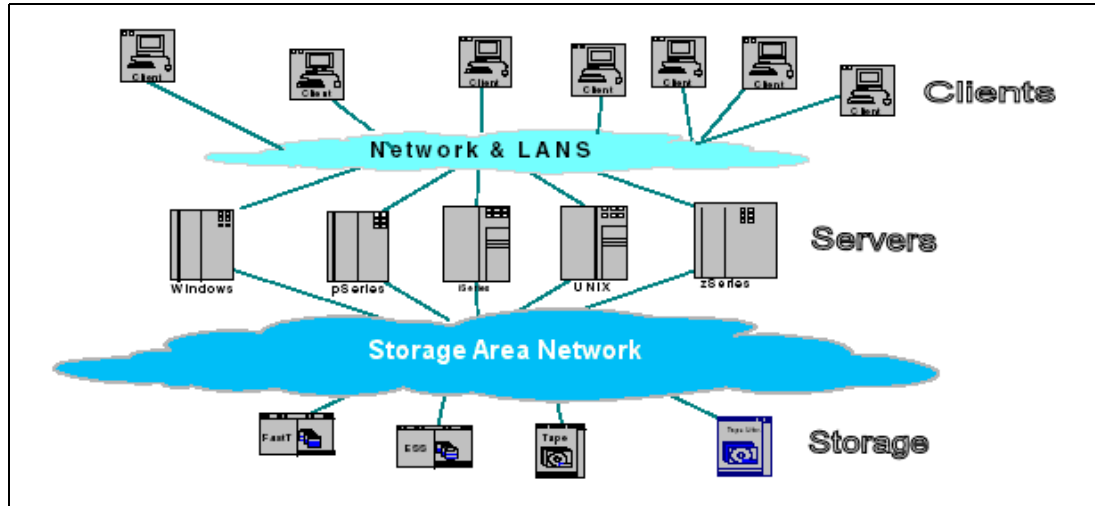


Figure 15-1 A picture of SAN

The major potential benefits of a SAN can be categorized as:

- ▶ **Improvement to application availability**
Storage is independent of applications and accessible through multiple data paths for better reliability, availability and serviceability.
- ▶ **Higher application performance**
Storage processing is off-loaded from servers and moved onto a separate network e.g. LAN-free and server-free backups.
- ▶ **Centralized and consolidated storage**
Simpler management, scalability, flexibility and availability.
- ▶ **Data transfer and vaulting to remote sites**
Remote copy of data enabled for disaster protection and against malicious attacks.
- ▶ **Simplified centralized management**
Single image of storage media simplifies management.

In this section, we explain important concepts involved in SAN.

15.2.1 Fibre Channel

Today, Fibre Channel is well established in the open systems environment as the underlining architecture of the SAN. Fibre Channel is a technology standard that allows data to be transferred from one network node to another at very high speed. The interconnections between nodes are not necessarily based on *fiber* optics, but can also be based on copper cables. In current implementations, Fibre Channel standard speed is 4 Gbps, and 10 Gbps products are being introduced. Most older 2 Gbps equipment is being replaced by faster kit.

This standard is backed by a consortium of leading vendors and has been accredited by the American National Standards Institute (ANSI).

Fibre Channel is structured with independent layers, as are other networking protocols. There are five layers, where 0 is the lowest layer. The physical layers are 0 to 2. These layers carry the physical attributes of the network and transport the data created by the higher level protocols, such as SCSI, TCP/IP, or FICON.

As shown in Figure 15-2, the top two layers (the session and transport layers) can be used by these protocols to move data segments. These segments are then rolled into a packet, which in turn are rolled into a frame. The originator creates the frame and sends it to the destination, which unravels the frame back to a segment.

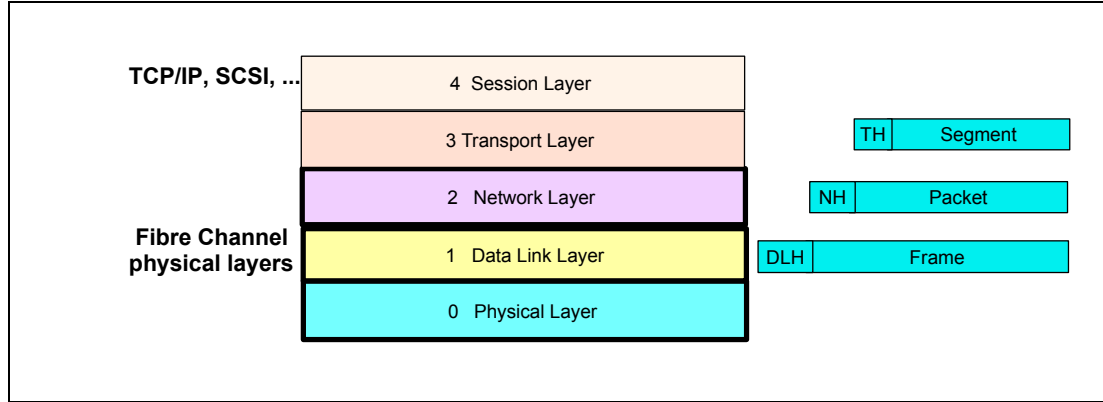


Figure 15-2 Fibre Channel protocol layers

15.2.2 SAN topologies

Fibre Channel based networks support three types of topologies:

- ▶ Point-to-point
- ▶ Loop (arbitrated)
- ▶ Switched

These can be implemented separately or interconnected to form a *fabric*. The fabric can also be extended to cover even greater distances. Currently, the distance limitation for synchronous data transfers is about 100 km and virtually unlimited for asynchronous data transfers (see 15.3, “IP Storage Networking technologies” on page 313).

Synchronous data transfers require an acknowledgement that the data arrived at the receiving before the data is discarded at the transmitting end, whereas asynchronous data transfers do not require acknowledgement before the data is discarded.

Point-to-point topology

The point-to-point topology is the easiest Fibre Channel configuration to implement, and it is also the easiest to administer. This simple link can be used to provide a high-speed interconnection between two nodes as shown in Figure 15-3 on page 307. A node is any device with one or more Fibre Channel ports.

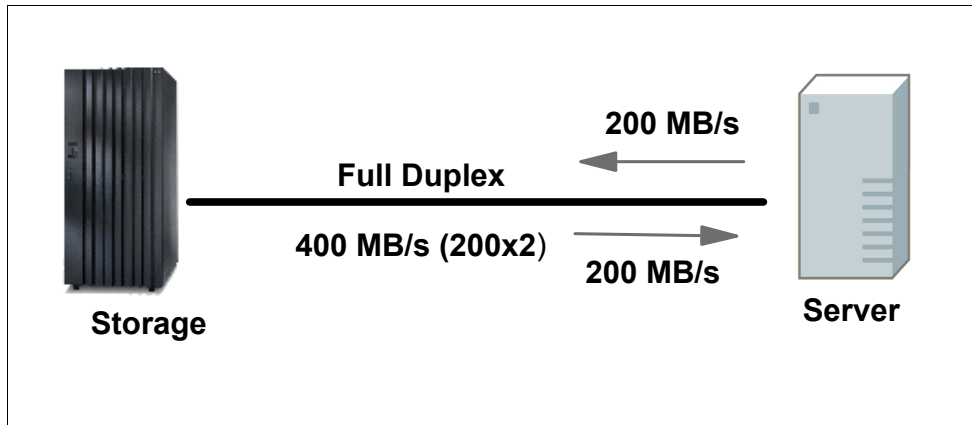


Figure 15-3 Fibre Channel point-to-point topology

Because connectivity is limited to two nodes, the exploitation of point-to-point in tape environments is limited; however the distance between nodes can be up to 10 km, which enables a tape library to be located at a different site.

When greater connectivity and performance are required, each device can be connected to a fabric without incurring any additional expense beyond the cost of the fabric itself.

Loop (arbitrated) topology

The Fibre Channel arbitrated loop offers relatively high bandwidth and connectivity at a low cost. For a node to transfer data, it must first arbitrate to win control of the loop. Once the node has control, it is now free to establish a virtual point-to-point connection with another node on the loop. After this point-to-point (virtual) connection is established, the two nodes consume all of the loop's bandwidth until the data transfer operation is complete. Once the transfer is complete, any node on the loop can now arbitrate to win control of the loop. The characteristics of a Fibre Channel arbitrated loop include:

- ▶ Support of up to 126 devices is possible on a single loop.
- ▶ Devices can be hot-swapped with the implementation of hubs and bypass ports.
- ▶ A loop is self-discovering; it finds out who is on the loop and tells everyone else.
- ▶ Logic in the port allows a failed node to be isolated from the loop without interfering with other data transfers.
- ▶ Virtual point-to-point communications are possible.
- ▶ A loop can be interconnected to other loops, essentially forming its own fabric.
- ▶ A loop can be connected to a suitable Fibre Channel switch to create fan-out, or the ability to increase the size of the fabric even more. Note that not all switches support direct attachment of loops.

FC hub devices support FC loop connections while offering some of the benefits of switches. Figure 15-4 on page 308 shows an FC loop using a hub. FC hub technology is now no longer marketed and is rare in production environments.

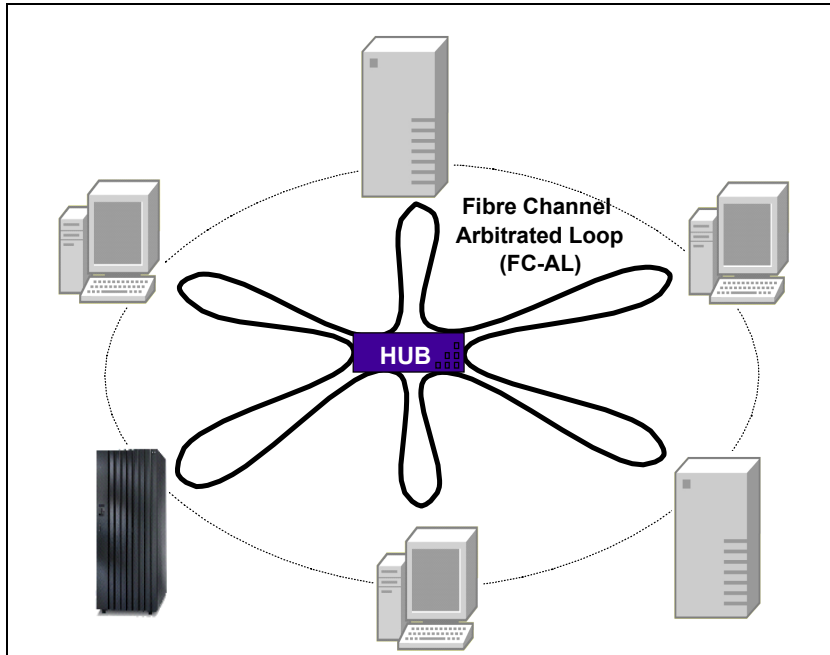


Figure 15-4 Fibre Channel loop topology

Switched topology

Fibre Channel switches function in a manner similar to traditional network switches to provide increased bandwidth, scalable performance, an increased number of devices, and, in some cases, increased redundancy. Fibre Channel switches vary in the number of ports and media types they support.

Note: This is the most widely used topology in current SAN implementations.

Multiple switches can be connected to form a switch fabric capable of supporting a large number of host servers and storage subsystems, as shown in Figure 15-5 on page 309. When switches are connected, each switch's configuration information has to be copied (*cascaded*) into all the other participating switches.

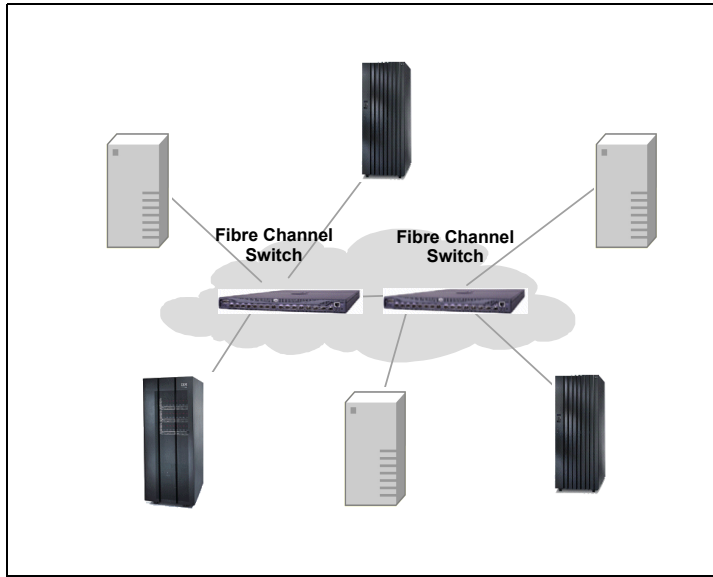


Figure 15-5 Fibre Channel switched topology

Switched fabrics provide the richest function for the tape environment, providing distance for remote tape vaulting and disaster tolerance, multi-server connectivity for tape library sharing, and multiple paths for redundancy. Switches can be connected together (as shown in Figure 15-5) to provide extra paths and larger networks.

Single switched fabric

A switched point-to-point fabric without redundancy can support a large number of servers. A configuration can be set up to allow every server to have access to every switch, and every controller to be connected to at least two switches. This allows any server to get to any controller or device, and it allows for continuous operation (although with degraded performance) in the event that a switch fails.

An example of a non-redundant fabric is shown in Figure 15-6 on page 310.

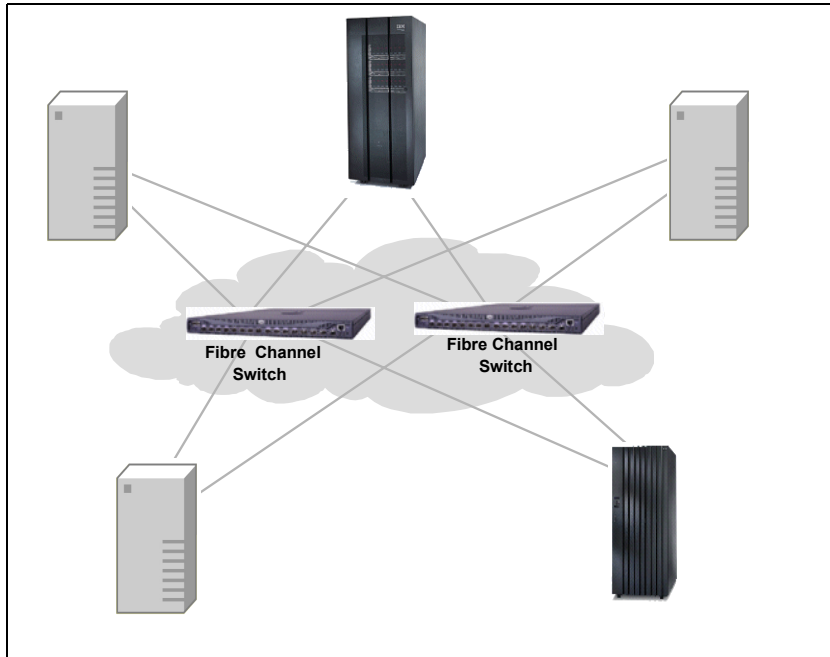


Figure 15-6 Fibre Channel single switched topology

Switched fabric with redundancy

A switched fabric with redundancy provides interconnections between switches, so that the collection of switches looks like one large, any-to-any switch. Fabric management becomes more complex than with basic switched point-to-point configurations and there are often limits imposed by product vendors on the number of switch-to-switch hops permitted when negotiating the fabric. Inter-switch links can fail and must be identified (many switch vendors do not yet support any reporting on inter-switch links).

Traffic can be routed in many ways. For technical, security, or other reasons, various levels of zoning or other mechanisms may be used to restrict the any-to-any access. Performance monitoring and configuration changes or upgrades needed to keep the network performing adequately are more complex.

Tip: Zoning splits the SAN into subnetworks. The servers within a zone have any-to-any connectivity, but anything outside the zone is not visible to them. VSANs (Virtual SANs) are used by Cisco, where the SAN is partitioned at the hardware level, into logical sections, for the purpose of isolating traffic, security, and error containment.

The primary advantage of a switched redundant fabric is that it looks like a very large logical switch, where a single connection provides access to any other port on the total set of switches, as shown in Figure 15-7 on page 311.

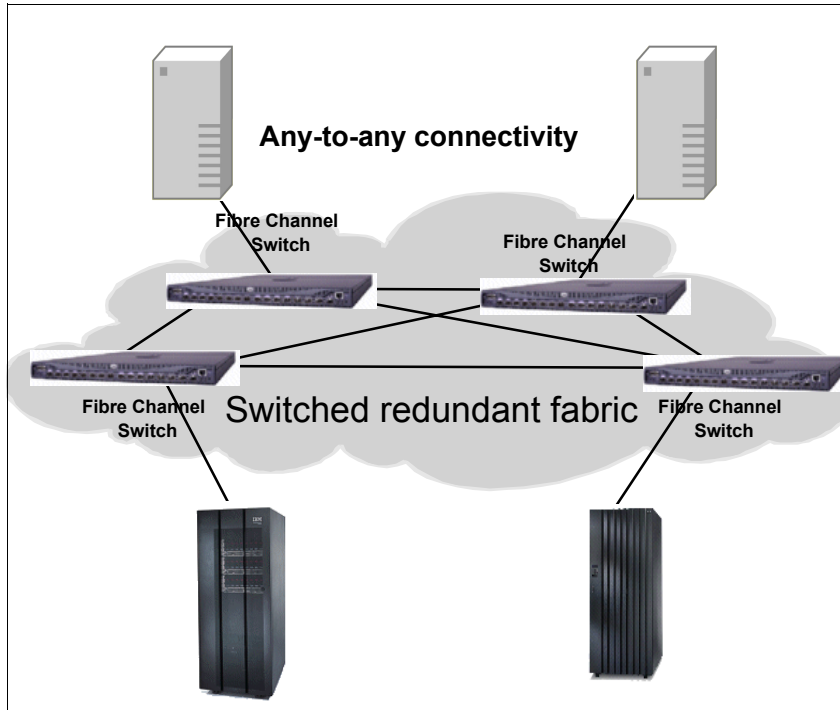


Figure 15-7 Fibre Channel switched topology (redundant switches)

15.2.3 Physical components of the SAN infrastructure

Here we will take a brief look at some of the components that are commonly encountered in the FC SAN implementations.

Cables and connectors

FC connectors come mainly in two types as LC and SC connectors. Figure 15-8 shows FC connectors and cables. LC connectors are now the standard.

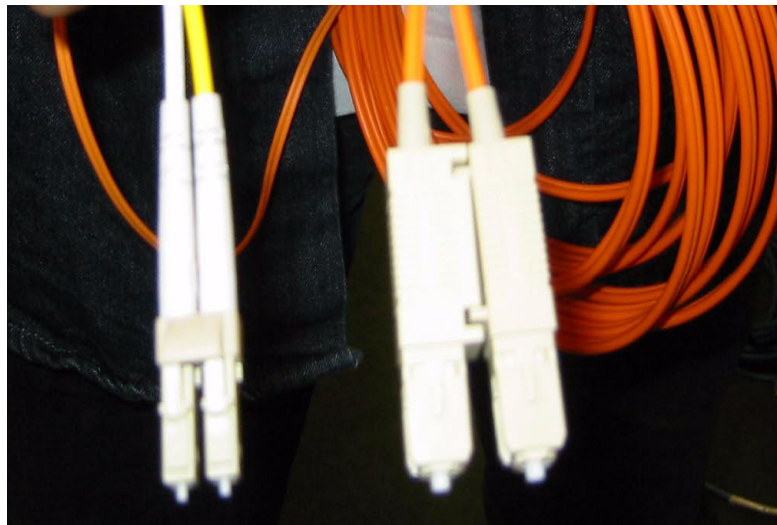


Figure 15-8 LC and SC connectors

Transceivers

Transceivers are plugged on switches or directors on each port and used to convert the internal communication transport to gigabit transport. For 1 Gbps, the transceivers are called GBICs (Gigabit Interface Converters) and for 2 Gbps and 4 Gbps, the transceivers are called SFPs (Small Form Factor Pluggable Media). Figure 15-9 shows SFPs and a GBIC.

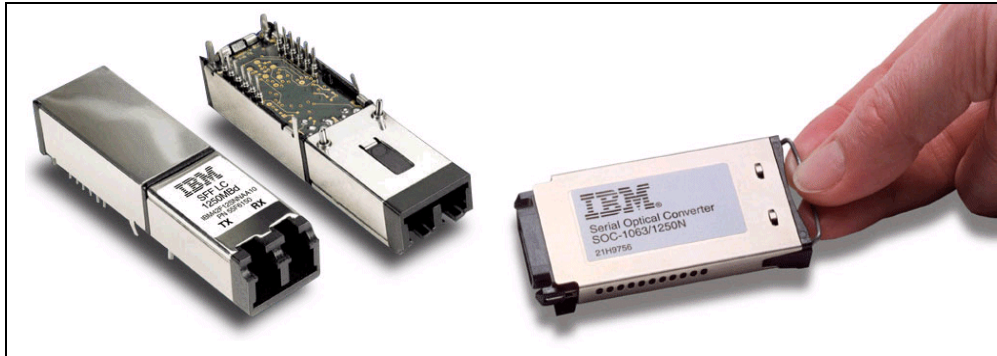


Figure 15-9 SFP (Small Form Factor Pluggable Media) and GBIC (Gigabit Interface Converter)

Host Bus Adapters

The Host Bus Adapter (HBA) connects to the bus of the host or storage system. It has some means of connecting to the cable leading to the SAN. The function of the HBA is to convert the parallel electrical signals from the bus into a serial signal to pass to the SAN. Figure 15-10 shows an HBA.

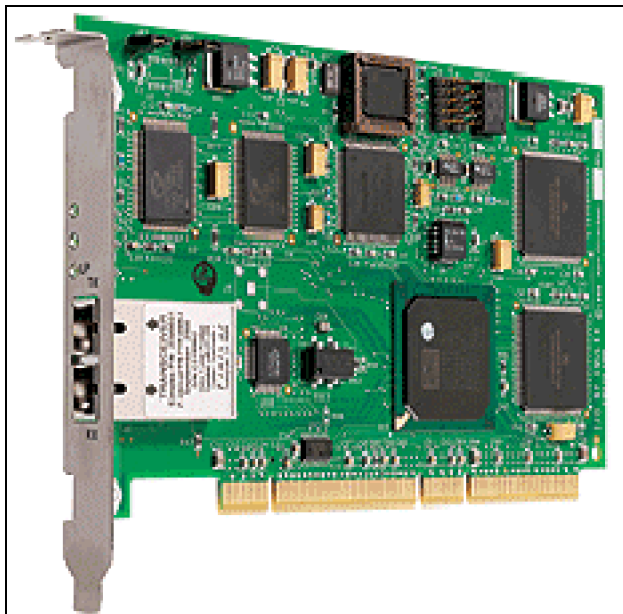


Figure 15-10 Host Bus Adapter

15.2.4 Naming and addressing

In various discussions, we will hear of different kinds of Fibre Channel ports and addresses. So it is important to understand important concepts involved in naming and addressing.

- ▶ A node is a communicating device.

- ▶ A node may contain one or more ports called N_Ports. Each N_port has a unique 8-byte port name or World-Wide-Name.
- ▶ During communication each N_Port is assigned a 24-bit port address called N_Port ID which is used for frame routing.
- ▶ A Fibre Channel fabric switch port is called F_Port. It is for attachment to N_Ports.
- ▶ An E_Port is a fabric switch expansion port. It is used to connect fabrics.

Note: For more information about naming and addressing, refer to the redbook *Introduction to Storage Area Networks, SG24-5470*.

15.2.5 FICON

FICON (Fibre Connectivity) is a high speed input/output interface for mainframe computer connections to storage devices. A FICON channel is a high bandwidth connection between processor and storage device within a relatively close proximity.

FICON channels increase I/O capacity through the combination of a new architecture and faster physical link rates to make them more efficient than ESCON (Enterprise System Connection), IBM previous fiber optic channel standard.

FICON is based on FCP, and runs on an FCP infrastructure - including all the cabling, switched, and directors.

Note: For more information about FICON, refer to the redbook *FICON Native Implementation and Reference Guide, SG24-6266*.

15.3 IP Storage Networking technologies

ISCSI, FCIP and iFCP are SAN extension technologies ideal for connecting smaller departmental and less I/O intensive servers into a SAN.

SAN deployment and its resulting benefits have primarily been focused on mission-critical islands of application servers within individual data centers. The difficulty and cost associated with migrating the large number of data center mid-range servers to Fibre Channel have made it impractical for IT managers to extend the benefits of SAN to mid-range applications. A basic diagram of these technologies is shown in Figure 15-11 on page 314.

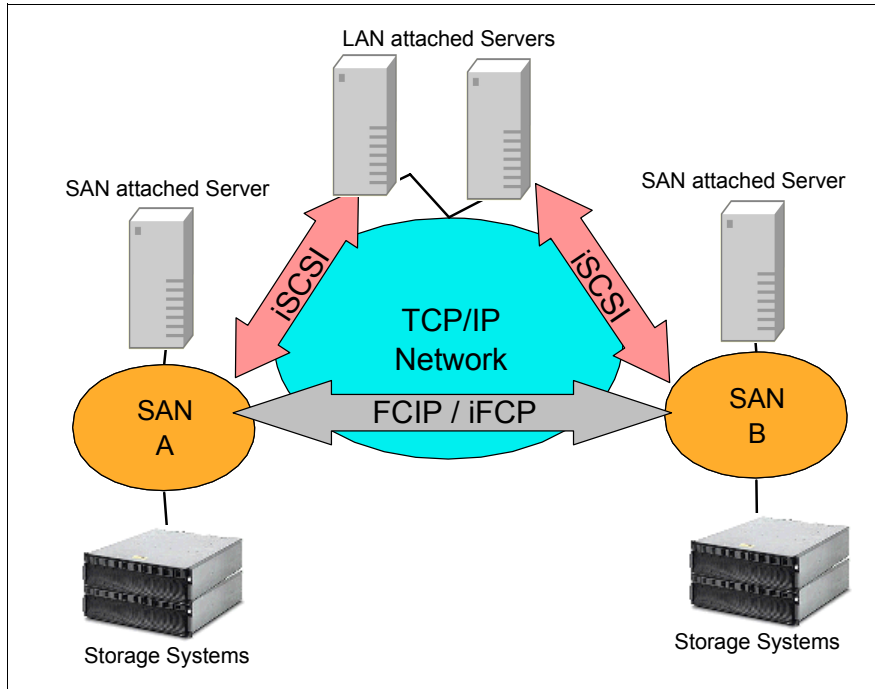


Figure 15-11 IP Storage networking technology

15.3.1 iSCSI

iSCSI allows storage to be accessed over a TCP/IP network as though it was locally attached. The server pushes SCSI commands out through the Ethernet NIC. As the SCSI commands exit out through the server, they are encapsulated within IP packet and are forwarded across the LAN to a LAN/SAN gateway interface. These iSCSI packets are then translated onto the Fibre Channel SAN through the TCP/IP transport and conversion protocol.

iSCSI is an ideal point to multipoint solution for connecting dispersed SAN islands. Using iSCSI, mid-range servers can gain access to consolidated storage while retaining their existing IP infrastructure, allowing a cost-effective extension of SAN benefits to mid-range applications residing on servers within the data center and on departmental servers located throughout the enterprise.

SCSI uses TCP/IP for reliable data transmission over potentially unreliable networks. The iSCSI layer interfaces to the operating system's standard SCSI set. The iSCSI layer includes encapsulated SCSI commands, data and status reporting capability. When, for example, the operating system or application requires a data write operation, the SCSI CDB (Command Descriptor Block) must be encapsulated for transport over a serial gigabit link and delivered to the target.

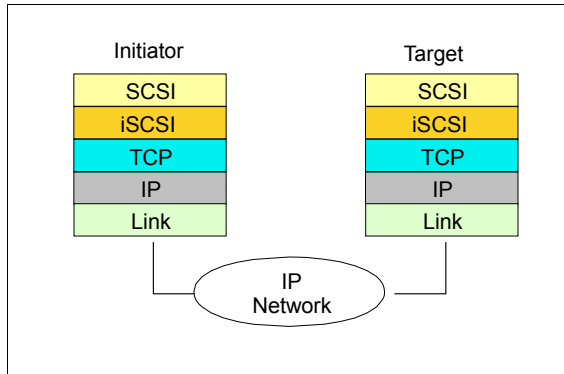


Figure 15-12 iSCSI protocol layering model

The iSCSI protocol monitors the block data transfer and validates completion of the I/O operation. This occurs over one or more TCP connections between initiator and target. In practical applications, an initiator may have multiple target resources over an IP network, and consequently multiple concurrent TCP connections active.

15.3.2 FCIP

Fibre Channel over IP (FCIP) uses a tunneling protocol to transport Fibre Channel frames over an existing IP infrastructure, and therefore is better suited for point-to-point solutions. Refer Figure 15-13. Using FCIP, customers can utilize their current wide-area networking infrastructure for connecting remote SAN islands over long distances.

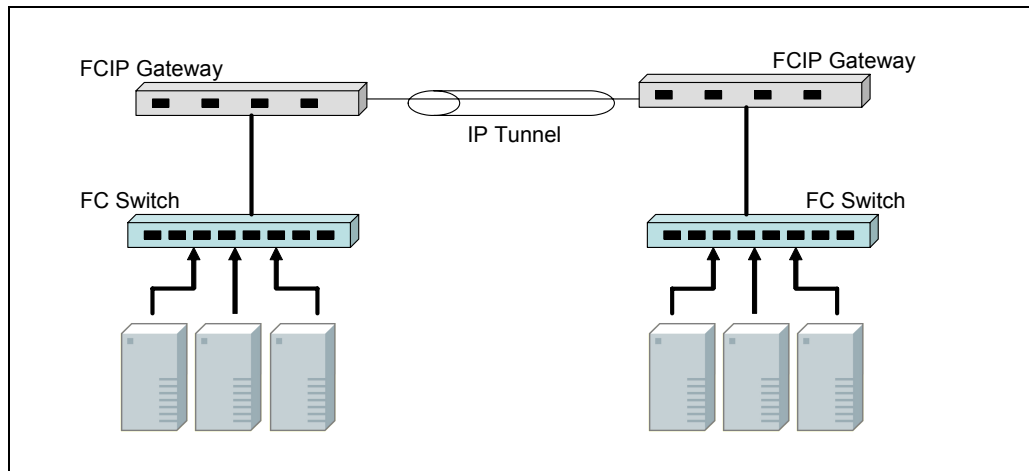


Figure 15-13 FCIP with IP tunneling

Fibre Channel over IP is an ideal combination of technologies to address the dual requirements of storage networking and networking over distance. Fibre Channel is a mature technology optimized for storage data movement within the campus and data center. It represents a major investment in software compatibility, interoperability, and proven applications for campus-based storage networking. Likewise, IP is a mature technology optimized for data movement across WAN distances. It represents a major investment in software compatibility, equipment interoperability, and proven applications for WAN-based data networking.

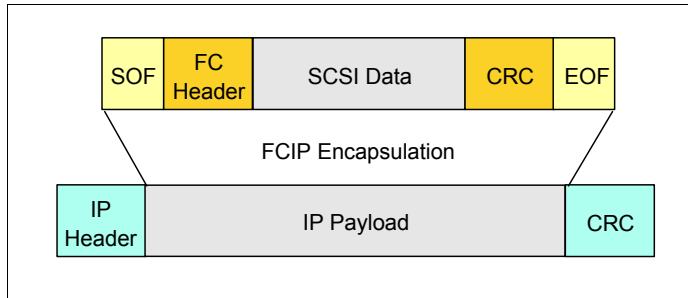


Figure 15-14 FCIP Encapsulation model

Today's Fibre Channel-over-IP solutions encapsulate Fibre Channel and transport it over a TCP socket (Figure 15-14). As in all IP networks, performance can vary based on the types of switches and routers, the number of hops the packets must traverse, and the level of congestion in the network. Today, storage transport performance over IP networks-especially over public networks- is limited due to the variable latency of service provider networks. As IP and Ethernet equipment continues to evolve, higher levels of Quality of Service (QoS), Cost of Service (CoS), provisioning, and circuit emulation should provide the latency guarantees required by synchronous storage applications. In controlled environments, these technologies might even improve the performance of IP networks. Regardless, Fibre Channel over IP is currently a very cost-effective technology for asynchronous applications such as remote data backup.

15.3.3 iFCP

Internet Fibre Channel Protocol (iFCP) is an emerging standard and TCP/IP-based method for interconnecting Fibre Channel SANs and SAN devices and IP networks. iFCP technology provides multipoint access to Fibre Channel devices. iFCP capitalizes on Internet protocol network services while leveraging the performance and interoperability capabilities of the Fibre Channel network. With iFCP, existing SCSI and Fibre Channel networks can be interconnected into the existing Internet Protocol environment. iFCP can be used in conjunction with Fibre Channel switching and routing protocols, or it can completely replace them.

iFCP is designed for customers who may have a wide range of Fibre Channel devices (that is, Host Bus Adapters, Subsystems, Hubs, Switches, and so on) and want the flexibility to interconnect these devices with IP network. iFCP can interconnect Fibre Channel SANs with IP, as well as allow customers the freedom to use TCP/IP networks in place of Fibre Channel networks for the SAN itself. Through the implementation of iFCP as a gateway-to-gateway protocol, these customers can maintain the benefit of their Fibre Channel devices while leveraging a highly scalable, manageable and flexible enterprise IP network as the transport medium of choice.

iFCP enables Fibre Channel device-to-device communication over an IP network, providing more flexibility compared to only enabling SAN-to-SAN communication. For example, iFCP has a TCP connection per N_Port to N_Port couple, and such a connection can be set to have its own Quality of Service (QoS) identity (Figure 15-15 on page 317).

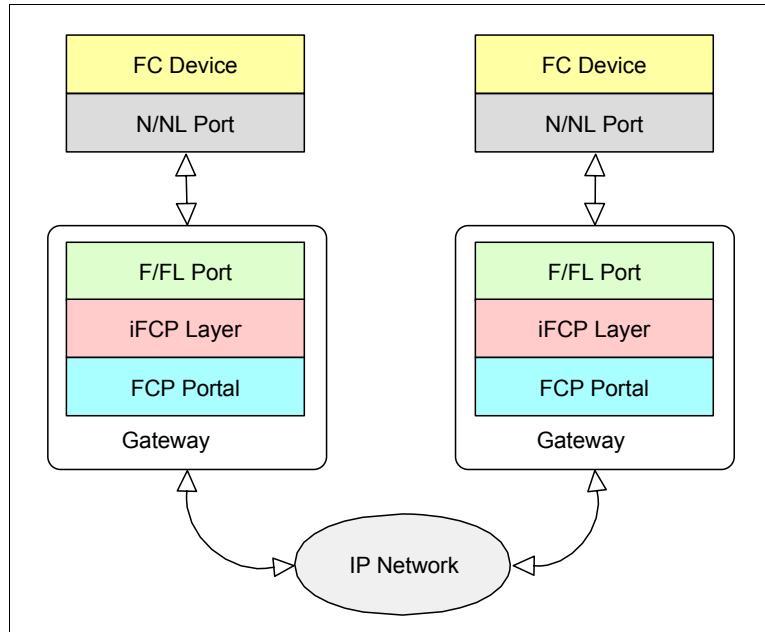


Figure 15-15 iFCP Gateway architecture

A drawback of the FCIP protocol which makes iFCP more attractive is that FCIP is a protocol that uses tunneling to encapsulate Fibre Channel data packets for forwarding over the TCP/IP network. This means that FCIP only works within a Fibre Channel infrastructure. Whereas, iFCP can handle both iSCSI and FCIP traffic. Applications developed for Fibre Channel SAN environments are supported over iFCP.

Some benefits of Internet Fibre Channel Protocol (iFCP) for storage networks are that scalability, distance and connectivity issues are virtually eliminated. Existing Ethernet protocol structure allows for rapid deployment of applications and solutions that already make use of the TCP/IP protocol layers. With one less routing protocol to support, network complexity and management is also potentially reduced. More importantly, the lower cost of Gigabit Ethernet switches as opposed to Fibre Channel switches enables for a lower total cost of ownership (TCO) of the enterprise SAN.

15.3.4 Comparison iSCSI - FCIP - iFCP

There is a short comparison provided here to show the differences between iSCSI, FCIP and iFCP. You can see in Figure 15-16 on page 318, the different structure of the protocol stacks as well as the different key attributes in Table 15-1 on page 318.

iSCSI

- ▶ A transport protocol for SCSI that operates on top of TCP
- ▶ A new mechanism for encapsulating SCSI commands on an IP network
- ▶ A protocol for a new generation of storage end-nodes that natively use TCP/IP

FCIP

- ▶ FCIP is a tunneling protocol for connecting geographically distributed Fibre Channel SANs transparently over LANs, MANs, or WANs
- ▶ Relies upon TCP for congestion control and management and upon both TCP and FC for data error and data loss recovery
- ▶ Uses TCP/IP as the transport while retaining Fibre Channel services intact

iFCP

- ▶ A gateway-to-gateway protocol for the implementation of a Fibre Channel fabric in which TCP/IP switching and routing elements supplement or replace Fibre Channel fabric components
- ▶ The protocol enables the attachment of existing Fibre Channel storage devices or Fibre Channel SANs to an IP network

The following benefits are common to all IP storage protocols:

- ▶ Built on SCSI and Ethernet technologies
- ▶ Provides more affordable SAN infrastructure
- ▶ Increase operating distance
- ▶ Improves availability of storage systems

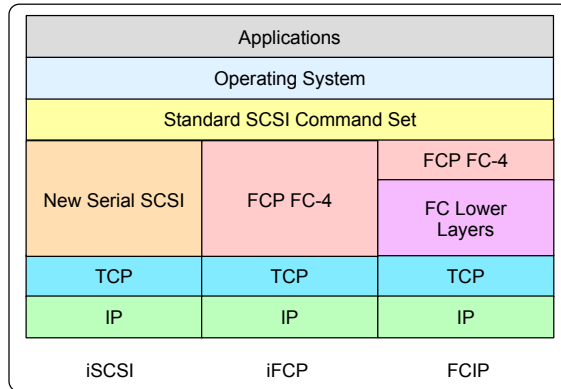


Figure 15-16 iSCSI, FCIP and iFCP protocol stacks

Table 15-1 Key attributes iSCSI, FCIP, iFCP

Protocol Attributes	iFCP	iSCSI	FCIP
Implementation	native IP transport	native IP transport	encapsulation, tunneling
SCSI encapsulation	FCP	new iSCSI layer	FCP
Prioritization based on port identification	yes	yes	no
End device interface	FC/FCP	IP/iSCSI	FC/FCP
End device routing	RIP, OSPF, BGP	RIP, OSPF, BGP	FSPF
Fibre Channel device support	yes	no	yes

15.4 Network Attached Storage (NAS)

A NAS is a device that resides on a network that may be shared with non-storage traffic. Today, the network is usually an Ethernet LAN, but could be any network that supports the IP-based protocols that NAS uses. Figure 15-17 on page 319 shows a picture of an NAS appliance. In contrast to “block I/O” used by SANs, NAS I/O requests are called “file I/Os”. File I/O is a higher-level type of request that, in essence, specifies the file to be accessed, an offset into the file (as though the file was a set of contiguous bytes), and a number of bytes to read or write beginning at that offset. File I/O requests are mainly of the type CIFS, NFS or specialized I/O protocols for file access and file sharing. Unlike block I/O, there is no

awareness of a disk volume or disk sectors in a file I/O request. Inside the NAS product, an operating system tracks where files are located on disk, and issues a block I/O request to the disks to fulfill the file I/O read and write requests it receives.

In contrast to SAN devices that can usually also be direct-attached (for example, by point-to-point Fibre Channel) as well as network-attached by SAN hubs and switches, an NAS device is generally only an NAS device and attaches only to processors over a LAN or WAN.

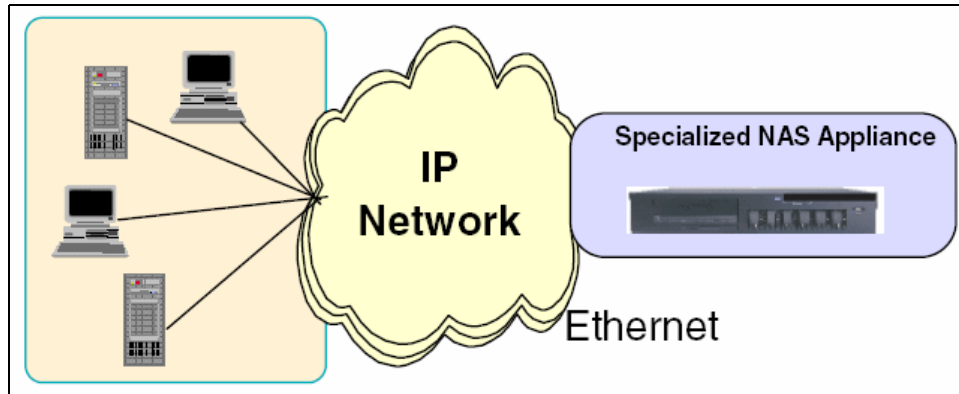


Figure 15-17 A picture of NAS appliance

The major potential benefits of an NAS device can be categorized as:

► **Ease-of-installation**

NAS is generally easier to install and manage than a SAN. A NAS appliance can usually be installed on an existing LAN/WAN network. NAS manufacturers often cite “up and running” times of 30 minutes or less. Hosts can potentially start to access NAS storage quickly, without needing disk volume definitions or special device drivers. In contrast, SANs take more planning, including design of a Fibre Channel network and selection/installation of SAN management software.

► **Resource pooling**

NAS allows capacity within the appliance to be pooled. That is, the NAS device is configured as one or more file systems, each residing on a specified set of disk volumes. All users accessing the same file system are assigned space within it on demand. That is certainly more efficient than buying each user their own disk volumes (DAS), which often leads to some users having too much capacity and others too little. So NAS pooling can minimize the need to manually reassign capacity among users. However, NAS pooling resides within an NAS appliance, and there is little if any sharing of resources across multiple appliances. This raises costs and management complexity as the number of NAS nodes increases. In contrast, an advantage of a SAN is that all devices on a SAN can be pooled—multiple disk and tape systems. So, at some point as total capacity grows, a SAN may be easier to manage and more cost-effective.

► **File sharing**

NAS provides file sharing using the NFS and CIFS protocol.

15.4.1 NAS gateways

A NAS gateway provides the function of a conventional NAS appliance but without integrated disk storage. The disk storage is attached externally to the gateway, possibly sold separately, and may also be a standalone offering for direct or SAN attachment. The gateway accepts a file I/O request (for example, using the NFS or CIFS protocols) and translates that to a SCSI

block-I/O request to access the external attached disk storage. The gateway approach to file sharing offers the benefits of a conventional NAS appliance, with additional potential advantages:

- ▶ Increased disk capacity scalability (compared to the capacity limits of an integrated NAS appliance).
- ▶ Ability to offer file sharing and block-I/O on the same disk system. Disk capacity in the SAN could be shared (reassigned) among gateway and non-gateway use. So a gateway can be viewed as an NAS/SAN hybrid, increasing flexibility and potentially lowering costs.

15.5 Wavelength Division Multiplexing

Wavelength Division Multiplexing (WDM) or Multiplexing by itself is not a storage networking protocol. Because it is used also to create storage networking infrastructure, we briefly explain the technology here.

Wavelength Division Multiplexing is deployed as part of the physical layer. It is therefore independent of protocol, simply passing signal information in the format it is received. WDM allows the simultaneous transmission of a number of data streams over the same physical fiber, each using a different optical wavelength. WDM receives incoming optical signals from many sources (Fibre Channel, IP, ESCON, FICON) which it converts to electrical signals, it then assigns them a specific wavelength (or lambdas) of light and retransmits them on that wavelength. This method relies on the large number of wavelengths available within the light spectrum. You can think about WDM as though each channel is a different color of light; several channels then make up a rainbow (Figure 15-18). In summary WDM enables many signals to be concentrated into a single fibre, all being sent at different wavelengths.

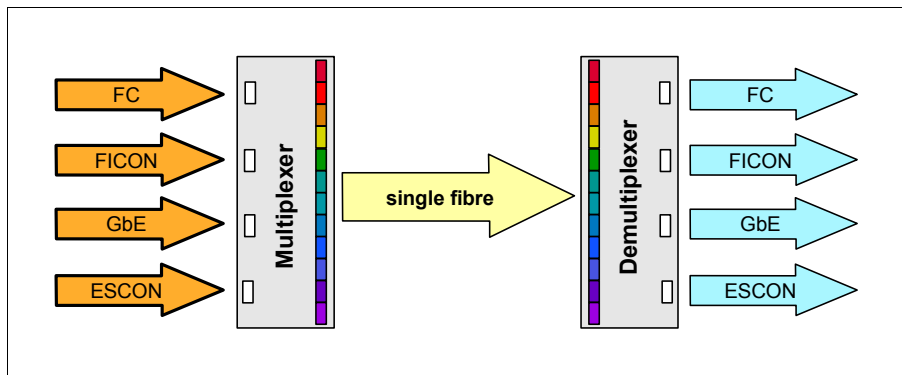


Figure 15-18 Wave Division Multiplexer concept

Each wavelength can carry a signal at any bit rate less than an upper limit defined by the electronics, typically up to several gigabits per second. Due to the nature of these boxes, they are often considered transparent to protocol and bit rate.

Today, WDM of itself, is rarely seen. Coarse Wave Division Multiplexing (CWDM), which combines up to 16 wavelengths onto a single fiber, and Dense Wave Division Multiplexing (DWDM) which combines up to 64 wavelengths onto a single fiber are the standard.

15.6 Selecting the best alternative

Which storage networking alternative is best for a given organization may be obvious based on organizational objectives, current storage infrastructure and what the alternatives provide.

Or, it may be a totally open question. Storage technology has clearly become more varied and sophisticated, and accordingly decisions have become more complex than ever. Choice means flexibility and that's good, but which choice to make is not always clear.

- ▶ If a group of individual users with PCs needs to share disk storage capacity and perhaps also share files in that storage, then NAS may be easiest to install and manage.
- ▶ If application servers need to share disk storage, and are each accessing independent (block I/O) databases, a FC based SAN may be appropriate.
- ▶ For a small number of servers where no SAN exists, iSCSI may be less expensive and less complex.
- ▶ If required, FCIP or iFCP provide a cost-effective way to achieve business protection, enabling such solutions as remote tape archiving.

15.7 More information

For a better understanding of Fibre Channel, SAN technology, FICON and NAS technology refer to these Redbooks:

- ▶ *Introduction to Storage Area Networks*, SG24-5470
- ▶ *Designing an IBM Storage Area Network*, SG24-5758
- ▶ *IBM SAN Survival Guide*, SG24-6143
- ▶ *FICON Implementation Guide*, SG24-6497
- ▶ *FICON (FCV Mode) Planning Guide*, SG24-5445
- ▶ *Introduction to IBM S/390 FICON*, SG24-5176
- ▶ *IBM S/390 FICON Implementation Guide*, SG24-5169
- ▶ *The IBM TotalStorage NAS Integration Guide*, SG24-6505



Storage Area Network products

Companies are searching for more efficient ways to manage ever-expanding volumes of data and to make that data accessible throughout the enterprise - this is propelling the move of storage into the network. The Storage Area Network (SAN) infrastructure offers simplified storage management, scalability, flexibility, availability, and improved data access, movement, and backup.

SAN switches interconnect multiple host servers with storage servers and devices to create a SAN. SAN switches can be used either as a standalone device to build a simple SAN fabric, or they can be interconnected with other switches to build a larger SAN fabric.

In this chapter we present the IBM System Storage SAN product portfolio, which includes Fibre Channel switches and directors with Fibre Channel and FICON ports.

The following entry, mid-range and enterprise level switch products are presented:

- ▶ IBM System Storage SAN b-type switches and directors
- ▶ IBM System Storage SAN m-type family
- ▶ IBM System Storage SAN n-type directors
- ▶ Cisco switches and directors

Note: These products can be used with disk storage systems, which are explained in Part 1, “Disk systems” on page 57. Additionally, these products can be used with tape drives and libraries, as discussed in Part 2, “Tape systems” on page 199.

For the latest SAN products and information, refer to:

<http://www.ibm.com/servers/storage/san/>

16.1 SAN switches and connectivity products

IBM System Storage SAN switch offerings provide integrated small and medium business (SMB) and enterprise solutions with multiprotocol local, campus, metropolitan and global storage networking. IBM provides the choice of Brocade (b-type), McDATA (m-type) and Cisco switches and directors.

IBM System Storage SAN switch products provide a broad range of storage networking options designed with a common architecture and integrated enterprise SAN management capabilities, and supported by the broadest range of IBM open server and storage devices. The interconnection of IBM and compatible switches can support the creation of scalable, dual redundant core-to-edge SAN fabrics that can support high performance, scalability, and fault tolerance required by e-business applications and enterprise storage management applications.

IBM System Storage SAN switch offerings can be positioned into four broad product groups: entry, mid-range, enterprise and multiprotocol router. Suitable products for every Storage Area Network project can be found among these.

16.1.1 Common characteristics

With all IBM System Storage SAN products you get:

- ▶ Industry standard performance with 1, 2, and 4 Gigabit per second (Gbps) throughput
- ▶ Broadest range of IBM open server and storage support including fabric, loop and private loop attachments
- ▶ Intelligent Fabric Services Architecture providing switch interoperability
- ▶ Enterprise-level scalability with fault-tolerant core-to-edge SAN fabrics containing thousands of devices
- ▶ Open fabric management allowing support for the widest range of solutions—from very small workgroup SANs up to very large enterprise SAN fabrics with thousands of devices
- ▶ Flexible management options including Tivoli Ready certification for centralized management of very large enterprise SAN fabrics
- ▶ Manageable by IBM TotalStorage Productivity Center for Fabric
- ▶ Common enterprise SAN fabric which simplifies deployment, management and network growth
- ▶ Common firmware which enables introduction of new switch technologies while protecting prior switch investments
- ▶ Flexible Fibre Channel connectivity provides connectivity to a host of IBM and non-IBM servers and storage products
- ▶ Pay-as-you-grow scalability provides scalable network growth in a modular, cost-effective, and nondisruptive manner

16.1.2 Other switch features

Here are other features available in selected IBM System Storage SAN switch products.

Advanced Web Tools

Advanced Web Tools is an intuitive graphical user interface (GUI) which allows network managers to monitor and manage SAN fabrics consisting of switches using a Java-capable Web browser from standard desktop workstations.

Advanced Performance Monitoring

Advanced Performance Monitoring enables SAN administrators to monitor transmit and receive traffic from the source device all the way to the destination device. This end-to-end visibility into the fabric enables SAN administrators to identify bottlenecks and optimize fabric configuration resources.

Advanced Security

Advanced Security (AS) significantly reduces the security holes left by traditional SAN implementations and greatly improves the ability to minimize SAN-specific vulnerabilities by providing a comprehensive, policy based security system for IBM SAN Switch fabrics.

Advanced Zoning

Advanced Zoning segments a fabric into virtual private SANs. It provides data exchange between devices in the same zone and prohibits exchange to any device not in the same zone. Advanced zoning of 1, 2, and 4 Gbps switches separately enlarges the range of hardware enforcement and so provides the switch with more security access control functions as before, preventing unauthorized devices from accessing the fabric.

Common IBM SAN Switch capabilities

IBM System Storage SAN Switches include universal ports that can determine the port type when connected to a fabric port (F_port), fabric loop port (FL_port) or expansion port (E_port).

Enterprise Fabric Connectivity Manager

Enterprise Fabric Connectivity Manager (EFCM) software is designed to support interconnection of multiple IBM System Storage SAN m-type switches and directors for the creation of enterprise-to-edge SANs.

Extended Fabric

The Extended Fabric feature provides extensions within the internal switch buffers. This maintains performance with distances greater than 10 km, and up to 120 km, by maximizing buffering between the selected switch interconnect links. With the Extended Fabric feature, the ISLs are configured with up to 60 buffer credits and optimize buffers for up to 120 km on 1 Gbps fiber optic link, and up to 60 km on 2 Gbps fiber optic link.

Fabric Manager

Fabric Manager provides a Java-based application that can simplify management of a multiple switch fabric. It administers, configures, and maintains fabric switches and SANs with host-based software.

Fabric Watch

Fabric Watch enables switches to continuously monitor the health of the fabrics, watching for potential faults based on defined thresholds for fabric elements and events, so making it easy to quickly identify and escalate potential problems.

FlexPort scalability

FlexPort technology is designed to support scalable switch upgrades in four-port increments without fabric disruption.

ISL Trunking

ISL Trunking feature allows up to four ISLs between the same pair of switches to be grouped and to act as a single, high speed “pipe” or trunk with a capacity of up to 8 Gbps.

Remote Switch

The Remote Switch feature enables two fabric switches to be connected over an asynchronous transfer mode (ATM) connection. This requires a compatible Fibre Channel to ATM gateway, and can have a distance of up to 10 km between each switch and the respective ATM gateway.

More information

- ▶ For the interoperability matrix, go to:
<http://www.ibm.com/servers/storage/support/san/index.html>
Select the product from list then click *Plan and Upgrade* under the desired switch product.
- ▶ For technical support information, go to:
<http://www.ibm.com/servers/storage/support/san/index.html>
- ▶ SAN-related Redbooks include:
 - *Introduction to Storage Area Networks*, SG24-5470
 - *IBM TotalStorage: SAN Product, Design, and Optimization Guide*, SG24-6384
 - *IBM SAN Survival Guide Featuring the IBM 3534 and 2109*, SG24-6127
 - *IBM TotalStorage: Implementing an Open IBM SAN*, SG24-6116

16.2 Entry SAN switches

The IBM System Storage SAN switch entry-level products are designed specifically to address the needs of small to medium-size SAN environments. They can be used to create a wide range of high-performance SAN solutions, from simple single-switch configurations to larger multi-switch configurations which support fabric connectivity and advanced business continuity capabilities. Infrastructure simplification solutions for IBM System x, IBM System i, and IBM System p include storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity solutions include data protection with IBM System Storage tape libraries and devices and IBM Tivoli Storage Manager data protection software. IBM entry fabric switches provide up to 4 Gbps fully non-blocking performance, and advanced intelligence features. Scalability ranges from 4 to 20 ports.

Entry-level products include:

- ▶ IBM TotalStorage SAN16B-2
- ▶ IBM TotalStorage SAN16M-2
- ▶ Cisco MDS 9020 Fabric Switch

In this section, we introduce IBM System Storage Switch solutions ideal for entry-level applications.

16.2.1 IBM TotalStorage SAN16B-2

The IBM TotalStorage SAN16B-2 (2005-B16), also marketed as the IBM TotalStorage SAN16B-2 Express Model (2005-16B) fabric switch is the Brocade SilkWorm 200E. It is intended for environments which are just beginning to implement Fibre Channel storage systems, and which are looking for storage consolidation benefits. A typical entry-level

configuration can consist of one or two Fibre Channel links to a disk storage array or to an LTO tape drive. An entry-level eight-port storage consolidation solution can support up to seven servers with a single path to either disk or tape. A small configuration such as four servers and one disk storage array is shown in Figure 16-1. The Ports on Demand feature allows the switch to grow to sixteen ports to support more servers and more storage devices without taking the switch offline.

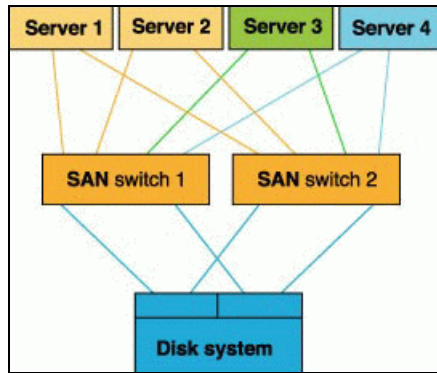


Figure 16-1 Sample Configuration for SAN16B-2

Highlights

- ▶ Simple-to-use SAN switch with ease-of-installation and ease-of-use features designed specifically for the needs of small to medium-size environments
- ▶ Foundation for new infrastructure simplification and business continuity solutions for servers running Microsoft Windows, UNIX, Linux, NetWare and OS/400 operating systems
- ▶ High-performance, 4 Gbps links (requires storage hardware that supports 4 Gbps throughput) with pay-as-you-grow “Ports on Demand” scalability enable growth from 8 to 16 ports

Figure 16-2 shows the IBM TotalStorage SAN16B-2.

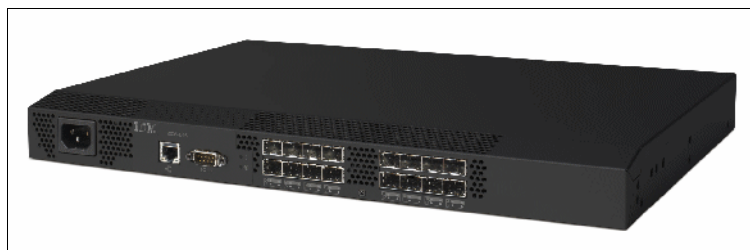


Figure 16-2 IBM TotalStorage SAN16B-2

More information

- ▶ For additional product details, go to:
http://www.ibm.com/servers/storage/san/b_type/san16b-2/

16.2.2 IBM TotalStorage SAN16M-2

An IBM TotalStorage SAN16M-2 (2026-16E or 2026-416) is the McDATA Sphereon 4400 Fabric Switch. The SAN16M-2 is intended for environments which are just beginning to implement Fibre Channel storage systems, and which are looking for storage consolidation benefits. A typical entry-level configuration can consist of one or two Fibre Channel links to a

disk storage array or to an LTO tape drive. An entry-level eight-port storage consolidation solution can support up to seven servers with a single path to either disk or tape. The FlexPort feature is designed to enable a base switch to grow to 16 ports, in four port increments, to support more servers and more storage devices without taking the switch offline.

Highlights

- ▶ Simple-to-use SAN switch with ease-of-installation and ease-of-use features designed specifically for the needs of small- to medium-sized environments
- ▶ Foundation for new infrastructure simplification and business continuity solutions for servers running Microsoft Windows, UNIX, Linux, NetWare and OS/400 operating systems
- ▶ High-performance, 4 Gbps links with pay-as-you-grow FlexPort scalability enable growth from 8 to 12 to 16 ports
- ▶ Designed for high availability with hot-swappable, dual power supplies and HotCAT online code activation



Figure 16-3 IBM TotalStorage SAN16M-2

More information

- ▶ For additional product details, go to:
http://www.ibm.com/servers/storage/san/m_type/san16m-2/

16.2.3 Cisco MDS 9020 Fabric Switch

Cisco MDS 9020 switch (2061-420) is intended for environments which are just beginning to implement Fibre Channel storage systems, and which are looking for storage consolidation benefits. A typical entry-level configuration can consist of one or two Fibre Channel links to a disk storage array or to an LTO tape drive. An entry-level eight-port storage consolidation solution can support up to seven servers with a single path to either disk or tape. as Figure 16-4 on page 329. An entry-level, eight-port storage consolidation solution could support up to seven servers with a single path to either disk or tape. The shortwave optical transceiver feature allows the base switch to grow to 20 ports, in single port increments, to support more servers and more storage devices without taking the switch offline.

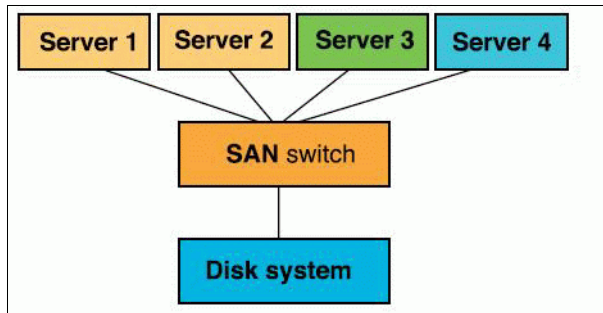


Figure 16-4 One disk storage array sample

Highlights

- ▶ SAN switch with Cisco Fabric Manager software that can help simplify management of multiple switch fabrics
- ▶ Foundation for new infrastructure simplification and business continuity solutions for servers running Microsoft Windows, UNIX, Linux, NetWare and OS/400 operating systems
- ▶ High-performance switch ready to support 4 Gbps Fibre Channel capable servers and storage
- ▶ Designed for high availability with support for nondisruptive firmware upgrades



Figure 16-5 Cisco MDS 9020 Fabric Switch

A wide range of IBM System Storage storage area network (SAN) infrastructure simplification and Business Continuity solutions can be created with the Cisco MDS 9020 Fabric Switch. Infrastructure simplification solutions for System x, System i, and System p servers include high-availability server clustering and storage consolidation with IBM System Storage disk storage arrays. Business Continuity solutions include data protection with IBM TotalStorage Linear Tape-Open (LTO) or IBM System Storage TS1030 and TS1120 Tape Drives and IBM Tivoli Storage Manager data protection software.

More information

- ▶ For additional product details, go to:
http://www.ibm.com/servers/storage/san/c_type/9020/

16.3 Mid-range SAN switches

IBM System Storage mid-range SAN solutions provide additional capability, features and benefits beyond the simple entry solutions. The IBM System Storage mid-range SAN switches provide 2 and 4 Gbps port-to-port non-blocking throughput with auto-sensing capability for connecting to older 1 Gbps host servers, storage, and switches. They come in 16, 24, 32, and 40 port models.

IBM System Storage SAN switches can be used to create dedicated, reliable, and high performance networks for storage products like disk subsystems, tape drives, and tape drive libraries. In addition, all of these models are fully interoperable with the previous IBM System Storage SAN Switches, and can be added to existing fabrics, enabling transition from existing Fibre Channel storage networks to the faster technology.

Mid-range products include:

- ▶ IBM TotalStorage SAN32B-2
- ▶ IBM TotalStorage SAN32M-2
- ▶ IBM TotalStorage SAN64B-2
- ▶ Cisco MDS 9120 and 9140 Fabric Switches
- ▶ Cisco MDS 9216A Fabric Switch

In this section, we introduce IBM System Storage Switch solutions ideal for mid-range applications.

16.3.1 IBM TotalStorage SAN32B-2

The IBM TotalStorage SAN32B-2 fabric switch (2005-B32), also marketed as the IBM TotalStorage SAN32B-2 Express Model (2005-32B) is the Brocade SilkWorm 4100 switch. It provides for creation of a wide range of IBM System Storage midrange and enterprise storage area network (SAN) infrastructure simplification and business continuity solutions. Infrastructure simplification solutions for IBM System i, System p, System x and System z include storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity solutions include data protection with IBM System Storage tape libraries and devices and IBM Tivoli Storage Manager data protection software.

Highlights

- ▶ Simple-to-use midrange and enterprise infrastructure simplification and business continuity solutions for IBM System i, System p, System x and System z servers
- ▶ 4 Gbps ports and enhanced Inter-Switch Link (ISL) trunking with up to 32 Gbps per data path
- ▶ Pay-as-you-grow scalability with Ports on Demand features, growing from 16 to 24 to 32 ports.
- ▶ Redundant, hot-swappable fans and power supplies and nondisruptive software upgrades
- ▶ Broad range of open server Fibre Channel and mainframe FICON switching including FICON/Fibre Channel Intermix, FICON CUP (Control Unit Port) and FICON cascading
- ▶ Multiple management options for first-time storage area network (SAN) users and complex enterprise SAN consolidation solutions
- ▶ Interoperable with IBM System Storage SAN b-type switch family helps protect switch investment

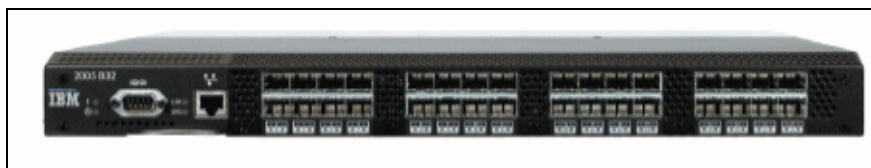


Figure 16-6 IBM TotalStorage SAN32B-2

A high-availability server clustering solution can be created with redundant switches, as shown in Figure 16-7. A midrange server clustering solution consists of up to 30 servers each

with dual Fibre Channel adapters crossconnected to redundant SAN32B-2 switches with two disk storage arrays, each with dual adapters.

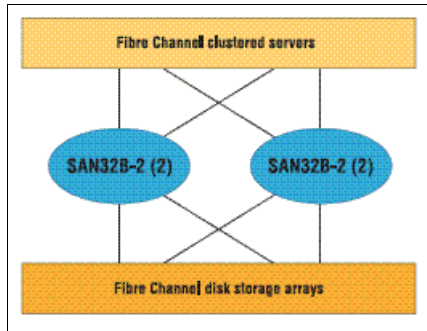


Figure 16-7 High-availability server clustering solution

More information

- ▶ For additional product details, go to:

http://www.ibm.com/servers/storage/san/b_type/san32b/

16.3.2 IBM TotalStorage SAN32M-2

The IBM TotalStorage SAN32M-2 is the McDATA Sphereon 4700 Fabric Switch. A single SAN32M-2 switch (2026-32E or 2026-432) is intended for medium-sized and enterprise SAN environments. An entry-level, 16-port storage consolidation solution can support up to 15 servers with a single path to either disk or tape. The FlexPort feature is designed to enable a base switch to grow to 32 ports, in eight port increments, to support more servers and more storage devices without taking the switch offline.

A high-availability solution can be created with redundant switches. This capability is ideal for server clustering environments. Such a configuration can support from 14 to 30 servers, each with dual Fibre Channel adapters cross-connected to redundant SAN32M-2 switches which are cross-connected to a dual-controller storage system.

While the SAN32M-2 can be the foundation of medium-sized SANs, it can be configured to participate as a full member in a tier enterprise SAN with other members of the IBM System Storage SAN m-type family. This capability helps provide investment protection as SAN requirements evolve and grow over time.

Highlights

- ▶ Simple-to-use SAN switch with ease-of-installation and ease-of-use features designed specifically for the needs of medium-sized and enterprise environments
- ▶ Foundation for new infrastructure simplification and business continuity solutions for servers running Microsoft Windows, UNIX, Linux, NetWare and OS/400, AIX, and z/OS operating systems
- ▶ 1, 2 and 4 Gbps links with pay-as-you-grow FlexPort scalability enables growth from 16 to 24 to 32 ports
- ▶ Hot-swappable, dual power supplies and HotCAT online code activation

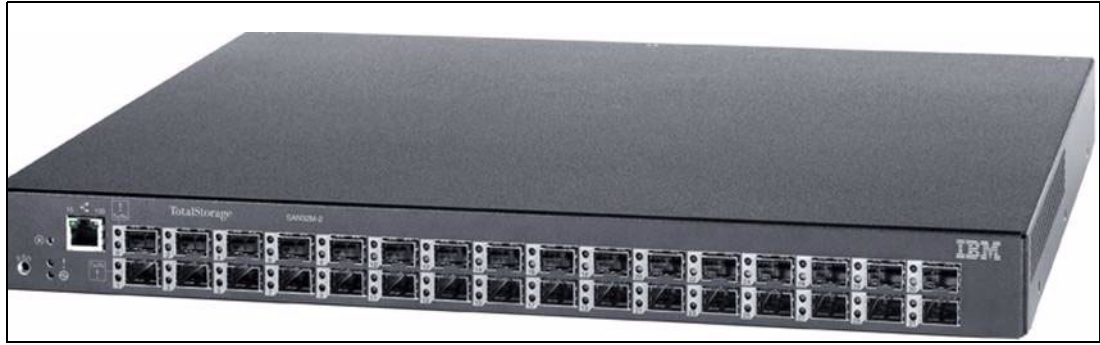


Figure 16-8 IBM TotalStorage SAN32M-2

Business continuity solution

Many small- and medium-sized companies want to implement a business continuity or remote backup capability to help address strict new regulatory requirements. The SAN32M-2 fabric switch can help provide the SAN connectivity required for these environments. The Figure 16-9 below is intended to represent two different sites. The blue side represents the production site and the green side represents the remote or backup site.

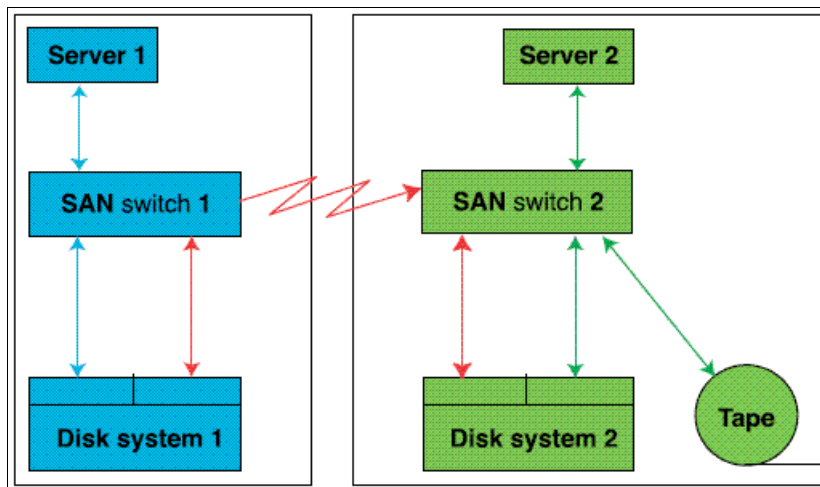


Figure 16-9 Business Continuity Solution with SAN32M-2

More information

- For additional product details, go to:

http://www.ibm.com/servers/storage/san/m_type/san32m-2/

16.3.3 IBM TotalStorage SAN64B-2

A wide range of IBM System Storage medium-size and enterprise storage area network (SAN) infrastructure simplification and business continuity solutions can be created with the IBM System Storage SAN64B-2 fabric switch. Infrastructure simplification solutions for the IBM System i, System p, and System x families of servers include storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity solutions include data protection with IBM System Storage tape libraries and devices and IBM Tivoli Storage Manager data protection software.

Highlights

- ▶ Simple-to-use midrange and enterprise infrastructure simplification and business continuity solutions for IBM System i, System p, System x and System z servers
- ▶ Designed for high-performance with 4 Gbps ports and enhanced Inter-Switch Link (ISL) trunking with up to 32 Gbps per data path
- ▶ Pay-as-you-grow scalability with Ports on Demand features
- ▶ Designed to support high availability with redundant, hot-swappable fans and power supplies and nondisruptive software upgrades
- ▶ Provides broad range of open server Fibre Channel switching
- ▶ Multiple management options for medium-size and complex enterprise SAN consolidation solutions
- ▶ Interoperability with IBM System Storage SAN b-type switch family helps protect switch investment

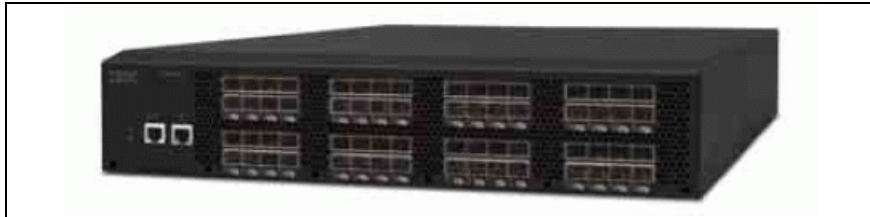


Figure 16-10 IBM TotalStorage SAN64B-2

More information

- ▶ For additional product details, go to:
http://www-03.ibm.com/servers/storage/san/b_type/san64b-2/

16.3.4 Cisco MDS 9120 and 9140 Multilayer Switches

The Cisco MDS 9120 (2061-020) and 9140 (2061-040) multilayer fabric switches are 1 RU (rack-unit) fabric switches that can support 20 or 40 shortwave or longwave SFP fiber optic transceivers at 2 Gbps for Windows and UNIX server clustering, infrastructure simplification and Business Continuity solutions.

The MDS 9120 has a total of 20 ports - the first group of four ports are full bandwidth ports, and the remaining four groups of ports are host optimized port groups.

The MDS 9140 has a total of 40 ports - the first group of eight ports are full bandwidth ports, and the remaining four groups of ports are host optimized port groups.

Virtual SAN (VSAN) and Cisco Fabric Manager are also included with these switches.

Highlights

- ▶ MDS 9000 inter-family compatibility supports scalability and consistent service as the SAN grows
- ▶ Up to 2 Gbps per port throughput and PortChannel support high performance core edge SAN deployments
- ▶ Host-optimized and target-optimized Fibre Channel ports designed to help reduce total cost of ownership
- ▶ Compact 20 and 40 port design with high availability capabilities

- ▶ Built-in intelligent network services can help simplify SAN management and reduce total cost of ownership
- ▶ Comprehensive security features support SAN consolidation
- ▶ Virtual SAN (VSAN) capability is designed to create virtual SAN islands on a single physical fabric
- ▶ Enterprise Package and Fabric Manager Server Package provide added intelligence and value

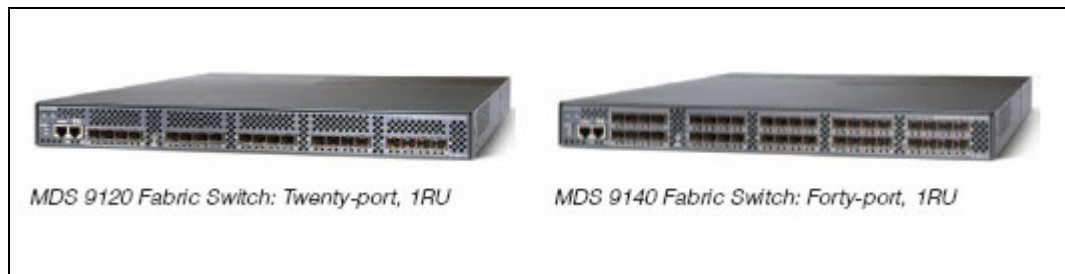


Figure 16-11 Cisco MDS 9120 and 9140 Fabric Switches

High performance and MDS 9000 family compatibility

The Cisco MDS 9120 and 9140 Multilayer Fabric Switches feature 1 Gbps and 2 Gbps speeds and intelligent network services for the security, performance and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN. These switches Cisco MDS 9120 and 9140 switches help address the needs for high performance and reliability in SAN applications ranging from mid-range to large enterprise SAN environments. Full functional compatibility with other members of the Cisco MDS 9000 family supports end-to-end service delivery in large data center core-edge deployments.

Flexible Fibre Channel connectivity

The Fibre Channel ports support an auto-sensing 1 Gbps and 2 Gbps interface for high-performance connectivity and compatibility with existing or earlier devices.

Target-optimized and host-optimized ports designed for reduced TCO

Target-optimized ports can help support the most demanding storage networking applications. These ports are designed to attach high performance servers and storage subsystems as well as connect to other switches using Inter Switch Link (ISL) connections.

Capabilities to help reduce TCO

VSAN capability allows more efficient SAN utilization by creating multiple isolated environments within a single SAN fabric. Each VSAN can be zoned as a typical SAN and maintains its own fabric services for added scalability and resilience. VSANs allow the cost of the SAN infrastructure to be shared among more users, while helping to segregate and secure traffic and retain independent control of configurations on a VSAN-by-VSAN basis.

More information

- ▶ For additional product details, go to:
http://www.ibm.com/servers/storage/san/c_type/9100/

16.3.5 Cisco MDS 9216A Fabric Switch

The Cisco MDS 9216A Fabric Switch (2062-D1A) provides 1, 2 and 4 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN. The Cisco MDS 9216A switch helps address the needs for high performance and reliability in SAN environments ranging from small workgroups to very large, integrated global enterprise SANs.

Highlights

- ▶ Supports Fibre Channel throughput of up to 4 Gbps per port and up to 64 Gbps with each Port Channel Inter-Switch Link (ISL) connection
- ▶ Scalable from 16 to 64 Fibre Channel ports
- ▶ 10 Gbps ISL ports for inter-Data Center links over metro optical networks
- ▶ Modular design with excellent availability capabilities
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric



Figure 16-12 Cisco MDS 9216A Fabric Switch

A modular switch designed for high availability

The Cisco MDS 9216A Fabric Switch uses Fabric Shortest Path First (FSPF) multipath routing, using intelligence to load balance across a maximum of 16 equal-cost paths and to dynamically reroute traffic if a switch fails.

The basic Cisco MDS 9216 configuration has redundant power supplies and cooling components. Hot-swappable components include an optional switching module, SFP optics, power supplies and a fan tray with integrated temperature and power management.

More information

- ▶ For additional product details, go to:
http://www.ibm.com/servers/storage/san/c_type/9216/

16.4 Enterprise SAN switches

The IBM System Storage SAN switches and directors are well suited to address enterprise SAN customer requirements for infrastructure simplification and improved Business Continuity.

The SAN Director is designed to be interoperable with other members of the IBM System Storage switch family. You can configure a wide range of highly scalable solutions that address demands for integrated, heterogeneous mainframe and open server enterprise SANs.

Enterprise products include:

- ▶ IBM TotalStorage SAN256B
- ▶ IBM TotalStorage SAN SAN140M
- ▶ IBM TotalStorage SAN256M
- ▶ Cisco MDS 9506 Multilayer Director
- ▶ Cisco MDS 9509 Multilayer Director
- ▶ Cisco MDS 9513 Multilayer Director

In this section, we introduce IBM System Storage Switch solutions ideal for enterprise applications. Advanced security with comprehensive, policy-based security capabilities can improve availability and simplify operation.

16.4.1 IBM TotalStorage SAN256B

The IBM TotalStorage SAN256B (2109-M48) is the Brocade SilkWorm 48000 switch. It delivers high performance, scalability, flexibility, functionality and availability. It provides full-duplex link speeds of 1, 2 and 4 Gbps capable of automatically negotiating to the highest speed supported by the attached server, storage or switch. The SAN256B is well suited to address enterprise SAN customer requirements for infrastructure simplification and improved business continuity.

The SAN256B director interoperates with other members of the IBM System Storage SAN b-type family. It can be configured with a wide range of highly scalable solutions that address demands for integrated IBM System z and open system server enterprise SANs.

Highlights

- ▶ High availability with built-in redundancy designed to avoid single points of failure
- ▶ Highly scalable director with 16 or 32 ports per port switch blade, and from 16 to 256 ports in a single domain
- ▶ Multiprotocol router blade with sixteen Fibre Channel (FC) ports and two Internet Protocol (IP) ports for SAN routing and distance extension over IP
- ▶ Port switch blades support FICON Director switching with Fibre Channel/FICON intermix, FICON CUP (Control Unit Port) and FICON cascading
- ▶ Interoperable with other IBM System Storage SAN b-type switches and directors
- ▶ Offers advanced security with comprehensive policy-based security capabilities
- ▶ Offers advanced fabric services such as end-to-end performance monitoring and fabric-wide health monitoring



Figure 16-13 IBM TotalStorage SAN256B

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/b_type/san256b/

16.4.2 IBM TotalStorage SAN140M

The IBM TotalStorage SAN140M (2027-140) is the McDATA Intrepid 6140 Director. The director is a 140-port product that provides dynamic switched connections between Fibre Channel servers and devices in a SAN environment. It is 12U high, so up to three can be configured in an SANC40M cabinet equipment cabinet, providing up to 420 ports in a single cabinet.

It provides 2 Gbps fabric switching for Windows and UNIX environments, and FICON switching for mainframe server clustering and Business Continuity solutions. You can interconnect with any m-type fabric switch to provide FC-AL tape attachment.

Infrastructure simplification solutions for System x, System i and System p Fibre Channel servers and System z mainframe FICON servers include storage consolidation and highest-availability server clustering with IBM System Storage DS4000 series and disk storage arrays.

Business continuity solutions include data protection with IBM TotalStorage LTO3 or IBM System Storage TS1030 and TS1120 Tape Drives and IBM Tivoli Storage Manager data protection software.

Highlights

- ▶ Easy-to-manage enterprise infrastructure simplification and Business Continuity solutions for System x, System i, System p and System z servers
- ▶ Highly-scalable 16- to 140-port switching backbone for advanced enterprise infrastructure simplification and Business Continuity solutions including mainframe FICON disk and tape storage
- ▶ Designed to provide high availability of all IBM SAN switches through redundancy of active components, including hot-swappable processors, fans and power supplies; HotCAT online code activation; and call-home capability with Enterprise Fabric Connectivity Manager (EFCM) software

- ▶ Enterprise Fabric Connectivity Manager, FICON Management Server (CUP) and Open Systems Management Server software designed to help simplify management of complex SAN infrastructures



Figure 16-14 IBM TotalStorage SAN 140M

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/m_type/san140m/

16.4.3 IBM TotalStorage SAN256M

IBM TotalStorage SAN256M high availability enterprise director (2027-256) is the McDATA Intrepid 10000 Director (also known as the i10K). It provides 10 Gbps backbone connections, 64-256 ports 2 Gbps fabric switching for Windows and UNIX; and FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. McDATA Enterprise Fabric Connectivity Manager provides integrated management of complex IBM SAN m-type (McDATA) tiered enterprise fabrics.

Highlights

- ▶ Easy-to-manage tiered enterprise infrastructure simplification and Business Continuity solutions for System x, System i, System p and System z.
- ▶ Highly-scalable 64 to 256-port switching backbone for tiered global enterprise storage area networks (SANs)
- ▶ Designed to provide high availability with concurrent hardware and firmware upgrades and call-home with McDATA Enterprise Fabric Connectivity Manager, EFCM
- ▶ Director FlexPar, designed to provide dynamic application network provisioning, can help simplify Fibre Channel and mainframe FICON SAN consolidation
- ▶ Helps to provide global Business Continuity solutions with 10 Gbps links up to 190 km
- ▶ EFCM and FICON Management Server (CUP) software can help simplify management of complex SAN infrastructures



Figure 16-15 IBM TotalStorage SAN256M

Configuration flexibility

The IBM SAN256M enterprise director provides scalability with two to eight **line module** features. A 64-port entry configuration consists of two 32-port 2 Gbps line modules. This configuration is designed to scale-up nondisruptively to 256 ports by adding 32-port 2 Gbps line modules. Each line module has four paddles with 2 Gbps SFP or 10 Gbps XFP optical transceivers.

Director FlexPar for SAN consolidation

The IBM SAN256M Director FlexPar (standard feature) enables the director to be segmented into sub-directors which can help simplify consolidation of application-based SAN island consolidation and provisioning. Each Director FlexPar acts like an independent director with separate fabric services and isolation of fabric events. Up to four Director FlexPars may be enabled, on a line module basis. Director FlexPar architecture is designed to provide improved granularity at the port level which can help support auto provisioning of network resources by application

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/m_type/san256m/

16.4.4 Cisco MDS 9506 Multilayer Director

The Cisco MDS 9506 Multilayer Director (2062-D04) supports 1, 2, and 4 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN. Administrators can use the Cisco MDS 9506 to help address the needs for high performance and reliability in SAN environments ranging from small workgroups to very large, integrated global enterprise SANs.

The Cisco MDS 9506 Multilayer Director utilizes two Supervisor-2 Modules designed for high availability and performance. The Supervisor-2 Module combines an intelligent control module and a high-performance crossbar switch fabric in a single unit. It uses Fabric Shortest Path First (FSPF) multipath routing, which provides intelligence to load balance across a maximum of 16 equal-cost paths and to dynamically reroute traffic if a switch fails.

Each Supervisor-2 Module provides the necessary crossbar bandwidth to deliver full system performance in the MDS 9506 director with up to four Fibre Channel switching modules. It is

designed to provide that loss or removal of a single crossbar module has no impact on system performance.

Highlights

- ▶ Provides Fibre Channel throughput of up to 4 Gbps per port and up to 64 Gbps with each PortChannel ISL connection
- ▶ Scalable from 12 to 192 Fibre Channel ports
- ▶ 10 Gbps ISL ports for inter-Data Center links over metro optical networks
- ▶ Gigabit Ethernet IP, GbE ports for iSCSI or FCIP connectivity over global networks
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ High-availability design with nondisruptive firmware upgrades
- ▶ Enterprise, SAN Extension over IP, Mainframe, Storage Services Enabler and Fabric Manager Server Packages provide added intelligence and value



Figure 16-16 Cisco MDS 9506 Multilayer Director

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/c_type/9506/

16.4.5 Cisco MDS 9509 Multilayer Director

The Cisco MDS 9509 Multilayer Director (2062-D07) supports 1, 2, and 4 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN. Administrators can use the Cisco MDS 9509 to help address the needs for high performance and reliability in SAN environments ranging from small workgroups to very large, integrated global enterprise SANs.

The Cisco MDS 9509 Multilayer Director utilizes two Supervisor-2 Modules designed for high availability and performance. The Supervisor-2 Module combines an intelligent control module and a high-performance crossbar switch fabric in a single unit. It uses Fabric Shortest Path First (FSPF) multipath routing, which provides intelligence to load balance across a maximum of 16 equal-cost paths and to dynamically reroute traffic if a switch fails.

Each Supervisor-2 Module provides the necessary crossbar bandwidth to deliver full system performance in the MDS 9509 director with up to four Fibre Channel switching modules. It is designed to provide that loss or removal of a single crossbar module has no impact on system performance.

Highlights

- ▶ Supports throughput of up to 4 Gbps per port and up to 64 Gbps with each PortChannel Inter-Switch Link (ISL) connection
- ▶ Offers scalability from 12 to 336 4 Gbps Fibre Channel ports
- ▶ Offers Gigabit Ethernet IP (GbE) ports for iSCSI or FCIP connectivity over global networks
- ▶ High-availability design with support for nondisruptive firmware upgrades Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Enterprise, SAN Extension over IP, Mainframe and Storage Services Enabler and Fabric Manager Server Packages provide added intelligence and value



Figure 16-17 Cisco MDS 9509 Multilayer Director

More information

For additional product details, go to:

http://www-03.ibm.com/servers/storage/san/c_type/9509/

16.4.6 Cisco MDS 9513 Multilayer Director

The Cisco MDS 9513 Multilayer Director (2062-E11) supports 1, 2 and 4 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate dispersed SAN islands into a large enterprise SAN. Administrators can use the Cisco MDS 9513 to help address the needs for high performance, scalability and availability in SAN environments ranging from single site environments to very large multiple site metropolitan environments.

The Cisco MDS 9513 utilizes two Supervisor-2 Modules, designed to support high availability. The Supervisor-2 Module is designed to provide industry leading scalability, intelligent SAN services, nondisruptive software upgrades, stateful process restart and failover, and redundant operation. Dual crossbar switching fabric modules provide a total internal switching bandwidth of 2.4 Tbps for inter-connection of up to eleven Fibre Channel switching modules.

Highlights

- ▶ Supports Fibre Channel throughput of up to 4 Gbps, per port and up to 64 Gbps with each PortChannel ISL connection
- ▶ Scalable from 12 to 528 4 Gbps Fibre Channel ports

- ▶ Gigabit Ethernet (GbE) IP ports for iSCSI or FCIP connectivity over global networks
- ▶ High-availability design with support for nondisruptive firmware upgrades Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Enterprise, SAN Extension over IP, Mainframe and Storage Services Enabler and Fabric Manager Server Packages provide added intelligence and value



Figure 16-18 Cisco MDS 9513 Multilayer Director

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/c_type/9513/

16.5 Multiprotocol routers

Since the introduction of Storage Area Networks, customers have built multiple SAN networks (or islands) for different applications, often with fabric switch components from different manufacturers. Some islands were built by different departments within a company, while other islands resulted from mergers, acquisitions, or reorganizations. Dissimilar SAN equipment with different capabilities or a desire to isolate important applications has constrained opportunities for enhanced infrastructure simplification and vital business continuity solutions.

The IBM System Storage multiprotocol router provides Fibre Channel FC-FC Routing Service, which allows the interconnection of multiple SAN islands without requiring that the separate fabrics be merged into a single large SAN. This capability can help create a tiered or extended enterprise SAN infrastructure without having to redesign or reconfigure the entire environment.

Figure 16-19 on page 343 shows a conceptual view of SAN multiprotocol routing, where SANs from different vendors can be connected.

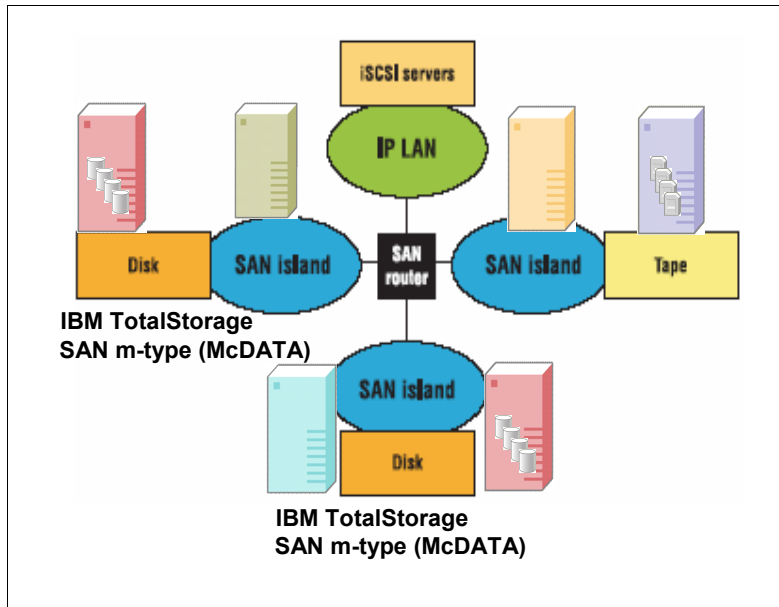


Figure 16-19 SAN multiprotocol routing

Enterprise products include:

- ▶ IBM System Storage SAN18B-R
- ▶ IBM TotalStorage SAN16B-R
- ▶ IBM TotalStorage SAN16M-R
- ▶ IBM TotalStorage SAN04M-R
- ▶ Cisco 9216i Fabric Switch

In this section, we introduce IBM System Storage SAN Multiprotocol Router solutions ideal for enterprises with a business need to route SAN traffic between different SAN islands.

16.5.1 IBM System Storage SAN18B-R multiprotocol router

IBM System Storage SAN18B-R multiprotocol router (2005-R18) is the Brocade SilkWorm Multiprotocol Router. The SAN18B-R router delivers up to 4 Gbps FC routing and 1 Gbps Ethernet in an easy-to-manage, compact design.

The I SAN18B-R supports sixteen Fibre Channel ports and two Gigabit Ethernet ports. Each Fibre Channel port is capable of self-negotiation to the highest speed supported by the attached SAN infrastructure, operating at either 1, 2 or 4 Gbps throughput. It supports distances up to 10 km or 35 km using the appropriate optional long wave laser transceiver and 9.0u fiber cables. Shortwave laser transceivers support distances up to 500m at 1 Gbps or up to 300 m at 2 Gbps or 150 m at 4 Gbps with 50.0u fiber cables.

A wide range of IBM System Storage mid-range and enterprise Storage Area Network (SAN) infrastructure simplification and business continuity solutions can be created with the IBM System Storage SAN18B-R multiprotocol router. Infrastructure simplification solutions for the IBM System i, System p and System x family include SAN island consolidation, while business continuity solutions include disaster tolerance over metropolitan and global IP networks with IBM System Storage disk arrays, tape libraries and IBM Tivoli Storage Manager data protection software.

Highlights

- ▶ Enables consolidation of Storage Area Network (SAN) islands for infrastructure simplification without compromising security
- ▶ Designed for high performance with 4 Gbps FC ports and hardware-assisted traffic processing for line-rate performance across Gigabit Ethernet IP ports
- ▶ Utilizes existing IP Metropolitan Area Network (MAN) or WAN infrastructures for metro and global SAN extension for business continuity solutions
- ▶ Hardware-based compression, large window sizes and selective acknowledgement of IP packets designed to optimize performance of SAN extension over IP networks
- ▶ Eight virtual FCIP tunnels per IP port help maximize scalability and utilization of MAN/WAN resources
- ▶ Integrated router and IBM System Storage SAN b-type (Brocade) switch management helps simplify installation and administration and helps provide fabric investment protection

Figure 16-20 shows the SAN18B-R.

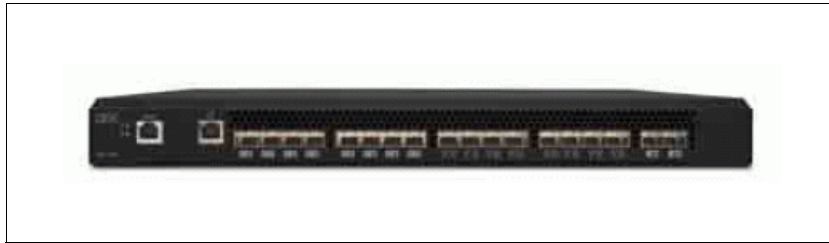


Figure 16-20 IBM System Storage SAN18B-R

The IBM System Storage SAN18B-R multiprotocol router provides Fibre Channel FC-FC Routing Service, which allows the interconnection of multiple SAN islands without requiring that the separate fabrics be merged into a single large SAN. As shown in Figure 16-21, this capability can help create a tiered or extended enterprise SAN infrastructure without having to redesign or reconfigure the entire environment.

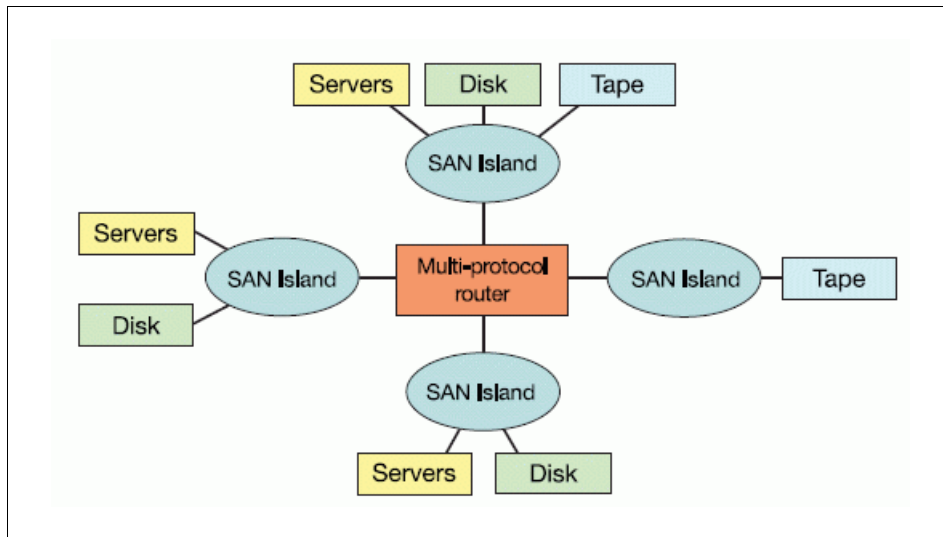


Figure 16-21 Inter-SAN routing

FC-FC Routing Service

FC-FC Routing Service feature is designed to allow devices located on separate SAN fabrics to communicate without merging the fabrics into a single large SAN environment. This routed network consists of multiple individual SAN fabrics that form one storage network connectivity model, known as a MetaSAN.

FCIP Tunneling Service

FCIP Tunneling Service feature is designed to allow organizations to extend Fibre Channel SANs over longer distances, using IP-based MAN/WAN infrastructure. This service can be integrated with the FC-FC Routing Service.

iSCSI Gateway Service

iSCSI Gateway Service is designed to allow the integration of low-cost Ethernet-connected servers into a Fibre Channel SAN by using iSCSI protocol, enabling IP-based servers to leverage Fibre Channel SAN resources. It enhances asset utilization and can help enable new applications such as centralized backup. This level of integration reduces the expense of connecting low-cost servers directly into centrally managed SAN storage.

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/b_type/san18b-r/

16.5.2 IBM TotalStorage SAN16B-R multiprotocol router

IBM TotalStorage SAN16B-R multiprotocol router (2109-A16) is the Brocade SilkWorm Multiprotocol Router, and provides 8 or 16 Fibre Channel and IP ports. Each port can operate at either 1 or 2 Gbps FC or 1 Gbps Ethernet. the SAN16B-R router can deploy services as needed on any active port including FC-FC Routing capability to enable increased SAN connectivity and resource sharing between fabrics, FCIP Tunneling to enable SAN extension over geographic distances, and iSCSI Gateway to enable host server storage consolidation. Full interoperability and integrated management with IBM System Storage SAN b-type (Brocade) switches and directors helps protect switch investments.

IBM TotalStorage SAN16B-R multiprotocol router provides improved scalability, security and manageability by enabling devices in separate SAN fabrics to communicate without merging fabrics into a single, large SAN fabric. This capability enables customers to initially deploy separate SAN solutions at the departmental and data center levels and then to consolidate them into large enterprise SAN solutions as their experience and requirements grow and change.

Highlights

- ▶ Integrated switch and router management helps simplify deployment and operation of large enterprise SAN infrastructure simplification, business continuity and information lifecycle management solutions
- ▶ 1 and 2 Gbps Fibre Channel and Gigabit Ethernet IP ports
- ▶ Pay-as-you-grow scalability with Ports on Demand feature
- ▶ Redundant, hot-swappable fans and power supplies and hot-pluggable optical transceivers



Figure 16-22 IBM TotalStorage SAN16B-R multiprotocol router

More information

- ▶ For additional product details, go to:

http://www.ibm.com/servers/storage/san/b_type/san16b-r/

16.5.3 IBM TotalStorage SAN16M-R multiprotocol SAN router

IBM TotalStorage SAN16M-R multiprotocol SAN router (2027-R16) is the McDATA Eclipse 2640 SAN Router. It provides 12 Fibre Channel and 4 IP ports with SAN Routing for SAN island consolidation, Internet Fibre Channel Protocol, iFCP for high performance metro and global Business Continuity solutions and Internet SCSI, iSCSI for low cost infrastructure simplification solutions. The 16-port router includes zoning, two SAN routing ports, iSCSI and SANvergence Manager. Optional features include SAN routing on twelve Fibre Channel SAN ports, iFCP with Fast Write and compression on four IP ports and SANvergence Manager Enterprise

Highlights

- ▶ Enables SAN island consolidation for secure data center infrastructure simplification solutions
- ▶ iSCSI server SAN connectivity for low cost infrastructure simplification solutions
- ▶ Provides SAN routing over distance for metro and global Business Continuity solutions
- ▶ 1 and 2 Gbps FC and 1 Gbps Ethernet with Fast Write and compression
- ▶ Interoperability with IBM System Storage SAN m-type (McDATA) family provides switch investment protection
- ▶ Includes SANvergence Manager for router and network management



Figure 16-23 IBM TotalStorage SAN16M-R multiprotocol SAN router

The IBM TotalStorage SAN16M-R multiprotocol SAN router provides IP LAN connectivity to existing data center SAN infrastructures. It also is designed to provide SAN routing between multiple SAN islands while preserving separate fabric fault isolation, security and management. This can help create a tiered data center SAN infrastructure with improved network stability, simplified network management and extended scalability.

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/m_type/san16m_r/

16.5.4 IBM TotalStorage SAN04M-R

IBM TotalStorage SAN04M-R multiprotocol SAN router (2027-R04) is the McDATA Eclipse 1620 SAN Router. It provides 4 ports - two FC 1 Gbps ports, and 2 IP ports in 1U rack space. Support is provided for SAN Routing for SAN island consolidation, Internet Fibre Channel Protocol, iFCP for high performance metro and global Business Continuity solutions and Internet SCSI, iSCSI for low cost infrastructure simplification solutions. Base functionality of the two user configurable ports provides SAN routing on up to two FC ports, FC fabric support, and iSCSI support on Gigabit Ethernet ports. Two optional firmware versions are available for iFCP Standard and iFCP Enterprise

Simplify storage infrastructure and protect business continuity with the IBM TotalStorage SAN04M-R multiprotocol SAN router. Cost-effective business continuity solutions over metropolitan and global IP networks include remote IBM System Storage tape libraries with IBM Tivoli Storage Manager data protection software and remote mirroring with IBM System Storage Resiliency Portfolio. Infrastructure simplification solutions for System x, System i, and System p include iSCSI server integration with IBM System Storage disk storage arrays.

Highlights

- ▶ Extends SAN infrastructure over IP networks for cost-effective metro and global business continuity solutions
- ▶ iSCSI server SAN connectivity for cost-effective infrastructure simplification solutions
- ▶ 1 Gbps FC and 1 Gbps Ethernet with Fast Write and compression to help improve performance over long distances
- ▶ Interoperability with IBM System Storage SAN m-type (McDATA) family helps provide switch investment protection
- ▶ Includes SANvergence Manager for router and network management



Figure 16-24 IBM TotalStorage SAN04M-R

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/m_type/san04m-r/

16.5.5 Cisco MDS 9216i Fabric Switch

The Cisco MDS 9216i Multilayer Fabric Switch (2062-D1H) includes fourteen 2 Gbps Fibre Channel ports, two GbE IP ports, SAN Extension of IP Package for integrated IP ports and one modular expansion slot. The expansion slot accepts one optional 12-, 24- and 48-port 4 Gbps Fibre Channel switching module, allowing the Cisco MDS 9216i to support up to 62 Fibre Channel ports. Optionally, a 4-port 10 Gbps Fibre Channel module is available for high performance inter-switch link (ISL) connections over metro optical networks. The expansion slot also accepts one IP, Multiprotocol or Storage Services Module. The optional modules are designed for hot-swap capability.

Highlights

- ▶ Supports Fibre Channel throughput of up to 4 Gbps per port and up to 64 Gbps with each Port Channel ISL connection
- ▶ 9216i model offers integrated 14 FC ports and two 1 Gbps Ethernet IP ports for iSCSI or FCIP connectivity over global networks
- ▶ 10 Gbps ISL ports for inter-Data Center links over metro optical networks
- ▶ Modular design with excellent availability capabilities
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Compatible with a broad range of IBM servers as well as disk and tape storage devices



Figure 16-25 Cisco MDS 9216i Fabric Switch

Multiprotocol support and traffic management features

The unique architecture of the Cisco MDS 9216i Fabric Switch allows integration of new transport protocols for greater flexibility. For example, the Cisco MDS 9216 is designed to support Fibre Channel, Internet SCSI (iSCSI) and Fibre Channel over IP (FCIP).

IP and Multiprotocol Storage Services Module features offer two to ten Gigabit Ethernet ports for iSCSI or FCIP connectivity, software configurable on a port-by-port basis. The IP Storage Services Module feature provides four or eight iSCSI ports. Multiprotocol Services Module feature offers two Gigabit Ethernet iSCSI ports and fourteen Fibre Channel ports. The Tri-Rate shortwave and longwave SPF Transceiver features are required for each port to be used. The SAN Extension over IP Package for IP Services Modules and for Multiprotocol Services Module features add Fibre Channel over IP (FCIP) support. SAN Extension over IP Package helps improve performance with FCIP Compression, Write Acceleration and Tape Acceleration and helps improve security with Inter-VSAN Routing for FCIP. Users may now

use iSCSI for cost-effective connectivity to shared storage pools and FCIP for secure metro and global connectivity between data centers.

More information

For additional product details, go to:

http://www.ibm.com/servers/storage/san/c_type/9216/

16.6 SAN solutions

IBM System Storage SAN solutions integrate IBM and other vendor server, storage, SAN switches and software components into solution templates which have been extensively tested to provide high availability, scalability, security, and simplicity of management. IBM SAN solutions are offered with worldwide IBM service and end-to-end solution support. These solutions templates may be customized by IBM Business Partners or IBM Global Services to address individual customer requirements.

IBM System Storage Infrastructure Simplification solutions and Business Continuity solutions are offered for Small and Medium Business (SMB) customers who require simplicity and affordability and for large enterprise customers who demand highest availability, scalability, security and management intelligence.

- ▶ Infrastructure simplification solutions consist of storage consolidation with shared disk storage and storage pooling with IBM SAN Volume Controller software.
- ▶ Business Continuity solutions consist of data protection with shared local or remote tape library technology and disk storage with point-in-time copy technology and automated backup and recovery software; and disaster tolerance with disk storage mirroring technologies and SAN fabric extension technologies which can provide metro and global distance solutions.

The following are just a few examples of solution possibilities using IBM System Storage SAN switches.

More information

- ▶ For additional IBM SAN solution information, go to:

<http://www.ibm.com/servers/storage/solutions/san/>

16.6.1 Infrastructure simplification solutions

IBM offers solutions to simplify IT infrastructures - a key first step in freeing the resources you need to begin developing an on demand operational model.

Complexity can prevent even the best organization from acting nimbly to meet ever-changing market and client demands. Complexity inherent in typical IT infrastructures can stall your business goals by making execution on new tasks too difficult or time consuming.

IBM solutions for infrastructure simplification are designed to help you improve efficiency, lower total cost of ownership, and reduce time-consuming and costly errors. These solutions can help you lower the cost of storing and managing data. IBM System Storage solutions for infrastructure simplification are designed to improve the effectiveness of your storage infrastructure and reduce long-term costs, while reducing the time and effort spent on managing data.

An *entry-level Infrastructure simplification solution* consists of up to fourteen servers attached to one 16-port SAN16M-2 or SAN16B-2 fabric switch with an IBM System Storage disk array.

An example *Infrastructure simplification with high-availability server solution* is shown on Figure 16-26. The topology is created with redundant switches, and each server, with dual Fibre Channel adapters is cross-connected via the separate switches to the IBM System Storage DS4000 disk array.

An easy way to migrate your SAN from *entry-level* to *mid-range* products is to shutdown one 16-port SAN16M-2 or SAN16B-2 fabric switch and replace it with one 32-port SAN32M-2 or SAN32B-2 fabric switch.

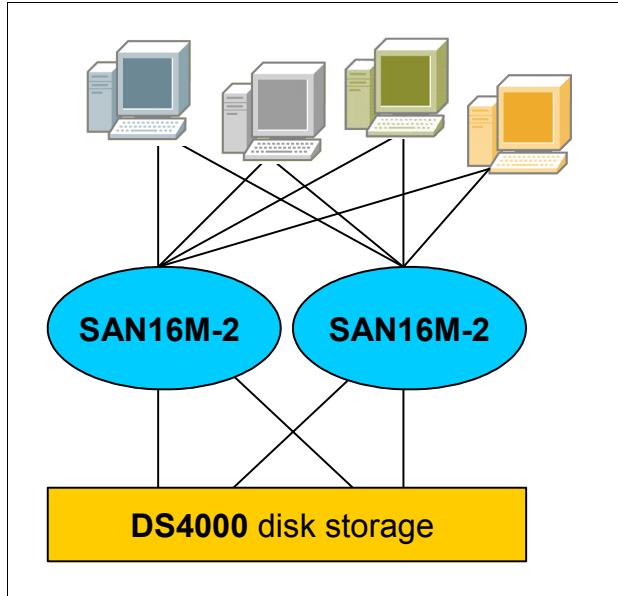


Figure 16-26 *Infrastructure simplification with high-availability server solution*

An expanded solution provides a higher data transfer rate by using 4 Gbps throughout - servers with 4 Gbps HBA card, for example: IBM BladeCenter, the IBM System Storage SAN fabric switches with 1, 2, and 4 Gbps auto-sensing capabilities, and the IBM System Storage DS4700 Express with 4 Gbps host connection function. This is shown in Figure 16-27 on page 351.

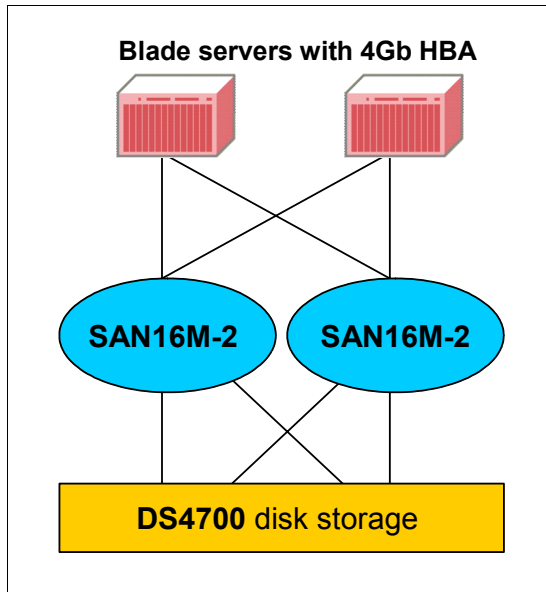


Figure 16-27 Infrastructure simplification with 4 Gbps solution

Local site infrastructure simplification solutions may be extended to one or many remote sites for enhanced data protection and disaster tolerance. The IBM System Storage multiprotocol routers provide Fibre Channel over IP and FCIP Tunneling Service for distance extension which can enable cost-effective and manageable metro and global business continuity solutions. This extended distance connectivity can help create consolidated remote tape vaulting data protection plus metro mirror and global mirror disk-based disaster tolerant solutions.

16.6.2 Business continuity solutions

Today's customers, employees, suppliers and business partners expect to be able to tap into your information any time of day, from any location. At the same time, your business must be increasingly sensitive to issues of customer privacy, data security and regulatory requirements.

To keep your operations running, and your business competitive, you need a comprehensive strategy that addresses three primary aspects of business continuity: high availability, disaster recovery and continuous operations.

IBM System Storage SAN Family solutions can help you protect critical business assets, and align recovery costs based on patented business impact and information value. A typical IBM System Storage SAN switch *Business Continuity solutions* example of backup site solution is shown in Figure 16-28 on page 352.

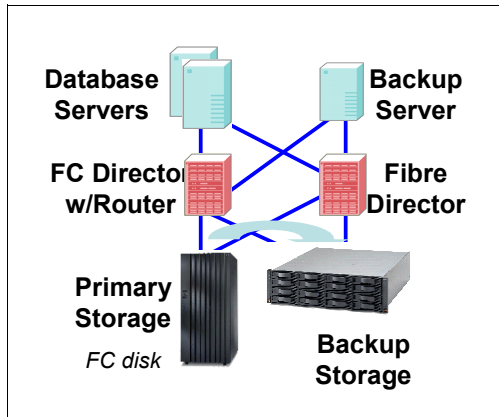


Figure 16-28 Typical backup site solution

When considering total cost of ownership, there are two choices: FCIP or FC direct route by DWDM or CWDM. See Figure 16-29 for FC over IP and Figure 16-30 for FC route through DWDM or CWDM.

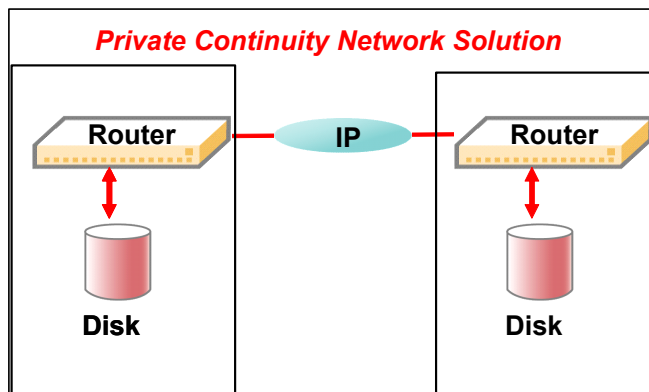


Figure 16-29 FC over IP

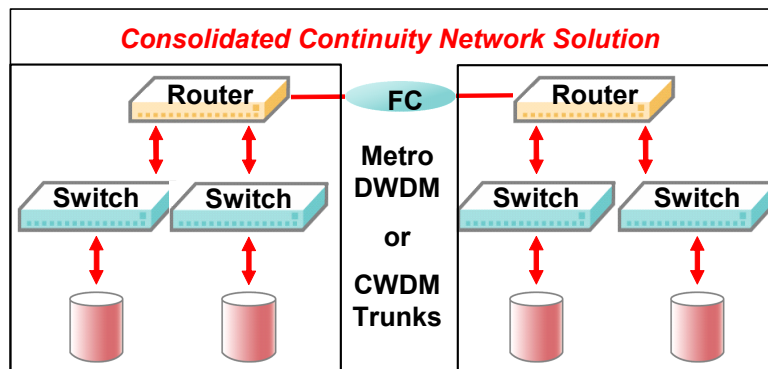


Figure 16-30 FC over DWDM or CWDM

Another IBM System Storage SAN switches *Business Continuity solutions* example is remote tape backup with the IBM SAN Multiprotocol router: As Figure 16-31 shows, an enterprise with a remote backup site can build a tape backup solution over distance with an IBM Multiprotocol router. An offsite Tivoli Storage Manager, via SAN routing, can pull data from a

production disk storage array then copy it to the tape drives. This allows a minimal number of tape drives to be used and also leverages electronic production of offsite media.

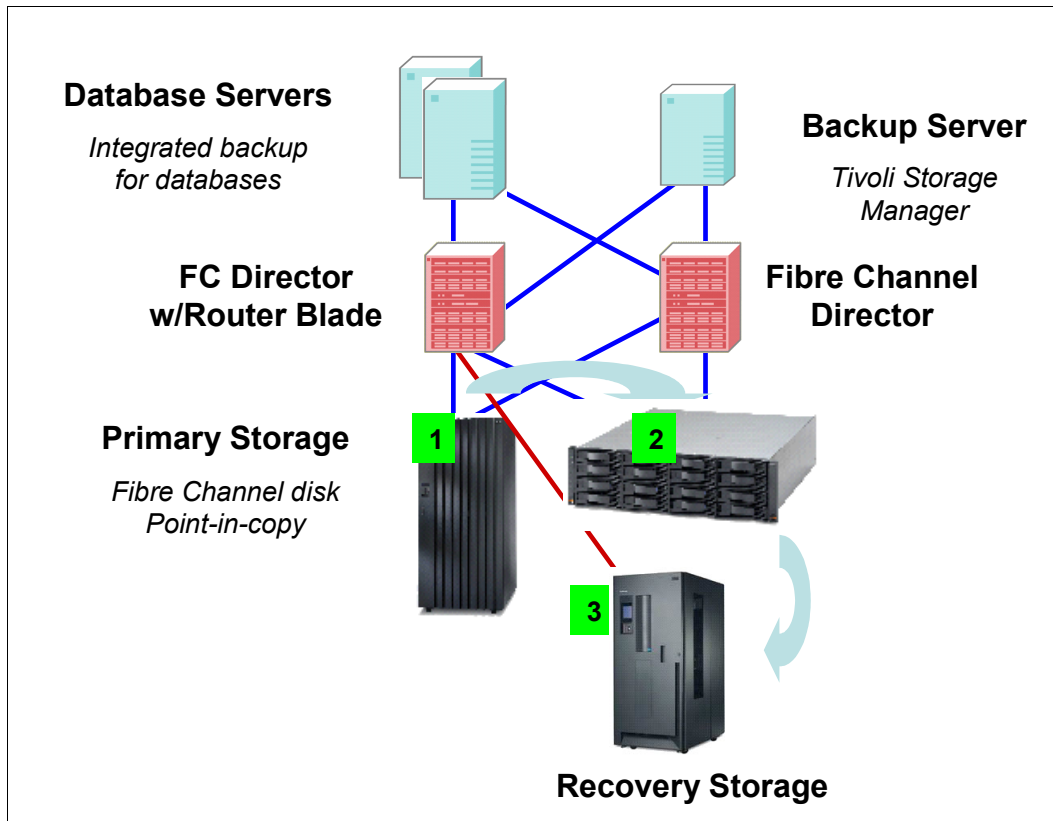


Figure 16-31 Remote tape backup solution

IBM System Storage open software

IBM System Storage software products are designed with the goal of helping customers drive down the cost and complexity of storage management while providing greater flexibility to address rapidly changing storage needs.

The IBM storage software portfolio consists of:

- ▶ IBM System Storage Virtualization:
 - IBM System Storage SAN Volume Controller
- ▶ IBM TotalStorage Productivity Center and Tivoli Provisioning Manager:
 - TPC data
 - TPC fabric
 - TPC disk
 - TPC replication
- ▶ IBM Tivoli Storage Manager, including:
 - IBM Tivoli Storage Manager
 - IBM System Archive Manager
 - IBM Tivoli Storage Manager for Space Management
- ▶ IBM TotalStorage Expert



IBM System Storage virtualization

Virtualization addresses the increasing complexity of managing storage and will help reduce its associated costs, as it will allow the full exploitation of the benefits promised by a SAN. Virtualization is an enabler for sharing data, ensuring higher availability, providing disaster tolerance, and improving performance. It represents the next step in the technological evolution for networked storage. Virtualization is one technique for enhancing on demand computing. Virtualization will also allow for consolidation of resources, provide policy-based automation, and enable several other benefits, which do not automatically result from the implementation of SAN hardware components.

In this chapter, we explain virtualization in relation to storage networking and discuss some of the benefits of implementing an IBM System Storage virtualization solution.

17.1 What storage “virtualization” is

Storage virtualization techniques are becoming increasingly more prevalent in the IT industry today. Storage virtualization forms one of several layers in a storage network, and can be described as the abstraction from physical volumes of data storage to a view of logical volumes of data storage. In practical terms, storage virtualization is the pooling of physical storage from multiple storage devices into what appears to be one single data storage device. This single storage is then managed from a central console.

The abstraction can be made on several levels of the components of storage networks and is not limited to the disk subsystem. The virtualization layer provides the same kind of services to the layer above (as the hidden layer below provides). Storage virtualization software separates the representation of storage to the operating system (and its users) from the actual physical components.

Note: Storage virtualization existed, and was taken for granted in the mainframe computing environment for many years.

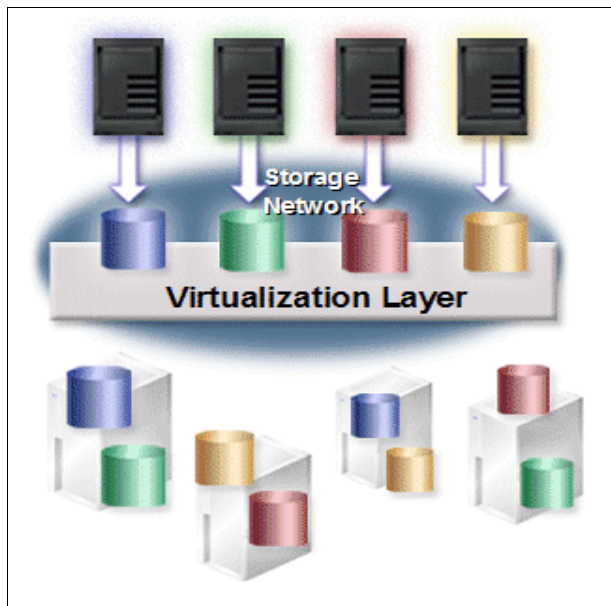


Figure 17-1 SAN with virtualization

SAN is making it easier for customers to spread their IT systems out geographically, but even in networks, different types of servers that use different operating systems do not get the full benefit of sharing storage. Instead, the storage is partitioned to each different type of server, which creates complex management and inefficient use of storage. When storage must be added, applications are often disrupted. At the same time, the reduced cost of storage and the technology of storage networks, with faster data transfer rates, have enabled customers to use increasingly sophisticated applications, such as digital media. This has caused even greater complexity and difficulty of management as the amount of storage required grows at unprecedented rates. Storage virtualization is aimed at eliminating these problems.

Some benefits of implementing virtualization in your SAN are:

- ▶ Reduced storage management complexity
- ▶ Significantly reduced downtime
- ▶ Enhanced productivity through improved data availability

- ▶ Consolidation to a single pool of storage
- ▶ Heterogeneous storage is presented as homogenous storage
- ▶ Better choice of storage, switch and SAN technologies
- ▶ Improved storage resource utilization
- ▶ A single, cost-effective set of advanced copy services
- ▶ Simplification and standardization of the IT infrastructure
- ▶ Ease transition to an on demand IT infrastructure

These benefits lead to lower TCO and better flexibility to react to business changes.

17.1.1 Levels of storage virtualization

Virtualization can be achieved at different levels in a storage network, as illustrated in Figure 17-2.

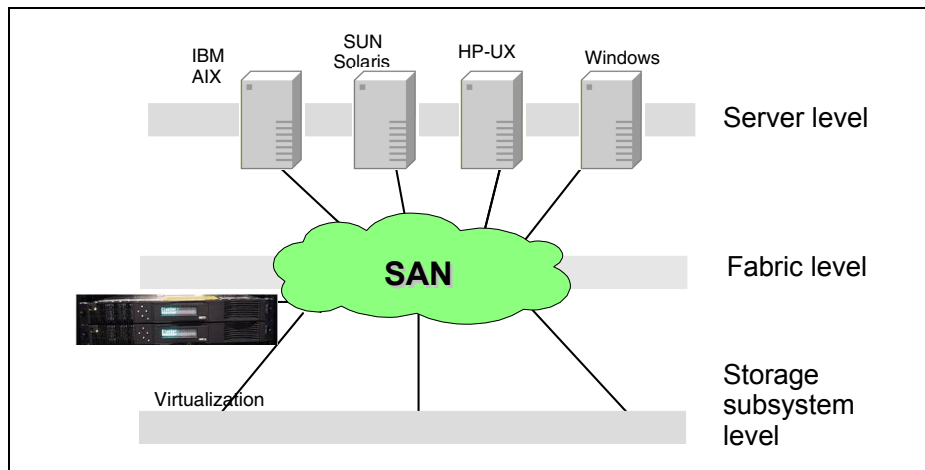


Figure 17-2 Levels of application of storage virtualization

Server level

Abstraction at the server level is achieved with logical volume management on the server operating systems. At first sight, increasing the level of abstraction on the server seems well suited for environments without storage networks, but this can be vitally important in storage networks too.

Fabric level

At the fabric level, virtualization can enable the independence of storage pools from heterogeneous servers. The SAN fabric would be zoned to allow the virtualization appliances to see the storage subsystems, and for the servers to see the virtualization appliances. Servers would not be able to directly see or operate on the storage subsystems.

Storage subsystem level

Disk storage systems can provide some level of virtualization already by subdividing disks into smaller virtual drives. Conversely, more storage devices could be consolidated together to form one large virtual drive. RAID subsystems are an example of virtualization at the storage level. Storage virtualization can take this to the next level by enabling the presentation, and the management, of disparate storage systems.

Server, fabric, or storage subsystem virtualization?

As we stated, virtualization can be implemented at any of these levels. The IBM strategy is to move the storage device management intelligence out of the server, reducing the

dependency of having to implement specialized software, like Logical Volume Managers (LVM), at the server level. IBM also intend to reduce the requirement for intelligence at the storage subsystem level, which will decrease the dependency on having to implement intelligent storage subsystems.

By implementing virtualization at a fabric level, storage control is moved into the network, which gives the virtualization opportunity to all, and at the same time reduces complexity by providing a single view of storage. The storage network can be used to leverage all kinds of services across multiple storage devices, including virtualization.

Fabric level virtualization models

In-band and out-of-band are two implementations of virtualization at the fabric level. Models can be drawn for these methods of storage virtualization, as illustrated in Figure 17-3. These models are not mutually exclusive. In many environments a combination of both may be desired.

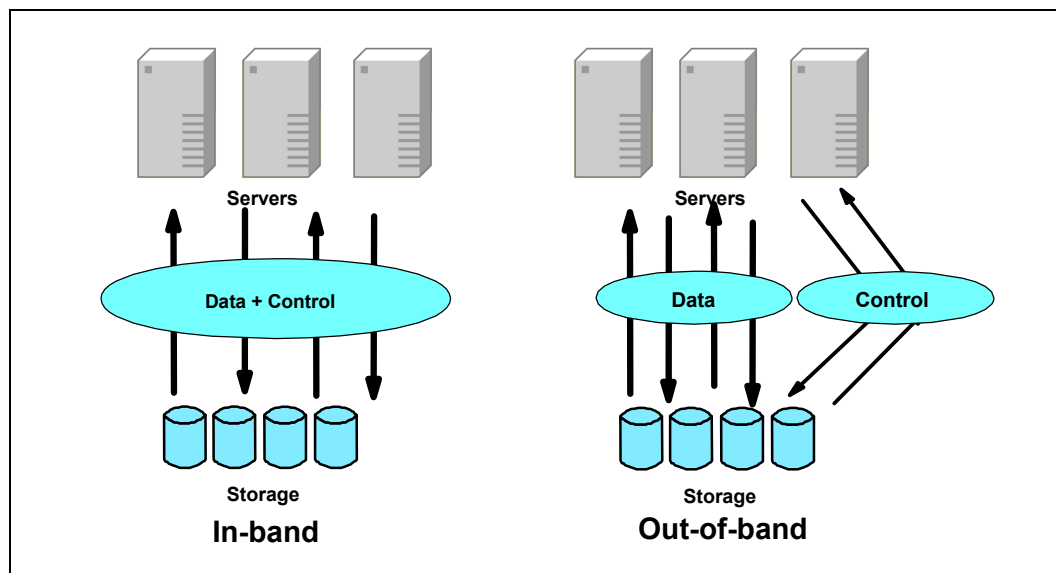


Figure 17-3 In-band and out-of-band virtualization models

In-band

When we implement an in-band virtual storage network, both data and control flow over the same path. Levels of abstraction exist in the data path, and storage can be pooled under the control of a domain manager. In general, in-band solutions are perceived to be simpler to implement, especially because they do not require special software to be installed in servers (other than conventional multi-pathing software). In-band solutions can also provide caching and advanced functions within the storage network. This can help to improve the performance of existing disk systems and can extend their useful life, and reduce the cost of new storage capacity by enabling the use of lower function, lower cost disk systems, without the loss of performance.

Other advantages include:

- ▶ Ability to off load function from the host
- ▶ Providing storage management for the SAN
- ▶ Performing performance optimizations in the data path
- ▶ Supporting host systems not in a cluster
- ▶ Supporting multiple heterogeneous hosts
- ▶ Releasing the customer from a particular vendor's storage

- ▶ Integrating with storage to create a better management picture
- ▶ Offering excellent scalability

The IBM System Storage **SAN Volume Controller** is a virtualization appliance solution that maps virtualized volumes visible to hosts and applications to physical volumes on storage devices. It is an example of an in-band solution.

Out-of-band

In an out-of-band implementation, the data flow is separated from the control flow. This is achieved by separating the data and metadata (data about the data) into different places. Out-of-band virtualization involves moving all mapping and locking tables to a separate server (the metadata controller) that contains the metadata of the files.

In an out-of-band solution the clients request authorization to data from the metadata controller, which grants it, handles locking, and so on. Once they are authorized, clients access the data directly without any metadata controller intervention. Once a client has obtained access to a file, all I/O will go directly over the SAN to the storage devices. For many operations, the metadata controller does not even intervene.

Separating the flow of control and data in this manner allows the I/O to use the full bandwidth that a SAN provides, while control could go over a separate network or routes in the SAN that are isolated for this purpose.

Other advantages include:

- ▶ Releasing the customer from a particular vendor's storage
- ▶ Providing storage management for the SAN
- ▶ Offering excellent scalability
- ▶ Off loading host processing
- ▶ Supporting storage management from multiple vendors
- ▶ Integrating well with storage management software
- ▶ Supporting multiple heterogeneous hosts
- ▶ Relatively low overhead in the data path

17.2 IBM System Storage approach to virtualization

The IBM approach to helping you with your storage needs is to address the entire problem. From a TCO perspective, the initial purchase price is becoming an increasingly small part of the equation - in the order of 10%. As the cost per megabyte of disk drives continues to decrease, the focus is shifting away from hardware toward software value add functions, storage management software, and services. This is not to downplay the importance of a highly reliable, high performance hardware solution such as the IBM System Storage DS6000 or DS8000. But software is emerging as the differentiating factor and offers a highly competitive advantage. This is due to the fact that advanced functionality provided by software and storage management software play a vital role in administering distributed IT assets, maintaining high availability, and minimizing downtime.

IBM has identified three major product areas where significant improvements can be achieved. The SNIA Storage Model describes these areas as *Block Aggregation*, *File Aggregation*, and *Management and Productivity*.

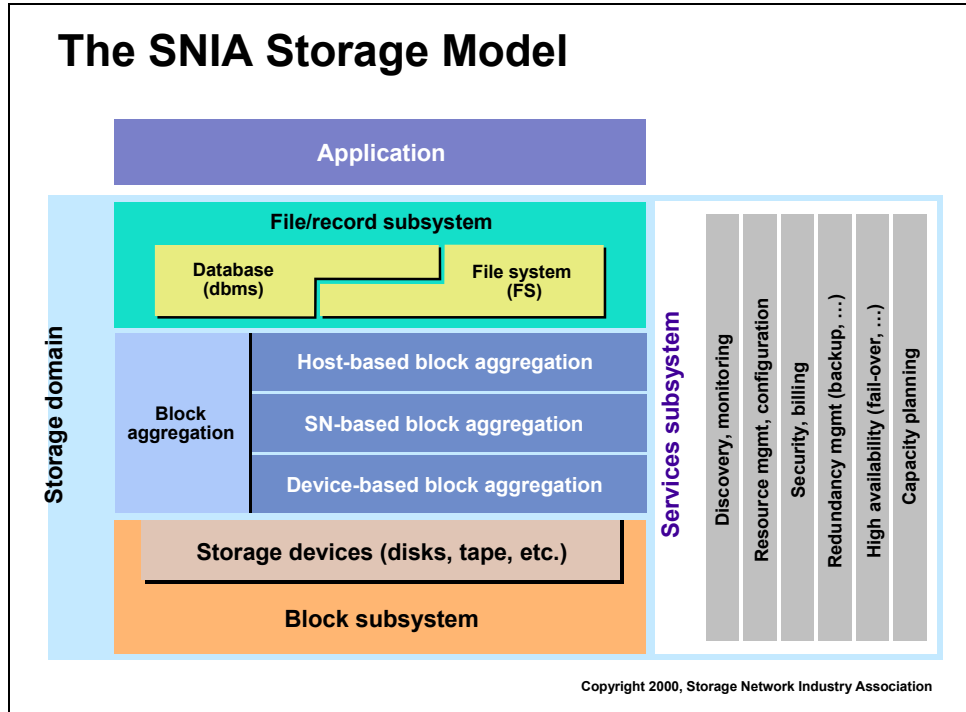


Figure 17-4 SNIA storage model

The following sections describe our product initiatives in each of these areas.

IBM block aggregation characteristics

Block aggregation in the SNIA model provides what is also referred to in the industry as block virtualization. Block level virtualization provides servers with a logical view of physical storage. The SNIA Block Aggregation Model, as seen in Figure 17-4, specifies that block aggregation can be performed within hosts (servers), in the storage network (storage routers, storage controllers) or in storage devices (intelligent disk arrays). The IBM System Storage SAN Volume Controller is a block aggregation product within the storage network.

IBM file aggregation characteristics

For file/record aggregation in the SNIA model, hosts get file metadata from file system or Network Attached Storage (NAS) controllers, then access the data directly. File aggregation can be used in conjunction with or independent from block aggregation. An IBM approach is through the use of a common file system available for different operating systems. It is a SAN-based file system with metadata controllers. The core benefit is a single point of management and a single name space.

17.3 IBM System Storage virtualization family

The IBM System Storage products are designed with the goal of helping customers drive down the cost and complexity of storage management while providing greater flexibility to address rapidly changing storage needs.

IBM System Storage Virtualization family of products can help you to:

- ▶ Simplify the deployment and administration, while increasing utilization, of physical storage asset:

IBM System Storage SAN Volume Controller

- ▶ Manage all their storage assets with a single, comprehensive management suite using:

IBM TotalStorage Productivity Center

In subsequent chapters, we talk about these products.

These virtualization products are part of an IBM commitment to the open standards adopted by the Storage Networking Industry Association (SNIA). They implement standard CIM-based API's to allow management applications from IBM, and other vendors to administer and monitor their activities.

Storage software solutions are a key element in the IBM overall storage strategy. This strategy addresses some of the most pressing needs currently facing Chief Information Officers (CIO) and IT managers. As part of its strategy, IBM intends to deliver industry-leading technologies that will help dramatically reduce total cost of ownership (TCO) for storage.

This strategy is the IBM System Storage Software Strategy. This strategy represents the next stage in the evolution of storage networking. It offers you the opportunity to fundamentally improve your company's effectiveness and efficiency in managing its storage resources.

17.4 More information

For more information about the IBM System Storage Virtualization family, refer to these links:

<http://www.ibm.com/systems/storage/virtualization>
<http://www.storage.ibm.com/software/roadmap/index.html>

For more information about SAN virtualization, refer to the Redpaper *Virtualization in a SAN*, REDP-3633, which can be found at:

<http://www.redbooks.ibm.com/redpapers/pdfs/redp3633.pdf>

For more information about *SNIA Shared Storage Model*, refer to:

http://www.snia.org/tech_activities/shared_storage_model/



SAN Volume Controller

IBM System Storage SAN Volume Controller (SVC) is an IBM storage virtualization solution which is designed to reduce the complexity and costs of managing storage networks. It provides storage virtualization and helps increase the utilization of existing capacity and centralize the management of multiple controllers in an open-system SAN environment. The IBM System Storage SAN Volume Controller (SVC) supports attachment to IBM and non-IBM storage systems. It enables storage administrators to reallocate and scale storage capacity and make changes to more underlying storage systems without disrupting applications.

In this chapter we cover the IBM System Storage SAN Volume Controller (SVC).

18.1 The need for storage virtualization

At the business level, clients are faced with three major storage challenges:

- ▶ **Managing storage growth:** Storage needs continue to grow at a rate that is normally higher than what has been planned for each year. As an example, storage subsystems may be purchased to last for 3 to 5 years; however, organizations are finding that they are filling to capacity much earlier than that.

To fill the growth, administrators are then either extending their current storage subsystems in chunks, or buying different types of storage subsystems to match their storage needs and budget.

- ▶ **Increasing complexity:** As storage needs grow, this need may be filled by multiple disk subsystems, which may not even be from the same vendor.

Together with the variety of server platforms and operating systems in a customer's environment, customers may have SANs with multiple and diverse storage subsystems and host platforms, as shown in Figure 18-1:

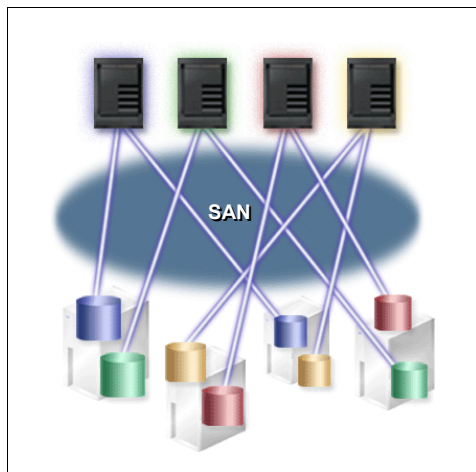


Figure 18-1 Today's environment

Combining this with the shortage of skilled storage administrators, the cost and risk of storage increases as the environment becomes more complex.

- ▶ **Maintaining availability:** With the increased range of storage options available, the storage growth rate, and no similar increase in storage budget, business has to manage more storage with no or minimal additional staff.

With the complexity highlighted above, and with requirements for higher business system availability, the potential for errors increases as each new storage subsystem is added to the infrastructure. Making changes to the storage infrastructure to accommodate storage growth runs the risk of outages that may not be acceptable to the enterprise.

Storage needs are rising, and the challenge of managing disparate storage systems is growing. The IBM TotalStorage SAN Volume Controller brings storage devices together in a *virtual pool* to make all storage appear as:

- ▶ One "logical" device to centrally manage and to allocate capacity as needed
- ▶ One solution to help achieve the most effective use of key storage resources on demand

Virtualization solutions can be implemented in the storage network, in the server, or in the storage device itself. The IBM storage virtualization solution is SAN-based, or in-band, which helps allow for a more open virtualization implementation. Locating virtualization in the SAN,

and therefore in the path of input/output (I/O) activity, helps to provide a solid basis for policy-based management. The focus of IBM on open standards means its virtualization solution supports freedom of choice in storage-device vendor selection.

The IBM System Storage SAN Volume Controller solution helps administrators:

- ▶ Simplify storage management
- ▶ Reduce IT data storage complexity and costs while enhancing scalability
- ▶ Extend on-demand flexibility and resiliency to the IT infrastructure
- ▶ Increase application availability by making changes in the infrastructure without having to shutdown hosts

18.2 IBM System Storage SAN Volume Controller (SVC)

The IBM System Storage SAN Volume Controller (SVC), machine type 2145, provides the ability to simplify the storage infrastructure, utilize storage resources more efficiently, improve personnel productivity and increase application availability. The SVC pools storage volumes from IBM and non-IBM disk arrays into a single reservoir of capacity which can be managed from a central point. The SVC allows data to be migrated between disk arrays without disrupting the applications and moves copy services into the network where they can be applied across the entire storage configuration.

18.2.1 Overview

Storage area networks (SAN) enable companies to share homogeneous storage resources across the enterprise. But for many companies, information resources are spread over a variety of locations and storage environments, often with products from different vendors. To achieve higher utilization of resources, companies need to be able to share their storage resources from all of their environments, regardless of vendor. While storage needs rise rapidly, and companies operate on lean budgets and staffing, the best solution is one that leverages the investments already made and that provides growth when needed. IBM System Storage SAN Volume Controller (SVC) offers a solution that can help strengthen existing SANs by increasing storage capacity, efficiency, uptime, administrator productivity and functionality.

The IBM System Storage SAN Volume Controller is designed to:

- ▶ Combine storage capacity from multiple vendors into a single repository of capacity with a central management point
- ▶ Help increase storage utilization by providing host applications with more flexible access to capacity
- ▶ Help improve productivity of storage administrators by enabling management of combined storage volumes from a single interface
- ▶ Support improved application availability by insulating host applications from changes to the physical storage infrastructure
- ▶ Enable a tiered storage environment, in which the cost of storage can be better matched to the value of data
- ▶ Support advanced copy services, from higher- to lower-cost devices and across subsystems from multiple vendors

The SVC combines hardware and software into a comprehensive, modular appliance. Using IBM eServer System x server technology in highly reliable clustered pairs, SVC is designed to

avoid single points of failure. SVC software is a highly available cluster optimized for performance and ease of use.

Storage utilization

SAN Volume Controller is designed to help increase the amount of storage capacity that is available to host applications. By pooling the capacity from multiple disk arrays within the storage area network (SAN), it helps enable host applications to access capacity beyond their island of SAN storage.

Scalability

A base configuration SVC consists of a single I/O group - an I/O group is a pair of high-performance, redundant Intel processor-based servers. Highly available I/O groups are the basic configuration of a cluster. Adding another I/O group can help increase cluster performance and bandwidth.

SVC can scale *out to* support up to four I/O groups, and it can scale *up to* support 1024 host servers. For every cluster, SVC supports up to 4096 *virtual disks* (see 18.2.2, “SVC commonly used terms” on page 369 for definitions).

Personnel productivity

SAN Volume Controller is managed at the cluster level, providing a single point of control over all the managed storage.

SAN Volume Controller provides a comprehensive, easy-to-use graphical interface for central management. This simple interface incorporates the Storage Management Initiative Specification (SMI-S) API, and further demonstrates the IBM commitment to open standards. With this single interface, administrators can perform configuration, management and service tasks over storage volumes from all the attached storage controllers. The GUI allows administrators to map disk storage volumes to virtual pooled volumes to help better use existing storage.

Link infrastructure performance to business goals

By pooling storage into a single reservoir, SVC helps insulate host applications from physical changes to the storage pool - so as to minimize disruption.

SVC simplifies the storage infrastructure by including a dynamic data-migration function, which migrates storage from one device to another, without taking it offline. Using this function, administrators can reallocate and scale storage capacity without disrupting applications, increasing application availability.

With SVC, businesses can build an infrastructure from existing assets that is simpler to manage, easier to provision, and can be changed without impacting application availability. Businesses can use their assets more efficiently and actually measure the improvements. They can allocate and provision storage to applications from a single view and know the affect on their overall capacity situation instantaneously. And they can quantify improvements to their application availability to enable better quality of service goals. These benefits help businesses manage their costs and capabilities more closely, linking the performance of their infrastructure to their individual business goals.

Tiered storage

In most IT environments, inactive data makes up the bulk of stored data. SVC helps administrators control storage growth more effectively by moving low-activity or inactive data into a hierarchy of lower-cost storage. Administrators can free disk space on higher-value storage for more-important, active data. This is achieved by creating different groups of

storage, corresponding to underlying storage with different characteristics, for example, speed and reliability. With SVC, you can match the cost of storage to the value of data.

Copy services

With many conventional SAN disk arrays, copy services can be performed within the array, or between identical arrays. SVC enables administrators to apply a single set of advanced copy services, such as the IBM FlashCopy and Global Mirror service across multiple storage subsystems from different vendors.

SVC continues to expand its ability to manage more diverse storage environments. With SVC V3.1 and higher, specific Virtual Disks (vdisks) can be configured individually to utilize the advanced copy services of the underlying disk storage arrays. This enables storage environments to realize the benefits of storage pooling and data migrations while continuing to use the copy services from their underlying disk arrays on a vdisk by vdisk basis.

With this latest software release V4.1, SVC now supports remote mirroring of data volumes at both “metropolitan” and “global” distances (IBM MetroMirror and GlobalMirror). This capability provides organizations more flexibility and choice to create solutions that meet their unique requirements for disaster recovery capabilities.

Technology for an on demand environment

Businesses are facing growing, critical application data supported by complex heterogeneous storage environments, while their staffs are overburdened. SVC is one of many offerings in our System Storage portfolio that are essential for an on demand storage environment. These offerings can help you simplify your IT infrastructure, manage information throughout its lifecycle and maintain business continuity.

18.2.2 SVC commonly used terms

Here is a short glossary of the most commonly used SVC terms (in alphabetic order).

Boss node

A single node acts as the boss node for overall management of the cluster. If the boss node fails, another node in the cluster will take over the responsibilities.

Configuration node

At any one time, a single node in the cluster is used to manage configuration activity. This configuration node manages a cache of the configuration information that describes the cluster configuration and provides a focal point for configuration commands. Similarly, at any one time, a single node acts as the boss node for overall management of the cluster.

Extent

An extent is a fixed size unit of data that is used to manage the mapping of data between MDisks and VDisks.

Front-end and back-end

SVC takes managed disks and represents these to application servers (hosts). The managed disks are looked after by the “back-end” application of the SAN Volume Controller. The virtual disks presented to hosts are looked after by the “front-end” application in the SVC.

Grain

A grain is the unit of data represented by a single bit in a FlashCopy bitmap, 256 K in SVC.

I/O group

An input/output (I/O) group contains two SVC nodes defined by the configuration process. Each SVC node is associated with exactly one I/O group. The nodes in the I/O group provide access to the VDisks in the I/O group.

LU and LUN

Strictly speaking, there is a difference between a logical unit (LU) and a logical unit number (LUN). A LUN is a unique identifier used on a SCSI bus that enables it to differentiate between up to eight separate devices (each of which is a logical unit). In practice, the two terms are used interchangeably. In this book, when we refer to a LUN, we refer to the unit of storage that is defined in a storage subsystem such as an IBM System Storage DS4000, DS6000 and DS8000 Storage Server or storage servers from other vendors.

Managed disk

Managed disk (MDisk) is a SCSI disk presented by a RAID controller and managed by the SAN Volume Controller. The MDisk must not be configured to be visible to host systems on the SAN.

Managed disk group

The managed disk group (MDG) is a collection of MDisks that jointly contain all the data for a specified set of VDisks.

Master console

The master console is the platform which runs the software used to manage the SVC.

Node

Node is the name given to the individual servers in an SVC cluster which run the SVC software.

SAN Volume Controller

The SAN Volume Controller (SVC) is a SAN appliance designed for attachment to a variety of host computer systems, which carries out block level virtualization of disk storage.

Virtual disk

Virtual disk (VDisk) is an SVC device that appears to host systems attached to the SAN as a SCSI disk. Each VDisk is associated with exactly one I/O group.

18.3 SVC architecture

SVC is based on the COMmodity PARTs Storage System, Compass architecture developed at the IBM Almaden Research Center. The overall goal of the Compass architecture is to create storage subsystem software applications which minimize the changes required when new storage hardware is deployed. The approach is to minimize the dependency on unique hardware, and to allow exploitation of, or migration to new SAN interfaces simply by plugging in new commodity adapters. Compass solutions can scale and expand over time, because they can be ported Compass to just about any platform and remain current with the latest processor and chipset technologies on each. The SVC compass architecture implementation uses Linux as a convenient development platform to deploy this function. This has, and will continue to enhance the ability of IBM to deploy robust function in a timely way.

The IBM System Storage SVC is a modular solution that consists of a Master Console for management, up to eight cluster nodes (added in pairs), and dual UPS, one for each node

pair, for write cache data protection (Figure 18-2). The nodes are the hardware elements of the SAN Volume Controller. SVC combines these nodes (servers) to create a high availability cluster. Each of the servers in the cluster is populated with 4 or 8 GB of high-speed memory, which serves as the cluster cache. A management card installed in each server monitors various parameters used by the cluster to determine the optimum and continuous data path.

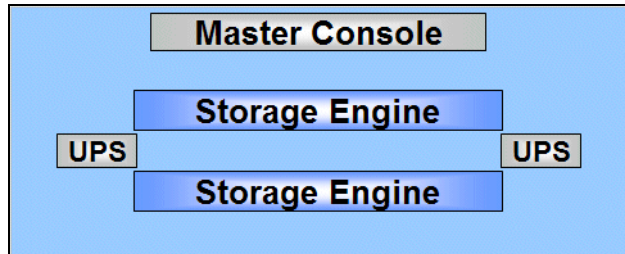


Figure 18-2 SVC components

The storage engines (or storage nodes) are always installed in pairs and combined into a high availability cluster. In simple terms, a cluster is a collection of servers that, together, provide a set of resources to a client. The key point is that the attached hosts (clients) using the SVC to access storage have no knowledge of the underlying physical hardware of the cluster. This means that the storage client is isolated and protected from changes to the physical hardware, which brings a number of benefits. Perhaps the most important of these benefits is high availability. Resources on clustered servers act as highly available versions of unclustered resources. If a node (an individual computer) in the cluster is unavailable, or too busy to respond to a request for a resource, the request is transparently passed to another node capable of processing it. Clients are, therefore, unaware of the exact locations of the resources they are using.

During normal cluster operation, the nodes keep in touch with each other. If a node is idle for a while (a few seconds) then a heartbeat signal is sent to assure connectivity with the cluster. If a node fails for any reason, the workload intended for it will be taken over by another node until the failed node has been restarted and re-admitted to the cluster (which happens automatically). If the microcode on a node becomes corrupted, resulting in a failure, the workload is transferred to another node. The code on the failed node is repaired, and the node is re-admitted to the cluster (again, all automatically).

Note: Although the SVC code is based on a Linux kernel, the clustering feature is not based on Linux clustering code. The clustering failover and failback feature is part of the SVC application software itself.

SVC nodes within the cluster are grouped into pairs (called I/O groups), with a single pair being responsible for serving I/O on a given VDisk (a VDisk is the SVC device that appears to a host system as a SAN attached disk). One node within the I/O Group represents the preferred path for I/O to a given VDisk; the other node represents the non-preferred path. This preference will alternate between nodes as each VDisk is created within an I/O Group to balance the workload evenly between the two nodes.

Each I/O group contains 8 GB (on SVC model 8F2 or 8F4) or 4 GB (on SVC model 4F2) of mirrored cache memory. Each also includes a host bus adapter (HBA):

- ▶ 4 Gbps (Standard for 8F4 and option for 8F2)
- ▶ 2 Gbps (model 4F2 and 8F2)

The HBA allows the SVC to connect and operate at up to 4 Gbps SAN speed.

Note: The 2145-4F2 SVC was withdrawn from marketing on April 14, 2006.

Beyond automatic configuration and cluster administration, data from attached storage clients is also transmitted reliably. When data is written by the host, the preferred node within the I/O Group stores a write in both its own write cache and the write cache of its partner (non-preferred) node before sending an “I/O complete” status back to the host application. To ensure that data is written if a node fails, the surviving node empties its write cache and proceeds in write-through mode until the cluster is returned to a fully operational state.

As yet another data protection feature, the SVC is supplied with UPS. In addition to voltage regulation to protect valuable electronic components within the SVC configuration, in the event of a main power outage, the UPS provides enough power to destage data to the SVC internal disk and shut down the nodes within the SVC cluster gracefully. This is a feature found in most high-end disk subsystems.

18.3.1 SVC virtualization

The SVC provides block aggregation and volume management for disk storage within the SAN. Essentially this means that the SVC manages a number of back-end storage controllers and maps the physical storage within those controllers to logical disk images that can be seen by application servers and workstations in the SAN. The SAN is zoned in such a way that the application servers cannot see the back-end storage, preventing any possible conflicts caused by the SVC and the application servers both trying to manage the back-end storage.

The SVC I/O Groups are connected to the SAN so that all back-end storage and all application servers are visible to all of the I/O Groups. The SVC I/O Groups see the storage presented to the SAN by the back-end controllers as a number of disks, known as Managed Disks or MDisks. Because the SVC does not attempt to provide recovery from physical disk failures within the back-end controllers, MDisks are usually, but not necessarily, part of a RAID array. The application servers on the other hand do not see the MDisks at all.

The zoning layout is shown in Figure 18-3 on page 373. Host zones connect the hosts to the I/O Groups, and the disk zone connects the I/O Groups to the storage controllers. Typically there are several host zones, since hosts can be grouped together in so-called “host groups” to better manage access.

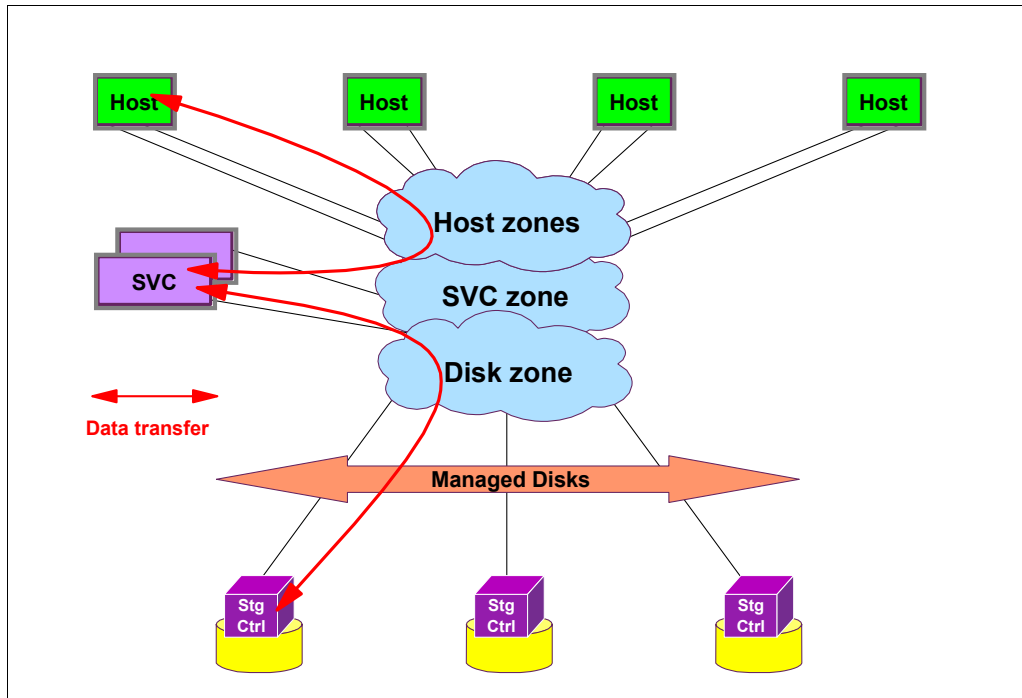


Figure 18-3 How to zone for correct SVC disk access

The MDisks are collected into groups, known as managed disk groups (MDGs). The MDisks that are used in the creation of a particular VDisk must all come from the same MDG. Each MDisk is divided up into a number of extents (default minimum size 16 MB, maximum size of 512 MB), which are numbered sequentially from the start to the end of each MDisk. Conceptually, this might be represented as shown in Figure 18-4.

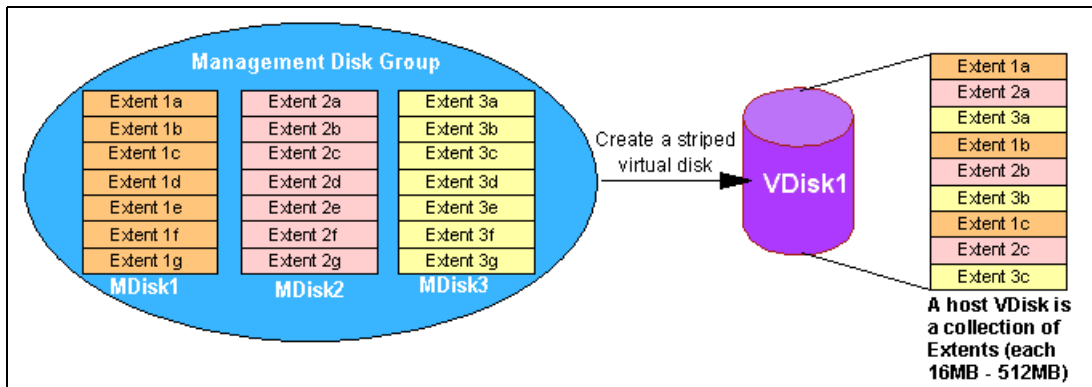


Figure 18-4 Extents being used to create a VDisk

The virtualization function in the SVC maps the VDisks seen by the application servers on to the MDisks provided by the back-end controllers. I/O traffic for a particular VDisk is, at any one time, handled exclusively by the nodes in a single I/O Group. Therefore, although a cluster could have many nodes within it, the nodes handle I/O in independent pairs. This means that the I/O capability of the SVC scales well (almost linearly), since additional throughput can be obtained by simply adding additional I/O Groups.

Each SVC node presents a VDisk to the SAN via multiple paths, usually four. Since in normal operation two nodes are used to provide redundant paths to the same storage, this means that depending on zoning, a single host HBA will see up to eight paths to each LUN presented

by the SVC. Because most operating systems are not able to resolve multiple paths back to a single physical device, IBM provides a multi-pathing device driver.

The multi-pathing driver supported by the SVC is IBM Subsystem Device Driver (SDD). Note that the SDD code supports both the SVC and the disk systems like the ESS, DS6000, and DS8000. Provided the latest version is used, a host can connect concurrently to both an SVC and a “native” DS6000 or Ds8000 environment.

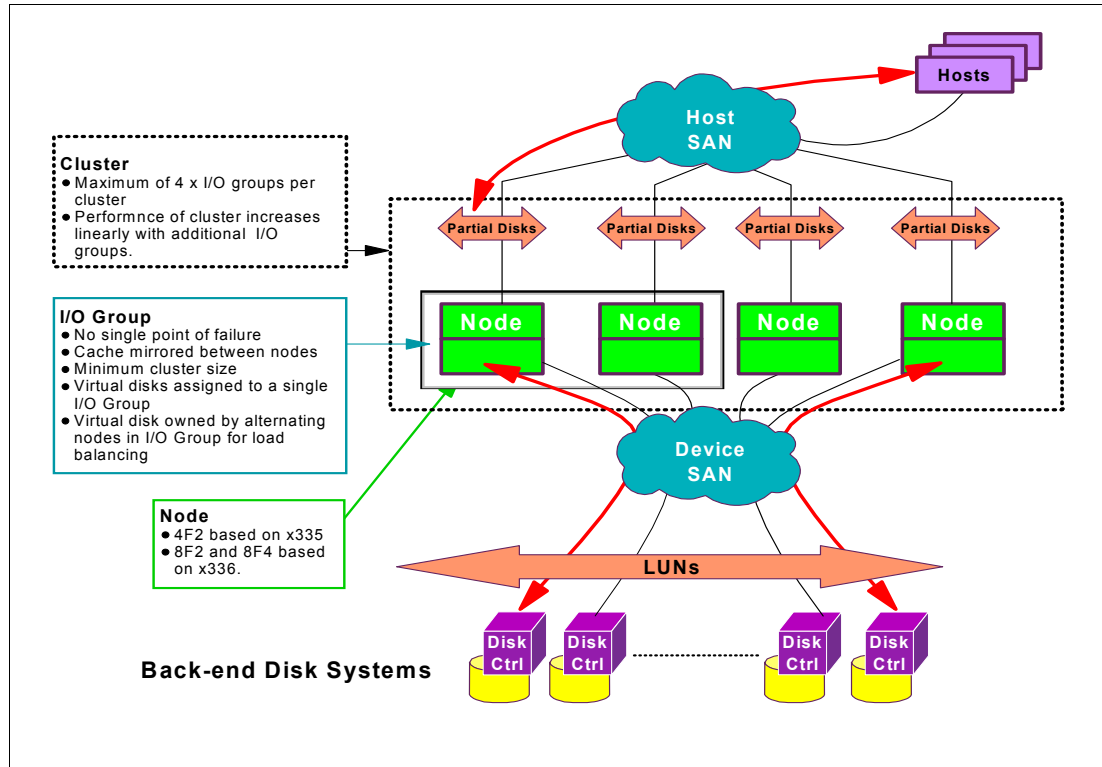


Figure 18-5 SVC logical view

Notes:

- ▶ Model 8F4 with 1 Gbps, 2 Gbps or 4 Gbps auto-negotiates the fabric speed on a per-port basis and is not restricted to run at the same speed as other node pairs in the cluster.
- ▶ To provide high availability, the SVC nodes should be configured in redundant SAN fabrics
- ▶ The Fibre Channel switches need to be zoned to permit the hosts to see the SVC nodes and the SVC nodes to see the storage Controllers. The SVC nodes within a cluster must be able to see each other and the master console. In addition, if there are two SVC clusters with Metro Mirror (formerly referred to as PPRC) services between them, zoning must be set so that all the nodes in both clusters see all the other nodes in both clusters.
- ▶ As well as a Fibre Channel connection or connections, each device has an Ethernet connection for configuration and error reporting, though only one of the nodes, the configuration node, binds an IP address to its Ethernet connection.

18.3.2 SVC software licensing

There are three parameters to consider when licensing the SAN Volume Controller software. These are:

- ▶ The amount of storage managed by the SAN Volume Controller cluster

The total amount of storage that is managed by the SAN Volume Controller cluster may be greater than the amount of storage that is virtualized by the cluster. Usage is charged based on the amount of Managed disk and not the amount of storage virtualized by a cluster. This is shown in Figure 18-6:

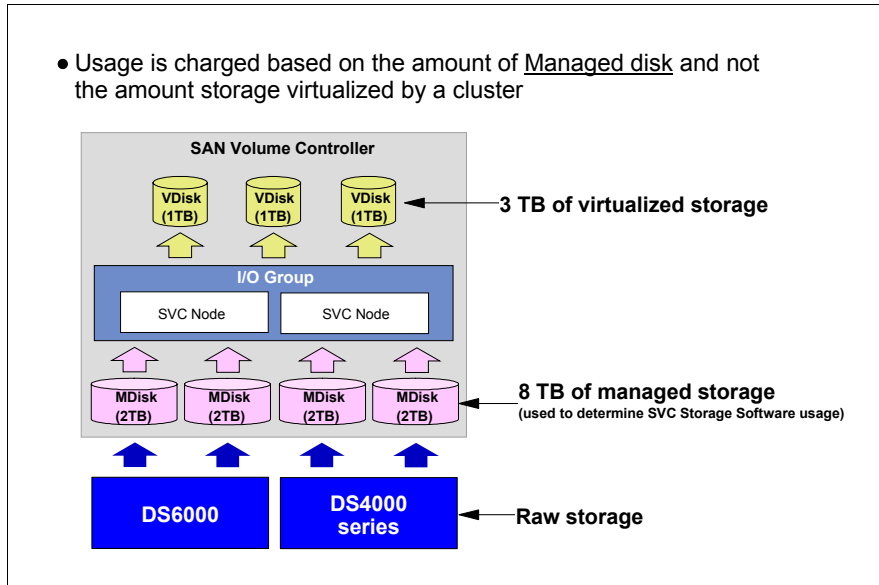


Figure 18-6 SVC managed storage-based licensing

- ▶ The amount of storage participating in FlashCopy

The amount of virtualized storage that you want to participate in simultaneous active FlashCopy relationships is priced based on the amount of storage using FlashCopy in a cluster. You may want to manage 8 TB of storage, but FlashCopy only 1 TB of storage. In this case, you need a 2 TB FlashCopy license since there is 1 TB of source and 1 TB of target FlashCopy volumes. See Figure 18-7 on page 376:

- Priced based on the amount of storage using FlashCopy in a cluster
- Storage using FlashCopy is derived from the total size of all source/target virtual disks in active FlashCopy relationships

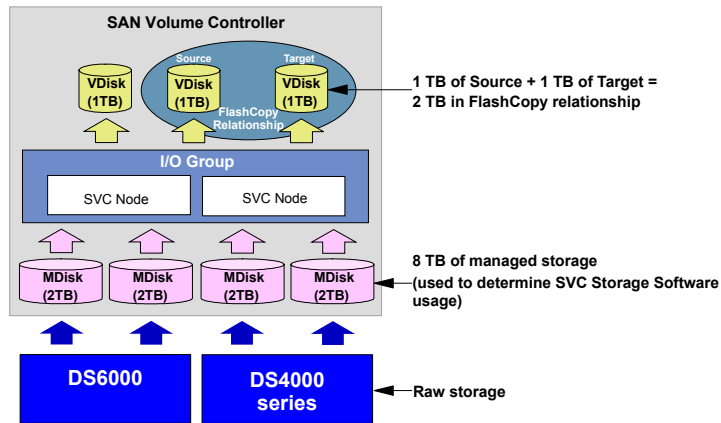


Figure 18-7 FlashCopy license illustration

- The amount of virtualized storage to Metro or Global Mirror (Metro and Global Mirror are delivered as a single feature) to another system.

The amount of virtualized storage that you want to Metro and Global Mirror to another system: You may be managing 8 TB of storage and FlashCopy copying 1 TB (2 TB FlashCopy license required) and Metro Mirror copying 1 TB. In this case, both clusters that are taking part in the Metro Mirror relationship require a 1 TB Metro Mirror license. In the case of intracluster Metro Mirror, where both primary and secondary volumes are in the same cluster, then the license must be large enough to cover both. This is shown in Figure 18-8:

- Intra Cluster MetroMirror primary and secondary exist in the same SVC cluster
 - For intracluster MetroMirror usage tracking similar to FlashCopy usage
- Priced based on the amount of storage using MetroMirror
- Storage using MetroMirror is derived from the total size of all primary and secondary virtual disks in active MetroMirror relationships

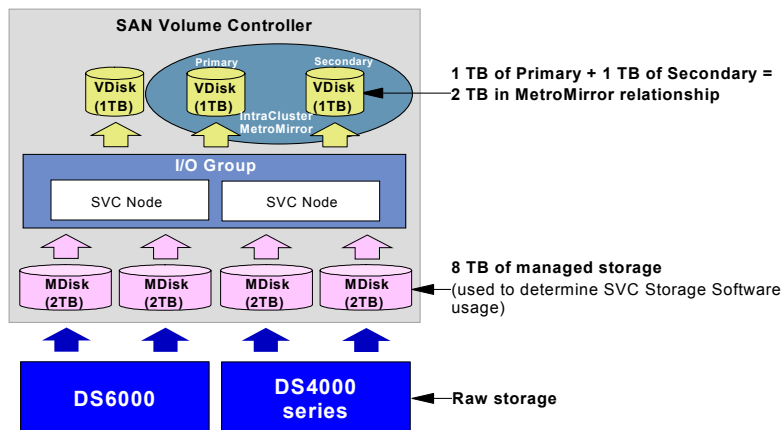


Figure 18-8 Metro Mirror and Global Mirror license illustration

As another example including Global Mirror, consider at site A: 10 TB of virtualization, 2 TB source for Metro Mirror, 2 TB target for Metro Mirror, and 2 TB source for Global Mirror; at Site B: Base virtualization of 4 TB, and 2 TB of target for Global Mirror (source is at Site A). This requires a Site A order for 10 TB base virtualization with 6 TB Metro/Global Mirror and a Site B order for 4 TB base virtualization and 2 TB of Metro/Global Mirror.

It is possible to increase any one of these three licenses independently of the other. That is, you can increase the total amount of managed storage without increasing the other licenses if the amounts of storage being copied remains unchanged. Similarly it is possible to change the copy licenses independently of each other.

18.4 SVC solutions

The SAN Volume Controller's unique features give it real advantages when used for storage tasks such as replication, data migration and storage capacity allocation.

18.4.1 Business continuity

Replicating data in a heterogeneous SAN environment from a high performance storage type at a central site to a lower cost set of storage at a remote site is always a performance concern. There are limits to the array based replication ability to have consistency groups across unlike devices. Also host level replication consumes precious server resources. SVC with its Metro and Global Mirror functions eliminates this problem making it significantly simpler to deploy for disaster recovery.

The latest SVC software V4.1 provides Global Mirror (consistent asynchronous secondary copy of data) at nearly unlimited distances with minimal performance impact on production servers. Figure 18-9 provides an overview of how this works:

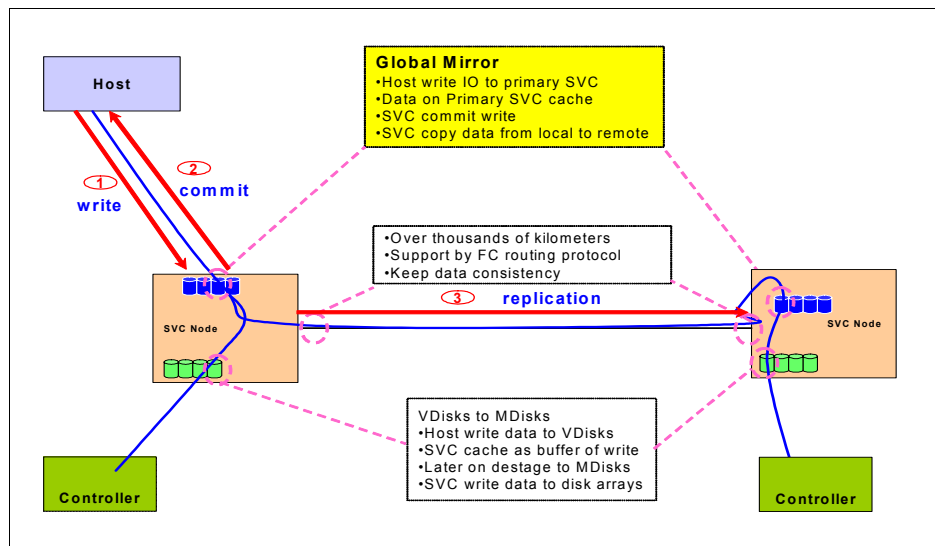


Figure 18-9 Business continuity with SVC Global Mirror

Metro Mirror provides synchronous data copy over shorter distances.

Both Metro Mirror and Global Mirror on SVC are supported by IBM TotalStorage Productivity Center for Replication.

18.4.2 Infrastructure simplification

SVC helps simplify the storage infrastructure by data migration and storage capacity utilization features.

Data migration

Data migration is the process of moving data from one storage volume to another. This might be required because you want to move less valuable data to cheaper storage, or perhaps a new faster storage system has been purchased, and you want to move some critical data onto it. In non-virtualized environments, this is a disruptive, time-consuming operation, requiring application downtime. Some storage systems do not support a direct remote copy function - which forces host-based volume drain, back up to tapes or restore from tapes during the migration. In addition data is unavailable during the migration process. Figure 18-10 shows a typical migration process without SVC.

Figure 18-10 shows a typical migration process without SVC.

- ▶ Step 1 — take the storage offline
- ▶ Step 2 — migrate data from old storage array to new storage array
- ▶ Step 3 — after migration, re-establish connections to host and bring the new storage system online
- ▶ Step 4 — return or dispose of old array

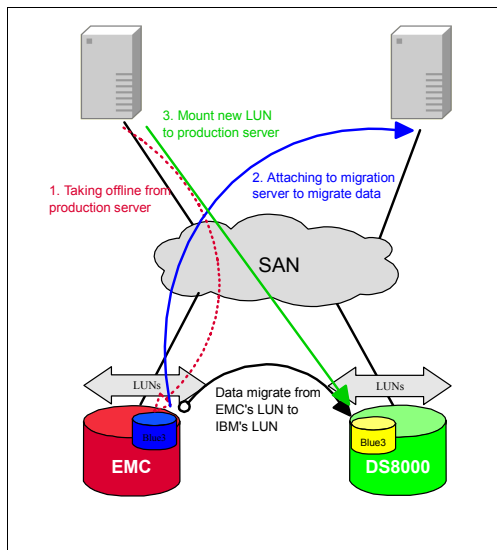


Figure 18-10 Data migration using SAN

Figure 18-11 on page 379 shows how you can migrate from a non-virtualized environment, to an SVC virtual environment.

1. The existing storage, applications and hosts are identified.
2. Next the SVC is installed, and the original volume to be migrated is mapped as a “hypothetical” or unmanaged SAN volume on the SVC. The application server now sees the virtual volume with all the existing data mapped to it. Operations can proceed as before.
3. Finally, if for example, more space is required for that data, you can nondisruptively move the data from the original volume to a new virtual volume. This will allow dynamic expansion of the volume and frees up the space used by the original application disk.

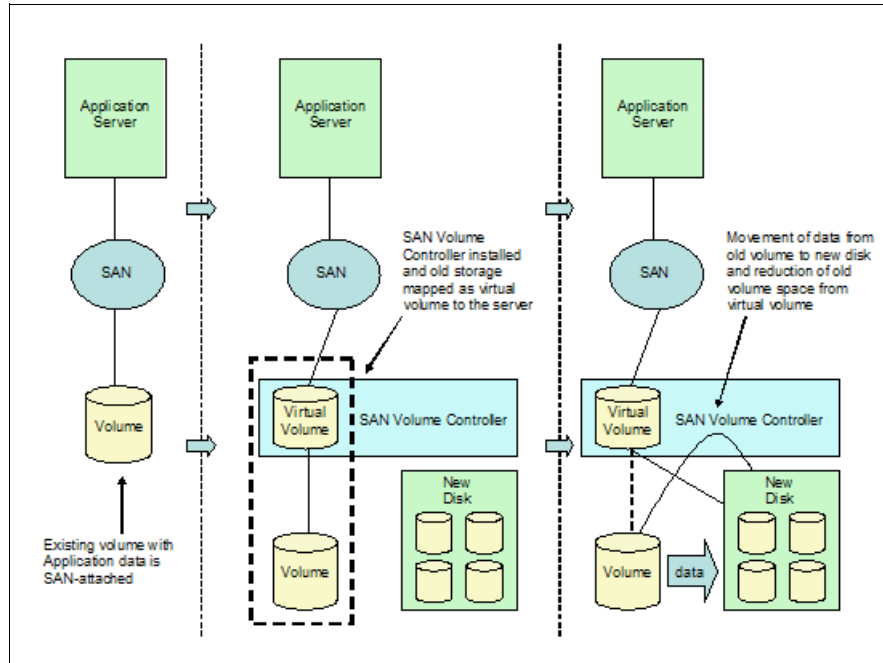


Figure 18-11 Data migration using SVC

To migrate existing SVC data to other storage, you simply tell SVC to remap the data on the virtual disk from one MDisk (managed disk) to another - online, without disrupting the application. The SVC manages all the physical infrastructure - and the attached hosts have no awareness of this, since the virtual disk (VDisk) mapping and structure remains unchanged.

Storage capacity utilization

Although traditional SAN heterogeneous environments share the physical network, but it offers very limited sharing capacity resulting in poor utilization of storage assets. It has procurement limitations due to the different purchasing requirements for each storage array island. Configuration changes are very complex as well as consolidating storage across data centers. It is hard to identify where storage pools capacity imbalances to move blocks around to improve and balance utilization.

SVC enables hosts to use virtual disks which consist of managed disks from any device within the storage pools. Unused storage capacity can be easily reallocated to help improve capacity utilization. The procurement of extra storage (in an environment using SVC) can be deferred until the total SAN capacity approaches utilization trigger points. Storage capacity balancing done at the virtualization layer transparently to servers and applications.

18.5 SVC compatibility

The SAN Volume Controller supports Windows NT, 2000, and 2003, AIX, SUSE SLES 8 and Red Hat EL AS Linux, Novell NetWare 6.5 with clustering, VMWare ESX, OpenVMS, Sun Solaris, and HP-UX hosts. The SAN switch support is also broad and includes the IBM System Storage SAN Switches and members of the Cisco families of SAN Switches and Directors.

Supported disk subsystems include: IBM System Storage DS4000, DS6000 and DS8000 series servers and N series. It also supports several models of Hitachi Thunder, Lightning,

and TagmaStore, EMC CLARiiON and Symmetrix, Network Appliances models, and Hewlett-Packard StorageWorks Modular Arrays and Enterprise Virtual Array.

For complete up-to-date SVC supported operating systems, hosts, HBAs, SAN switches and storage subsystems, see the links for *Supported hardware list* and *Recommended software levels* for the appropriate SVC code level under the Migration and upgrade session on the Web at:

<http://www-03.ibm.com/servers/storage/support/software/sanvc/planning.html>

18.6 More information

For more information, refer to these Web sites and Redbooks.

- ▶ For the latest SVC product information, go to:

<http://www-03.ibm.com/servers/storage/software/virtualization/svc/>

- ▶ *IBM System Storage SAN Volume Controller, SG24-6423*

- ▶ Support for SAN Volume Controller Web link:

<http://www-03.ibm.com/servers/storage/support/software/sanvc/>



IBM TotalStorage Productivity Center

The IBM TotalStorage Productivity Center is part of the Information Assets and Systems segment of the Information On Demand infrastructure. TPC provides single point of control and monitoring for subsystems, fabric components, and virtualization devices.

The IBM TotalStorage Productivity Center is comprised of:

- ▶ IBM TotalStorage Productivity Center for Data
- ▶ IBM TotalStorage Productivity Center for Fabric
- ▶ IBM TotalStorage Productivity Center for Disk
- ▶ IBM TotalStorage Productivity Center for Replication

19.1 Overview

The IBM TotalStorage Productivity Center is an open storage infrastructure management solution designed to help reduce the effort of managing complex storage infrastructures, to help improve storage capacity utilization, and to help increase administrative efficiency. It is designed to enable storage infrastructure to have the ability to respond to “on demand” storage needs. Figure 19-1 shows an ideal infrastructure to deploy TPC.

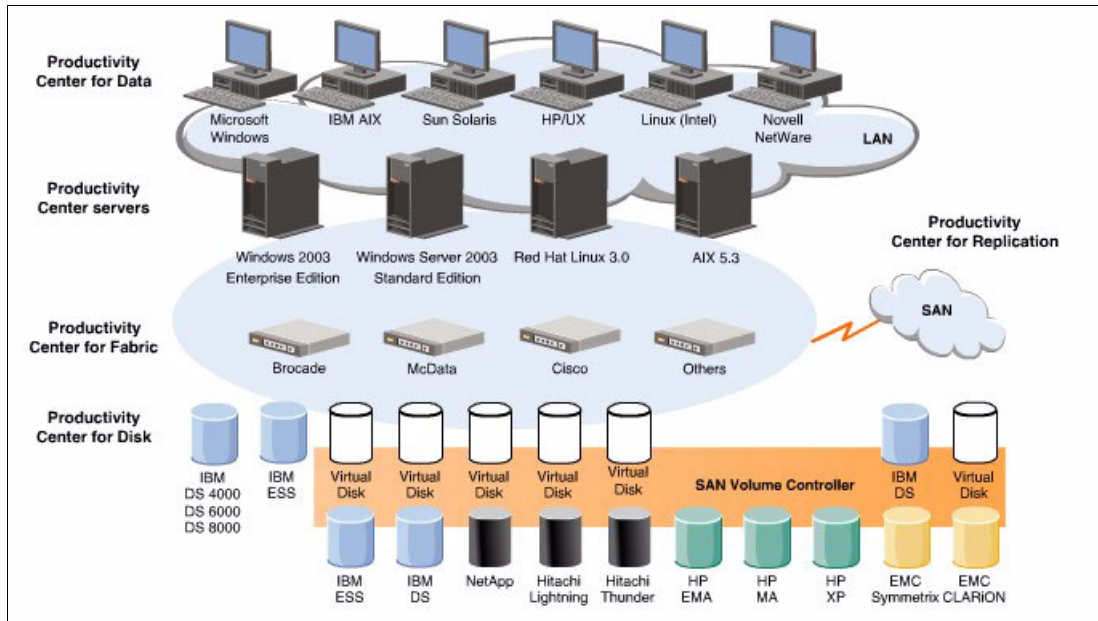


Figure 19-1 IBM TotalStorage Productivity Center infrastructure

IBM TotalStorage Productivity Center consists of the following four products:

IBM TotalStorage Productivity Center for Data

TPC for Data helps monitor, manage and automate capacity utilization of your file systems and databases. It includes enterprise reporting, policy based management, automated file system extension, Tivoli Storage Manager integration, database capacity reporting, and chargeback capabilities. For details, see Chapter 20, “IBM TotalStorage Productivity Center for Data” on page 389.

IBM TotalStorage Productivity Center for Fabric

TPC for Fabric monitors and manages the SAN fabric. It helps the administrator:

- ▶ Simplify the task of SAN management and configuration
- ▶ Maintain SAN availability
- ▶ Monitor SAN performance
- ▶ Improve SAN ROI

For details, see Chapter 21, “IBM TotalStorage Productivity Center for Fabric” on page 405.

IBM TotalStorage Productivity Center for Disk

TPC for Disk is designed to help reduce the complexity of managing SAN storage devices by allowing administrators to configure, manage and performance monitor storage from a single console. TPC for Disk is designed to:

- ▶ Configure multiple storage devices from a single console

- ▶ Monitor and track the performance of SAN attached Storage Management Interface Specification (SMI-S) compliant storage devices
- ▶ Enable proactive performance management by setting performance thresholds based on performance metrics and the generation of alerts
- ▶ Committed to open standards This solution is architected to SNIA's (Storage Networking Industry Association) SMI-S and supports compliant SAN components

For details, see Chapter 22, “TotalStorage Productivity Center for Disk” on page 421.

Note: For a list of IBM and non-IBM SMI-S Compliant devices TPC supports, visit:

<http://www.snia.org/ctp/conformingproviders/>

IBM TotalStorage Productivity Center for Replication

TPC for Replication provide management of FlashCopy, Metro Mirror and Global Mirror capabilities for the IBM ESS Model 800, IBM DS6000, and IBM DS8000. It also manages FlashCopy and MetroMirror for IBM SAN Volume Controller. TPC for Replication simplifies management of advanced copy services by:

- ▶ Automating administration and configuration of these services with wizard-based session and copy set definitions.
- ▶ Providing simple operational control of copy services tasks, including starting, suspending and resuming.
- ▶ offering tools for monitoring and managing copy sessions.

For details, see Chapter 23, “TotalStorage Productivity Center for Replication” on page 429.

19.2 IBM TotalStorage Productivity Center Version 3.1 enhancements

IBM TotalStorage Productivity Center Version 3.1 includes the following enhancements to Version 2.x:

- ▶ A single interface for administrative tasks. The interface centralizes the tasks for planning, monitoring, configuring, reporting, and problem determination. The interface provides consistent operations across host, fabric, and storage subsystems. You can troubleshoot and enforce storage service levels by providing performance and connectivity information about hosts, the fabric, and storage subsystems.
- ▶ Performance management support for select IBM System Storage, Brocade, Cisco, and McDATA fabric switches and directors.
- ▶ Support for IBM System Storage DS4000 and enhanced support for the IBM System Storage DS6000 and DS8000.
- ▶ Support for IBM TS3500 and IBM 3494 tape libraries.
- ▶ Role-based task authentication that assists in implementing storage management that conforms to government initiatives.
- ▶ Support for third-party disk array systems that include Storage Management Interface Specification (SMI-S) Providers certified by the SNIA Conformance Test Program (CTP) to be SMI-S 1.02 or SMI-S 1.1 compliant. This support includes asset and capacity reporting.

- ▶ A single database for IBM TotalStorage Productivity Center for Data, IBM TotalStorage Productivity Center for Fabric, and IBM TotalStorage Productivity Center for Disk.
- ▶ New management server platform support AIX 5.3, Red Hat AS 3.0, Windows 2003
- ▶ Management Database on DB/2 Import/Export of TPC Database and ESS Expert Database
- ▶ SAN and Disk Subsystem failure logging and audit logging
- ▶ Productivity Center replaces ESS Expert

TPC Master Console allows:

- ▶ Single Management Server
- ▶ Single Database
- ▶ Single Management Interface
- ▶ Single Point of Control
- ▶ Master Console
 - New, improved SAN and Storage topology viewer
 - Single sign on and role based administration
- ▶ Leading edge Topology Viewer
 - Allows for layered drill down capabilities without complexity
 - Relationships between hosts, fabric components and storage systems

TPC 3.1 improves the user profile for Operators and Administrators

- ▶ Roles are provided for administration (full access) and operation (read only access)
- ▶ Super user role provided
- ▶ Users can belong to more than one role, the complete set of roles are:
- ▶ TPC Super user
- ▶ TPC Administrator
- ▶ Data Administrator
- ▶ Data Operator
- ▶ Disk Administrator
- ▶ Disk Operator
- ▶ Fabric Administrator
- ▶ Fabric Operator
- ▶ Tape Administrator
- ▶ Tape Operator

New Installation PROGRAM:

- ▶ installs Agent Manager and the embedded version of IBM WebSphere Application Server Express, V5.0.
- ▶ installs IBM TotalStorage Productivity Center and the embedded version of IBM WebSphere Application Server Express, V5.0.

Topology Viewer

- ▶ Productivity Center's topology viewer is a powerful tool for visualizing the storage network, the relationships between devices within the network and the state of those devices
- ▶ Entirely new interface for version 3.1

Figure 19-2 on page 385 shows a TPC master console.

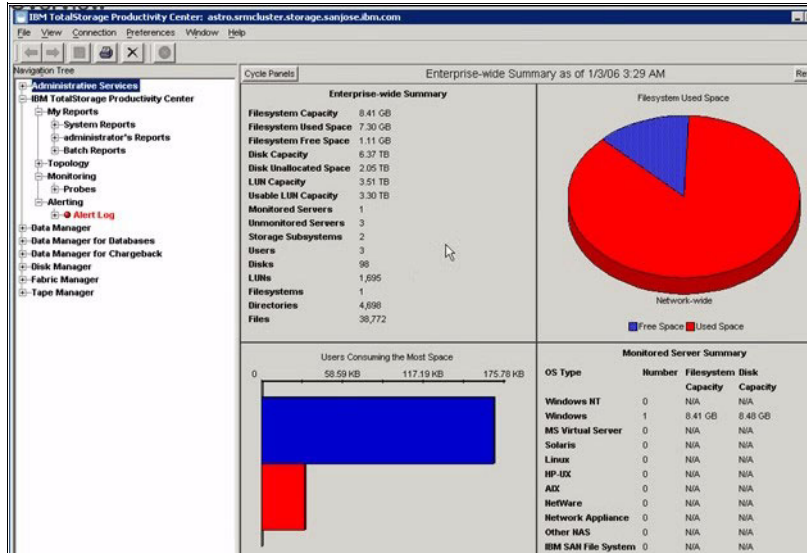


Figure 19-2 TPC Master Console

19.3 IBM TotalStorage Productivity Center Limited Edition

IBM TotalStorage Productivity Center Limited Edition is packaged with various IBM System Storage disk products, including the DS8000, DS6000 and DS4000. It also is available with the SAN Volume Controller and select IBM tape libraries. TPC Limited Edition is designed to:

- ▶ Discover and configure IBM and heterogeneous SMI-S supported devices
- ▶ Perform event gathering, error logging and launch device element managers
- ▶ Provide basic asset and capacity reporting
- ▶ Display an end-to-end topology view of your storage infrastructure and health console
- ▶ Enable a simple upgrade path to IBM TotalStorage Productivity Center Standard Edition (or single priced modules)

19.4 IBM TotalStorage Productivity Center Standard Edition

IBM TotalStorage Productivity Center Standard Edition provides all the management capabilities to better manage your heterogeneous storage infrastructure from application to back-end storage system disk at a single bundled price. It consists of:

- ▶ IBM TotalStorage Productivity Center for Data
- ▶ IBM TotalStorage Productivity Center for Fabric
- ▶ IBM TotalStorage Productivity Center for Disk

19.5 IBM TotalStorage Productivity Center for Replication

TPC for Replication is packaged and priced separately from TPC Standard Edition. In addition, TPC for Replication is available in two complementary packages: TPC for Replication and TPC for Replication Two Site BC. Figure 19-3 on page 386 shows the TPC products.

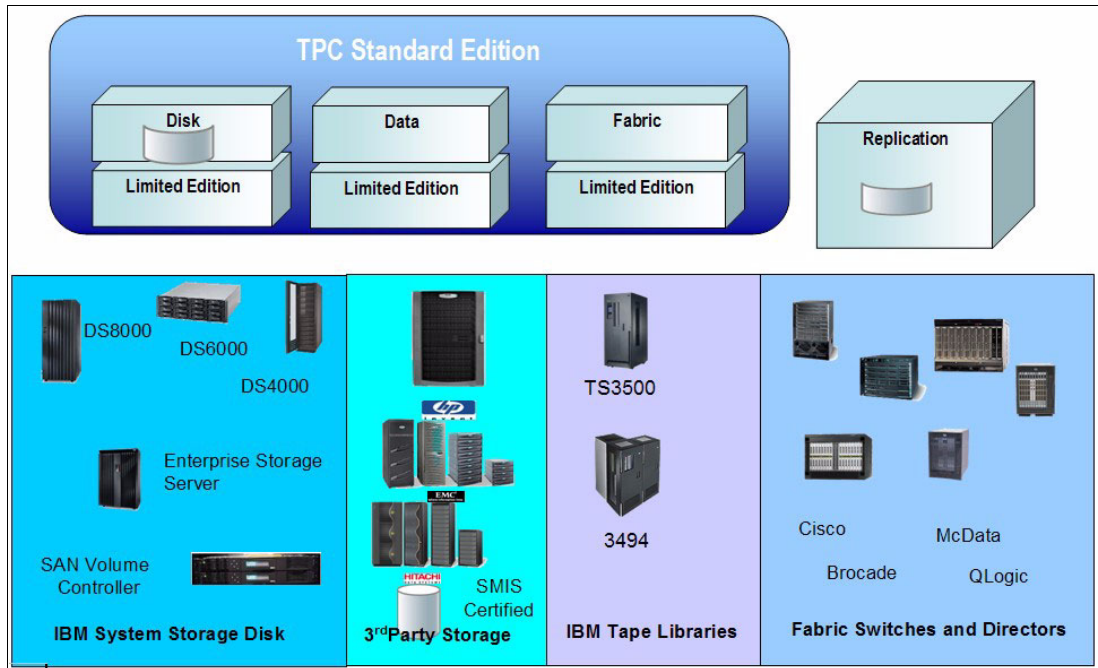


Figure 19-3 IBM TotalStorage Productivity Center Products

19.6 SMI-S

Storage Management Initiative – Specification by Storage Networking Industry Association provides an industry-standard interface for the discovery, monitoring and management of storage devices. Includes Tape, Storage Systems, fabric components, anything that can be attached to a SAN. SNIA provides a certification test for devices that claim support for the standard. Device manufacturers and management application developers submit their agents and their applications to be certified.

The API provided by SMI-S makes it easier for management applications to collect data from devices. The specification also provides consistent definitions of the data fields in the API, which in theory will give consistent views and comparisons between devices. In theory, because vendors are still free to interpret the fields as they see fit. So, it is important to know your devices and understand just what is meant by the performance metrics returned by the device.

TPC uses SMI-S for:

- ▶ Storage Array discovery and reporting
- ▶ Storage Array provisioning
- ▶ Storage Array performance monitoring and reporting
- ▶ Fabric (switch port) performance monitoring and reporting
- ▶ Tape discovery and reporting
- ▶ SVC discovery and reporting (IBM only)

19.7 More Information

Refer to these Redbooks for more information about IBM TotalStorage Productivity Center:

- ▶ *Exploring Storage Management Efficiencies and Provisioning - Understanding IBM TotalStorage Productivity Center and IBM TotalStorage Productivity Center with Advanced Provisioning*, SG24-6373
- ▶ *IBM TotalStorage Productivity Center V3.1: The Next Generation*, SG24-7194

For the latest information about IBM TotalStorage Productivity Center, refer to following link:

<http://www-1.ibm.com/servers/storage/software/center/index.html>

Latest documentation is available from the TPC Infocenter at:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

The supported platform matrix for TPC is at:

<http://www-1.ibm.com/support/docview.wss?rs=1133&uid=ssg1S1002814>

Supported device hardware matrix for TPC is at:

<http://www-1.ibm.com/support/docview.wss?uid=ssg1S1002800&aid>



IBM TotalStorage Productivity Center for Data

As a component of the IBM TotalStorage Productivity Center, IBM TotalStorage Productivity Center for Data can help you improve your storage ROI by:

- ▶ Improving storage utilization
- ▶ Enabling intelligent capacity planning
- ▶ Helping you manage more storage with the same staff
- ▶ Supporting high application availability
- ▶ Detecting failed backup and archive by file
- ▶ Providing enterprise-wide reports on capacity via role-based management
- ▶ Assisting customers in support of data classification, ILM assessments and ITIL storage practices

This chapter introduces and positions the IBM TotalStorage Productivity Center for Data and discusses its architecture and key features.

20.1 TotalStorage Productivity Center for Data overview

IBM TotalStorage Productivity Center for Data (Data Manager or TPC for Data) is a comprehensive file and capacity management solution for heterogeneous storage environments. It includes enterprise-wide reporting and monitoring, policy-based management and automated capacity provisioning for Direct Attached Storage (DAS), Network Attached Storage (NAS), and SAN environments. TPC for Data lets you improve storage utilization, plan for future capacity, and ensure availability by providing storage on demand for file systems.

TPC for Data includes the basic Data Manager, Data Manager for Databases, and Data Manager for Chargeback.

Use TPC for Data to perform the following functions:

- ▶ Discover and monitor disks, partitions, shared directories, and servers.
- ▶ Monitor and report on capacity and utilization across platforms to help you to identify trends and prevent problems.
- ▶ Monitor storage assets associated with enterprise-wide databases and issues notifications of potential problems.
- ▶ Provides a wide variety of standardized reports about file systems, databases (using the Data Manager for Databases function), and storage infrastructure to track usage and availability.
- ▶ Provide file analysis across platforms to help you to identify and reclaim space used by non-essential files.
- ▶ Provide policy-based management and automated capacity provisioning for file systems when user-defined thresholds are reached.
- ▶ Generate invoices that charge back for storage usage on a departmental, group, or user level (using the Data Manager for Chargeback function).

These functions that are available with TPC for Data are designed to help you lower your storage costs by:

- ▶ Improving storage utilization
- ▶ Enabling intelligent capacity planning
- ▶ Helping you manage more storage with the same staff
- ▶ Supporting application availability through computer uptime reporting and application database monitoring.

20.1.1 TPC for Data and data management

Every storage administrator would like to have an unlimited amount of storage. Unfortunately in the real world, this is not possible. The best you can hope for is to manage effectively the storage which is available, and accurately predict what is needed for the future. An important tool to assist this is knowing how much storage is actually needed, and to do that, you need to know what you actually need to store. Data stored electronically essentially can be divided into the categories shown in Figure 20-1 on page 391.

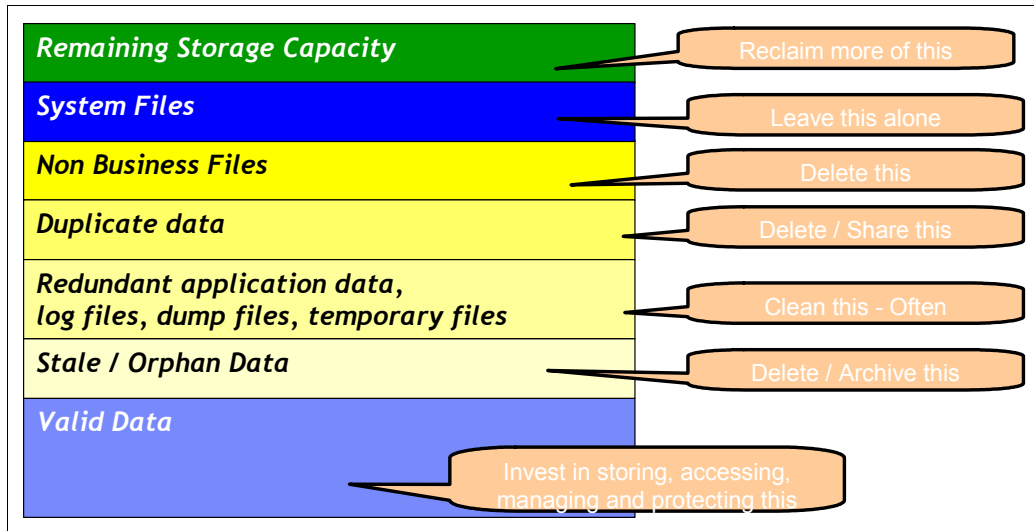


Figure 20-1 Types of data typically found in an enterprise and what to do with them

TPC for Data helps you identify how much of each type of data that you have, and automate actions to deal with each type. A more detailed discussion of this topic is in the IBM Redbook *ILM Library: Techniques with Tivoli Storage and IBM TotalStorage Products*, SG24-7030.

20.1.2 TPC for Data product highlights

TPC for Data includes many advanced features to help manage and automate capacity utilization of file systems and databases. This section describes the product highlights:

► Enterprise reporting

Over 300 comprehensive enterprise-wide reports (and the ability to customize your own reports). These reports can help administrators to make intelligent capacity management decisions based on current and trended historical data.

► Policy-based management

TPC for Data can enable you to define and enforce storage policies through user-defined alerts, quotas, and constraints, notifying the user by e-mail, pager, or the event log, or a systems management console for events such as when a quota has been exceeded or a constraint violated.

However, finding a problem is not enough. You need a way to fix problems — or potential problems as they are discovered. TPC for Data can provide automated solutions through event management.

For example, if TPC for Data discovers data that has not been accessed in more than a year, it can trigger Tivoli Storage Manager or another archive utility.

► Automatic file system extension

TPC for Data's automated file system extension capability can allow the user to specify a policy to automatically extend a file system when a threshold has been reached.

For example, if a file system's threshold is set at 78% and, through monitoring, TPC for Data identifies that this threshold has been exceeded, it can automatically initiate a file system extension to reduce the possibility of a storage-related outage. The feature supports both manual and automated initiated extension. Once you are comfortable with the manual process, you can turn over all the steps to TPC for Data. A probe will run on agents and send file system statistics to the server. The server will compare the current

file system statistics against the policy, and invoke provisioning and extension as necessary.

▶ **Direct Tivoli Storage Manager integration**

With this feature, TPC for Data can automatically invoke Tivoli Storage Manager to archive and delete files. This can free up space in a file system and can allow you to more effectively manage storage utilization. For example, a policy can be created to archive all files over 365 days old to tape using Tivoli Storage Manager, and then delete the files to free up the disk space.

▶ **The Database capacity reporting feature** (Data Manager for Databases) enables you to see how much storage is being consumed by users, groups of users and OS's within the database application. It can help you manage and automate capacity utilization of your databases.

▶ **Chargeback capabilities** (Data Manager for Chargeback) enable you to provide usage information by department, group or user making data owners aware of and accountable for their data usage.

20.2 TPC for Data architecture

The architecture of TPC for Data lets you see all of the storage assets including direct-attached storage and network-attached storage. This comprehensive view of the entire storage map allows the administrators to manage much larger environments, but also get the information about utilization and usage that is typically required in large environments. The information collected by TPC for Data can help you to make intelligent decisions optimizing the utilization of your open system environments.

The scope of TPC for Data is not just limited to files and their attributes, but also includes relational database managers such as Oracle, Sybase, SQL Server, and DB2.

The data collected by TPC for Data helps you understand what is really going on with the data that resides on your servers. This includes views as to when files are created, accessed and modified and by what group or user. This type of information enables you to map the actual storage resource to the consumers of that resource. The ability to map storage consumption to storage hardware has become increasingly important as the size of open systems environments have increased. In addition to understanding the current consumption and usage of data within the enterprise, TPC for Data keeps track of this information over time. Not only does this historical view of storage consumption and utilization allow you to see usage trends over time, it also enables you to see a projected use of storage into the future. This allows you to prepare for the need to purchase additional capacity in a planned proactive manner rather than just reacting to being out of space.

The major components of IBM TotalStorage Productivity Center for Data are:

▶ **Data Manager**

The manager controls the discovery, reporting, and alert functions. It does the following:

- Receives information from the agents and stores that information in the central repository
- Issues commands to agents for jobs
- Receives requests from clients for information and retrieves the requested information from the central data repository

▶ **Data agents on managed systems**

An agent resides on each managed system. Each agent performs the following functions:

- Runs probes and scans.

- Collects storage-related information about the volumes or file systems that are accessible to the managed systems.
 - Forwards information to the manager to be stored in the database repository.
- ▶ Web server
The optional Web server permits remote Web access to the server.
 - ▶ Clients
Clients communicate directly to TPC for Data to perform administration, monitoring, and reporting. A client can be a locally installed interface to TPC for Data, or it can use the Web server to access the user interface through a Web browser.

20.3 TPC for Data functions

TPC for Data provides you with a number of functions that enable you to collect information about and perform tasks against the storage resources within your enterprise. This section provides a brief overview of the main functions that can help you manage and automate capacity utilization of your file systems and databases.

20.3.1 Dashboard

The dashboard, shown in Figure 20-2, appears whenever you start TPC or close all the active windows in a running session. The dashboard provides a concise, yet detailed overview of the health of your storage environment. It enables you to quickly point out potential problem areas that need further investigation.

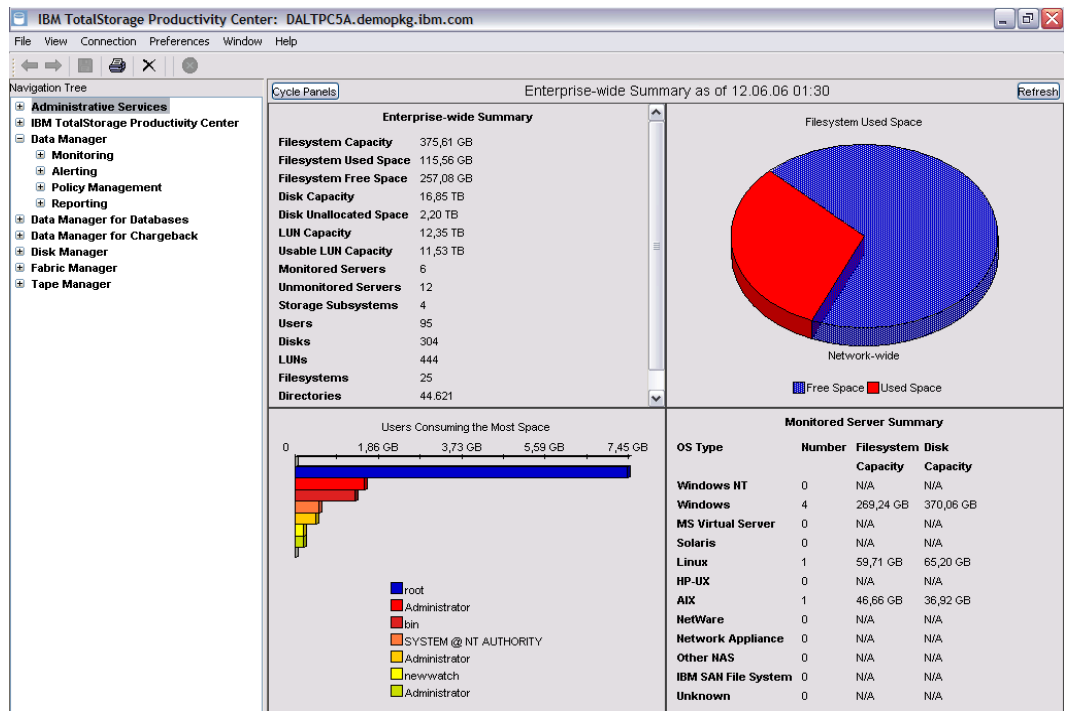


Figure 20-2 TPC for Data - dashboard

The dashboard contains four display areas and seven panels that you can cycle through to view your environment's storage information. Use the **Refresh** button to update the display.

Enterprise-wide summary

Provides an enterprise-wide summary of your storage. This includes:

- ▶ File system capacity: total amount of allocated storage space
- ▶ File system used space: total amount of used storage space
- ▶ File system Free space: total amount of available storage space
- ▶ Disk Capacity: total amount of allocated and unallocated storage space
- ▶ Disk Unallocated Space: amount of unused space unallocated to file systems
- ▶ Monitored Servers: number of monitored servers
- ▶ Unmonitored Servers: number of unmonitored servers
- ▶ Users: number of users
- ▶ Disks: number of disks
- ▶ File systems: number of file systems
- ▶ Directories: number of directories
- ▶ Files: number of files

File system used space

Displays a pie chart showing the total amount of used storage space across the entire network. Different chart types can be selected here. This provides a quick snapshot of you file system space utilization efficiency.

Users consuming the most space

By default displays a bar chart (different chart types can be selected) of the users who are using the largest amount of file system space.

Note: The space usage of administrative users within your environment is not included as part of this dashboard report.

Monitored server summary

Shows a table of total disk file system capacity for the monitored servers, sorted by operating system type.

File systems with least free space percentage

Shows a table of the monitored machines with the least percentage of free space, including the percent of free space on each machine, the total file system capacity on each machine, and the name of the machines and file systems/mount points.

Users consuming the most space report

Displays a table with a list of the users consuming the most space within your environment, and the total amount of storage each user is consuming.

Alerts pending

Displays the active alerts that are still pending. These alerts are organized by all, computer, disk, file system, directory, user quota, OS user group quota, Directed to user, storage subsystem, switch, fabric, endpoint device, external, and tape library.

Note: Alerts for failed jobs (pings, probes, and scans) are not listed in separate rows on this panel, but are included in the total displayed in the All row.

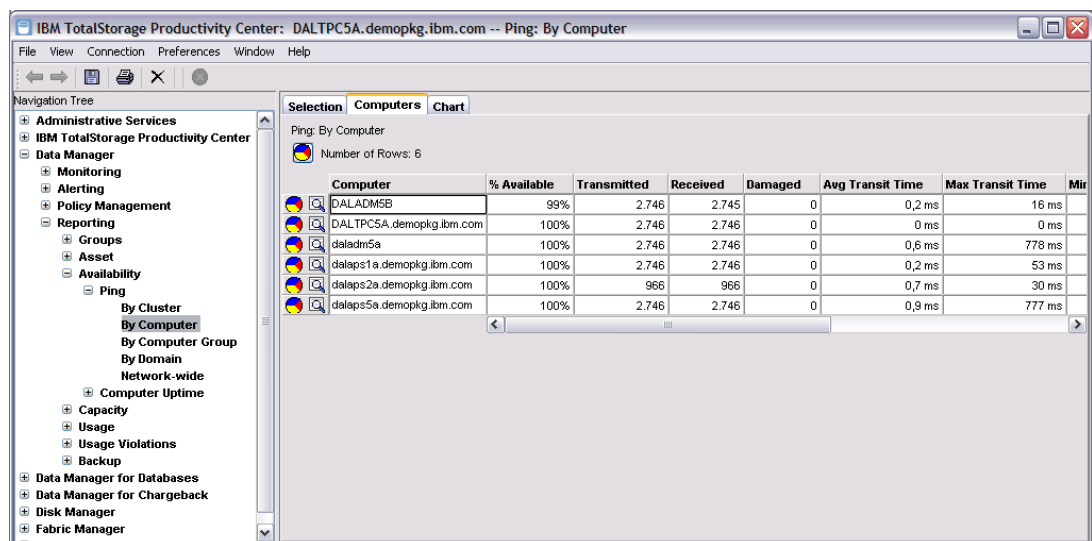
20.3.2 Discover and monitor information

Data Manager uses three methods to discover information about the assets in the storage environment: pings, probes, and scans. These are typically set up to run automatically as scheduled tasks. You can define different ping, probe and scan jobs to run against different Agents or groups of Agents (for example, to run a regular probe of all Windows systems), according to your particular requirements.

Pings

A ping is a standard ICMP (Internet Control Message Protocol) ping which collects information about the availability of the storage assets in your storage environment. By running pings and viewing the results of those pings in Reporting > Availability reports, you can monitor and report on the availability of your storage, both from a network point of view and from a computer uptime perspective. This allows you to see what percentage of the time the storage in your enterprise (or on particular server or group of servers), is off-network due to network problems or perhaps is unavailable due to system downtime. You can define any number of pings to check the availability of different storage resources at different times.

An example of an Availability Report for Ping is shown in Figure 20-3.



Computer	% Available	Transmitted	Received	Damaged	Avg Transit Time	Max Transit Time	Mir
DALADMSB	99%	2,746	2,745	0	0,2 ms	16 ms	
DALTPCSA.demopkg.ibm.com	100%	2,746	2,746	0	0 ms	0 ms	
daladm5a	100%	2,746	2,746	0	0,6 ms	778 ms	
dalaps1a.demopkg.ibm.com	100%	2,746	2,746	0	0,2 ms	53 ms	
dalaps2a.demopkg.ibm.com	100%	966	966	0	0,7 ms	30 ms	
dalaps5a.demopkg.ibm.com	100%	2,746	2,746	0	0,9 ms	777 ms	

Figure 20-3 Availability report - ping

If an agent does not respond to a ping (or a predefined number of pings) you can set up an alert to take some action. The actions could be one, any or all of SNMP Trap, TEC Event, Login Notification, Windows Event Log, UNIX Syslog, running a script, and e-mail.

Scans

Scan jobs are used to collect detailed statistics on the usage and trending of the storage consumption within your environment. By running scans and viewing the results of those scans in Capacity, Usage, Usage Violations, and Backup Reports, you can do such things as the following:

- ▶ View when files were created accessed and modified and by what group or user
- ▶ Map the actual storage resource to the consumers of that resource
- ▶ Generate a historical view of storage consumption and utilization that enables you to see usage trends over time and enables the system administrator to see a projected use of storage into the future

By using the directory groups option you can specify directories within a file system to scan, rather than having to scan an entire file system.

Profiles enable you to control what files are scanned and what statistics are gathered by a scan. Using profiles is a powerful method for determining the statistics you want to gather during scans. Keep in mind that the more profiles you include within a scan, the longer the scan will take to gather the statistics specified by those profiles.

You can trigger the same actions by failed scans as with failed pings.

Figure 20-4 shows a sample of a report produced from data collected in scans.

Computer	Mount Point	Used Space	Percent Used Space	Free Space	Capacity	File Count	Directories
TOTAL		102,47 GB	27%	269,58 GB	375,02 GB	501,367	
dalaps5a.demopkg.ibm.com	E:/	35,31 GB	94%	1,95 GB	37,26 GB	53,435	
dalaps5a.demopkg.ibm.com	C:/	5,29 GB	90%	583,04 MB	5,86 GB	18,391	
dalaps2a.demopkg.ibm.com	/usr	2,50 GB	87%	380,62 MB	2,88 GB	69,381	
DALADMSB	C:/	4,12 GB	70%	1,74 GB	5,86 GB	26,556	
DALTPC5A.demopkg.ibm.com	C:/	3,55 GB	60%	2,31 GB	5,86 GB	24,060	
DALTPC5A.demopkg.ibm.com	E:/	21,46 GB	57%	15,80 GB	37,26 GB	55,780	
dalaps5a.demopkg.ibm.com	G:/	246,55 MB	49%	247,61 MB	494,16 MB	519	
dalaps2a.demopkg.ibm.com	/opt	228,70 MB	44%	283,30 MB	512,00 MB	5,258	
dalaps1a.demopkg.ibm.com	/boot	37,73 MB	40%	55,89 MB	98,72 MB	52	
dalaps2a.demopkg.ibm.com	/	12,71 MB	19%	51,29 MB	64,00 MB	1,915	
DALTPC5A.demopkg.ibm.com	D:/	4,13 GB	18%	17,95 GB	22,08 GB	35,689	
DALADMSB	E:/	6,61 GB	17%	30,64 GB	37,25 GB	3,359	
dalaps5a.demopkg.ibm.com	D:/	3,26 GB	14%	18,81 GB	22,08 GB	10,768	
daladm5a	C:/	8,50 GB	11%	63,51 GB	72,01 GB	59,214	
dalaps2a.demopkg.ibm.com	/storage	3,84 GB	10%	31,16 GB	35,00 GB	5,767	

Figure 20-4 Capacity report - file system used space

This report shows a list of the file systems on each agent, the amount of space used in each, expressed in bytes and as a percentage, the amount of free space, and the total capacity available in the file system.

Probes

Probes are used to gather information about the assets and system resources of monitored servers, such as processor count and speed, memory size, disk count and size, file systems, and so on. Probes also gather information about the files, instances, logs and objects that makeup the monitored databases. Data collected by probes is used in the Asset Reports. Figure 20-5 on page 397 shows an Asset Report for discovered LUNs.

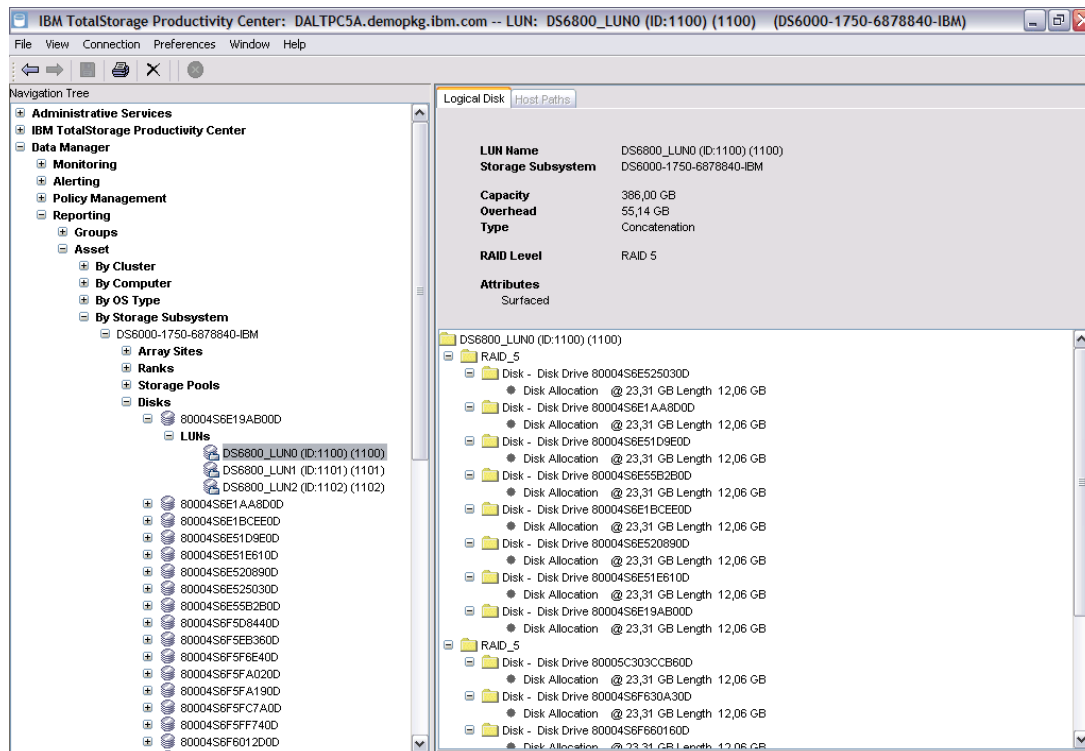


Figure 20-5 Asset report - discovered LUNs

20.3.3 Alerts

An alert defines an action to be performed if a particular event occurs or condition is found. Alerts in TPC for Data can be set on computers, file systems, and directories. Alerts in Data Manager for Databases can be set on instances, database-tablespaces, and tables. Alerts can tell you, for instance, if a file system or database is approaching capacity.

An Alert will register in the Alert log, plus you can also define one, some or all of the following actions to be performed in addition:

- ▶ Send an SNMP trap
- ▶ Send an alert to the Tivoli Enterprise™ Console (TEC).
- ▶ Generate a login notification next time a specified user logs in to TPC for Data.
- ▶ Write out alert messages to the OS log (Windows Event Log or UNIX Syslog).
- ▶ Run a script in response to an alert (you can execute any third-party tools for actions, such as archiving, backup and recovery, or provisioning).
- ▶ Send an e-mail or a page indicating the nature of the alert.

Figure 20-6 on page 398 shows the Alert Log. Entries in red signify that an Alert threshold has been reached.

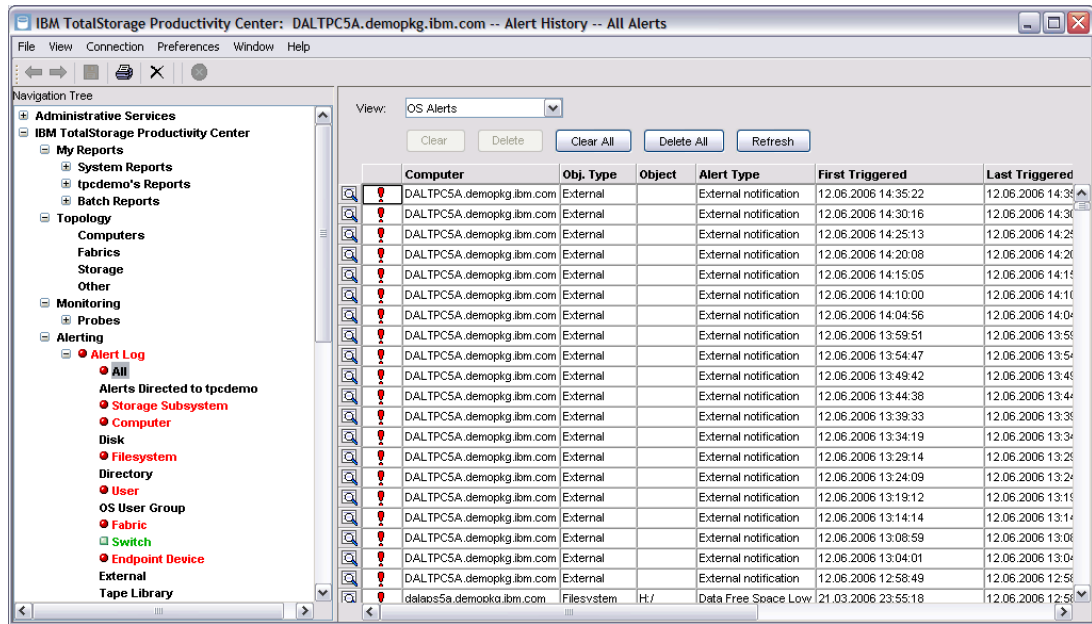


Figure 20-6 TPC 3.1 Alert Log

20.3.4 Reporting

Reporting in TPC for Data is very rich, with over 300 predefined views, and the capability to customize those standard views, save the custom report, and add it to your menu for scheduled or ad hoc reports. You can also create your own individual reports according to particular needs and set them to run as needed, or in batch (regularly). Reports can be produced in table format or a variety of charting (graph) views. You can export reports to CSV (Comma Separated Values) or HTML file formats for external usage.

Reports are generated against data already in the repository. A common practice is to schedule scans and probes just before running reports.

Reporting can be done at almost any level in the system, from the enterprise down to a specific entity and any level in between. Figure 20-2 on page 393 shows a high-level summary report. Or, you can drill down to something very specific. Figure 20-7 on page 399 is an example of a lower-level report, where the administrator has focused on a particular Agent, to look at a particular disk on a particular controller.

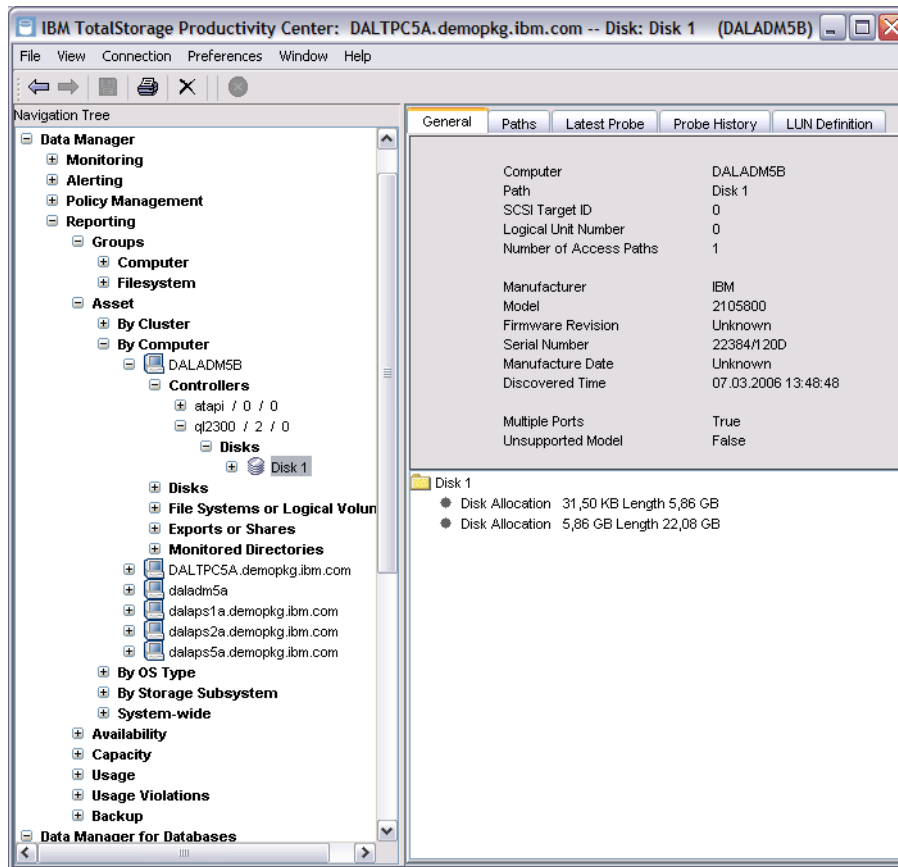


Figure 20-7 Asset report - By computer

Reports can be produced either system-wide or grouped into views, such as by computer, or operating system type.

TPC for Data allows you to group information about similar entities (disk, file systems, and so on) from different servers or business units into a Summary Report so that business and technology administrators can manage an enterprise infrastructure. Or, you can summarize information from a specific server. The flexibility and choice of configuration is entirely up to the administrator.

Reporting categories

Major reporting categories for file systems and databases are:

- ▶ Assets Reporting uses the data collected by probes to build a hardware inventory of the storage assets. You can then navigate through a hierarchical view of the assets by drilling down through computers, controllers, disks, file systems, directories, and exports. For databases, information on instances, databases, tables and data files is presented for reporting.
- ▶ Availability Reporting shows responses to ping jobs, as well as computer uptime
- ▶ Capacity Reporting shows how much storage capacity is installed, how much of the installed capacity is being used and how much is available for future growth. Reporting is done by disk and file system, and for databases, by database.
- ▶ Usage Reporting shows the usage and growth of storage consumption, grouped by file system, and computers, individual users, or enterprise-wide.

- ▶ Usage Violation Reporting shows violations to the corporate storage usage policies, as defined through Data Manager. Violations are either of Quota (defining how much storage a user or group of users is allowed) or Constraint (defining which file types, owners and file sizes are allowed on a computer or storage entity). You can define what action should be taken when a violation is detected; for example, SNMP trap, e-mail or running a user-written script.
- ▶ Backup Reporting identifies the most at risk files in your enterprise that are not backed up properly.

20.3.5 Data Manager for Databases

Data Manager for Databases provides a set of policy-driven automated tools for managing storage capacity, availability, events, performance and assets in your relational databases. It can help you identify, evaluate, control and predict the storage needs of Relational Database Management Systems (RDBMSs), which include Oracle, Sybase SQL Server, Microsoft SQL Server, and UDB/DB2.

Data Manager for Databases is policy-based and through its autonomic self-healing capabilities it can detect potential problems and automatically make adjustments based on the policies and actions you have established. For example, it can notify you when your database tables are running out of storage space or warn you of a dropped tablespace. By alerting you to these and other issues related to your stored data, it enables you to prevent unnecessary system and application downtime.

Figure 20-8 shows an example of a Data Manager for Databases report.

Computer	RDBMS Instance	Database	Free Space	Percent Free	Used Space	Total Size	Number Containers
TOTAL =>			0	0%	78,43 MB	78,43 MB	21
tpcsrv.almaden.ibm.com	DB2	IBMCDB	0	0%	20,43 MB	20,43 MB	13
tpcsrv.almaden.ibm.com	DB2	TPCDB	0	0%	58,00 MB	58,00 MB	8

Figure 20-8 Data Manager for Databases DB2 free space report

Note: Data Manager for Databases does not currently support the monitoring of clustered database applications.

20.3.6 Data Manager for Chargeback

Today's IT environments are a significant but essential overhead in the operations of businesses and organizations. Because of this, the costs of an organization's storage resources can quickly get out of control without proper monitoring and management.

TPC for Data provides the ability to produce chargeback information for storage usage. It uses the storage usage information gathered by Data Manager, and Data Manager for Databases to generate invoices that chargeback for storage usage. It enables administrators to provide usage information on a departmental, group, or user level, making the data owners aware of and accountable for their data usage. You can allocate costs by storage usage by user, disk capacity by computer, tablespace, or file system/physical device, shown in Figure 20-9 on page 401. You can create cost centers by creating user, computer, or

tablespace groups, allowing organization to chargeback individuals or departments for their storage usage.

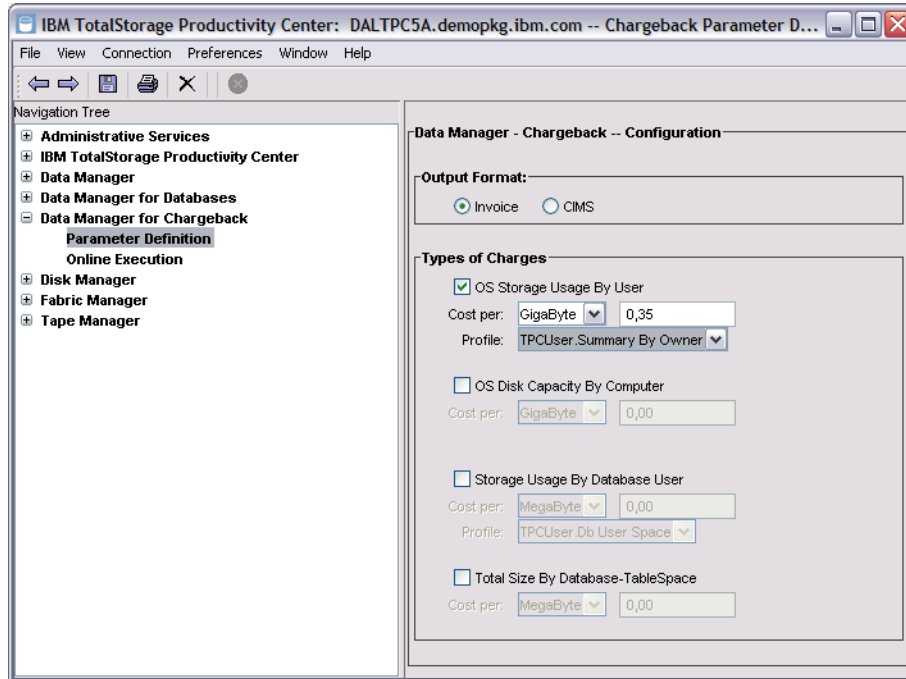


Figure 20-9 Data Manager for Chargeback

In addition to providing invoicing for storage usage, Data Manager for Chargeback also integrates with the chargeback systems already implemented in your environment. It provides you with a higher level, application-specific CIMS output format which can be imported into CIMS applications. With Data Manager for Chargeback, you can export chargeback data for direct import into CIMS, giving you the ability to integrate your storage usage data with other enterprise chargeback information and processes.

For more information about CIMS, refer to the following Web site:

<http://www.cims.com>

Example 20-1 shows a chargeback report based on disk usage by user.

Example 20-1 Chargeback report based on disk usage by user

TPC for Data for Chargeback page 1
 User Usage Invoice 12.06.2006

TPCUser.Default User Group

NAME	SPACE GB	COST 0,35/GB
aaron	20	7,00
adm	1	0,35
Administrator	3	1,05
Administrators	84	29,40
byrne	1	0,35
bin	2	0,70
ckrax	4	1,40

daemon	1	0,35
desktop	1	0,35
greenee	1	0,35
guest	0	0,00
higuera	1	0,35
ident	1	0,35
invscout	1	0,35
ipsec	1	0,35
jkuo	1	0,35
joseph_hill	2	0,70
LOCAL SERVICE	1	0,35
lotti	2	0,70
netdump	1	0,35
newwatch	1	0,35
ntp	1	0,35
nuucp	1	0,35
NETWORK SERVICE	1	0,35
root	8	2,80
roy	2	0,35
rpcuser	1	0,35
rpm	1	0,35
simon	3	1,05
snapp	1	0,35
sshd_server	1	0,35
sys	2	0,35
S-1-5-21-448539723-839522115-725345543-1002	4	1,40
SYSTEM	1	0,35
uucp	1	0,35
vic	1	0,35
watcher	1	0,35
xfs	2	0,70
group total	162	56,70
TPC for Data for Chargeback		page 2
Run Summary		12.06.2006
User Usage Invoice	162 GB	56,70
run total		56,70

20.4 Supported platforms

We list here some of the platforms for agents, servers, repository databases virtualization products, and storage subsystems that are supported by TPC for Data.

Supported Platforms for agents

- ▶ Windows 2000 Server, Advanced Server, DataCenter Server, or Professional with Service Pack 2, or later
- ▶ Windows XP
- ▶ Windows 2003 Server (Standard Edition, Enterprise Edition or DataCenter Edition)
- ▶ HP-UX 11.0, 11i
- ▶ Solaris 8 or 9
- ▶ Red Hat 3.0 on Intel and Power5

- ▶ AIX version 5.1, 5.2 or 5.3
- ▶ Novell NetWare 4.0, or later
- ▶ SUSE Linux Enterprise Server 7, 8 on zSeries
- ▶ SUSE 9 on Intel and Power5 and zSeries
- ▶ United Linux 1 / SUSE Linux Enterprise Server 8

Supported Platforms for servers

- ▶ Windows 2000 Server, Advanced Server, DataCenter Server, or Professional with Service Pack 2, or later
- ▶ Windows XP
- ▶ Windows 2003 Server (Standard Edition, Enterprise Edition or DataCenter Edition)
- ▶ HP-UX 11.0, 11i
- ▶ Solaris 8 or 9
- ▶ Red Hat 3.0 on Intel and Power5
- ▶ AIX version 5.1, 5.2 or 5.3
- ▶ United Linux 1 / SUSE Linux Enterprise Server 8
- ▶ SUSE 9 on Intel and Power5

Repository Database Support

- ▶ Microsoft SQL Server 7.0, or later (no service pack required), SQL2000 (32-bit)
- ▶ Oracle 8i, 9i Rel1, 9i Rel2 and 10G
- ▶ Sybase SQL Server (Adaptive Server Enterprise) 12.5, or later
- ▶ DB2/UDB 7.2, 8.1, or later

Database support (Data Manager for Databases)

- ▶ Microsoft SQL Server 7.0, or later (no service pack required), SQL2000 (32-bit)
- ▶ Oracle 8i, 9i Rel1, 9i Rel2 and 10G
- ▶ Sybase SQL Server (Adaptive Server Enterprise) 12.5, or later
- ▶ DB2/UDB 7.2, 8.1, or later

Subsystem/Virtualization Reporting Support

- ▶ IBM System Storage DS6000, DS8000, DS4000, and ESS
- ▶ IBM TotalStorage SAN Volume Controller
- ▶ EMC Symmetrix
- ▶ HP StorageWorks storage subsystem - see Web site for complete details
- ▶ Hitachi Data Systems storage subsystems - see Web site for complete details

Note: For more details on supported agent platforms, storage devices, and managed databases, refer to these Web sites:

<http://www.ibm.com/support/docview.wss?rs=1133&uid=ssg1S1002814>
<http://www.ibm.com/support/docview.wss?rs=1133&uid=ssg1S1002813>
<http://www.ibm.com/support/docview.wss?rs=1133&uid=ssg1S1002713>

20.5 More information

For more information about TPC for Data, see the following documents and Web sites:

- ▶ *IBM TotalStorage Productivity Center V3.1: The Next Generation*, SG24-7194
- ▶ *Exploring Storage Management Efficiencies and Provisioning - Understanding IBM TotalStorage Productivity Center and IBM TotalStorage Productivity Center with Advanced Provisioning*, SG24-6373

<http://www.ibm.com/servers/storage/software/center/index.html>

<http://www.ibm.com/software/tivoli/products/totalstorage-data/>



IBM TotalStorage Productivity Center for Fabric

This chapter explains the functionality and benefits of the IBM TotalStorage Productivity Center for Fabric. This product can help enterprises to meet their complex systems requirements and storage management challenges involved in designing, implementing, maintaining, and protecting Storage Area Networks (SAN).

IBM TotalStorage Productivity Center for Fabric (TPC for Fabric) provides automatic device discovery and allows multiple SAN views, including physical, logical and zone views. Because administrators can gather configuration, network statistics and other status information from switches, they can view and analyze multiple aspects of the storage environment, including capacity, utilization, assets and availability. Administrators can use this information to determine which systems, for example, are approaching maximum capacity and which systems are near minimums, so that they can perform load balancing as needed. TPC for Fabric monitoring capability is designed to detect storage events and generate the appropriate alerts to the administrator.

21.1 TPC for Fabric - overview

TPC for Fabric helps manage the SAN fabric that connects the host systems and applications to the storage devices. TPC for Fabric is a comprehensive management solution for multivendor SANs. It includes automatic resource and topology discovery, monitoring and alerts, and zone control. TPC for Fabric is an enterprise scalable solution architected to ANSI SAN standards, allowing you to choose the products right for your environment. TPC for Fabric helps:

- ▶ Simplify the task of SAN management and configuration
- ▶ Ensure SAN availability
- ▶ Improve SAN return on investment

TPC for Fabric provides SAN fabric performance and capacity management reporting and monitoring. The ability to do SAN fabric performance and capacity management reporting and monitoring can help customers determine if more bandwidth is needed. Wide-area and local-area IP networks, and storage area networks (SANs) move data from one place to another. The management of the bandwidth is needed to continually monitor for link utilization and errors. Data needs to be gathered to tune resources, balance workloads, and do capacity planning.

With TPC for Fabric, the comprehensive management of multivendor SANs can help simplify the IT infrastructure by providing automatic resource and topology discovery, monitoring, and alerts, and zone control. TPC for Fabric brings all the sources of information about SAN topology and configuration into a single place, and creates topology mapping of the SANs. This topology offers both host-centric and device-centric views that can be displayed graphically or in a hierarchical format. The SAN topology display tracks all topology and configuration changes through both in-band and out-of-band monitoring. Without this type of centralized topology, information from a number of sources, such as Element Management tools, device logs, and SNMP traps would have to be continually monitored and manually correlated to determine the current SAN configuration and topology.

With TPC for Fabric, you can view events and changes happening in your SAN environment. The events are displayed in a color-coded fashion and can be further customized to reflect organizational priorities. TPC for Fabric will forward events signaling topology changes or updates to the IBM Tivoli Enterprise Console®, another SNMP manager, or both. Filtering allows you to control what events get sent to your centralized console. Filtering helps ensure that the administrators get the information they need and are not overwhelmed by extraneous information. TPC for Fabric supports host bus adapters (HBAs), disk subsystems, tape systems, SAN switches, routers and gateways. For a complete list of the supported devices, refer to the following Web site. See the Supported products link under the **Install and use** tab.

<http://www.ibm.com/servers/storage/support/software/tpcfabric>

Additionally, TPC for Fabric provides basic diagnostic capabilities to show which resources are impacted by an availability or performance issue in the SAN. Today, it can monitor performance at the port and switch level, and display this information in the common topology viewer for IBM TotalStorage Productivity Center.

21.1.1 TPC for Fabric - components

Here are the components of TPC for Fabric.

Fabric Manager

The manager performs the following functions:

- ▶ Discovers SAN components and devices.
- ▶ Gathers data from agents on managed hosts, such as descriptions of SANs and host information.
- ▶ Generates SNMP events when a change is detected in the SAN fabric.
- ▶ Forwards events to the Tivoli Enterprise Console or an SNMP console.
- ▶ Monitors switch performance by port and by constraint violations.

Fabric agents on managed hosts

To get the information collected by TPC for Fabric, you must install Fabric agents on the computers you want to monitor. There are two types of Fabric agents: in-band and out-of-band. The in-band Fabric agents use in-band discovery protocols and are supplied with the TPC for Fabric installation package. The out-of-band Fabric agents are SNMP agents which use SNMP queries through a TCP/IP connection to gather topology and attribute information. The out-of-band agents are not installed in the same way as in-band agents. To use out-of-band agents, you enable and configure SNMP at the device level to report to TPC for Fabric. Each agent performs the following functions:

- ▶ Gathers information about the SAN by querying switches and devices for attribute and topology information.
- ▶ Gathers event information detected by host bus adapters (HBAs).

CIMOM

To get fabric performance data, you must enable or install a CIMOM for the switch. For information about enabling or installing CIMOMs for switches, contact your switch vendor.

For information about the devices and switches supported, see the following Web site:

http://www.ibm.com/software/sysmgmt/products/support/IBM_TSANM_Device_Compatibility.html

21.2 SAN management

SAN management automatically discovers SAN components and devices, and the topology of the storage area network environment. You can monitor storage utilization on the SAN and determine the availability of SAN components. The following sections define some of the terms used in SAN Management.

21.2.1 Standards and SAN management tools

For the storage networking community (both vendors and users), standards form the basis for compatibility and interoperability. Standards enable users to pick the solutions they want to implement with the knowledge that today's solution will be interoperable with tomorrow's solution, and that existing hardware investments will be protected as the environments are extended. For vendors, standards give the confidence that a wide market exists for their solutions, and lowers the costs of compatibility testing. Today, there are at least ten organizations involved in creating standards for storage, storage management, SAN management, and interoperability.

Industry organizations, such as the Storage Networking Industry Association (SNIA) and the Fibre Alliance, have taken a leading role in facilitating discussions among vendors and users.

Developments by these organizations are considered de facto standards. Recommendations from these organizations are submitted to the officially recognized standards bodies (IETF, ISO and ANSI) for consideration as a formal standard.

A key standard is contained in the FC-MI (Fibre Channel - Methodologies for Interconnects) technical report published by the ANSI T11 standards committee. However, more recently SNIA finalized the first storage management standard - the Storage Management Initiative Specification (SMI-S) version 1.0. By early 2005, 17 vendors and more than 120 products had passed SMI-S conformance tests. SMI-S is based on the Web Based Enterprise Management (WBEM) architecture and the Common Information Model (CIM), and pioneered by the Distributed Management Task Force (DMTF). It uses an object-oriented approach to describe management information, and the description (data model) is platform- and vendor-independent. CIM profiles have already been developed for many devices, such as Fibre Channel switches, and NAS devices. As this standard is still evolving we cannot expect that all devices will support the native CIM interface, and because of this the SMI-S is introducing CIM agents and CIM object managers (CIMOMs). The agents and object managers bridge proprietary device management to device management models and protocols used by SMI-S. The agent is used for one device and an object manager for a set of devices. This type of operation is also referred to as the proxy model. The CIM Agent or Object Manager will translate a proprietary management interface to the CIM interface. An example of CIM Agent is the IBM CIM agent for DS open API compatibility matrix. See the Web site:

http://www-1.ibm.com/support/docview.wss?rs=1118&context=STC4NKB&dc=DB500&uid=ssg1S1002714&loc=en_US&cs=utf-8&lang=en

For more information about these evolving standards refer to Appendix B, "Introduction to SMI-S and CIM/WBEM" on page 509 and the following Web sites:

<http://www.snia.org>

<http://www.ansi.org>

<http://www.t11.org>

21.2.2 Discovery

The process of finding resources within an enterprise, including detection of network topology, is called *discovery*. TPC for Fabric uses the following two methods to discover the network:

In-band Monitors only the Fibre Channel switches using Fibre Channel protocols. In the case of TPC for Fabric, an agent loaded onto the target server queries a standard HBA API loaded onto the managed host, which then queries reachable devices in the SAN. The information obtained is returned to the Manager.

Out-of-band Monitors the switches, managed hosts, and storage over an IP network via standardized Management Information Base (MIB) files, which typically are loaded only onto the managed switches. TPC for Fabric gathers SNMP-collected information from out-of-band agents.

TPC for Fabric stores the results of inband and out-of-band discoveries through the agents in its database, co-relates it to look for duplication, and uses the information to draw or re-draw the topology map.

The process of discovery is as follows:

- ▶ *Inband discovery*: This is the process of discovering information about the SAN, including topology and attribute data, through the Fibre Channel data paths. Inband discovery uses the following general process:

- The agent sends commands through its Host Bus Adapters (HBA) and the Fibre Channel network to gather information about the switches.
 - The switch returns the information through the Fibre Channel network and the HBA to the Agent.
 - The agent queries the endpoint devices using Request Node Identification (RNID) and SCSI protocols.
 - The agent returns the information to the Manager over the IP network.
 - The manager then responds to the new information by updating the database and redrawing the topology map if necessary.
- ▶ *Out-of-band discovery*: This is the process of discovering SAN information, including topology and device data, without using the Fibre Channel data paths. out-of-band discovery uses SNMP queries, invoked over IP network. Out-of-band management and discovery is normally used to manage devices such as switches and hubs which support SNMP.

21.2.3 Zones and zone control

A Storage Area Network *zone* is a grouping of multiple ports to form a virtual private storage network. A *zone set* is a collection of zones that belong to a single SAN and can be activated or deactivated as a single entity across all switches in the fabric. A zone set can contain one or more zones, and a zone may be a member of more than one zone set. A zone may contain one or more *zone members* (aliases), and an alias may belong to more than one zone. Ports that are members of a group or zone can communicate with each other, but are isolated from ports in other zones. Devices, however, can belong to more than one zone. Using zoning, you can automatically or dynamically arrange fabric-connected devices into logical groups, or zones, across the physical fabric. Support for aliases is also provided. An alias is a name assigned to a device so that the name can be meaningful and familiar. Also, an alias can be a group of devices that can be managed together to make the zoning process easier.

To configure zones, you can use TPC for Fabric's Fabric Manager, or the management application for the devices. Advantages to zoning through Fabric Manager are that with Fabric Manager you can perform zoning tasks from a single interface. Fabric Manager supports industry zone management standards, which means that you can work with devices from multiple vendors using consistent, familiar methods.

TPC for Fabric topology viewer console lets you view zones in SANs. Fabric Manager lets you create, update, and delete zones and zone sets within a SAN. In addition, you can:

- ▶ Assign zones to zone sets.
- ▶ Activate and deactivate zone sets.
- ▶ View zone membership.
- ▶ Add and remove zone members.

21.2.4 Topology views

Within TPC for Fabric, the Topology Viewer provides an extended graphical topology view; a graphical representation of the physical and logical resources (for example, computers, fabrics, and storage subsystems) that have been discovered in the storage environment. In addition, the Topology Viewer depicts the relationships among resources (for example, the disks comprising a particular storage subsystem). Detailed, tabular information (for example, attributes of a disk) is also provided. With all the information that topology viewer provides, you can easily and more quickly monitor and troubleshoot your storage environment.

The Topology Viewer provides a central location to view a storage environment, quickly monitor and troubleshoot problems, and gain access to additional tasks and function within the TPC user interface without users losing their orientation to the environment. A sample Topology Viewer is shown in Figure 21-1:

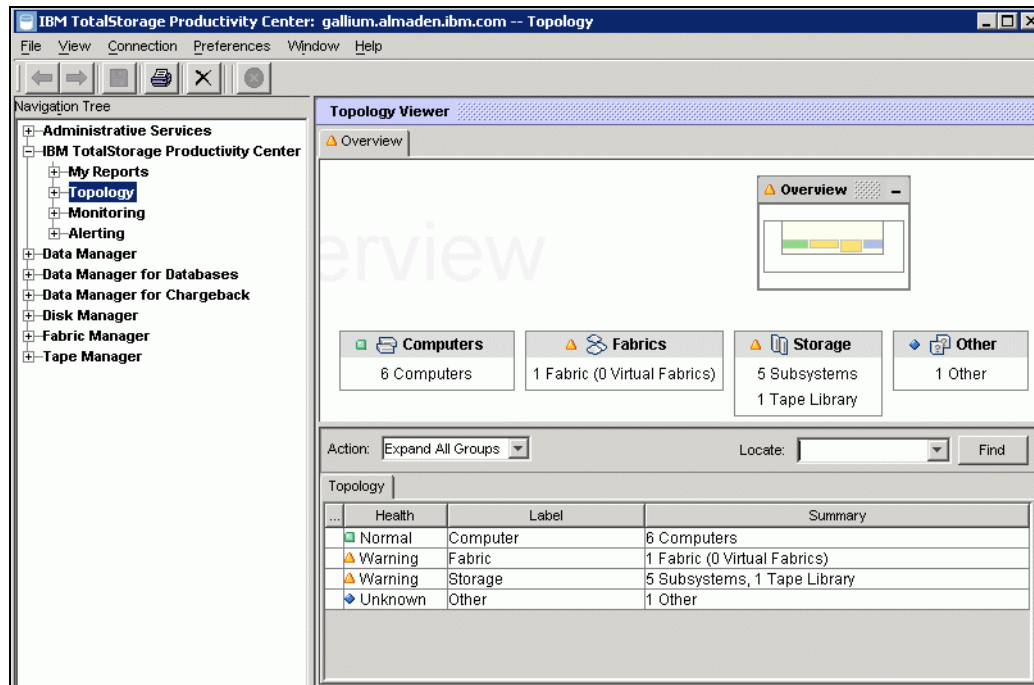


Figure 21-1 Topology Viewer

The Topology Viewer uses the TPC database as the central repository for all data it displays. It actually reads the data in user definable intervals from the database and updates, if necessary, the displayed information automatically.

21.2.5 SNMP events

TPC for Fabric can send events, which represent a change in the state of the fabric, to any event console within the enterprise that TPC for Fabric participates in. These events are generated in both SNMP and Tivoli Enterprise Console format.

The SNMP traps that are sent by the devices and by the Device Server can be sent to the TPC GUI. The events can be displayed in the event browser, or they can be sent to another trap console for monitoring. You then can provide error recovery or other advanced processing, such as paging.

21.3 Functions

In this section, we illustrate the functions provided by TPC for Fabric.

21.3.1 Viewing topology

The primary launch point for the Topology Viewer is the standard navigation area on the left side, i.e the node tree, of the TPC interface, as shown Figure 21-2 on page 411. Here you can see the tabs on the top of the window, the graphical view and the tabular view.

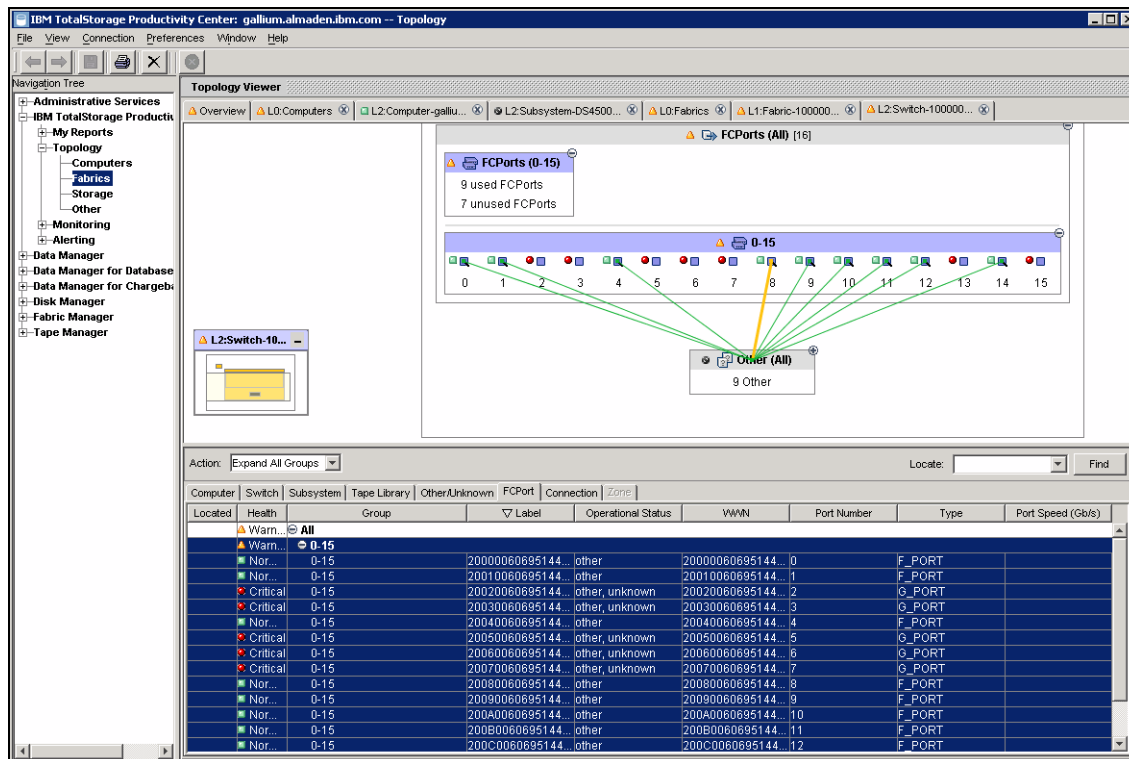


Figure 21-2 Topology Viewer for Fabric

The Topology Viewer provides several views of the environment. Each of the views is split into two synchronized views of the environment. Two subviews are displayed in the Topology Viewers panel. These are:

- ▶ Graphical Topology View - displayed in the upper part of the panel
- ▶ Tabular View - displayed in the lower part of the panel

The Graphical and Tabular subviews are displayed together as one view and are always synchronized - that is, they both display the same entities. Changes in one subview will update the other one and vice versa.

After launching the Topology Viewer, the top level view is always open. If multiple views are open, only one is displayed in the panel. Each view creates its own Tab as shown in Figure 21-3 on page 412. By selecting one of the Tabs you can quickly switch between the single views.

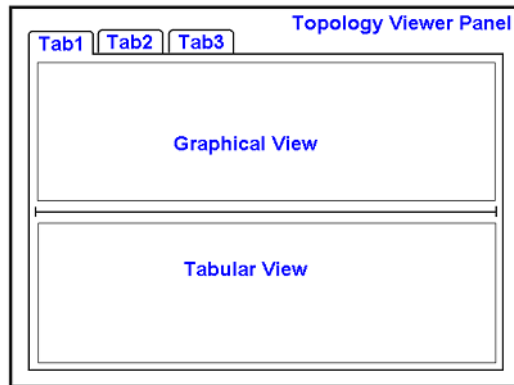


Figure 21-3 Topology Viewer Panel Layout

Graphical View

Essentially the Graphical View displays visual renderings of the environment with icons, boxes, and lines showing the entities and their logical and physical connections spatially.

The *Graphical View* is basically a visual and spatial rendering of the entities and their relationship to each other in the environment. Entities in the environment are rendered as icons with labels. Groups of entities are rendered as boxes with labels. Depending on whether the group is collapsed or expanded, the individual entities in the group are visible, or a summary of the group's content is provided. Logical and physical connections among entities are rendered on demand as lines connecting the relevant entities or groups of entities.

Tabular View

The *Tabular View* displays a text view of the environment, using a set of tables organized in tabs by class of entities. Each table is organized in rows and columns showing different attributes of the entities in the environment.

Mini-map

To further facilitate effective navigation the graphical view contains a mini-map window that represents an abstract overview of the whole topology view thereby providing an environmental and work context for the user. The mini-map indicates the visible portion of the topology, and facilitates user interaction to pan the visible portion. It is shown in Figure 21-4 on page 413:

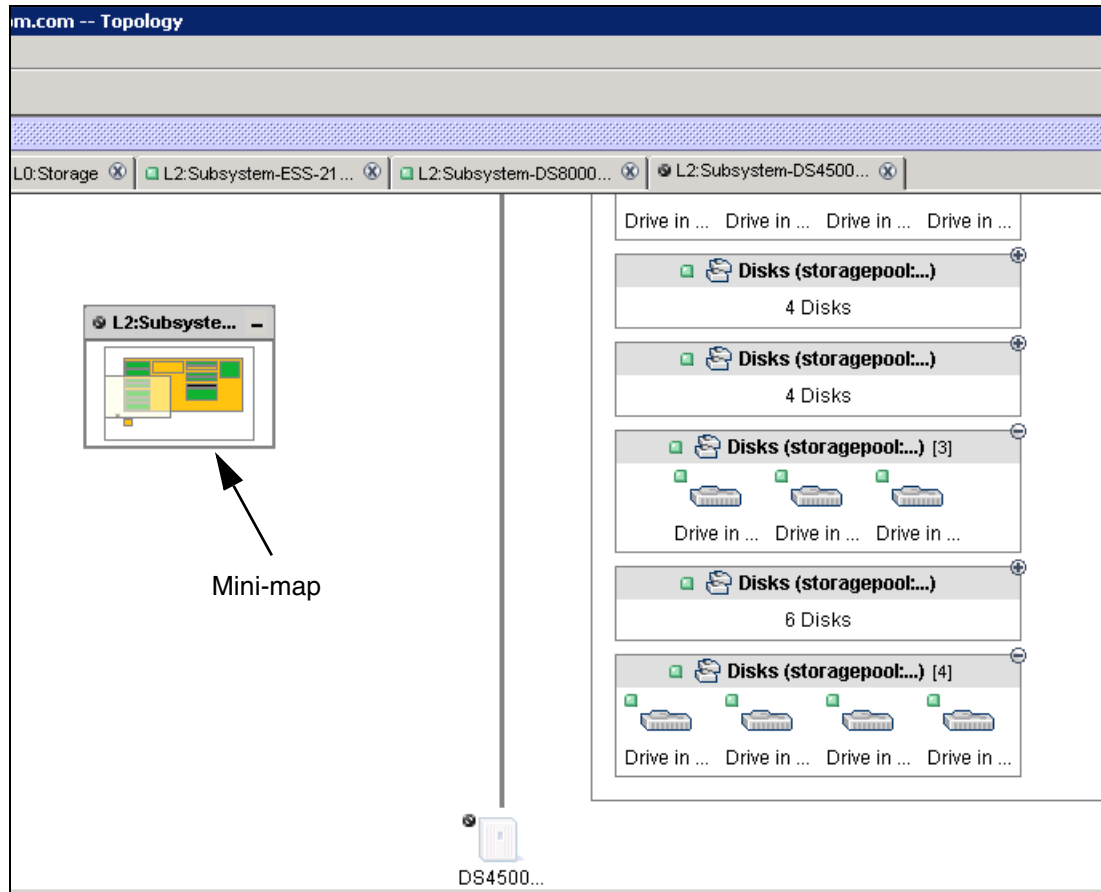


Figure 21-4 Topology Viewer MiniMap

The mini-map window renders the whole topology view in abstract form, reducing the level of detail and ensuring that the complete mini-map is visible at all times. Essentially, the map shows only the top level and second level of detail of the current view and groups are represented as colored boxes based on the aggregated health value of that group.

21.3.2 Semantic Zooming

Graphical topology displays appear simple and manageable when the environment is relatively simple with few numbers of entities and connections. Scaling is an issue in all kinds of complex displays. *Semantic Zooming* is essentially a scaling technique applied to information abstractions, as opposed to graphical features, as in graphical zooming. While graphical zooming changes the scale of the graphical representation of the objects in a view, semantic zooming, changes the level of information abstractions.

Semantic Zooming is used in the Topology Viewer to solve the scaling issue and is key to the design of the Topology Viewer. It is implemented with four distinct zoom levels or levels of abstractions. The four levels are defined as:

1. **Overview, *everything***
A global view which shows a highly aggregated view of the entire environment.
2. **L0 (level 0) *large groups of similar stuff***
A groups view focusing on a class of entities. It shows several groups of entities that correspond to the topology class selected. The entities are Computers, Fabrics, Storage Subsystems and other.

3. **L1 (level 1), related groups of stuff**

A group view which focuses on one group of selected entities and shows their immediate neighbors such as a group of computers.

4. **L2 (level2), one entity**

Detail view, the most detailed view that focuses on a selected entity and its neighbors.

The goal of this graphical view navigation class is to provide the user with Fabric and switch information. A Fabric is a network of entities that are attached through one or more switches. In addition, TotalStorage Productivity Center is to support the display of Virtual SANs (VSANs) and Logical SANs (LSANs). Although SANs, VSANs, and LSANs are implemented and managed differently in the user's environment, they are conceptualized in the same manner via the Topology Viewer. Thus, because of the conceptual similarities they are grouped under the larger category of Fabrics in the Topology Viewer for ease in categorization and visualization.

In general, the semantic zoom levels within this class allow the user to view Fabric components, view relationships and connections between switches and the entities of the Fabrics and view details of switches.

L0: Fabric

This is the default view upon entering the graphical view via selection of Fabric in the navigation menu. The view will render all known Fabrics and Virtual Fabrics as groups (that is, Fabrics being physically networked entities versus being defined by fabric or zone relationships). Each group represents one Fabric depending on what is discovered in the environment, therefore there may be a mix of Fabrics displayed at this level. Shown Figure 21-5:

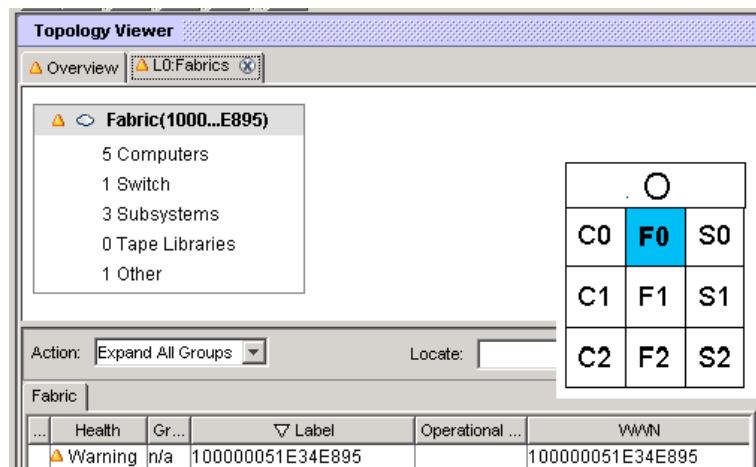


Figure 21-5 Fabric-L0 (F0)

L1: Fabric

At the L1: Fabric level both Fabrics and Virtual Fabrics are represented in the same manner. Computer groups and Other groups are shown on the left of the switches, vertically stacked and left-aligned. By default computer groups are shown collapsed. Storage groups are shown on the right of the switches, vertically stacked and left-aligned. By default storage groups are shown collapsed. Groups of Tape Libraries are stacked below Subsystem groups. Groups of Other type entities are stacked below computer groups.

Connections between the single groups, or if expanded, between the entities can be displayed. Figure 21-6 on page 415 shows the L1 fabric view, and all groups are expanded.

The Other group shows an entity with a user defined name of Helium. At this point no type has been assigned.

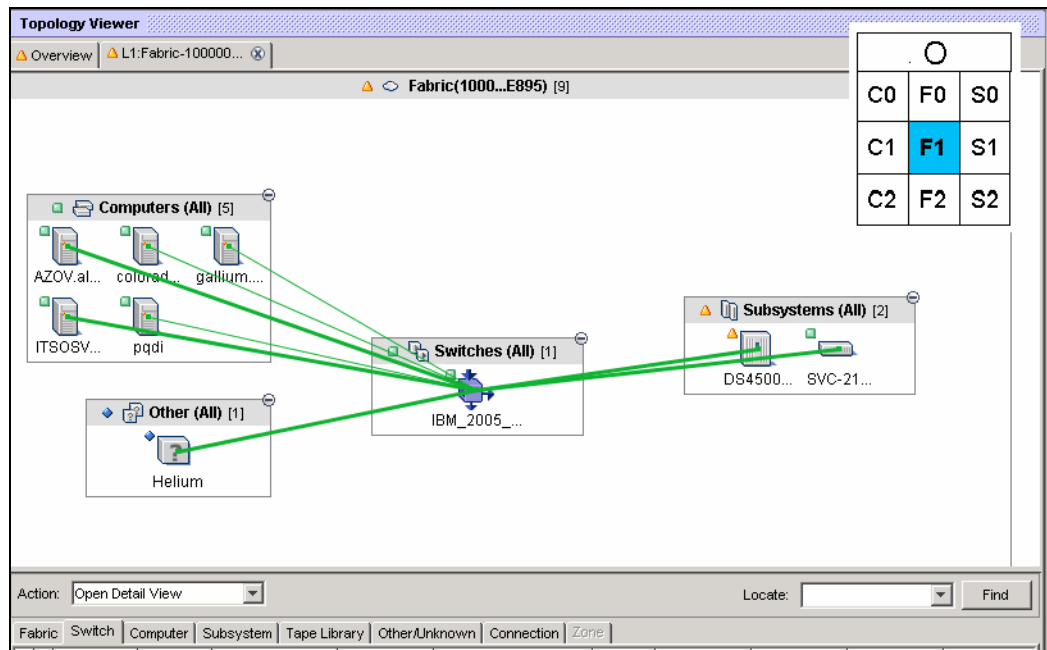


Figure 21-6 Fabric-L1 (F1)

L2: Fabric (Switch)

At this level the focus is on one switch. When the user navigates to the L2 level, the view will render an individual switch, its ports, and all known and related environmental entities.

Entities connected to the switch are shown as groups under the switch. By default these groups are collapsed. The ordering of the groups is determined by entity types with Computers on the left, followed by Switches, followed by Storage entities (first Subsystems and then Tape Libraries) and last by Others. Depending on the selected grouping algorithm, each of these groups might be split into several groups in this view.

Connections between the single groups, or if expanded, between the entities can be displayed. Figure 21-7 on page 416 shows the L2 Fabric view, with an expanded Storage Subsystem group showing the two fabric connections of a DS4500 and the eight connections of a 2-node SVC Cluster to the switch. The Other group shows an entity with a user defined name of Helium, but no type has been assigned so far.

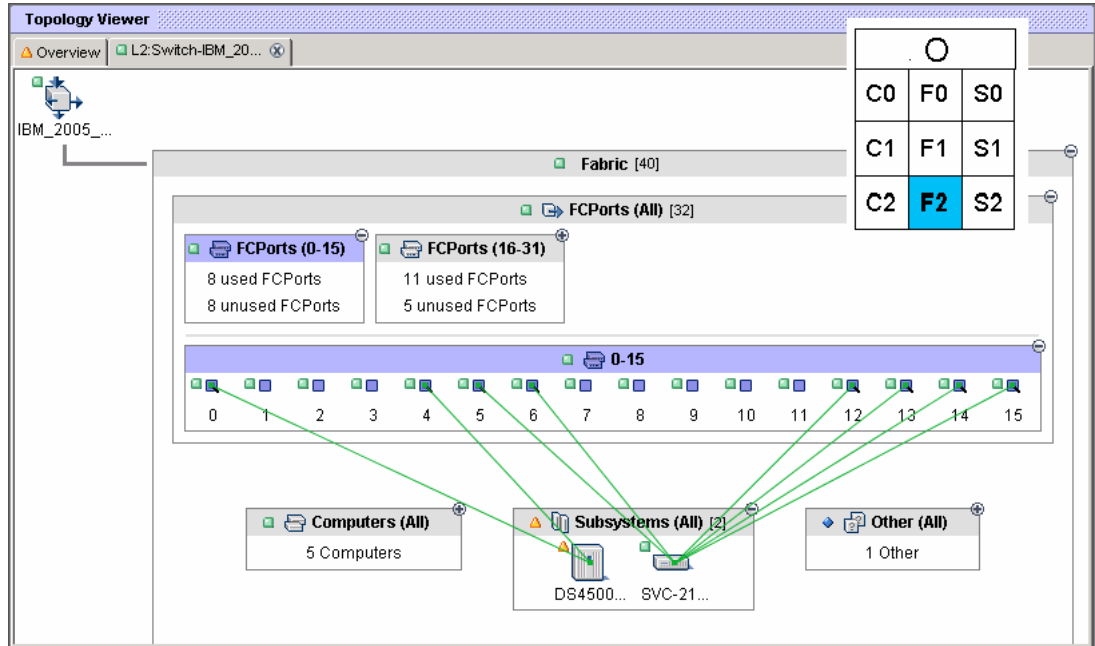


Figure 21-7 Fabric-L2 (F2) Switch

21.3.3 Zone and zonesets

The Topology Viewer also allows users to examine zones and zonesets. Zone information is shown in a Zone tab in the table view. When you select a zone or zoneset in the zone tab, the zone and zoneset members are highlighted in the graphical view.

Figure 21-8 shows how zoning information is displayed in the Topology Viewer tabular view for a server (AZOV) with two HBAs connected to the SVC (ITSOSVC01).

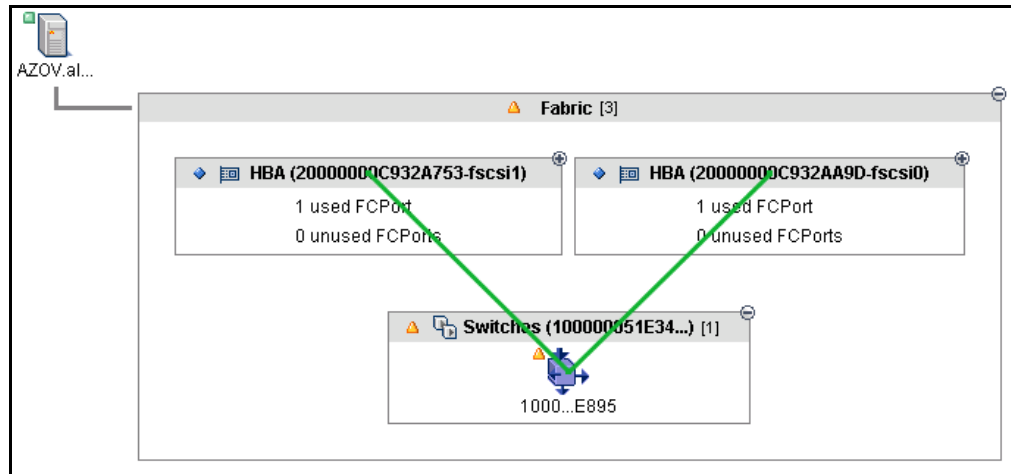


Figure 21-8 Zoning example graphical view

The Zoning Information for this configuration will be shown in the tabular view of the Topology Viewer as shown in Figure 21-9 on page 417.

Computer	Switch	Subsystem	LUN	Volume	FCPort	Connection	Zone
Group	Label	Entity	FCPort				
Fabric: 100000051E34E895							
(Active) Zone Set: B32CFG_0							
(Active) Zone: AZOV_HBA1_ITSOSVC01							
(Active) Zone: AZOV_HBA1_ITSOSVC01		10000000C932AA9D	Computer (AZOV.almaden.ibm.com)		10000000C932AA9D		
(Active) Zone: AZOV_HBA1_ITSOSVC01		500507680140188E	Subsystem (SVC-2145-ITSOSVC01-IBM)		500507680140188E		
(Active) Zone: AZOV_HBA1_ITSOSVC01		5005076801101883	Subsystem (SVC-2145-ITSOSVC01-IBM)		5005076801101883		
(Active) Zone: AZOV_HBA1_ITSOSVC01		5005076801401883	Subsystem (SVC-2145-ITSOSVC01-IBM)		5005076801401883		
(Active) Zone: AZOV_HBA1_ITSOSVC01		500507680110188E	Subsystem (SVC-2145-ITSOSVC01-IBM)		500507680110188E		
(Active) Zone: AZOV_HBA2_ITSOSVC01							
(Active) Zone: AZOV_HBA2_ITSOSVC01		10000000C932A753	Computer (AZOV.almaden.ibm.com)		10000000C932A753		
(Active) Zone: AZOV_HBA2_ITSOSVC01		500507680120188E	Subsystem (SVC-2145-ITSOSVC01-IBM)		500507680120188E		
(Active) Zone: AZOV_HBA2_ITSOSVC01		5005076801201883	Subsystem (SVC-2145-ITSOSVC01-IBM)		5005076801201883		
(Active) Zone: AZOV_HBA2_ITSOSVC01		500507680130188E	Subsystem (SVC-2145-ITSOSVC01-IBM)		500507680130188E		
(Active) Zone: AZOV_HBA2_ITSOSVC01		5005076801301883	Subsystem (SVC-2145-ITSOSVC01-IBM)		5005076801301883		

Figure 21-9 Zoning tabular view

Zone configuration via TPC

In the following sequence of figures we show you some steps, but not full procedures, about how to modify an existing ZoneSet via TPC, so that there will be a new zone included. This is meant to be for an example only.

To start with Zone Configuration, navigate to the Fabric Manager -> Fabrics on the left hand side, as shown Figure 21-10:

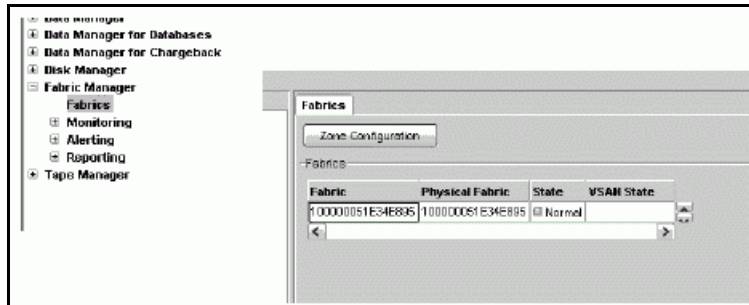


Figure 21-10 Entry panel for Zone Configuration

Click the Zone Configuration button. The current configuration displays, as shown in Figure 21-11 on page 418:

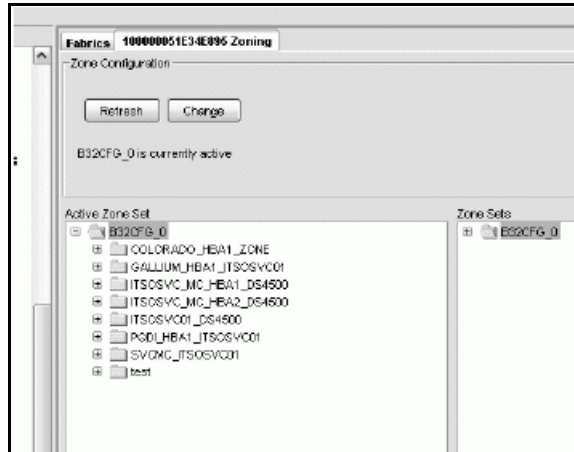


Figure 21-11 Existing zone configuration

Select the Zone you want to change and click **Change**. In this case we selected the zone set, as in Figure 21-12:

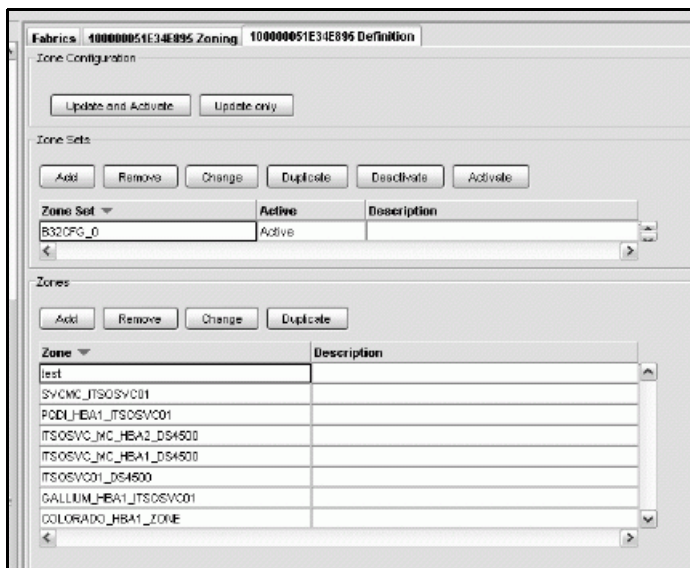


Figure 21-12 The active zone-definition panel

Here you can see a list of the zones contained in the active zoneset. Click **Add** to start the definition of the new zone.

21.3.4 Monitoring the fabric

Groups

A group represents a user-defined set of related objects against which you want to run monitoring jobs. An object represents a storage resource, such as a computer, fabric, storage subsystem, tablespace, and so on. For example, if you want to scan all of the SAN switches in your Payroll department, create a Fabric Group named Payroll and populate that group with the individual SAN switch from the Payroll department. Then, whenever you want to collect information about those switches, you can simply select the Fabric Group named Payroll

when defining a monitoring job. As switches are added or removed from the Payroll department, you can simply update the Payroll Fabric Group to reflect those changes.

Running probes

Run probes to collect statistics on the storage assets in your enterprise, such as computers, disk controllers, hard disks, clusters, fabrics, storage subsystems, LUNs, tape libraries, file systems, etc. Probes can also discover information about new or removed disks and file systems. You can define any number of probes to collect statistics for different storage resources at different times. The results of probes are stored in the repository and are used to supply the data necessary for generating a number of reports.

Performance monitor

You can monitor the performance of those switches in your fabrics from which Fabric Manager can collect performance data and you can generate reports from the results of those monitors. Data is collected on data rates, operations rates, switch port errors, and other performance statistics.

Figure 21-13 shows a sample historical performance report for a number of switches.

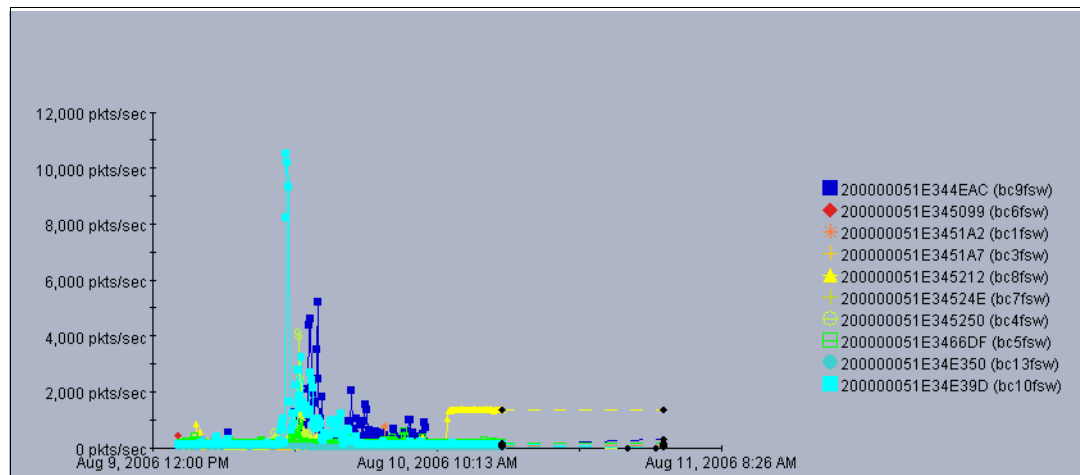


Figure 21-13 Fabric performance monitor

Alerting

The alerting feature lets you define alerts for events that occur in your environment. You can define alerts as part of monitoring jobs, or use the alerting nodes under the Fabric Manager component. For example, you can define an alert that will send you an e-mail when a switch goes offline or comes online. You can also define an alert to be triggered if one of your monitoring jobs fails. There are three type of Fabric Manager alerts:

- ▶ Fabric
- ▶ Switch
- ▶ Endpoint device

Reporting

In addition to performance data collection and performance thresholds, TPC allows the user to display the collected data and the recognized threshold exceptions in a user-friendly format. Data can be displayed in table format, giving the actual values of the various collected metrics, or can be displayed in a graphical format, primarily as line-graphs. Reports can be displayed on recent data, historical data, or on exception conditions. They allow the user to pick a device, a device component, to pick a particular metric, and to specify a time range,

and the performance data for the given device component and the given metrics will be plotted over the given time range. Exception reports allow the user to view a list of any threshold violations in a given time range. They are usually displayed as bar graphs, showing the total number of threshold violations. There are some predefined reports of TPC for Fabric:

- ▶ SAN Switch Report
- ▶ Switch Port Error Report
- ▶ Top 25 Switch Ports Ops Rate Report
- ▶ Top 25 Switch Data Rate Report

21.4 More information

- ▶ For more information about SAN software, check the following Web site:
<http://www.ibm.com/servers/storage/software/>
- ▶ *IBM TotalStorage Productivity Center V3.1: The Next Generation*, SG24-7194
<http://www.redbooks.ibm.com/redpieces/abstracts/sg247194.html?Open>
- ▶ *IBM TotalStorage Productivity Center Planning and Installation Guide*, GC32-1774
<http://publibfp.boulder.ibm.com/epubs/pdf/c3217740.pdf>
- ▶ The TPC Information Center can be found at the following Web site:
<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>
- ▶ Support for TotalStorage Productivity Center for Fabric can be found at:
<http://www-03.ibm.com/servers/storage/support/software/tpcfabric/>



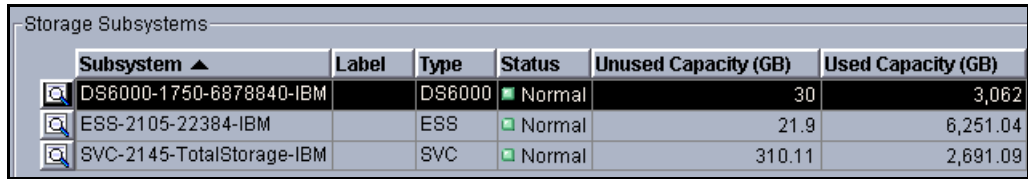
TotalStorage Productivity Center for Disk

This chapter describes the functionality of the TotalStorage Productivity Center for Disk.

TotalStorage Productivity Center for Disk is designed to:

- ▶ Manage heterogeneous storage system from a single console
- ▶ Discover and configure storage systems
- ▶ Monitor and manage performance
- ▶ Configure performance thresholds
- ▶ Generate alerts
- ▶ Store performance statistics
- ▶ Generate reports
- ▶ Volume contention analysis
- ▶ Graphical performance reports
- ▶ Volume performance advisor (ESS Only)
- ▶ Advanced performance management – IBM DS4000/6000/8000/SVC

22.1 Overview



Subsystem ▲	Label	Type	Status	Unused Capacity (GB)	Used Capacity (GB)
DS6000-1750-6878840-IBM		DS6000	Normal	30	3,062
ESS-2105-22384-IBM		ESS	Normal	21.9	6,251.04
SVC-2145-TotalStorage-IBM		SVC	Normal	310.11	2,691.09

Figure 22-1 TotalStorage Productivity Center for Disk (Disk Manager)

As a component of the IBM TotalStorage Productivity Center, TotalStorage Productivity Center for Disk improves you to manage SANs and heterogeneous storage from a single console. Disk Manager lets you manage network storage components based on SMI-S, including:

- ▶ IBM System Storage Disk Subsystems (DS4000, DS6000, and DS8000 Series)
- ▶ IBM System Storage SAN Volume Controller
- ▶ IBM TotalStorage Enterprise Storage Server (ESS)
- ▶ Other storage subsystems that support the SMI-S standards

Through the use of data collection, setting of thresholds and use of performance reports, performance can be monitored for the ESS, DS4000, DS6000, DS8000, SAN Volume Controller, and any other storage subsystem that supports the SMI-S block server performance subprofile. The performance function starts with the data collection task, which captures performance statistics for the devices and stores the data in the database.

You can set thresholds for certain performance metrics depending on the type of device. Threshold checking is performed during data collection, and when performance is outside the specified bounds, alerts can be generated.

After performance data has been collected, you can configure Disk Manager to present graphical or text reports on the historical performance behavior of specified devices, or of specific internal components of those devices. The performance reports provide information about the performance metrics and can display past or current performance data in graphical form.

22.2 TPC for Disk - what's new in the V3.1

TPC for Disk V3.1 is substantially enhanced over previous versions. Here is a summary of what is new:

- ▶ DS8000/DS6000 port statistics
- ▶ DS4000 Performance Management
- ▶ 3rd party storage management
- ▶ Disk Configuration
- ▶ Basic Performance Management
- ▶ New Reporting Capabilities
- ▶ Ability to drill up/drill down
- ▶ Report on multiple storage systems, multiple metrics
- ▶ Table and graphical reports
- ▶ Web based and exportable
- ▶ New Threshold Capabilities
- ▶ Ability to set threshold for low traffic conditions

22.3 TPC support for SAN Volume Controller

TPC for Disk provides the following support for SVC:

- ▶ Asset and Capacity Reporting
 - Physical characteristics such as the manufacturer, model, serial number, capacity, etc
 - Show the allocated and free capacity of every SVC on the network
- ▶ Configuration Reporting and Management
 - Reports on SVC's storage allocated to logical host volumes (which appear to hosts as disk drives) and the managed disks being used on the backside
 - Display the physical managed disks behind what the host sees as a disk drive
 - List all SVC volumes which have been allocated but aren't in use
 - Show which hosts have access to a given SVC volume
 - Show which hosts have access to a given disk drive (within the SVC)
 - Show which SVC volume (and managed disks) a host has access to Discovery, Show the Storage Controllers (ex. ESS/DS4000) that provide volumes to SVC
 - Historical Reporting Historical SVC Subsystem occupancy data will be maintained (e.g. assigned/unassigned space within the SVC).
 - Measure and monitor SVC performance

22.4 Monitoring

Monitoring functions run regularly scheduled or as one-time data collection jobs. These jobs gather detailed statistics about the storage subsystem performance. This data is later used in reports, probes, and so forth. The rules of monitoring are:

- ▶ Threshold violations (exceptions) are always logged and saved in the TPC database
- ▶ The user can view the exceptions that were detected in the TPC alert log, or can generate an exception report. The exception report allows the user to drill-down into the details of individual exceptions.

22.5 Alerting

TPC for Disk allows configured alerts like e-mail, SNMP trap, SMTP, windows Event Log, run a scrip, for all devices managed from TPC. Use the Alerting feature of Disk Manager to set alerts that notify you of storage-related events that occur within your environment. Once you have defined the events or conditions for which you want to be alerted, you can let Disk Manager monitor your storage so that you don't have to. Figure 22-2 on page 424 shows a window with the Alerts configuration menu.

Alert Storage Subsystems

Triggering-Condition

Condition: Disk Utilization Percentage Threshold

Critical Stress: 90

Warning Stress: 75

Suppress alerts when Sequential I/O exceeds 80 %

Triggered-Actions

SNMP Trap

TEC Event

Login Notification Login ID: []

Windows Event Log Event Type: Warning

Run Script [Define...]

Email

Email Recipients: [Add] [Del] [Edit e-mail]

tpcdemo@daltpc5a.demopkg.ibm.com

Figure 22-2 TPC for Disk configuring alerts

22.6 Reporting

TPC for Disk can provide performance monitoring, customization of thresholds based on your storage environment, and generation of events if thresholds are exceeded. In addition, TPC for Disk can help the IT administrators select LUNs for better performance.

The reporting data can be displayed in table format, giving the actual values of the various collected metrics, or can be displayed in a graphical format, primarily as line-graphs. Reports can be displayed on recent data, historical data, or on exception conditions. Exception reports allow the user to view a list of any threshold violations in a given time range. They are usually displayed as bar graphs, showing the total number of threshold violations.

TPC for Disk has defined the following performance reports:

- ▶ Storage Subsystem Performance
- ▶ Storage Subsystem Controller Performance
- ▶ Storage Subsystem Controller Cache Performance
- ▶ I/O Group Performance
- ▶ Array Performance
- ▶ Managed Disk Group Performance
- ▶ Port Performance
- ▶ Top 25 Volumes I/O Rate
- ▶ Top 25 Volumes Data Rate
- ▶ Top 25 Volumes Cache Hit
- ▶ Top 25 Volumes Response Time
- ▶ Top 25 Volumes Disk
- ▶ SAN Switch Report

- ▶ Switch Port Error Report
- ▶ Top 25 Switch Ports Ops Rate Report
- ▶ Top 25 Switch Ports Data Rate Report

However, the user can define his own reports:

- ▶ The By Controller report (user-defined) has all possible metrics available, and the user can select whatever they want. And they can save their customizations.
- ▶ They can make it act like a system-defined report (it's a "My Report") if they want.
- ▶ "Customizing" here means: selecting which components are in the report, which columns/metrics to include, what order to show them in. Reports can also be scheduled to run in a 'batch' on a schedule and generate, say, HTML files that can be viewed later. Figure 22-3 on page 426 has a window from TPC disk by array report.

Performance Management capabilities

- ▶ "By Volume", "By Array", "By Port", "By Subsystem" Reports
- ▶ By Volume, they are viewing a (textual) report of volume-level performance statistics across multiple volumes
- ▶ A set of default thresholds are in effect for all devices of a particular device type. In addition, thresholds can be user-defined for individual devices, overriding the default thresholds for those individual devices
- ▶ Thresholds can be enabled or disabled, and the upper (stress) and lower (idle) boundaries of the threshold can be set
- ▶ Performance data collection must be active for thresholds to be monitored
- ▶ All thresholds defined by PM are bi-level thresholds, meaning that two different boundary conditions can be defined, one to indicate a warning condition, and another to indicate a critical condition
- ▶ Each threshold can allow a total of 4 boundary values to be defined, the critical stress, warning stress, warning idle, and critical idle boundaries, if so desired.

Configuring performance thresholds

You can use the TotalStorage Productivity Center for Disk to set performance thresholds for each device type. Setting thresholds for certain criteria enables TotalStorage Productivity Center for Disk to notify you when a certain threshold has been exceeded, so that you can take action before a critical event occurs.

- ▶ You can specify what action should be taken when a threshold-exceeded condition occurs. The action may be to log the occurrence or to trigger an event. The threshold settings can vary by individual device. Figure 22-3 on page 426 shows an alert which was triggered based on disk utilization.

Alert Storage Subsystems

Triggering-Condition

Condition: Disk Utilization Percentage Threshold

Critical Stress: 90

Warning Stress: 75

Suppress alerts when Sequential I/O exceeds 80 %

Triggered-Actions

SNMP Trap

TEC Event

Login Notification Login ID: []

Windows Event Log Event Type: Warning

Run Script [Define...]

Email

Email Recipients: [Add] [Del] [Edit e-mail]

tpcdemo@daltpc5a.demopkg.ibm.com

Figure 22-3 TotalStorage Productivity Center for Disk Reporting by Array

You can also do general storage configuration reporting. Figure 22-4 an extract of a report showing storage volume assignments.

Volume HBA Assignment: By Capacity
Number of Rows: 173

Storage Subsystem	Volume Name	Volume Capacity	Volume WWN	HBA Port WWN	SMIS Host Alias
TOTAL =>		7.28 TB			
SVC-2145-TotalStorage-IBM	AIX161_SPACE_1	2.00 TB	Unknown	2100001125934954	AIX161-2100001125934954
SVC-2145-TotalStorage-IBM	AIX161_SPACE_2	597.00 GB	Unknown	2100001125934954	AIX161-2100001125934954
ESS-2105-22384-IBM	1702	364.05 GB	Unknown	1000000C944DBB9	DEMOAIX2
ESS-2105-22384-IBM	160F	93.13 GB	Unknown	2100000D60D3852E	BC78LOT10
ESS-2105-22384-IBM	160E	93.13 GB	Unknown	2100000D60D386BA	KQBG2K
ESS-2105-22384-IBM	1417	27.94 GB	Unknown	210000112592409C	BC10SLOT10
ESS-2105-22384-IBM	1418	27.94 GB	Unknown	2100001125923FE2	BC10SLOT11
ESS-2105-22384-IBM	1419	27.94 GB	Unknown	2100000D60D386E4	BC10SLOT12
ESS-2105-22384-IBM	141A	27.94 GB	Unknown	2100000D60D386EC	BC10SLOT13
ESS-2105-22384-IBM	1410	27.94 GB	Unknown	2100000D60D385EC	BC10SLOT3
ESS-2105-22384-IBM	1412	27.94 GB	Unknown	210000112593A894	BC10SLOT5
ESS-2105-22384-IBM	1413	27.94 GB	Unknown	210000112593A7FE	BC10SLOT6
ESS-2105-22384-IBM	1414	27.94 GB	Unknown	210000112593A7F4	BC10SLOT7
ESS-2105-22384-IBM	1416	27.94 GB	Unknown	2100001125923FFA	BC10SLOT9
ESS-2105-22384-IBM	1500	27.94 GB	Unknown	210000112593A7CA	BC11SLOT1

Figure 22-4 TPC for Disk Volume HBA assignment report

22.7 More information

For more information about TotalStorage Productivity Center for Disk check this Web site:

<http://www.ibm.com/servers/storage/software/center/disk/index.html>

See these Redbooks for more details about TotalStorage Productivity Center for Disk:

- ▶ *IBM TotalStorage Productivity Center V3.1: The Next Generation*, SG24-7194

- ▶ *Exploring Storage Management Efficiencies and Provisioning - Understanding IBM TotalStorage Productivity Center and IBM TotalStorage Productivity Center with Advanced Provisioning, SG24-6373*



TotalStorage Productivity Center for Replication

In this chapter we describe the functionality as well as advantages of TotalStorage Productivity Center for Replication and TotalStorage Productivity Center for Replication Two Site Business Continuity.

23.1 TPC for Replication overview

The IBM TotalStorage Productivity Center (TPC) for Replication helps administrators manage the Advanced Copy services provided by many IBM storage systems. It is available in two complementary packages:

- ▶ TPC for Replication
- ▶ TPC for Replication Two Site Business Continuity

The architecture of TPC for Replication is shown in Figure 23-1.

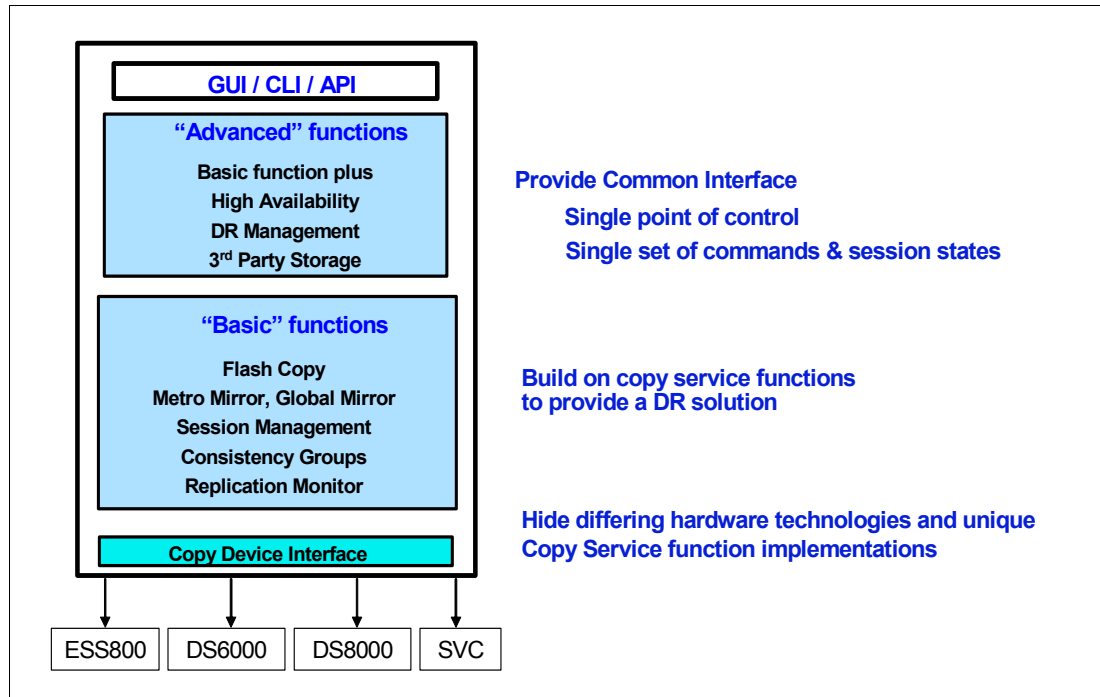


Figure 23-1 TPC for Replication architecture

The basic functions of TPC for Replication provide management of FlashCopy, Metro Mirror, and Global Mirror capabilities for the IBM ESS Model 800, DS6000, and DS8000. It also manages FlashCopy and MetroMirror for IBM SAN Volume Controller.

TPC for Replication simplifies management of advanced copy services by:

- ▶ Automating administration and configuration of these services with wizard-based session and copy set definitions.
- ▶ Providing simple operational control of copy services tasks, including starting, suspending and resuming.
- ▶ Offering tools for monitoring and managing copy sessions.

In addition to the basic functions of TPC for Replication, TPC for Replication Two Site Business Continuity (BC) provides disaster recovery management through planned and unplanned failover and failback automation for the IBM ESS Model 800, IBM DS6000, and IBM DS8000. TPC for Replication is a prerequisite for TPC for Replication Two Site BC.

TPC for Replication Two Site BC helps you manage replication to a remote backup site through Metro Mirror or Global Mirror. The software allows you to monitor the progress of the

copy services so you can verify the amount of replication that has been done as well as the amount of time needed to complete the replication.

Automated failover can keep critical data online and available even if your primary site fails. When the primary site comes back online, the software manages failback to the default configuration as well.

TPC for Replication Two Site BC offers also a high availability capability, so it manages replication even if the main TPC for Replication server experiences a failure. With a second server operating as an active standby, services can switch quickly to the backup server to maintain copy services operations if the primary server goes off-line.

23.2 Consistency management

To maintain integrity of data in any operating system, the sequence in which updates are being written is crucial. If that sequence is changed, data corruption will occur. The correct sequence must be maintained within a volume, across volumes, and across multiple storage systems.

In remote disk mirroring environments, the order of dependant writes across volumes and across multiple storage systems must be maintained at the remote location. During a real disaster (fire, flooding, earthquake), all the components in your environment will almost certainly not fail simultaneously. Failures will more likely be intermittent and gradual, and the disaster time frame will take many seconds or even minutes. This is known as a rolling disaster.

Advanced Copy Services integrated with automation software such as TPC for Replication can protect data from being a mirror of a so-called “dying scenario” - that is, where multiple failures are occurring during a rolling disaster.

TPC for Replication FREEZES affected sessions at a known point instead of mirroring literally hundreds of time-offset failures in a short amount of time, as shown in Figure 23-2.

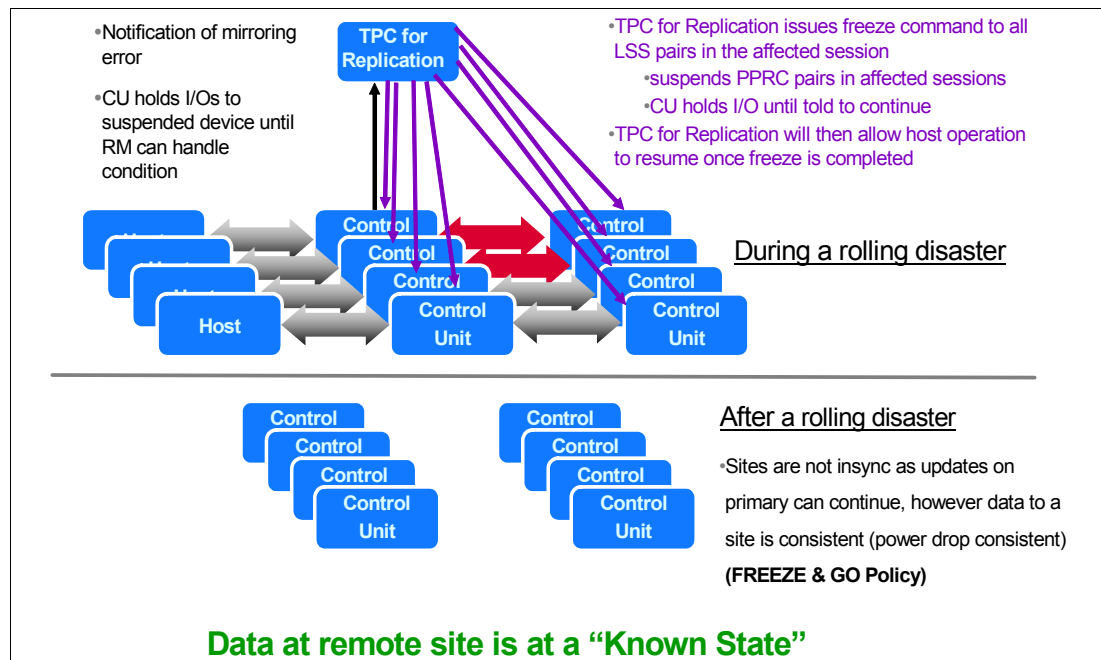


Figure 23-2 TPC for Replication freeze command

The heartbeat functionality on managed LSSes (logical subsystems) is used to avoid losing the alerting and freeze functionality of isolated subsystems (e.g. when the TCP/IP connection is broken), shown in Figure 23-3.

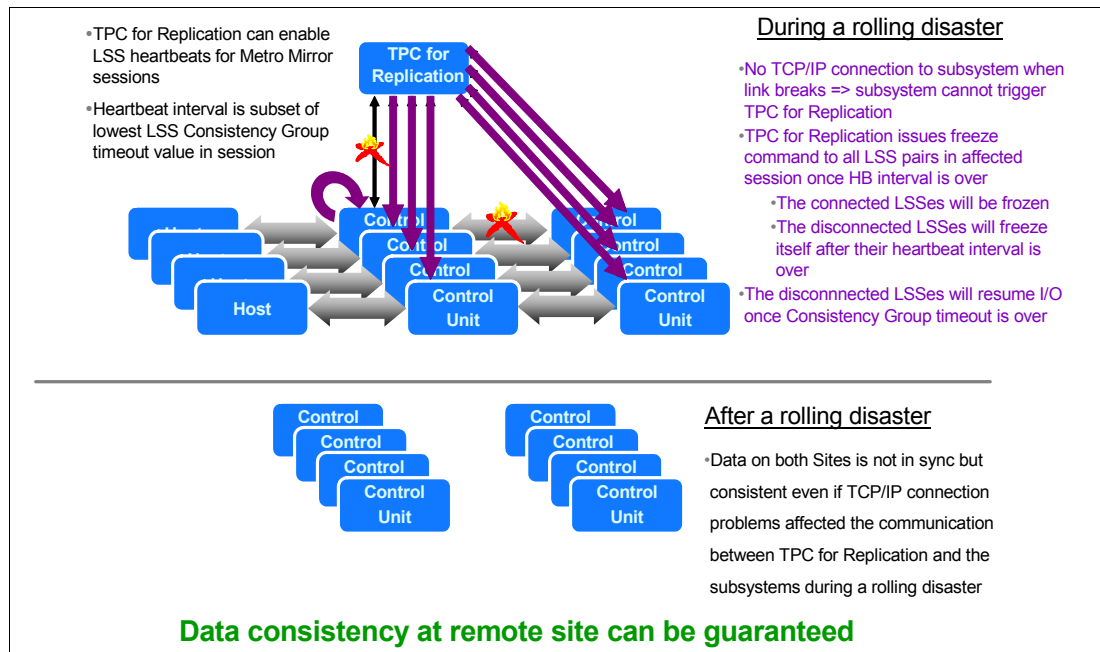


Figure 23-3 TPC for Replication heartbeat functionality

23.3 Replication management

TPC for Replication uses a simple graphical user interface (GUI) as well as a command line interface (CLI) to configure automation, manage ongoing activities and monitor progress of all key tasks. Different types of storage systems can be managed by a single integrated tool.

The functions of the GUI are:

- ▶ Manage the entire Copy Services environment
 - Add, modify and remove subsystems
 - Add and remove Paths
 - Add, modify and remove Sessions
 - Add and remove Copy Sets in Sessions
 - Manage Sessions with simple commands depending on Session type and state
- ▶ Monitor the entire Copy Services environment
 - Health Check screen
 - All sessions overview screen
 - Individual session views with Copy sets overview and details
 - Progress Bars
- ▶ Diagnostics
 - Error overview
 - Detail view of each error
 - Error messaging and history console

The GUI runs within a Web browser, such as Mozilla Firefox 1.5 or greater or Internet Explorer® 6.0 or greater. All communications between the server, clients, and hardware components are secured using SSL.

To start the GUI, perform the following steps:

1. Launch the Web browser.
2. Type the URL for the GUI (<https://servername:9443/CSM>) and press Enter.
3. Type your user name and password and click Login.
Optionally, select a standby server hyperlink at sign-on (if applicable).

23.3.1 Health Overview

The first panel that loads after logging into the GUI is the Health Overview, shown in Figure 23-4. It displays a top-level view of TPC for Replication Status, with a summary of Session status (Number with errors, warnings, and number normal), a summary of connection to Storage Subsystems, and a summary of Management Servers (High Availability) status. A miniature version of the Health Overview is always visible on the lower left-hand side of window.

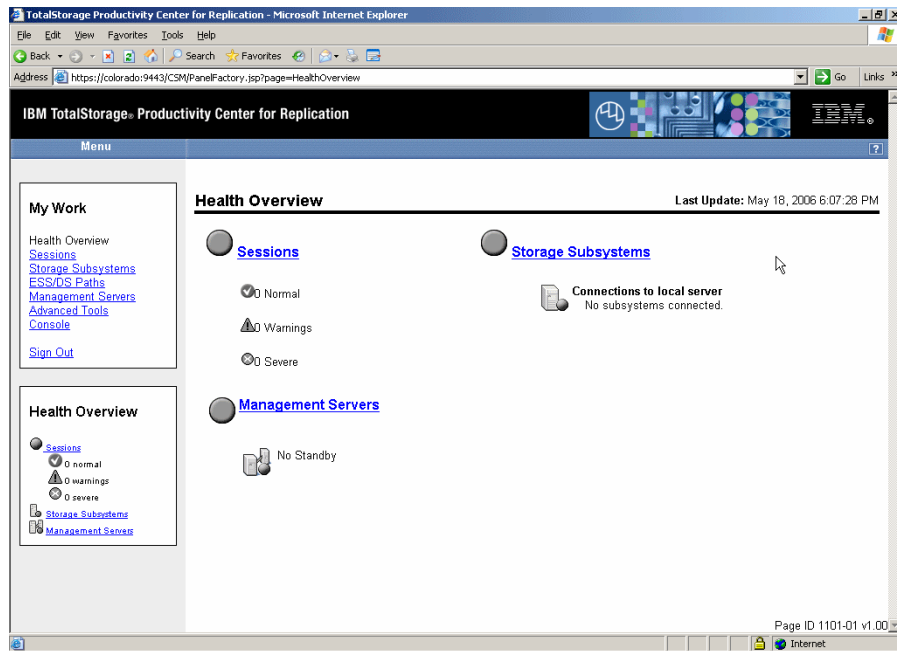


Figure 23-4 TPC for Replication Health Overview

23.3.2 Storage subsystems

ESS/DS series or SVC storage subsystems can be easily added, modified or deleted by using the Storage Subsystem panel of the GUI.

Figure 23-5 on page 434 shows an example for adding a storage subsystem.

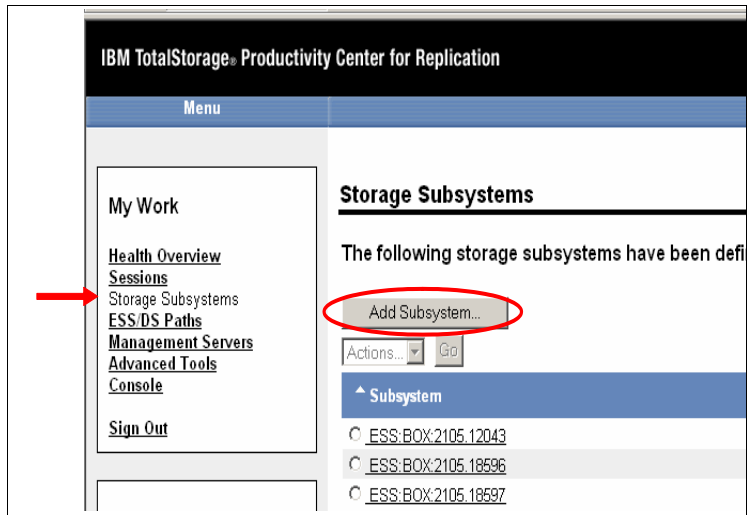


Figure 23-5 TPC for Replication adding storage subsystem

To add an ESS/DS series storage subsystem, enter the IP addresses or domain names, Ports, User names, and Password.

To add an SVC storage subsystems, enter the Common Information Model object manager (CIMOM) server IP or domain name, CIMOM communication port, CIMOM username, and CIMOM password.

23.3.3 Logical paths

ESS/DS logical paths can be added, removed, and reestablished by clicking the associated hyperlink in the ESS/DS Paths panel, shown in Figure 23-6.

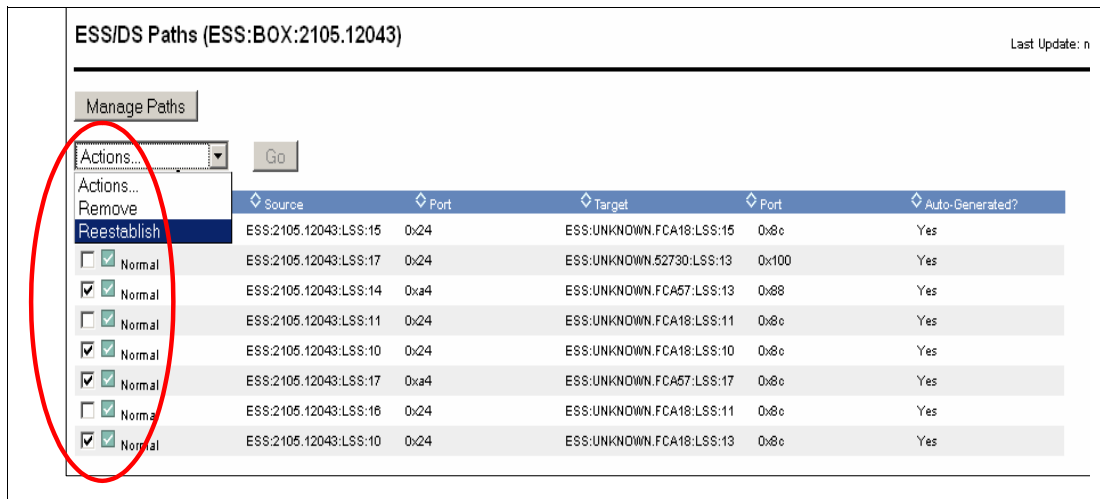


Figure 23-6 TPC for Replication configuring logical paths

SVC paths can be configured by using the SVC GUI.

23.3.4 Session and copy sets

A copy set is a set of volumes that contain copies of the same data, shown in Figure 23-7 on page 435. All the volumes in a copy set are the same format (count key data (CKD) or fixed

block) and size. In a replication session, the number of volumes in a copy set and the role that each volume in the copy set plays are determined by the copy type.

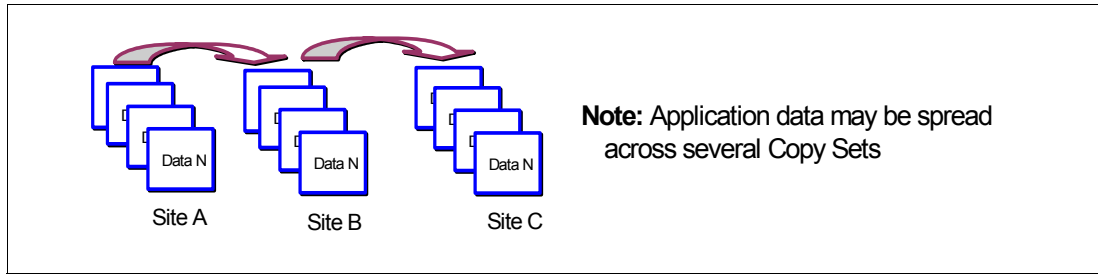


Figure 23-7 TPC for Replication copy set

A session is a container of multiple copy sets making up a consistency group, shown in Figure 23-8.

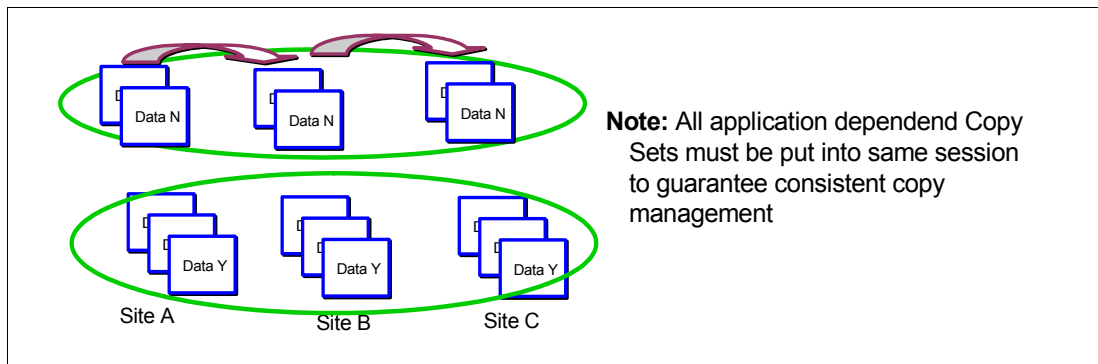


Figure 23-8 TPC for Replication sessions

Three different types of sessions can be created, corresponding to the type of the copy service required:

- ▶ **FlashCopy** - A FlashCopy is a point-in-time local copy. FlashCopy replication creates the target volume with a copy of the data that represents the data at the time the relationship was established. Both source and target can be referenced and updated independently.
- ▶ **Metro Mirror** - A Metro Mirror session is synchronous remote replication. In this context, remote means the source is in one physical subsystem and the target is in a different physical subsystem. Synchronous means that when a write is issued to change the source, the change is propagated to the target before the write is completely posted. This method of replication *maintains* identical data in both the source and target (unlike FlashCopy where the copy is static). The advantage of Metro Mirror is that when a disaster occurs, there is no data loss at the recovery site because both writes must complete before signaling completion of a write to the source application.
- ▶ **Global Mirror** - A Global Mirror session is continuous asynchronous remote replication. As with Metro Mirror, remote means the source is in one physical subsystem and the target is in a different subsystem. Asynchronous means that when a write is issued to the source copy, the change is not synchronously propagated and applied to the target. The updates to the target occur after control is given back to the application. Asynchronous replication greatly reduced (and may eliminate) performance impact of the copy on the application and is preferred for copying data at distances of greater than 200 km. However, because the write is not applied synchronously, there is a chance for data loss in the case of a disaster.

Once you have decided which method of replication you want to use (copy type), you set up the copy sets. In a replication session, the number of volumes in a copy set and the role that each volume in the copy set plays is determined by the copy type.

The following terms are used in describing sessions:

- ▶ **Host volume** - A host volume is a volume that an application, such as a database, reads data from and writes data to. A host volume can be the source for a copy function. A host volume can also be the target of the copy function, in which case you cannot write to the volume directly - it is only written by the copy service.
- ▶ **Journal volume** - A journal volume holds a consistent copy of the data until a new consistent copy is formed. The journal volume restores the last consistent point during a recovery.
- ▶ **Role** - A volume's role is the function it assumes in the copy set, and is composed of the intended use and, for Global Mirror and Metro Mirror, the volume's site location. Every volume in a copy set is assigned a role. A role can assume the functions of a host volume, journal volume, or target volume.
- ▶ **Role pair** - A role pair is the association of two roles in a session. For example, in a Metro Mirror session, the role pair could be the association between volume roles Host 1 and Host 2.
- ▶ **Target volume** - A target volume is an intermediate volume that receives data from a source. Depending on the session type, that data might or might not be consistent.

Copy sets as well as sessions can be created by using TPC for Replication wizards. An example for an Add Copy Sets wizard is shown in Figure 23-9.

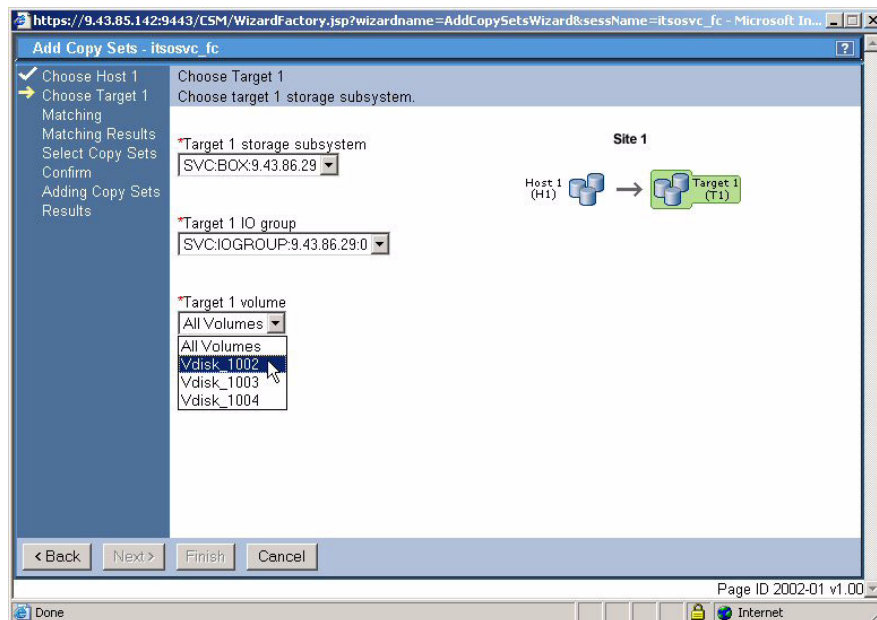


Figure 23-9 TPC for Replication Add Copy Sets wizard

23.4 More information

For more information about TotalStorage Productivity Center for Replication, refer to:

- <http://www.ibm.com/servers/storage/software/center/replication/index.html>
- <http://www.ibm.com/software/tivoli/products/totalstorage-replication/>



IBM Tivoli Storage Manager

IBM Tivoli Storage Manager is a storage management application built for the enterprise. Tivoli Storage Manager provides an enterprise solution for data protection, disaster recovery, space management, and record retention. Tivoli Storage Manager facilitates flexible and scalable storage management policies to support complicated business needs for storage management and disaster recovery. Most importantly, Tivoli Storage Manager automates storage management tasks by eliminating labor and cost intensive manual procedures for backup, archive, disaster recovery planning and bare machine recovery. Introduction to IBM Tivoli Storage Manager

IBM Tivoli Storage Manager is the core product of the IBM Tivoli Storage Management product set. IBM Tivoli Storage Manager can help protect your organization's data from hardware failures and other errors by storing backup and archive copies of data on offline storage. Scalable to protect thousands of computers running a dozen operating system platforms, its intelligent data movement and store techniques, and complete policy-based automation, work together to reduce administration costs while increasing data protection and service levels.

24.1 IBM Tivoli Storage Manager family of products

IBM Tivoli Storage Manager is available as either a basic or extended version. There is also IBM Tivoli Storage Manager Express which is an entry level data protection solution.

- ▶ IBM Tivoli Storage Manager (base) provides:
 - Data backup and restore
 - Managed data archiving and retrieval
 - Optional 24x7 critical-application protection modules
 - Optional Hierarchical Space Management (HSM) and SAN extensions
 - Optional bare machine recovery extensions
 - Small tape library management
- ▶ IBM Tivoli Storage Manager Extended Edition provides:
 - Data backup and restore
 - Managed data archiving and retrieval
 - Disaster-preparation planning and recovery
 - Automatic Disaster preparation and recovery scripts
 - NDMP backups of network-attached storage appliances
 - Optional 24x7 critical-application protection modules
 - Optional HSM and SAN extensions
 - Optional bare machine recovery extensions
 - Small and large tape library management
 - Tape library sharing between multiple data protection servers
 - Expanded administration to handle thousands of client computers
 - Data protection for NAS appliances through NDMP
- ▶ IBM Tivoli Storage Manager Express provides:
 - fast installation - around 30 minutes to your first backup
 - automatic deployment of client software
 - Web browser based interfaces for easy administration
 - pre-configured best practices for policies and automated schedules
 - intended for small to medium business with between 5 to 20 computers to back up
 - protection for Windows clients - including options Microsoft SQL Server and Microsoft Exchange

24.1.1 IBM Tivoli Storage Manager core functionality

The core functions of Tivoli Storage Manager provide centralized data backups and restores, managed archiving and retrieval, data protection for business critical applications, disaster recovery, and storage management functionality for the enterprise.

- ▶ **Centralized data backup and restore:** Data backups store a copy of your active online data on offline storage to protect against the operational loss or destruction of file or application data. If an online storage device fails, a data error occurs or someone accidentally deletes a file, the offline copy of that data can be restored quickly to online storage. Tivoli Storage Manager uses multiple techniques to make data backups and

restores fast and flexible while helping reduce the impact of data protection on computers, networks and storage.

- ▶ **Managed data archiving and retrieval:** Data archives are copies of active or inactive data from online storage that are stored offline for a specified amount of time. Files can remain on the local storage media or can be deleted. Especially useful for regulatory or bookkeeping purposes, archives provide point-in-time copies of important data. Managed data archives implemented by Tivoli Storage Manager are managed for a defined lifetime through the policy-based automation engine. The Tivoli Storage Manager client can quickly and easily retrieve the archived data, giving you a secure, easy way to store important data for long periods.
- ▶ **Data protection for business-critical applications:** IBM Tivoli Storage Manager provides virtually 24x7 online data protection for business-critical applications. Some business-critical application programs must remain operational continuously. Many of these around the clock applications have built-in capabilities that control external data protection applications and allow data backups and restores to occur while the application continues operating.

IBM Tivoli Storage Manager offers separate add-on software modules to interface with these applications, including:

- IBM Tivoli Storage Manager for Application Servers
- IBM Tivoli Storage Manager for Databases
- IBM Tivoli Storage Manager for Enterprise Resource Planning
- IBM Tivoli Storage Manager for Copy Services
- IBM Tivoli Storage Manager for Advanced Copy Services
- IBM Tivoli Storage Manager for Mail
- IBM Tivoli Storage Manager for System Backup and Recovery

Tivoli Storage Manager Data Protection for Business-critical Applications modules are separate software products which connect business applications to the Tivoli Storage Manager environment. These applications allow application specific storage management controls for backup and restore operations.

Oracle, Lotus Notes, Lotus Domino, Microsoft Exchange, Microsoft SQL Server, SAP, and WebSphere Application Server are some of the applications that have their own storage management interface or disaster protection application which integrates with the Tivoli Storage Manager data management API in each Tivoli Storage Manager data protection application.

In addition, DB2, Parallel Edition (DB2PE), DB2 Universal Database (UDB), and Informix from IBM integrate the Tivoli Storage Manager API directly, without requiring a separately purchased disaster protection product. Some of the Tivoli Storage Manager data protection applications leverage IBM advanced FlashCopy functions on the SAN Volume Controller, DS6000, and DS8000. This functionality bridges Tivoli Storage Manager and high-availability storage infrastructures to maximize application availability.

IBM Tivoli Storage Manager for Application Servers - works with the WebSphere Application Server software to provide an applet GUI to do reproducible, automated online backup of a WebSphere Application Server environment, including the WebSphere administration database (DB2 Universal Database), configuration data, and deployed application program files. Changes to the WebSphere environment, such as the addition of applications, are automatically detected and included in the data backup schedule.

IBM Tivoli Storage Manager for Databases - protects a wide range of application data via the protection of the underlying databases management systems holding that data. It exploits the backup-certified utilities and interfaces provided for Oracle and Microsoft SQL Server. In conjunction with Tivoli Storage Manager, this module automates data protection

tasks and allows database servers to continue running their primary applications while they backup and restore data to and from offline storage.

IBM Tivoli Storage Manager for Enterprise Resource Planning (ERP) - helps better protect the infrastructure and application data and improve the availability of mySAP servers. It provides automated data protection, reduces the CPU performance impact of data backups and restores on the mySAP server, and greatly reduces the administrator workload necessary to meet data protection requirements.

IBM Tivoli Storage Manager for Copy Services - helps protect Microsoft Exchange databases which require 24 x 7 availability. It integrates VSS based snapshot capabilities with IBM Tivoli Storage Manager and its data protection component for Microsoft Exchange, to provide enhanced backup and recovery features that are integrated with existing Exchange backup and restore capabilities. VSS snapshot support provides for very fast backup with minimal impact on the production Exchange server with a variety of VSS snapshot providers. When used together with the SVC VSS provider, IBM Tivoli Storage Manager for Copy Services provides “near instant” restore of Exchange storage groups through FlashCopy from a VSS shadow copy image to the production volumes.

IBM Tivoli Storage Manager for Advanced Copy Services - improves the data protection of business-critical databases and ERP applications that require 24x365 availability. Support is provided for backup of IBM DB2 UDB, Oracle, and mySAP databases. This software module helps IBM Tivoli Storage Manager and its other data protection modules to perform high-efficiency data backups and archives of your most business-critical applications while eliminating nearly all performance impact on database or ERP servers. This elimination of server performance impact is accomplished by coupling the FlashCopy function of IBM ESS, SAN Volume Controller, DS6000, and DS8000 with Tivoli Storage Manager and its database protection capabilities. The Tivoli Storage Manager for Advanced Copy Services module works with and requires either Tivoli Storage Manager or Tivoli Storage Manager Enterprise Edition and the corresponding Data Protection module - IBM Tivoli Storage Manager for Databases or IBM Tivoli Storage Manager for ERP- to provide a “near zero-impact” data backup and “near instant” recovery solution.

IBM Tivoli Storage Manager for Mail - is a software module for IBM Tivoli Storage Manager that automates the data protection of e-mail servers running either Lotus Domino or Microsoft Exchange. This module utilizes the application program interfaces (APIs) provided by e-mail application vendors to perform online *hot* backups without shutting down the e-mail server and improve data-restore performance.

- ▶ **Data Retention:** *IBM System Storage Archive Manager* (formerly IBM Tivoli Storage Manager for Data Retention) -Facilitates compliance with the most stringent regulatory requirements in the most flexible and function-rich manner. It helps manage and simplify the retrieval of the ever-increasing amount of data that organizations must retain for strict records retention regulations. Many of the regulations demand the archiving of records, e-mails, design documents and other data for many years, in addition to requiring that the data is not changed or deleted
- ▶ **Bare Machine Recovery:** Expedites your ability to restore a computer by reducing the complexity and time needed to restore a computer's operating system, modifications, patches and network options. Tivoli has various options to assist with different operating system's recovery. *IBM Tivoli Storage Manager for System Backup and Recovery* empowers you with a flexible backup method for your AIX systems. It helps to protect your data and to provide bare machine recovery capabilities. It offers a comprehensive system backup, restore and reinstallation tool.
- ▶ **Continuous Data Protection:** *IBM Tivoli Continuous Data Protection for Files* provides real-time protection for files on file servers and other systems such as notebooks, which are only sometimes connected to the network. Instead of waiting for a scheduled interval, IBM Tivoli Continuous Data Protection for Files backs up selected important files each

time they are saved, in real time, effortlessly and transparently in the background, in some cases eliminating the backup window. The module is lightweight, to minimize the impact on the protected systems. High-priority files can be backed up to up to three targets - local disk, a file server or Network Attached Storage (NAS) appliance, and an IBM Tivoli Storage Manager server. These will capture every save of a file when it occurs to help protect against corruption, file loss or system loss. For more information, see:

<http://www.ibm.com/software/tivoli/products/continuous-data-protection>

- ▶ **Hierarchical Space Management:** provides the automatic and transparent movement of operational data from the user system disk space to a central storage repository. If the user needs to access this data, it is dynamically and transparently restored to the client storage. To the client, the files always appear to be online for easy retrieval if required. This optimizes storage utilization by freeing up unneeded space on the clients.

IBM Tivoli Storage Manager for Space Management provides HSM function for UNIX systems — AIX, Linux, Solaris, and HP-UX. Its core functions include:

- Hierarchical Storage Manager
Automates the movement of seldom used files to and from near line storage.
- Pre-Migration Tools
Makes storage space available fast and can allow for scheduled data transfer over the network.
- Coordination with Backups
Storage Manager can clone an existing version of a backup to create the HSM version creating less network traffic.
- Adjustable Stub File Size
Enables you to decide the stub file size eliminating recall of entire file for programs that browse first part of file only.

IBM Tivoli Storage Manager HSM for Windows provides HSM function for Windows systems and provides similar HSM functions to IBM Tivoli Storage Manager for Space Management, with an intuitive GUI. Files can be either explicitly or automatically migrated, based on file name, as well as on dates of creation, last access and modification.

- ▶ **LAN-free data transfer:** *IBM Tivoli Storage Manager for Storage Area Networks* — Increases data transfer bandwidth and lessens the impact on the local area network (LAN). Computers connected to a SAN can take advantage of Tivoli Storage Manager to move backup and archive data over the SAN instead of the LAN.

24.1.2 IBM Tivoli Storage Manager key attributes

Many attributes of Tivoli Storage Manager make it efficient, scalable and automatic. Some of these include centralized management, broad hardware support, intelligent data movement, intelligent data storage, and policy automation.

Centralized, comprehensive management

A client/server application, Tivoli Storage Manager allows one server to handle hundreds or even thousands of clients. The administration control of Tivoli Storage Manager is extremely flexible and allows management by a hierarchy of administrators, each designated with different management authority levels.

- ▶ Multiple administrators can manage Tivoli Storage Manager simultaneously or one administrator can manage multiple Tivoli Storage Manager servers from one location simultaneously.

- ▶ Multiple Tivoli Storage Manager servers can be administered from a single server which is running Integrated Service Console with Administration Center installed.
- ▶ Data restore requests can be initiated and controlled from the Tivoli Storage Manager client, helping reduce administrator workload.
- ▶ Flexible client backup and administrator command schedules can automate many procedures.

The flexibility of Tivoli Storage Manager can simplify data protection tasks for IT administrators.

Broad hardware support

Tivoli Storage Manager software supports a wide range of hardware:

Platforms — More than a dozen different operating systems are supported as data protection clients. Eight different operating systems are supported as data protection systems. Platforms vary from laptops to mainframes, and a server can work with a client — cross-platform. Figure 24-1 shows many of the client and server platforms supported by Tivoli Storage Manager.

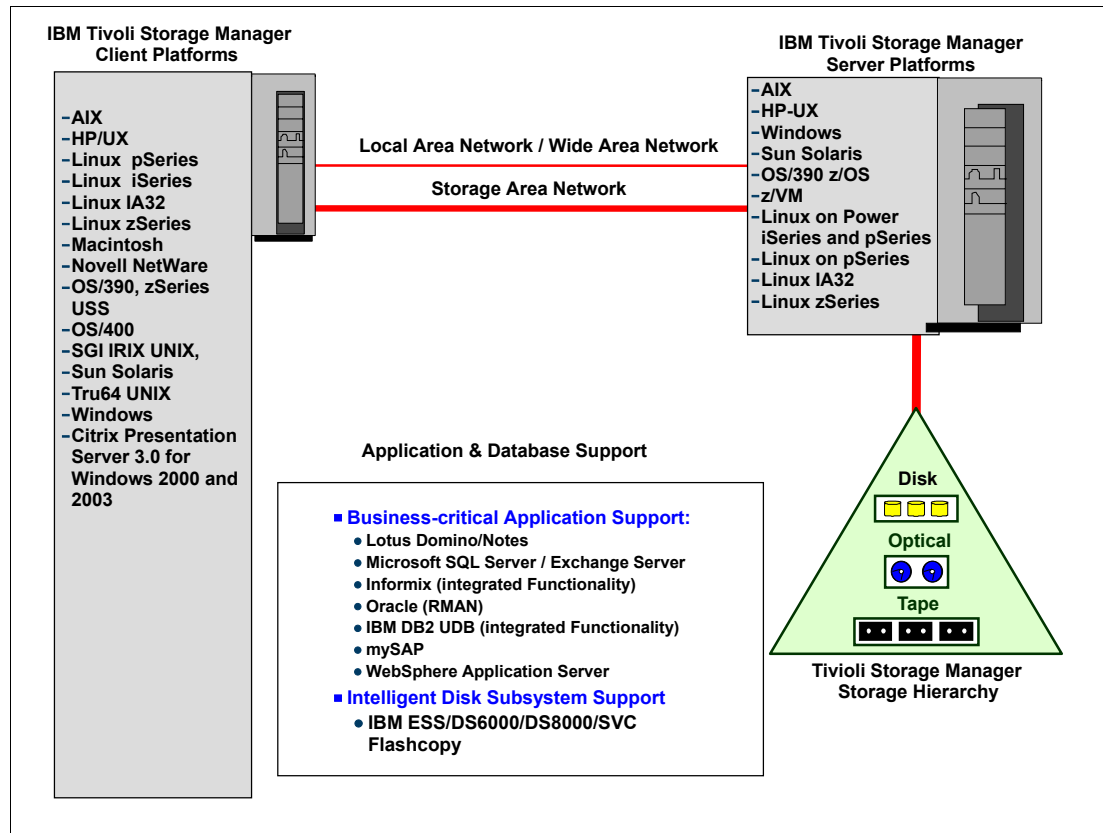


Figure 24-1 Tivoli Storage Manager Extended Edition client, server, application, and database support

Attention: Refer to this link for a complete list of all currently supported platforms:

<http://www.ibm.com/software/tivoli/products/storage-mgr-extended/platforms.html>

Networks — Many different types of networks can connect data protection clients to their servers and connect both to their storage. Remote and mobile users can use dialup lines, the

Internet and wide area networks (WANs) just as transparently as local users can use LANs or SANs for data movement.

Offline storage devices — Tivoli Storage Manager supports hundreds of magnetic disk, optical and tape storage devices. Tape libraries with more than three tape drives and 40 slots are supported by Tivoli Storage Manager Extended Edition.

Intelligent data movement

Tivoli Storage Manager uses various techniques that can significantly reduce data transfer sizes and total time required for data backups and restores.

Disk-to-disk data movements Tivoli Storage Manager was designed around the concept of offline “storage pools” that use disk storage in addition to optical and tape storage devices. Tivoli Storage Manager’s flexible storage hierarchy allows backup directly to disk or to tape and migration from one type of media to another. Disk-to-disk data transfers can save considerable time for each individual client, reducing backup windows across an organization. After an administrator-specified threshold is reached, the data in the disk storage pool automatically migrates to other less-expensive offline storage devices such as optical disk or tape. Because many requests for data restores happen within a few hours or days of a backup, the disk storage pool also provides rapid disk-to-disk data restores.

Multiple parallel session data backups and restores Multiple sessions from a single client expedite the backup of data. Likewise, restores to a single client are speeded by using multiple streams of data from various tapes.

Progressive backups Tivoli Storage Manager limits the bandwidth needed for backups and restores with its progressive incremental backup, which can decrease the amount of data that travels over the network by copying only files that are new or have changed since the last backup. This can help avoid the high-bandwidth use of weekly full backups and daily incremental backups. When a data restore is required, Tivoli Storage Manager transfers only the necessary version rather than the full backup and associated incrementally. This reduces the data movement load on networks and servers.

Adaptive subfile differencing Through adaptive subfile differencing, remote and mobile users with limited bandwidth connections to a Tivoli Storage Manager server over a dialup phone line, the Internet or WANs can make regular backups quickly and transparently. This technique takes progressive backups a step further by calculating an efficient level of backup file, block or byte. It then transfers the individual files, blocks or bytes of data that changed since the last backup, making the transfer over the network as efficient as possible.

Checkpoint restart It can frustrate remote and mobile workers to lose a network connection and have to start a data transfer over again. Tivoli Storage Manager recalls the point at which the data movement was disconnected, allowing the backup or data restore to continue from the point of interruption.

Data encryption If data security over a network connection between a client and a Tivoli Storage Manager server is a concern, Tivoli Storage Manager can encrypt the data.

Cyclical redundancy check To further protect data integrity, Tivoli Storage Manager can perform an optional automatic cyclical redundancy check on data movements.

Nondisruptive online image backup and restore — For data protection instances where speed is more important than network bandwidth or individual file restore, Tivoli Storage Manager incorporates comprehensive image backups and restores that can transfer entire disk volume images. For LAN- or SAN-connected clients with plenty of bandwidth to the Tivoli Storage Manager server and disk storage pool, this process can be a fast method of data protection.

Intelligent data storage

Relational database — An integrated relational database is at the heart of the Tivoli Storage Manager server. Borrowing the methods of “systems managed storage” from IBM mainframes, Tivoli Storage Manager performs lifetime management for data that has been backed up or archived. Down to the file level, this data is cataloged, assigned expiration and migration attributes and linked to the original data files it came from, as well as its predecessors and siblings on offline storage. The power of the Tivoli Storage Manager relational database becomes apparent with its intelligent data movement, intelligent data storage and flexible policy-based automation.

Collocation — The relational database can keep the data of a client or file system grouped in logical sets on individual tapes to help improve the performance of restores by reducing the number of tape mounts and positionings that occur.

Note: Collocation by group is now supported. Groups of nodes can be defined, and the server can then collocate data based on these groups. Collocation by group can yield the following benefits:

- ▶ Reduce unused tape capacity by allowing more collocated data on individual tapes.
- ▶ Minimize mounts of target volumes.
- ▶ Minimize database scanning and reduce tape passes for sequential-to-sequential transfer.

Tape reclamation — The relational database can automatically delete expired files on tapes and consolidate the remaining unexpired files to other tape volumes. The resulting empty tapes can be reused for new data, potentially offering substantial annual savings on tape media.

Operational reporting — Provides Tivoli Storage Manager server(s) status reports sent directly to the administrator. These customizable reports are sent daily and give the administrator a simple place to go to check the status of backups and of the Tivoli Storage Manager server. Additional monitoring reports and missed backup reports can also be sent. The information can be sent via emails, desktop alerts, pages or exported to a Web site.

Policy-based automation

Powered by the relational database inside Tivoli Storage Manager, the granular policy engine can automate capabilities down to the individual file level. Policies control such things as how many backup versions of a file to keep, where to direct the initial backup (disk/tape etc), what to do when a file is deleted, and how to handle files which change during backup. Policies can be set as simple defaults or highly tuned and customized to help keep costs low, while still meeting specific service-level agreements. The end result is a “set it and forget it” experience. Once configured, Tivoli Storage Manager can manage data protection for thousands of computers with little administrator assistance.

Benefits

The attributes of Tivoli Storage Manager combine with the broad data protection provided in a single integrated application — backup/restore, archive/retrieve, HSM, SAN operation, business-critical application protection and bare machine recovery. It can help reduce the:

- ▶ Amount of offline storage space required for a given amount of data
- ▶ Network bandwidth required for data movements
- ▶ Impact of data movements on the LAN
- ▶ Time required for data movements
- ▶ CPU impact on application servers

- ▶ IT manpower resources required to protect virtually any given amount of data

24.2 How Tivoli Storage Manager works

Tivoli Storage Manager is implemented as a client-server software application. The Tivoli Storage Manager server software component coordinates the movement of data from Tivoli Storage Manager backup-archive clients across the network or SAN to a centrally managed storage hierarchy. The classic Tivoli Storage Manager hierarchy includes disk, tape, and in some cases optical devices for data storage. Figure 24-2 shows the general movement of data in a Tivoli Storage Manager environment.

Tivoli Storage Manager client data is moved via the SAN directly to the disk or tape primary storage pool or via the LAN to the Tivoli Storage Manager server and then written directly to disk or tape primary storage pool (and optionally simultaneously to a copy storage pool in Tivoli Storage Manager 5.1 and later), migrated to other storage primary storage pools, and copied as many times as necessary to additional copy storage pools.

Tivoli Storage Manager manages and stores all data about policies, operations, locations of data, and Tivoli Storage Manager component definitions. It uses an internal relational database as the repository for all its data. This component of Tivoli Storage Manager server architecture makes Tivoli Storage Manager extremely scalable for large implementations. Although the Tivoli Storage Manager database requires minimal database administration, an understanding of the backup and archive concept in Tivoli Storage Manager terms is useful.

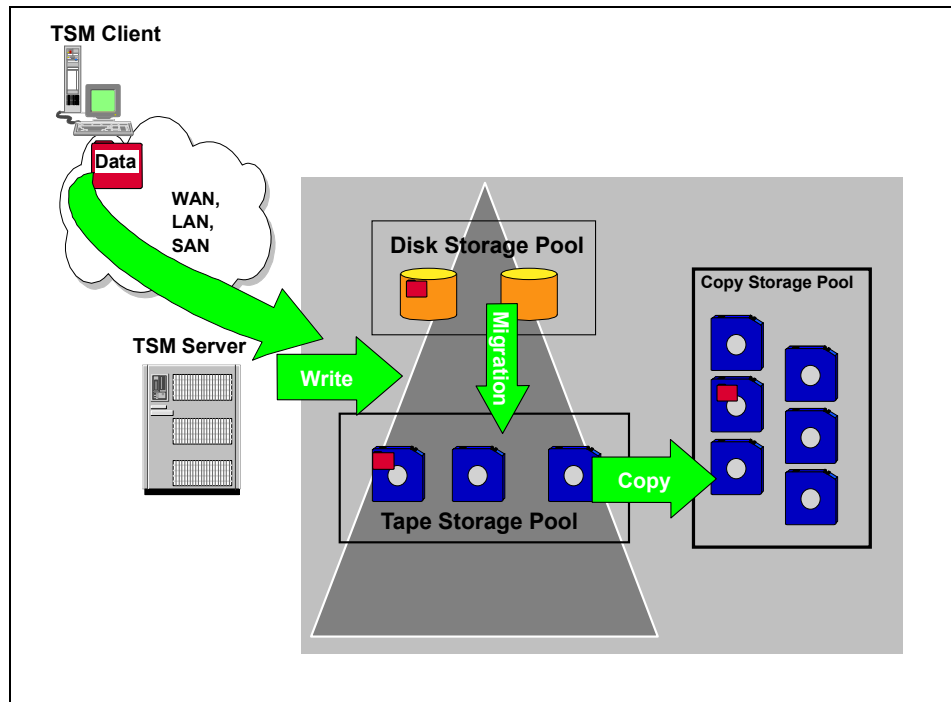


Figure 24-2 Data movement with Tivoli Storage Manager and the storage hierarchy

24.2.1 Tivoli Storage Manager backup and archive concepts

Backup, in Tivoli Storage Manager terms, means creating a copy of a data object to be used for recovery. A Tivoli Storage Manager data object can be a file, a part of a file, a volume

image, a directory or a user-defined data object like a database table. The backup version of this data object is stored separately in the Tivoli Storage Manager server storage hierarchy. Tivoli Storage Manager policy tools allow great flexibility for the way data is managed for each client. Backup frequency, retention, and copy policies are easily implemented on the Tivoli Storage Manager client.

In addition to data backup, *archive* copies of data can also be created using Tivoli Storage Manager. Archive creates an additional copy of data and stores it for a specific amount of time, known as the retention period. Tivoli Storage Manager archives are not expired until the retention period is past, even if the original files are deleted from the client system.

Therefore, the difference between *backup* and *archive* is that backup creates and controls multiple backup versions, whereas archive creates an additional file that is retained for a specific period of time.

24.2.2 Progressive incremental backups

One of the key differentiators between Tivoli Storage Manager and other data protection products is the *progressive incremental backup methodology*. Tivoli Storage manager only backs up new or changed files. It tracks all of the backups at a file level. It has no concept of a full backup with dependent incrementals or differentials. Because of Tivoli Storage Manager's powerful relational database, it does not require periodic full backups. Incremental backup by date is also available. This methodology reduces network and storage resource consumption and lowers the overall cost of storage management. Tivoli Storage Manager's file level progressive backup methodology is far superior to other traditional backup methods such as Full+Incremental or Full+Differential, because progressive incremental backups are never redundant

24.2.3 Tivoli Storage Manager Administration Center

The Administration Center is a Web-based interface that can be used to centrally configure and manage IBM Tivoli Storage Manager Version 5.3 servers. This new interface replaces the administrative Web interface found in the previous versions of Tivoli Storage Manager. The Administration Center is installed as an IBM Integrated Solutions Console component.

What is the Integrated Solutions Console?

IBM Tivoli Storage Manager's new user interface consists of a number of different components which will assist the administrator in managing multiple IBM Tivoli Storage Manager Servers within a single, integrated console.

Integrated Solutions Console Infrastructure

The ISC builds on top of the Web Sphere Application Server and Web Sphere Portal base and includes lightweight versions of both in the ISC runtimes. It looks for common problems, actions, and subtasks across the range of ISC components in order to provide reusable services. Basing the ISC on a lightweight portal infrastructure provides the ability to aggregate independent tasks and information into a single organized presentation.

Figure 24-3 on page 447 is an example of how the Welcome page looks after installing ISC and the Administration Center.

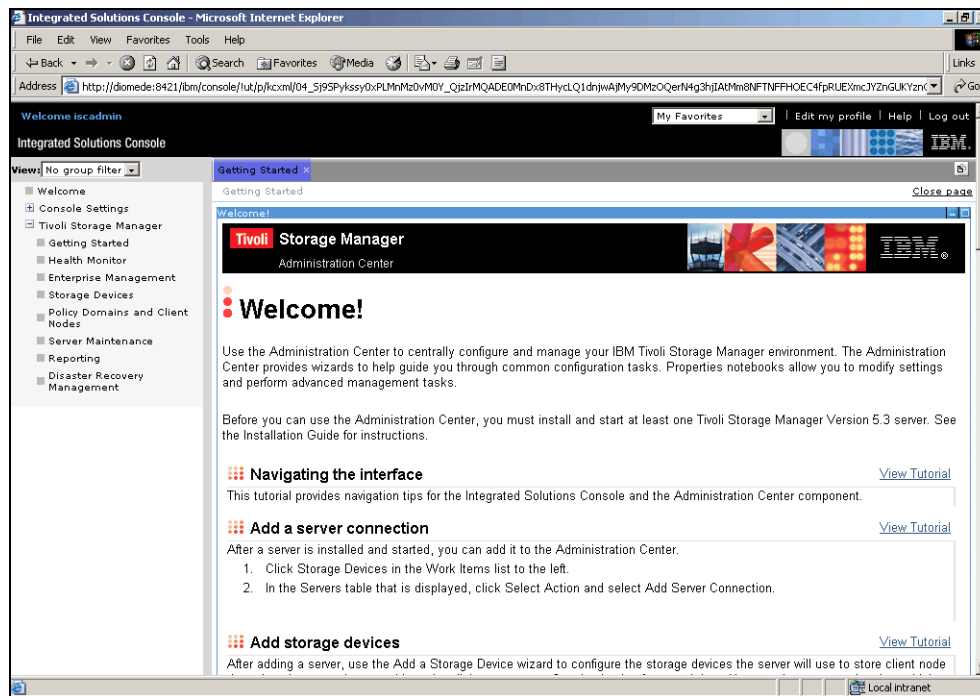


Figure 24-3 Integrated Solutions Console Administration Center

Important: Notice the **View Tutorial** links on the Welcome page. The tutorials provide vital instruction in performing the tasks listed in the Welcome page. It is strongly advised that the Tivoli Storage Manager Administrator follow these tutorials before using the new Administration Center.

The Administration Center provides several other features that can help you monitor and manage your storage management environment:

- ▶ To centrally manage multiple Tivoli Storage Manager servers from a single server.
- ▶ To monitor Tivoli Storage Manager server status.
- ▶ To generate usage and security reports for your Tivoli Storage Manager servers.

Note: Tivoli Storage Manager Administration Center replaces the Web interface that was previously available in the Tivoli Storage Manager versions before v5.3. The Tivoli Storage Manager server can still be administered through the command line but not anymore through the Admin Web GUI interface.

Administration Center main functions

The first level functions available within the Administration Center are: *Health Monitor*, *Enterprise Management*, *Storage Devices*, *Policy Domains and Client nodes*, *Server Maintenance*, *Reporting*, and *Disaster Recovery Management*.

Health Monitor

The Health Monitor section, as shown in Figure 24-4 on page 448, provides easy monitoring of the operations of your Tivoli Storage Manager server.

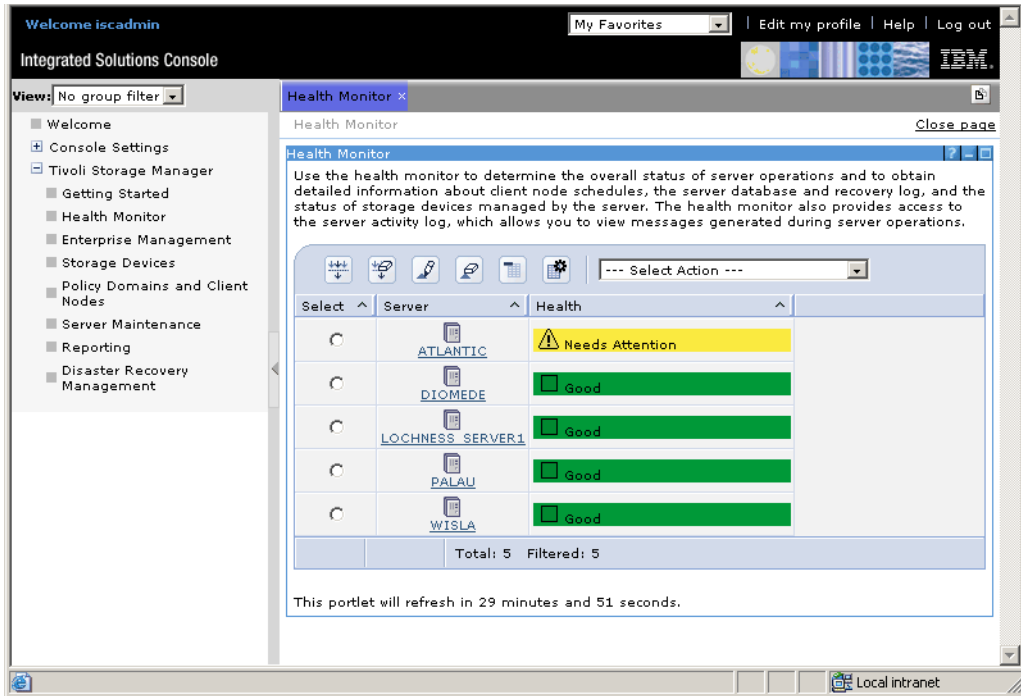


Figure 24-4 Health Monitor

The health monitor provides color coded status information about the overall status of server operations. You can drill down to obtain detailed information about:

- ▶ Scheduled events
- ▶ Server database and recovery log
- ▶ Activity log
- ▶ Storage device status

Enterprise Management

Figure 24-5 shows the enterprise management console.

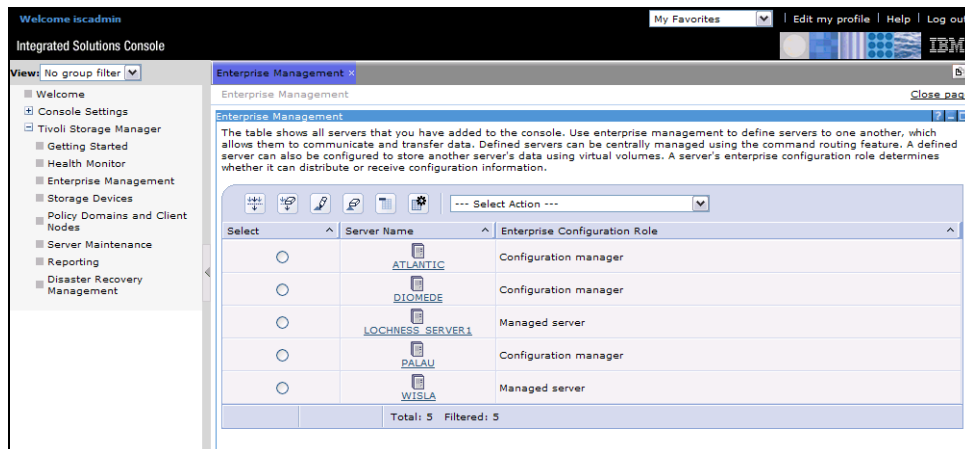


Figure 24-5 Enterprise Management

This console is used to set up a configuration manager and managed servers. A configuration manager is a Tivoli Storage Manager server which propagates configurations (for example, policies, client option sets, schedules, etc.) to other servers. A managed server is a Tivoli

Storage Manager server which receives these configurations from a configuration manager. The configurations are defined in *profiles* on the configuration manager. The managed servers subscribe to one or more profiles so that if any changes to the profile are automatically propagated from the configuration manager.

Storage devices

Figure 24-6 shows the Storage devices console.

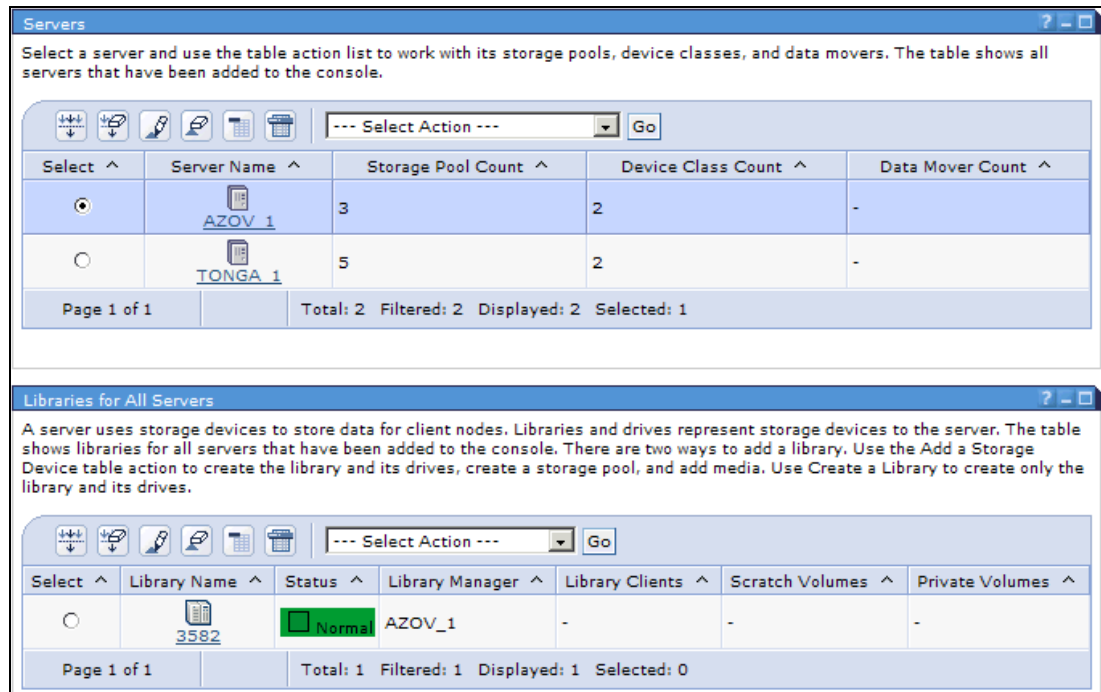


Figure 24-6 Storage devices

On this panel, you can add/modify/delete storage devices - including tape libraries, tape drives, paths, device classes, and storage pools.

Policy domains and client nodes

Figure 24-7 shows the policy domain section of the ISC.

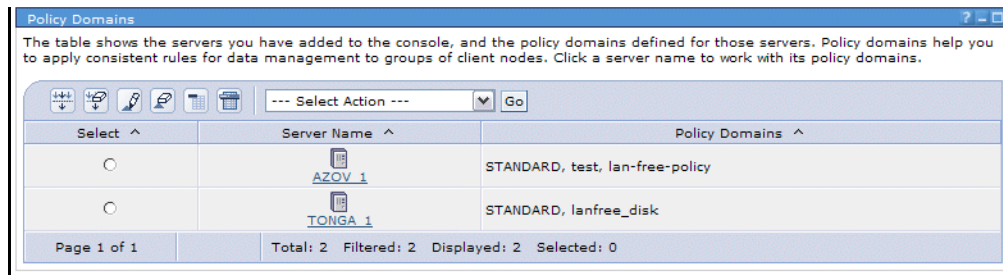


Figure 24-7 Policy Domains

On this panel, you can add/modify/delete policies, policy sets, management classes, and copy groups, and assign client nodes to a policy domain. You also define schedules for automated client backups in this section.

Server maintenance

Figure 24-8 shows the server maintenance console.

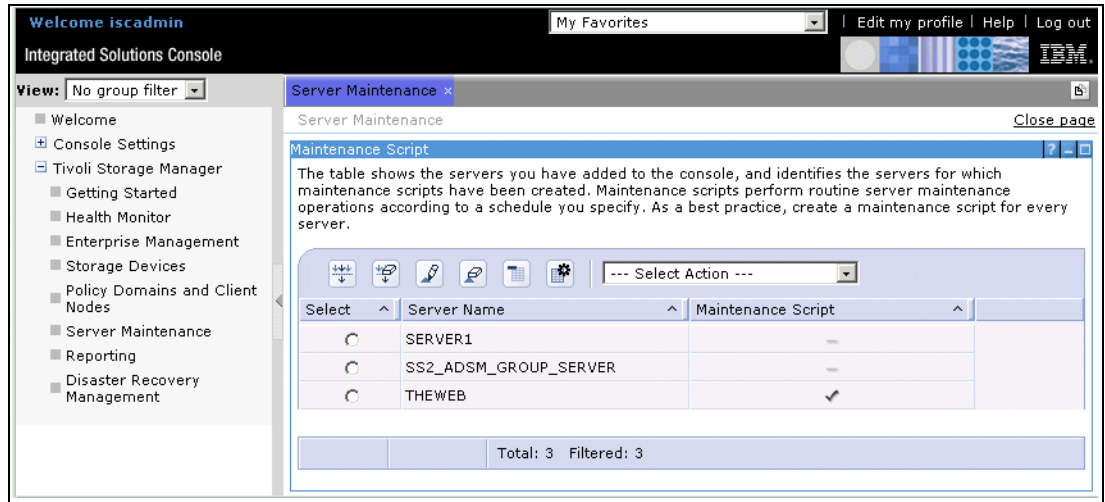


Figure 24-8 Server maintenance

From this console, you can create and run maintenance scripts to automatically run the essential processes that protect the server database and storage pools.

Reporting

The Reporting section, provides customized reports on the Tivoli Storage Manager environment. You can run a Usage report - on space utilized by the client nodes for backup, archive, and space management. Figure 24-9 shows an extract from a Usage report.

Reporting

The table shows the reports you can run against one or more servers. Use reports to view detailed information about your Tivoli Storage Manager environment.

Select	Report Name	Description
<input type="radio"/>	Usage Report	Number of stored files and amount of storage space occupied by client nodes
<input type="radio"/>	Security Report	Access and password information for administrators and client nodes

Total: 2 Filtered: 2

Usage Report for SS2_ADSM_GROUP_SERVER

Client nodes backup usage

Name	Physical (MB)	Logical (MB)	Number of Files
ANTWONG.SANJOSE.IBM.COM	20,601	20,601	77,594
N7340--A	68,084	60,482	194,291
LANEBTEST	0	0	4
BROCKMAN	209,270	201,087	1,525,548
YOUNG	11,628	11,628	6,304
MARKHAYE	66,214	59,149	605,202
NAS300B	98	98	50
TPCNOV74	21,105	21,105	160,654
GENOVA	56,399	56,180	202,968
3A6929	40,923	40,913	223,088

Figure 24-9 Reporting

There is also a security report, which displays for client nodes and administrators, statistics on how long since the password was changed, number of invalid sign-on attempts, and the password expiration period.

Disaster Recovery Management

The Disaster Recovery Management window, shown in Figure 24-10, has options to configure DRM, to generate a disaster recovery plan, and to display information about and track media used for disaster recovery.

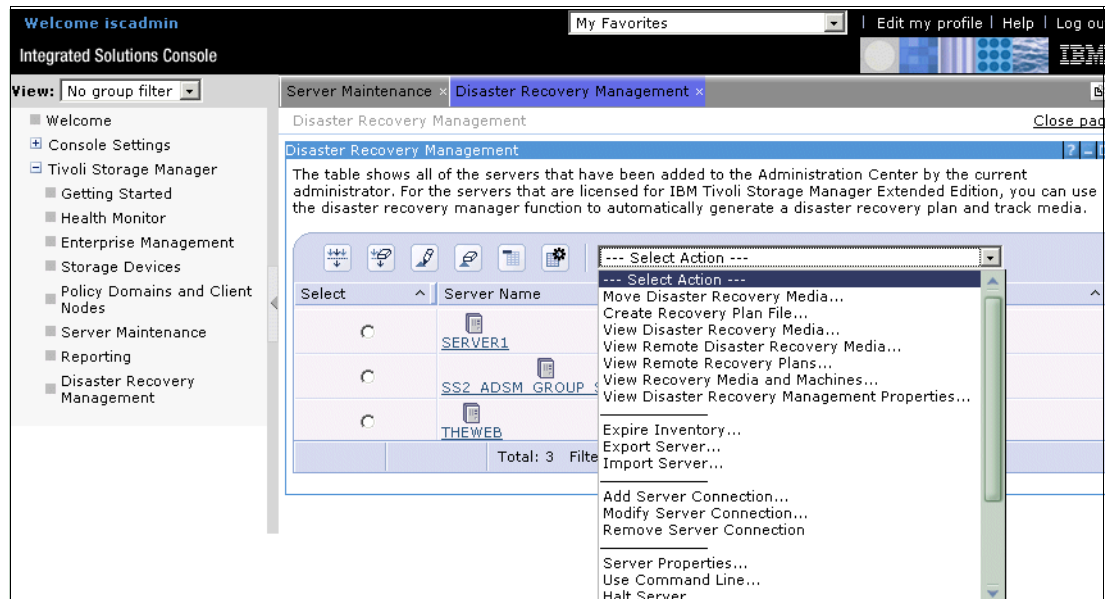


Figure 24-10 Disaster Recovery Manager

Tip: For more information and an in depth review of the Tivoli Storage Manager Administration Center, see the redbook *IBM Tivoli Storage Manager Version 5.3 Technical Guide*, SG24-6638.

24.2.4 IBM Tivoli Storage Manager backup-archive client

The Tivoli Storage Manager backup-archive client interface provides full backup/restore and archive/retrieve functionality in a user-friendly format. Users can select any combination of full, individual directory or individual file backups, as well as backup sets and image backups, via the interface. Tivoli Storage Manager end users can also use the interface to restore files, without requiring administrator intervention. The backup-archive Client also provides the archive and retrieve function. Depending on the client platform, the backup-archive Client may be available as a command-line, graphical, or Web-based interface. The client Web GUI allows for help desk or administrators to access the client remotely to assist with backups/restores or archives/retrieves if needed.

Figure 24-11 on page 452 shows the Windows Tivoli Storage Manager Client GUI. The GUI is either a native GUI for Windows client platforms, or a Java client, for other client platforms. A

Web-based GUI is also available for most client platforms which allows remote access to backup and restore functions.

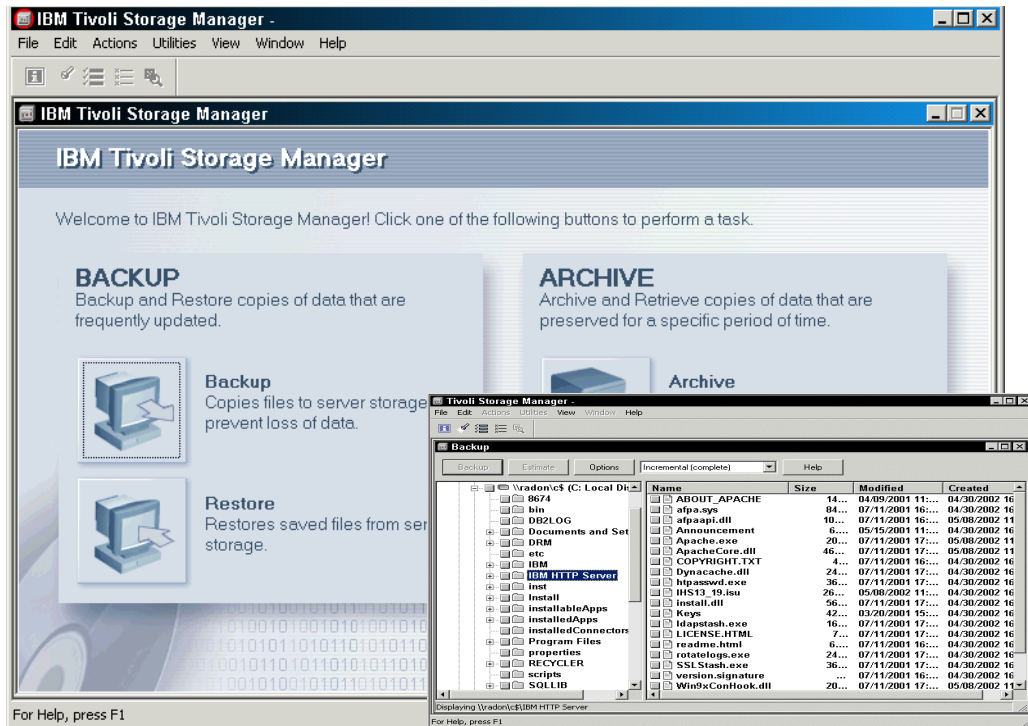


Figure 24-11 The Tivoli Storage Manager client interface

Tivoli Storage Manager implements the patented *progressive backup methodology* and *adaptive subfile backup technology*. Most of the backup-archive Clients are able to exploit the multi-threading capabilities of modern operating systems. Tivoli Storage Manager supports parallel backup and recovery processes to allow expedient movement of data to and from the Tivoli Storage Manager client.

Note: There is now a shared memory communications option between the Tivoli Storage Manager server for Windows and the backup-archive client for Windows. It can be used to perform backups, archives, restores, and retrieves. You can also enable the shared memory communication protocol in the Windows storage agent for communication with the Windows backup-archive client.

24.2.5 IBM Tivoli Storage Manager storage agent

The IBM Tivoli Storage Manager storage agent supports LAN-free backup solutions using a SAN infrastructure. The storage agent dynamically shares SAN connected tape libraries and disks with the Tivoli Storage Manager server, and it has the ability to write and read client data directly to and from server-owned storage media.

The storage agent receives data objects via the Tivoli Storage Manager API and communicates with the Tivoli Storage Manager server over the LAN using TCP/IP to exchange control information and metadata about the objects being backed up. The data movement itself utilizes the LAN-free path over the SAN to write directly to the storage media. Therefore the data movement is removed from both the LAN and the Tivoli Storage Manager server processor for potentially greater scalability.

The storage agent is available for selected backup-archive clients as well as for backing up popular databases and applications such as MS SQL-Server, MS Exchange, Oracle, DB2, SAP, and Lotus Domino.

24.2.6 Tivoli Storage Manager security

Security is a vital aspect for enterprise storage management. Data must be protected, available, and secure. From the moment data is backed up from the client, IBM Tivoli Storage Manager provides a secure storage management environment. Tivoli Storage Manager is the only interface to your backup and archive data.

Before a communication session between the Tivoli Storage Manager client and the Tivoli Storage Manager server begins, an authentication *handshaking* process occurs with authentication tickets and a *mutual suspicion algorithm*. The Tivoli Storage Manager security protocol is modeled after the Kerberos network authentication protocol, which is a highly respected method for secure sign-on cryptography. The client uses its password as part of an encryption key, and does not send the password over the network. Each session key is unique, so replaying a session stream will not result in a sign-on to the Tivoli Storage Manager server. This significantly lowers the chance of a Tivoli Storage Manager session being hijacked by an outside user.

To heighten security for Tivoli Storage Manager sessions, data sent to the Tivoli Storage Manager server during backup and archive operations can be encrypted with standard DES 56-bit encryption. For WAN implementations of Tivoli Storage Manager across public networks, data encryption complements and completes data security for Tivoli Storage Manager.

IBM Tivoli Storage Manager has enhanced support for environments with firewalls in which communication originating from outside the firewall is to be restricted. Clients normally contact the server but with the new firewall support, you can choose to restrict session initiation to the server (*sessioninitiation=serveronly*). Scheduled backup-archive client operations can be restricted to server-initiated sessions.

24.2.7 Tivoli Storage Manager API

The Tivoli Storage Manager APIs are used for Tivoli's own Data Protection products, and they are also documented and published. This allows ISVs to adapt their solutions to integrate with Tivoli Storage Manager to extend its functionality. In particular, various vendors have used the APIs to provide bare metal recovery solutions for various platforms. Among the vendors exploiting these APIs for disaster recovery include Cristie, UltraBac Software, and Symantec.

The IBM System Storage DR550 is a storage solution, which leverages applications using Tivoli Storage Manager API. For an overview of the DR550, refer to Chapter 10, "IBM System Storage DR550" on page 187 or to *Understanding the IBM System Storage DR550*, SG24-7091.

Figure 24-12 on page 454 shows how the API fits into the DR550 solution.

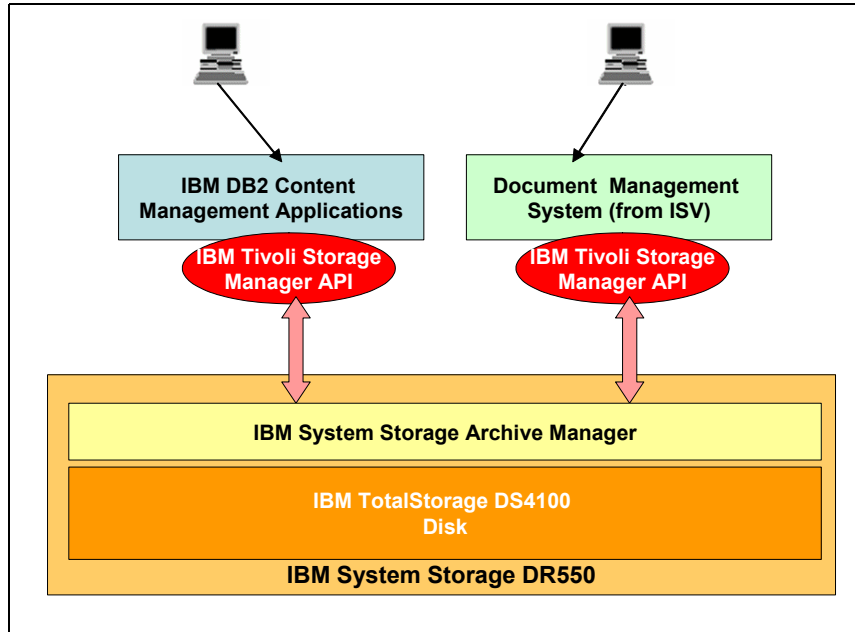


Figure 24-12 How the API fits into the DR550 solution

24.3 Client backup and restore operations

It is important to understand Tivoli Storage Manager options for client backup and restore operations. It is also important to understand the characteristics of each of these operations because each method may have an effect on backup and restore efficiency, retention periods, portability, CPU utilization, connection time, and network utilization. The standard backup method that Tivoli Storage Manager uses is called progressive incremental backup (see 24.2.2, “Progressive incremental backups” on page 446). It is a unique and efficient method for backup.

24.3.1 Traditional LAN and WAN backup topology

In a traditional LAN and WAN environment the Tivoli Storage Manager backup and archive client or application reads data from locally attached disks and sends it over the LAN to the Tivoli Storage Manager backup server as shown in Figure 24-13 on page 455. The server receives the data, and then writes it out to its storage pool — tape, disk, or optical media — based on predefined policies and server configuration. Data is read and written by both the Tivoli Storage Manager Client and Tivoli Storage Manager Server machines. In addition, control information is also sent over the LAN to the Tivoli Storage Manager Server.

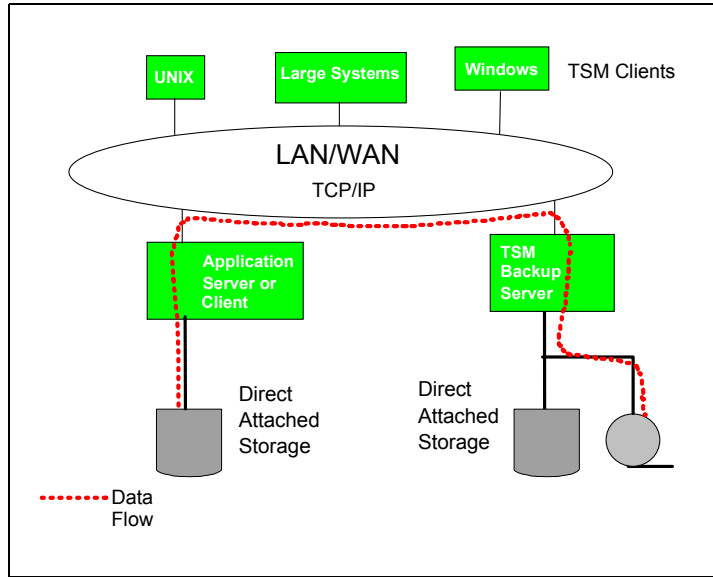


Figure 24-13 Tivoli Storage Manager LAN and WAN backup

24.3.2 SAN (LAN-free) backup topology

SAN technology provides an alternative path for data movement between the Tivoli Storage Manager client and the server. Shared storage resources (disk, tape) are accessible to both the client and the server through the Storage Area Network. Data movement is off-loaded from the LAN and from the server processor and allows for greater scalability. LAN-free backups decrease the load on the LAN by introducing a Storage Agent. The Storage Agent can be thought of as a small Tivoli Storage Manager server (without a database or recovery log) which is installed and run on the Tivoli Storage Manager client machine. The Storage Agent handles the communication with the Tivoli Storage Manager server over the LAN but sends the data directly to SAN attached tape devices or disks, relieving the Tivoli Storage Manager server from the actual I/O transfer.

A LAN-free backup environment is shown in Figure 24-14 on page 456.

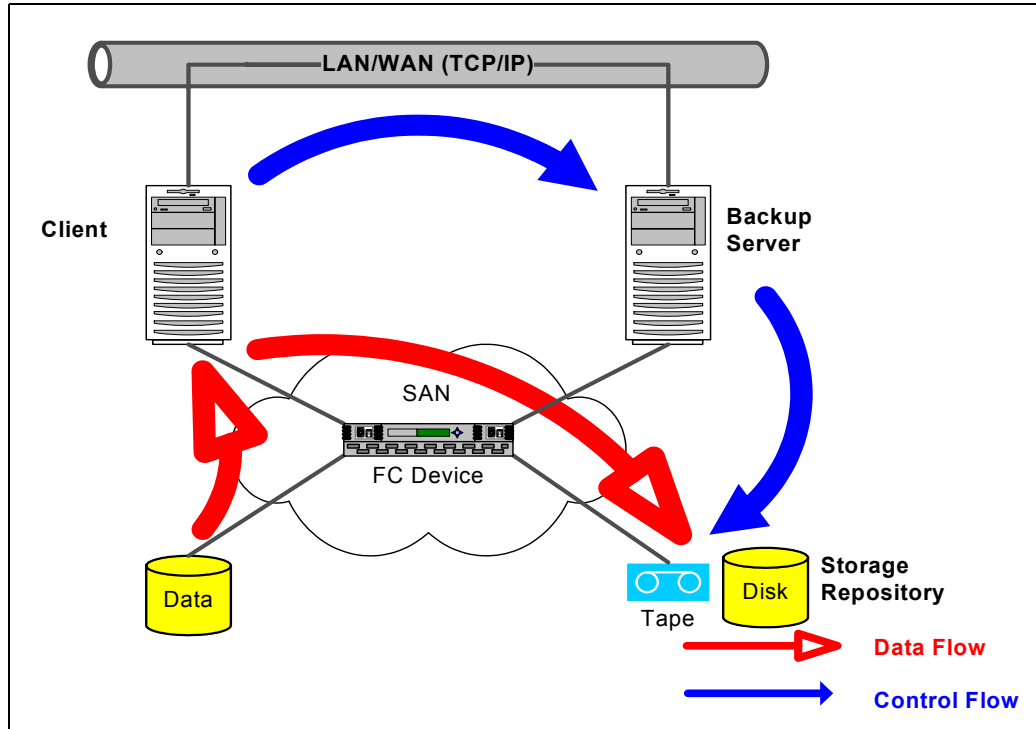


Figure 24-14 Tivoli Storage Manager LAN-free backup

24.3.3 Split-mirror/point-in-time copy backup using SAN

A split-mirror/point-in-time backup occurs when a copy volume generated by Operating System mirroring or a hardware assisted instant copy function (as found on many of today's high-end storage systems) is backed up to a Tivoli Storage Manager server, as shown in Figure 24-15 on page 457. Such a backup method virtually eliminates the backup-related performance impact on the production host. This approach is facilitated and automated with the Tivoli Storage Manager for Advanced Copy Services product, which integrates IBM FlashCopy on disk systems with Tivoli Storage Manager and its database protection capabilities for Oracle, SAP, and DB2. This *Copy-Backup* procedure adds value to storage and backup procedures, because it helps ensure that essential applications can continue to run 24x7 with minimal backup-related impact.

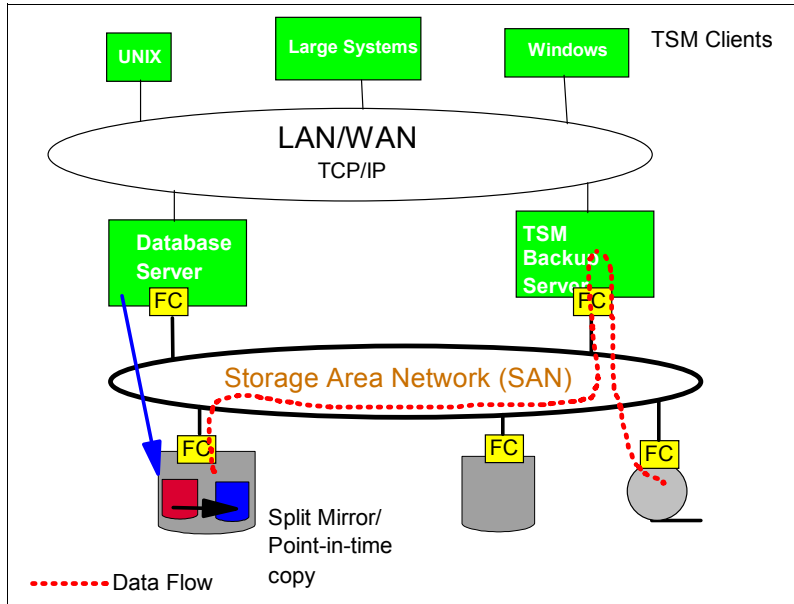


Figure 24-15 Tivoli Storage Manager split-mirror/point-in-time copy backup

24.3.4 NAS and N Series backup and restore

NAS appliances supporting NDMP (Network Data management Protocol), including IBM System Storage N series, can be backed up with Tivoli Storage Manager Extended Edition. Tivoli Storage Manager Extended Edition provides backup and recovery support for NAS file servers from various vendors.

NAS file servers require special handling for backup and recovery services, because these file servers are not typically intended to run third-party software. With Tivoli Storage Manager, the NAS file server does not need to install any extra software. Instead, the Tivoli Storage Manager server uses NDMP to connect to the NAS file server to initiate, control, and monitor a file system backup or restore operation. The NDMP protocol communicates with and provides backup and recovery services for NAS file servers. NDMP is an industry-standard protocol that allows a network storage-management application to control the backup and recovery of an NDMP-compliant file server without installing third-party software on that server. The implementation of the NDMP server protocol enables the NAS file servers to be backup-ready and enables high performance backup to tape devices without moving the data over the LAN.

Tivoli Storage Manager NDMP support provides file and directory level backup and restore. The backup destination is to a tape library which should be attached to the NAS file server, and can also optionally be attached to the Tivoli Storage Manager server. This leads to two possible configurations.

Option 1 - library attached only to NAS server

In Figure 24-16 on page 458, the library robotics and drives are physically connected directly to the NAS file server, with paths defined paths must be defined from the NAS data mover to the library and drives. No physical connection is required between the Tivoli Storage Manager server and the SCSI library.

The Tivoli Storage Manager server controls the library robotics by sending library commands cross the network to the NAS file server. The NAS file server passes the commands to the tape library. Any responses generated by the library are sent to the NAS file server, and

passed back across the network to the Tivoli Storage Manager server. This configuration supports physical distance between the Tivoli Storage Manager server and the NAS file server. For example, the Tivoli Storage Manager server could be in one city, while the NAS file server and tape library are in another location.

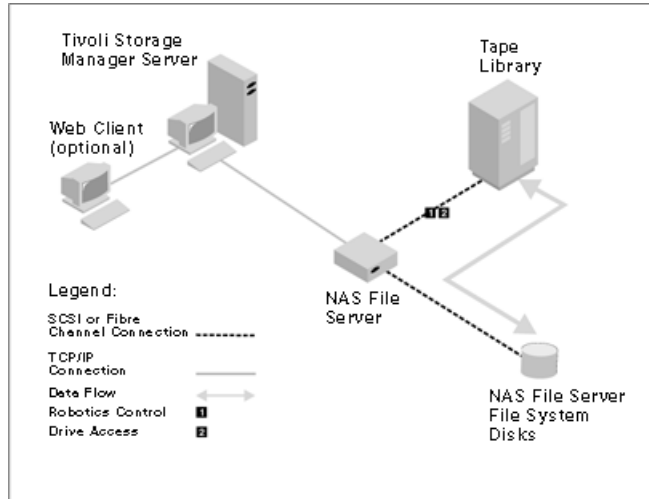


Figure 24-16 Tape library attached only to NAS server

Option 2 - library attached to both NAS and Tivoli Storage Manager servers

In this configuration, shown in Figure 24-17, the Tivoli Storage Manager server controls the SCSI library through a direct, physical connection to the library robotics control port. For NDMP operations, the drives in the library are connected directly to the NAS file server, and a path must be defined from the NAS data mover to each of the drives to be used. The NAS file server transfers data to the tape drive at the request of the Tivoli Storage Manager server. The tape library uses separate ports for robotics control and for drive access. In addition, the library must be within Fibre Channel or SCSI bus range of both the Tivoli Storage Manager server and the NAS file server.

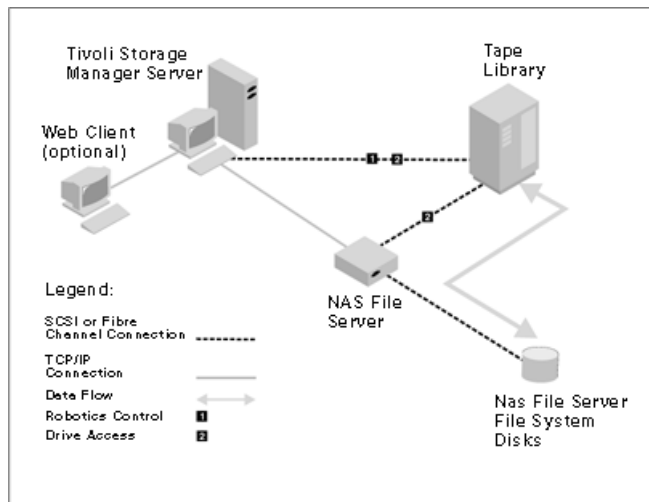


Figure 24-17 Tivoli Storage Manager and NDMP backup

For more information about backing up IBM N series, see *Using the IBM System Storage N Series with IBM Tivoli Storage Manager*, SG24-7243. For the current list of supported NDMP-capable NAS file servers, see the Network Attached Storage section of:

http://www-306.ibm.com/software/sysmgmt/products/support/IBM_TSM_Supported_Devices_for_AIXHPSUNWIN.html

24.3.5 Image backup

An image backup is a block-by-block copy, single object backup of a volume (typically a UNIX file system or raw logical volume, or Windows drive) on a Tivoli Storage Manager client. Being able to restore an entire volume as one object can lead to faster recoveries. Image backup is available at the time of writing on AIX, HP, Sun, Linux, and Windows client platforms. Specific image backup requirements for each platform should be reviewed in their associated installation guides found at:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v1r1/index.jsp>

With image backup, the Tivoli Storage Manager server does not track individual files in the file system image. File system images are tracked as individual objects and the management class policy are applied to the file system image as a whole. An image backup provides the following benefits.

- ▶ Can provide a quicker backup and restore than a file-by-file backup as there is no overhead involved in creating individual files.
- ▶ Conserves resources on the server during backups since only one entry is required for the image.
- ▶ Provides a point-in-time picture of your file system, which may be useful if your enterprise needs to recall that information.
- ▶ Restores a corrupt file system or raw logical volume. Data is restored to the same state it was when the last logical volume backup was performed.

Figure 24-18 on page 460 illustrates the process for image backup with Tivoli Storage Manager.

On Windows client platforms a Logical Volume Storage Agent (LVSA) is used, which can take snapshot of the volume while it is online. Optionally only occupied blocks can be copied. If the snapshot option is used (rather than static) then any blocks which change during the backup process are first kept unaltered in an Original Block File. In this way the client can send a consistent image of the volume as it was at the start of the snapshot process to the Tivoli Storage Manager server.

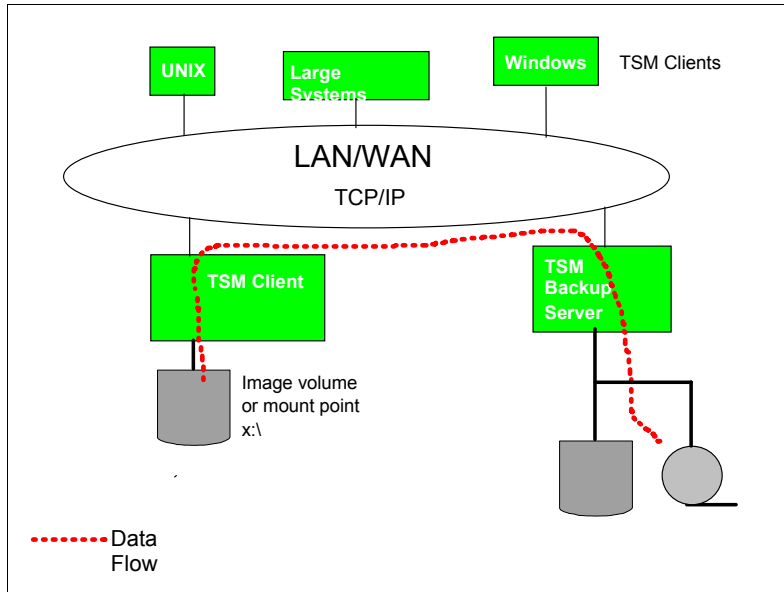


Figure 24-18 Tivoli Storage Manager image backup

This section has provided an overview of many of the common backup methods supported by Tivoli Storage Manager. These methods can be integrated with DR strategies being considered in your environment.

24.4 IBM System Storage Archive Manager

IBM Tivoli Storage Manager's existing policy-based data management capabilities helps meet the needs of most regulated industries today through an open application programming interface (API). IBM System Storage Archive Manager implements data retention policies that help meet regulatory requirements.

24.4.1 Overview

IBM System Storage Archive Manager provides enhancements to Tivoli Storage Manager Extended Edition to help meet additional requirements defined by regulatory agencies for the retention and disposal of data. These enhancements include new functionality and new device support and are delivered in four key areas:

- ▶ Data Retention Protection
- ▶ Event-based Retention Management
- ▶ Expiration/Deletion Suspension (Deletion hold)
- ▶ Additional Worm Tape, DVD library support

IBM System Storage Archive Manager facilitates compliance with the most stringent regulatory requirements in a flexible and function-rich manner. It helps manage and simplify the retrieval of the ever-increasing amount of data that organizations must retain for strict records retention regulations.

Content management and archive applications can interface with using the IBM Tivoli Storage Manager client API (see 24.2.7, "Tivoli Storage Manager API" on page 453) to apply business policy management for ultimate deletion of archived data at the appropriate time.

IBM System Storage Archive Manager is sold as a separately licensed software product integrated into Tivoli Storage Manager Extended Edition, and it requires a stand-alone Tivoli Storage Manager Extended Edition server be dedicated for its use.

Attention: For more information about data retention, refer to Chapter 10, “IBM System Storage DR550” on page 187.

Device support enhancements

IBM Tivoli Storage Manager (including IBM System Storage Archive Manager) supports:

- ▶ **WORM media:** Supports the 3592 and LTO WORM media in drives and libraries.
- ▶ **DVD:** Supports selected high-capacity, non-erasable DVD. DVD support provides another storage device that helps meet the regulatory agencies requirements as well as extending the Tivoli Storage Manager managed storage hierarchy with a new storage device.

Support for these devices is important to help meet additional regulatory agencies requirements for a subset of data that must be stored on media that cannot be altered. This is in addition to other software or hardware controls in place.

24.5 IBM Tivoli Storage Manager disaster recovery

Disaster preparation, planning and recovery (formerly Tivoli Disaster Recovery Manager) is a component of IBM Tivoli Storage Manager Extended Edition. This component provides disaster recovery, planning and scripting for the Tivoli Storage Manager server and assists in disaster recovery for clients.

The Disaster Recovery component offers various options to configure, control, and automatically generate a Disaster Recovery Plan file containing the information, scripts, and procedures needed to automate Tivoli Storage Manager server restoration and help ensure quick recovery of data after a disaster. It also manages and tracks the media on which the data is stored, whether onsite, in-transit, or in a vault, so that required data can be easily located if disaster strikes. The scripts can help document a basic IT recovery strategy, the steps to rebuild core systems, as well as the critical machines that must be recovered.

A key feature of Tivoli Storage Manager’s Disaster Recovery component is the ability to track media in all states that it could possibly be, such as onsite, in transit or in a vault. Because of the criticality of data in the production environment, controls are needed to make sure that all previously backed up data can be found and restored in a reasonable amount of time.

The Disaster Recovery function helps maintain business continuance by:

- ▶ Establishing and helping to automate a thorough Tivoli Storage Manager *server* Disaster Recovery Plan (DRP) — clients can then subsequently restore their data from the server if required.
- ▶ Ensuring that customer-provided information is available in the same plan
- ▶ Automating vital recovery steps to bring the Tivoli Storage Manager server and backup environment back to normal operation
- ▶ Managing and identifying off-site media needed for recovery
- ▶ Tracking and reporting systems destroyed, in the event of a disaster
- ▶ Storing client configuration information and assigning client recovery priorities

With the Disaster Recovery component, you can recover at an alternate site, on replacement computer hardware, recover using different hardware configuration at the recovery site, and with people who are not familiar with the applications. You can also use the DRP for audits to certify the recoverability of the server. The DRP can be easily recreated daily so that it is up to date. The main functions are illustrated in Figure 24-19.

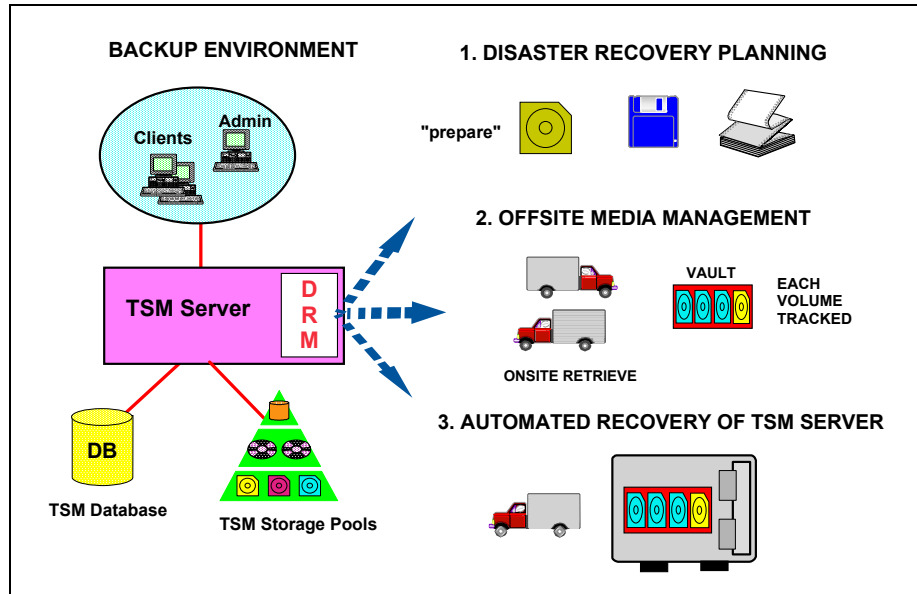


Figure 24-19 Tivoli Storage Manager Disaster Recovery (DRM) functions

During a real disaster, errors commonly encountered include:

- ▶ The DRP was not tested
- ▶ A skilled technical team performed the testing, who *filled-in* missing steps
- ▶ The recovery plan is out of date
- ▶ Disk volume definitions for the recovery site are not known
- ▶ Location of recovery tapes is not known
- ▶ It is not known which tapes are to be applied first

DRM will help answer questions, such as:

- ▶ Where is the current server configuration information located?
- ▶ What are the current server database volumes?
- ▶ What is my recovery sequence?
- ▶ Is my recovery plan current, is this guaranteed?
- ▶ What was the client and server machine configuration like?
- ▶ Who should be contacted in a disaster?
- ▶ Where is the recovery media located?
- ▶ Can I restore my environment to any point in time?

During recovery from a disaster, the Disaster Recovery Management component (DRM) automates the following procedures to restore the Tivoli Storage Manager servers:

- ▶ Restores Tivoli Storage Manager server's key option files
- ▶ Copies files from alternate locations to production locations
- ▶ Initializes Tivoli Storage Manager database and log volumes
- ▶ Matches sizes and locations of Tivoli Storage Manager database and log volumes
- ▶ Launches current DB restore automatically
- ▶ Tracks media needed and availability
- ▶ Registers Tivoli Storage Manager server features installed

- ▶ Returns server state to a valid license configuration
- ▶ Updates Tivoli Storage Manager volume catalog information
- ▶ Marks volume information for recovery, that is, is it destroyed or not?
- ▶ Rebuilds Tivoli Storage Manager hierarchical storage configuration
- ▶ Re-creates customer backup environment

A detailed description, recovery scenario, and recovery plan built with DRM can be found in the redbook *Disaster Recovery Strategies with Tivoli Storage Management*, SG24-6844. Also, recommendations and examples of using DRM to store client machine information in the DRM plan file for use during a client disaster recovery are given in the same redbook.

In summary, DRM will systematically re-build the storage management server environment and ensure current application data for the entire enterprise is available for recovery. This is all possible from a single scripted command, automatically.

24.6 Tivoli Storage Manager server-to-server communications

With a solution that includes multiple Tivoli Storage Manager servers, you can use server-to-server communications and virtual volumes to enhance management and improve disaster recoverability. Tivoli Storage Manager server-to-server communications provide the capability to:

- ▶ Configure and manage multiple servers with enterprise administration
- ▶ Distribute a consistent configuration for Tivoli Storage Manager servers through a configuration manager to managed servers. By having consistent configurations, you can simplify the management of a large number of servers and clients
- ▶ Perform tasks simultaneously on multiple servers by using command routing, enterprise logon and enterprise console
- ▶ Send server and client events to another server for logging
- ▶ Monitor events of many servers and clients from a single server
- ▶ Store data on another server using virtual volumes

Figure 24-20 on page 464 summarizes the server-to-server capabilities of Tivoli Storage Manager.

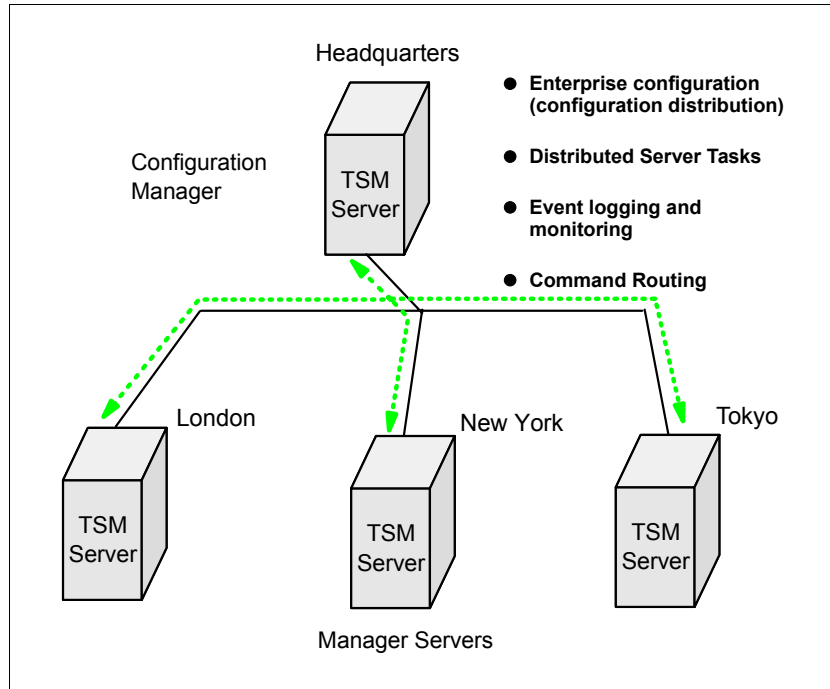


Figure 24-20 Tivoli Storage Manager server-to-server communications

24.6.1 Using server-to-server virtual volumes for disaster recovery

Server-to-server virtual volumes can be used as a strategy for site disaster protection and disaster recovery.

For site disaster protection, server-to-server virtual volumes (with Tivoli Disaster Recovery Manager) can be used to back up storage pools from one Tivoli Storage Manager server to another and back up the Tivoli Storage Manager database of one server to another. This form of virtual vaulting can occur on a peer-to-peer basis, between any number of servers, and at different data centers or different sites. For example, two servers can back up each other in an equal level relationship.

For disaster recovery, server-to-server virtual volumes can be used to store the Disaster Recovery Plan file remotely. In this strategy, the source server creates the Disaster Recovery Manager plan files, then stores the files on a remote target server. You can display information about recovery plan files from the server that created the files (the source server) or from the server on which the files are stored (the target server). You can easily display a list of all recovery plan files that have been saved on a target server.

24.7 Tivoli Storage Manager and high availability clustering

High availability clustering can be used to configure either or both a highly available Tivoli Storage Manager server and/or clients. A highly available Tivoli Storage Manager server provides highly available services to its clients without significant disruptions; and a highly available Tivoli Storage Manager client provides highly available backup or restore services to an application and data.

For a highly available Tivoli Storage Manager server, critical resources such as the database and storage pools must be shared between the clustered nodes. A highly available Tivoli

Storage Manager client will be able to back up file systems that are commonly available to all nodes in the cluster.

Tivoli Storage Manager supports clustering using supporting software such as IBM High Availability Cluster Multi-Processing (HACMP), Microsoft Cluster Server (MSCS), VERITAS Storage Foundation, and Tivoli System Automation. This section provides some overview information about some of these environments; for more details, see *IBM Tivoli Storage Manager in a Clustered Environment*, SG24-6679.

24.7.1 IBM High Availability Cluster Multi-Processing (HACMP)

A Tivoli Storage Manager server can use IBM High Availability Cluster Multiprocessing (HACMP) software for high availability. HACMP is an AIX-based clustering solution, which allows automatic system recovery on system failure detection. Using HACMP together with Tivoli Storage Manager ensures server availability. HACMP offers local or campus disaster survivability with real-time automated failover and reintegration within distance limitations. In an HACMP environment, TCP/IP is the communications method used to support the checking of status and availability of the production and failover server, also commonly referred to as the *heartbeat* connection.

HACMP detects system failures and manages failover to a recovery processor with a minimal end-user downtime. You can set up a 2-way active-passive HACMP clustered configuration so that, if the production Tivoli Storage Manager server fails, the server will be brought up on the standby system in the cluster. This is shown in Figure 24-21. Scripts are provided with Tivoli Storage Manager to automate the failover and failback operations. The Tivoli Storage Manager database, recovery log, and storage pools are stored on shared storage, available to both servers in the cluster.

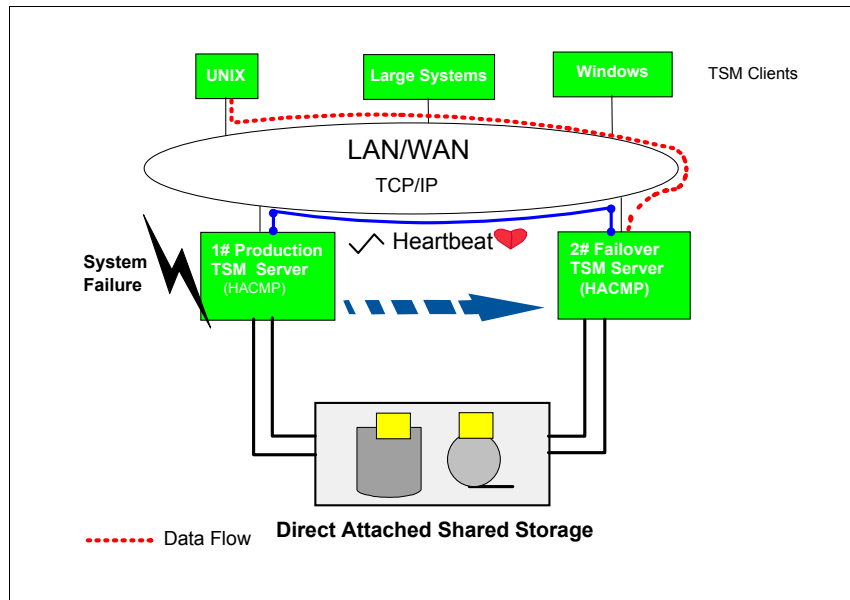


Figure 24-21 HACMP and Tivoli Storage Manager server high availability configuration

Both failover and fallback act as though a Tivoli Storage Manager server has crashed or halted and was then restarted - the fact that the server is running on different hardware is transparent to the clients. Any transactions that were in-flight at the time are rolled back, and all completed committed transactions are still complete. Tivoli Storage Manager clients see this as a communications failure and try to re-establish the connection. The backup-archive clients can usually restart from the last committed transaction.

Note: For a detailed discussion on supported environments, prerequisites, install, setup, and testing of an HACMP and Tivoli Storage Manager server failover environment, see *Tivoli Storage Manager for AIX Quick Start*, GC32-0770, and *IBM Tivoli Storage Manager in a Clustered Environment*, SG24-6679.

24.7.2 Backup-archive client support with HACMP

The Tivoli Storage Manager backup-archive client itself (including the administrator, backup-archive, HSM and API pieces) can also be clustered using HACMP to provide highly available backup-restore operations. This configuration allows scheduled Tivoli Storage Manager client operations to continue processing in the event of a system failure on a redundant clustered failover server. See Figure 24-22 for an illustration of how this works.

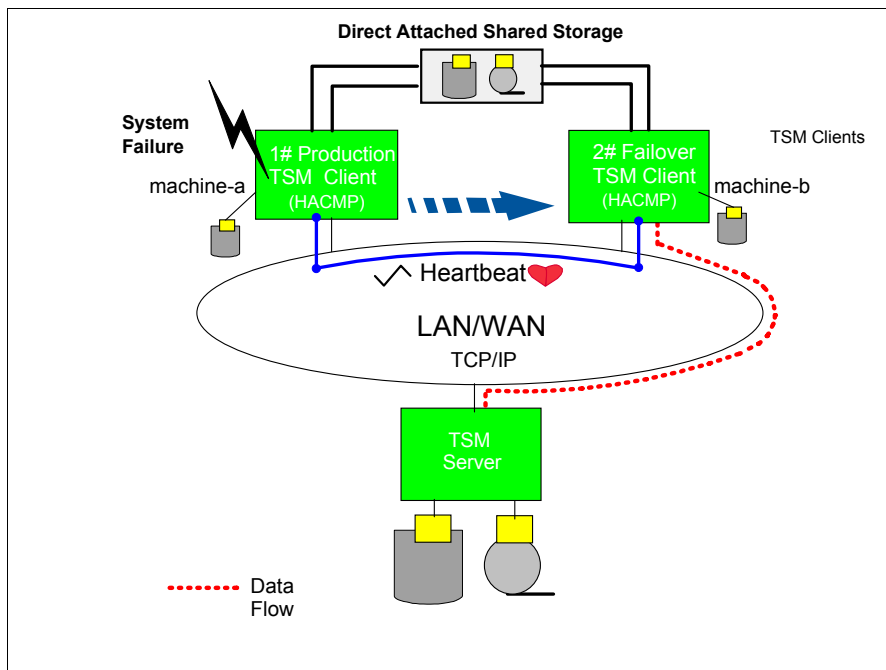


Figure 24-22 HACMP and Tivoli Storage Manager client high availability configuration

In this configuration, the data to be backed up is stored on shared disk which is available to either of the clustered client nodes.

If a scheduled incremental backup of a clustered volume is running on machine-a and a system failure causes a failover to machine-b, machine-b then reconnects to the server. If the re-connection occurs within the start window for that event, the scheduled command is restarted. This scheduled incremental backup will reexamine files sent to the server before the failover. The backup will then *catch up* to where it terminated before the failover situation.

If a failover occurs during a user initiated (that is, non-scheduled) client session, the Tivoli Storage Manager client starts on the node that is handling the takeover. This allows it to process scheduled events and provide Web client access. You can install the Tivoli Storage Manager client locally on each node of an HACMP environment. You can also install and configure the Tivoli Storage Manager Scheduler Service for each cluster node to manage all local disks and each cluster group containing physical disk resources.

HACMP support for Hierarchical Storage Management (HSM) clients on AIX provides support for HACMP failover on AIX so that HSM managed file systems can continue to operate in the case of an HACMP node failover and fallback.

24.7.3 Tivoli Storage Manager server and Microsoft Cluster Server (MSCS)

Tivoli Storage Manager is a cluster-aware application and can be configured in a MSCS high availability environment. Clustered systems are connected to the same disk subsystem and provide a high-availability solution that minimizes or eliminates many potential sources of downtime. Microsoft Cluster Server (MSCS) is software that helps configure, monitor, and control applications and hardware components that are deployed on a Windows cluster. When you use cluster configurations, you enhance the availability of your servers.

In the MSCS failover environment shown in Figure 24-23, a clustered Tivoli Storage Manager server called TSMSERVER1 runs on node A and a clustered Tivoli Storage Manager server called TSMSERVER2 runs on node B. This is an *active/active* configuration. The clients connect to either TSMSERVER1 or TSMSERVER2 as virtual servers. The clients are not aware of which physical node currently hosts their server - that is, the server's location is transparent to client applications.

If one of the software or hardware resources fails, failover occurs. Resources (for example, applications, disks, or an IP address) migrate from the failed node to the surviving node. The surviving node takes over the Tivoli Storage Manager server resource group, restarts the Tivoli Storage Manager service, and provides access to administrators and clients. If the TSMSERVER1 node fails, the other node assumes its role, so that it runs both servers, TSMSERVER1 and TSMSERVER2. To a client, it is exactly as though the Tivoli Storage Manager server was turned off and immediately turned back on again.

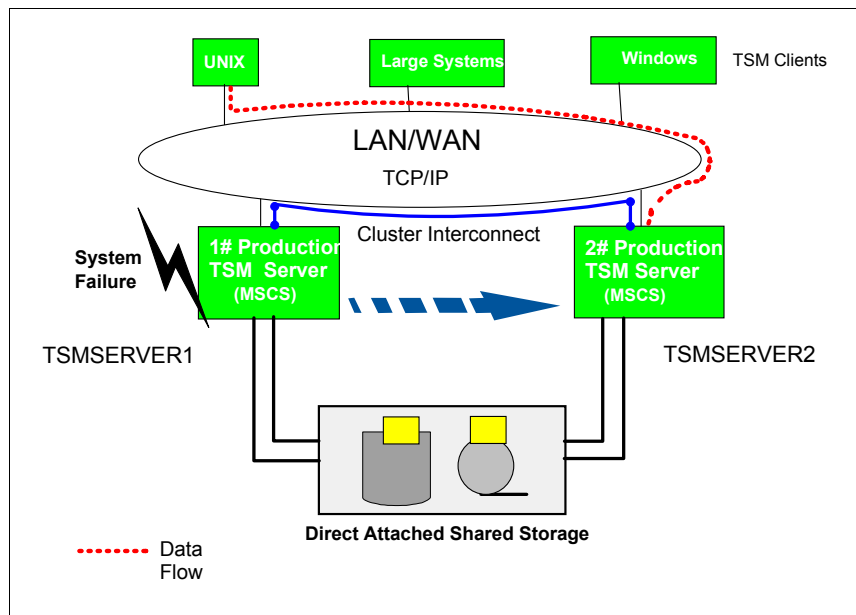


Figure 24-23 MSCS and Tivoli Storage Manager active/active configuration

Clients experience the loss of all connections to TSMSERVER1 and all active transactions are rolled back to the client. Clients must reconnect to TSMSERVER1 after this occurs, which is normally handled as an automatic attempt to reconnect by the Tivoli Storage Manager client. The location of TSMSERVER1 is transparent to the client. A node can host physical or logical units, referred to as resources. Administrators organize these cluster resources into functional units called groups and assign these groups to individual nodes. If a node fails, the

server cluster transfers the groups that were being hosted by the node to other nodes in the cluster. This transfer process is called failover. The reverse process, failback, occurs when the failed node becomes active again and the groups that were failed over to the other nodes are transferred back to the original node.

Tivoli Storage Manager can also run in an *active/passive* configuration. In the active/passive configuration you create one instance of a Tivoli Storage Manager server that can run on either node. One system runs actively as the production Tivoli Storage Manager server, while the other system sits passively as an online (hot) backup.

MSCS does not support the failover of tape devices. However, Tivoli Storage Manager can handle this type of a failover pattern with the correct set up. Tivoli Storage Manager uses a shared SCSI bus for the tape devices. Each node (two only) involved in the tape failover must contain an additional SCSI adapter card. The tape devices (library and drives) are connected to the shared bus. When failover occurs, the Tivoli Storage Manager server issues a SCSI bus reset during initialization. In a failover situation, the bus reset is expected to clear any SCSI bus reserves held on the tape devices. This allows the Tivoli Storage Manager server to acquire the devices after the failover.

For a detailed discussion on supported environments, prerequisites, install, setup, and testing of a MSCS and Tivoli Storage Manager server failover environment, see *Tivoli Storage Manager for Windows Administrator's Guide GC32-0782*, and *IBM Tivoli Storage Manager in a Clustered Environment, SG24-6679*.

24.7.4 Tivoli Storage Manager backup-archive client support with MSCS

The Tivoli Storage Manager client is supported in an MSCS cluster environment. This configuration allows scheduled Tivoli Storage Manager client operations to continue processing in the event of a system failure on a redundant clustered failover server as shown in Figure 24-24.

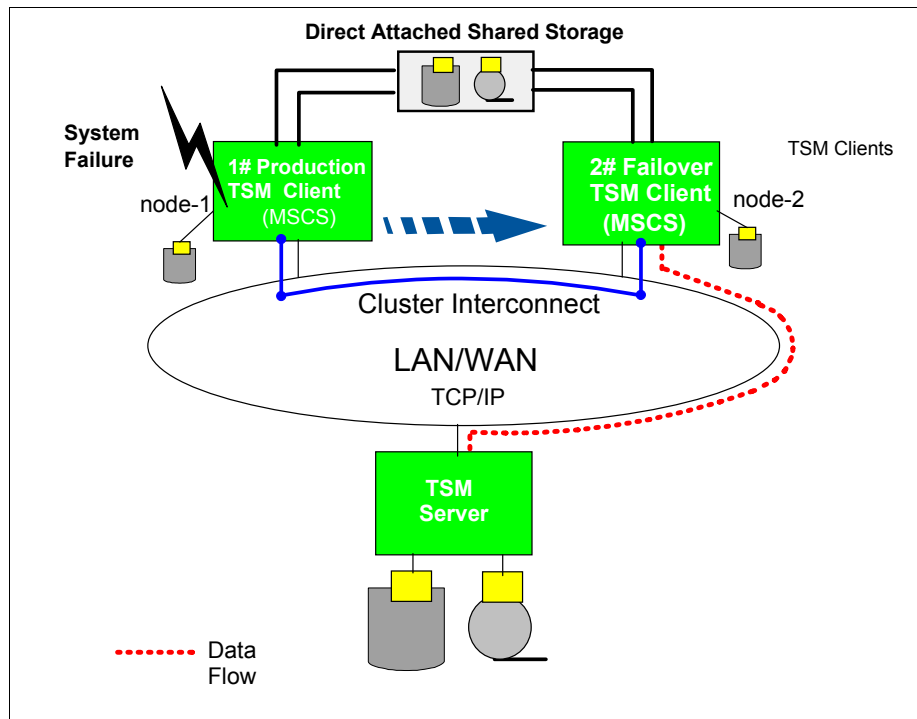


Figure 24-24 MSCS and Tivoli Storage Manager client high availability configuration

In this example, the cluster contains two nodes: node-1 and node-2, and two cluster groups containing physical disk resources. In this case, an instance of the Tivoli Storage Manager Backup-Archive Scheduler Service is installed for each node: node-1 and node-2. This ensures that proper resources are available to the Backup-Archive client when disks move (or fail) between cluster nodes. The CLUSTERNODE option in the client option file ensures that Tivoli Storage Manager manages backup data logically, regardless of which cluster node backs up a cluster disk resource.

For a detailed discussion on supported environments, prerequisites, install, setup, and testing of a MSCS and Tivoli Storage Manager client failover environment, see *Tivoli Storage Manager for Windows Backup-Archive Clients Installation and User's Guide Version, GC32-0788* and *IBM Tivoli Storage Manager in a Clustered Environment, SG24-6679*.

24.8 Tivoli Storage Manager and tape vaulting

Traditionally, disaster recovery plans include daily off-site tape backups that are picked up from the local site and transported via a courier to a secure facility, which is often a tape vaulting service provider. Vaulting of tapes at off-site locations can provide a secure means to protect data in the event of a disaster at the primary site. To recover from a disaster, you must know the location of off-site recovery media. Tivoli Storage Manager DRM helps determine which volumes to move off site and back on site and tracks the location of the volumes. With tape vaulting you can back up primary storage pools to a copy storage pool and then send the copy storage pool volumes off site. You can track these copy storage pool volumes by changing their access mode to off site, and updating the volume history to identify their location. If an off-site volume becomes expired, the server does not immediately return the volume to the scratch pool. The delay prevents the empty volumes from being deleted from the database, making it easier to determine which volumes should be returned to the onsite location. DRM handles all of this automatically.

24.8.1 Electronic tape vaulting

Using Tivoli Storage Manager with electronic tape vaulting provides additional data protection capabilities, with backups made to remote tape drives over communication links. Electronic vaulting can enable shorter recovery times and reduced data loss should the server be damaged. An electronic tape vaulting solution combined with Tivoli Storage Manager is fundamental to achieving Tier 3 and above RPO and RTOs, that is, less than 24 hours. With electronic tape vaulting the Tivoli Storage Manager server will have an alternate location to store primary and copy storage pools as though they are directly attached. The Tivoli Storage Manager server can first write a copy of disk storage pool data to tape pools at the remote site (Datacenter #2), then the data can be migrated to the tape storage pools at the primary site (Datacenter #1). See Figure 24-25 on page 470.

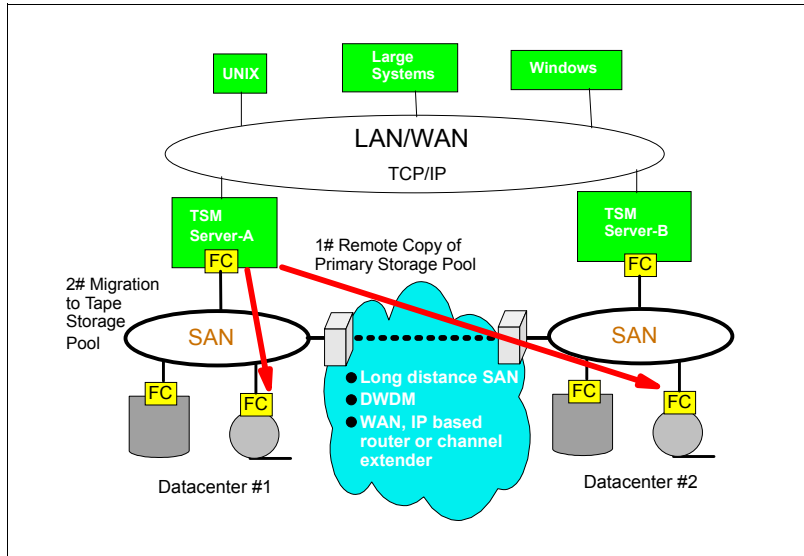


Figure 24-25 Tivoli Storage Manager and electronic tape vaulting

Depending on your configuration (and whether or not remote disk replication is being used in conjunction with electronic tape vaulting) you may choose to backup the Tivoli Storage Manager database and configuration files by this method. This ensures a copy of the data is stored at both sites and that the Tivoli Storage Manager server can rapidly recover at the remote site. If remote disk replication is used for mirroring of the Tivoli Storage Manager database and storage pools, the Tivoli Storage Manager server could be recovered very quickly without any loss of client data. A peer-to-peer configuration could be used to balance the load of Tivoli Storage Manager services in the enterprise and provide data protection and rapid recovery for a failure at either site. Some advantages for using electronic tape vaulting with Tivoli Storage Manager include:

- ▶ Critical data can be frequently and rapidly vaulted to remote off-site locations.
- ▶ In the event of a disaster, up-to-date data can be restored at a hot-site, therefore improving recovery time and recovery point objectives.
- ▶ Eliminates physical tape handling which could result in: damaged tapes, lost tapes, tapes delayed in transit, or data that is sabotaged. Increased reliability.
- ▶ Eliminates costs associated with couriers and off-site vaulting vendors.
- ▶ Government off-site vaulting regulations are satisfied.
- ▶ Lowers cost of downtime and storage management.
- ▶ Increases company control and data security.
- ▶ Peer solutions eliminate or reduce costs associated with hot-site service providers.

24.9 Remote disk mirroring and tape vaulting solutions

A variety of technologies exist for remote disk mirroring and electronic tape vaulting, these include. These include:

- ▶ Long distance SANs
- ▶ Dense Wavelength Division Multiplexing (DWDM)
- ▶ Fibre extenders
- ▶ WAN based channel extension using telco and IP protocols
- ▶ Newly emerging NAS and iSCSI gateways

Table 24-1 summarizes some of these. The use of these various technologies also may depend on a particular vendor's replication or vaulting solution. For example, the IBM System Storage DS6000 and DS8000 use PPRC to achieve data replication. Metro Mirror is supported via ESCON links, which can be further extended via DWDM or WAN channel extension.

Table 24-1 Extended electronic vaulting technologies

Electronic vaulting technology	Commonly supported distances between sites	Common product vendors	Relative technology costs
Long Distance SAN (shortwave/longwave Fibre Channel)	up to 11 km	Brocade (OEM) McDATA (OEM)	Low
DWDM (MAN) and fibre extenders	up to 180 km	Cisco Nortel Ciena Finisar CNT	Medium
WAN, IP Based Routers and Channel Extension	up to 1000's km	CNT Cisco (OEM)	Low to High (depending on solution)

SANs were designed to overcome the distance limitations of other storage channel protocols, such as SCSI. Longwave laser GBICs available on most SAN hubs, switches and directors enable a transmission distance of up to 10 kilometers (11 kilometers when including switch to host connections) when used with 9 micron diameter single-mode optical fibre. Shortwave GBICs use multi-mode fibre and is the ideal choice for shorter distance (less than 500 meters from transmitter to receiver or vice versa).

DWDM is a way to open up the conventional optical fibre bandwidth by breaking it up into many channels, each at a different optical wavelength (a different color of light). Each wavelength can carry a signal at any bit rate less than an upper limit defined by the electronics, typically up to several gigabits per second. DWDMs are implemented in areas that have dark fibre available through telcos and service providers. The DWDM is deployed as part of the physical layer. It is therefore independent of protocol, simply passing signal information in the format it is received. Examples of the protocols it can support are ATM, Gigabit Ethernet, ESCON, FICON and Fibre Channel.

WAN and IP based channel extenders typically use telecommunication lines for data transfer and therefore enable application and recovery sites to be located longer distances apart. The use of WAN and IP channel extenders provides the separation for disaster recovery purposes and avoids some of the barriers imposed when customers do not have a "right of way" to lay their fibre cable. WAN and IP channel extenders generally compress the data before sending it over the transport network, however the compression ratio needs to be determined based on the application characteristics and the distance.

Network attached storage (NAS) and iSCSI solutions are beginning to offer low cost IP based storage. Copies of Tivoli Storage Manager storage pools and the Tivoli Storage Manager database can be storage at a remote site using IP based storage to offer a low cost implementation while utilizing existing infrastructure. Configurations can include Tivoli Storage Manager clients attached to iSCSI based data backing up to a Tivoli Storage Manager server or Tivoli Storage Manager servers using iSCSI based storage as storage pools.

For a detailed overview of technologies, products, costs and best practices with distance solutions we recommend you review *Introduction to SAN Distance Solutions*, SG24-6408.

24.9.1 Collocation considerations for off-site vaulting

With collocation, large numbers of files belonging to a client node can be restored, retrieved and recalled more quickly. However, using collocation on copy storage pools requires special consideration. Primary and copy storage pools perform different recovery roles. Normally you use primary storage pools to recover data to clients directly. You use copy storage pools to recover data to the primary storage pools. In a disaster where both clients and the server are lost, the copy storage pool volumes will probably be used directly to recover clients. The types of recovery scenarios that concern you the most will help you to determine whether to use collocation on your copy storage pools.

You may also want to consider that collocation on copy storage pools will result in more partially filled volumes and probably increased off-site reclamation activity. Collocation typically results in a partially filled sequential volume for each client or client file space. This may be acceptable for primary storage pools because these partially filled volumes remain available and can be filled during the next migration process. However, for copy storage pools this may be unacceptable because the storage pool backups are usually made to be taken off site immediately. If you use collocation for copy storage pools, you will have to decide between:

- ▶ Taking more partially filled volumes off site, thereby increasing the reclamation activity when the reclamation threshold is lowered or reached.
- ▶ Leaving these partially filled volumes onsite until they fill and risk not having an off-site copy of the data on these volumes.

With collocation disabled for a copy storage pool, typically there will be only a few partially filled volumes after storage pool backups to the copy storage pool are complete. Consider carefully before using collocation for copy storage pools. Even if you use collocation for your primary storage pools, you may want to disable collocation for copy storage pools. Or, you may want to restrict collocation on copy storage pools to certain critical clients, as identified by the Business Impact Analysis.

24.9.2 Reclamation considerations for off-site vaulting

Space on a sequential volume becomes reclaimable as files expire or are deleted from the volume. For example, files become obsolete because of aging or limits on the number of versions of a file. In reclamation processing, the Tivoli Storage Manager server rewrites files on the volume being reclaimed to other volumes in the storage pool, making the reclaimed volume available for reuse.

When an off-site volume is reclaimed, the files on the volume are rewritten to another copy storage pool volume which is onsite. The Tivoli Storage Manager server copies valid files contained on the off-site volumes being reclaimed, from the original files in the primary storage pools. In this way, the server can reclaim off-site copy storage pool volumes without having to recall and mount these volumes. Logically, these files are moved back to the onsite location. The new volume should be moved off site as soon as possible. However, the files have not been physically deleted from the original off-site volume. In the event of a disaster occurring before the newly written copy storage pool volume has been taken off site, these files could still be recovered from the off-site volume, provided that it has not already been reused and the database backup that you use for recovery references the files on the off-site volume.

The server reclaims an off-site volume which has reached the reclamation threshold as follows:

1. The server determines which files on the volume are still valid.
2. The server obtains these valid files from a primary storage pool, or if necessary, from an onsite volume of a copy storage pool.
3. The server writes the files to one or more volumes in the copy storage pool and updates the database. If a file is an aggregate file with unused space, the unused space is removed during this process.
4. A message is issued indicating that the off-site volume was reclaimed.
5. The newly written volumes are then marked to be sent off site, and after this has occurred, the reclaimed volume can be returned to an onsite scratch pool.

Volumes with the access value of offsite are eligible for reclamation if the amount of empty space on a volume exceeds the reclamation threshold for the copy storage pool. The default reclamation threshold for copy storage pools is 100%, which means that reclamation is not performed.

If you plan to make daily storage pool backups to a copy storage pool, then mark all new volumes in the copy storage pool as offsite and send them to the off-site storage location. This strategy works well with one consideration — if you are using automatic reclamation (the reclamation threshold is less than 100%). Each day's storage pool backups will create a number of new copy storage pool volumes, the last one being only partially filled. If the percentage of empty space on this partially filled volume is higher than the reclaim percentage, this volume becomes eligible for reclamation as soon as you mark it offsite. The reclamation process would cause a new volume to be created with the same files on it. The volume you take off site would then be empty according to the Tivoli Storage Manager database. If you do not recognize what is happening, you could perpetuate this process by marking the new partially filled volume offsite.

If you send copy storage pool volumes off site, we recommend that you control copy storage pool reclamation by using the default value of 100. This turns reclamation off for the copy storage pool. You can start reclamation processing at desired times by changing the reclamation threshold for the storage pool.

Depending on your data expiration patterns, you may not need to do reclamation of off-site volumes each day. You may choose to perform off-site reclamation on a less frequent basis. For example, suppose you send copy storage pool volumes to and from your off-site storage location once a week. You can run reclamation for the copy storage pool weekly, so that as off-site volumes become empty they are sent back for reuse.

When you do perform reclamation for off-site volumes, the following sequence is recommended:

1. Back up your primary storage pools to copy storage pools.
2. Turn on reclamation for copy storage pools by lowering the reclamation threshold below 100%.
3. When reclamation processing completes, turn off reclamation for copy storage pools by raising the reclamation threshold to 100%.
4. Mark any newly created copy storage pool volumes as offsite and then move them to the off-site location.

This sequence ensures that the files on the new copy storage pool volumes are sent off site, and are not inadvertently kept onsite because of reclamation.

Attention: If collocation is enabled and reclamation occurs, the server tries to reclaim the files for each client node or client file space onto a minimal number of volumes.

24.10 More information

- ▶ More information about Tivoli Storage Manager can be found on the following Web site:
<http://www.ibm.com/software/tivoli/products/storage-mgr/>
- ▶ These Redbooks give good insight into Tivoli Storage Manager as well:
IBM Tivoli Storage Manager Implementation Guide, SG24-5416
IBM Tivoli Storage Management Concepts, SG24-4877
Disaster Recovery Strategies with Tivoli Storage Management, SG24-6844
IBM Tivoli Storage Manager in a Clustered Environment, SG24-6679
- ▶ Product manuals are:
Tivoli Storage Manager Windows Backup-Archive Client Installation and User's Guide, GC32-0788
Tivoli Storage Manager V5.3 for Windows Administrator's Guide, GC32-0782



TotalStorage Expert

TotalStorage Expert is a tool that gives administrators powerful yet flexible storage asset, capacity, and performance management capabilities to centrally manage Enterprise Tape Library systems located anywhere in the enterprise.

Using the information provided by the IBM TotalStorage Expert helps you get the maximum potential from your tape subsystems.

Note: The IBM TotalStorage Expert V2.1 ESS components were withdrawn from marketing on August 12, 2005. It is replaced by IBM TotalStorage Productivity Center. Existing customers may trade up to TotalStorage Productivity Center - see your IBM representative for details.

25.1 Overview

IBM TotalStorage Expert is an application in the IBM System Storage software family which helps you manage your Enterprise Tape Library (ETL) using a Web browser based user interface.

The two features can be separately licensed to accommodate users who may only be interested in one component.

The TotalStorage Expert is designed to augment commonly used IBM host-based performance tools - while these tools provide performance statistics from the host system's perspective, the TotalStorage Expert provides statistics from the ETL system perspective.

The ETL Expert provides performance, asset, and capacity management for three ETL solutions; IBM TotalStorage 3494 Tape Library, IBM TotalStorage 3494 Virtual Tape Server, and IBM TotalStorage 3494 Peer-to-Peer Virtual Tape Server.

The ETL Expert uses a database to store performance collected data, and a Web browser interface to that tape library resources can be efficiently monitored from any location within the enterprise.

The ETL Expert provides capabilities for performance management, asset management, and capacity management. The information provided within these categories varies significantly because of the different characteristics of disk and tape products, and are described in separate sections in order to differentiate them clearly.

ETL Expert uses a DB2 database to store data, and is available to run on Windows 2000 Server and Advanced Server, as well as AIX.

25.2 Key features

This section summarizes the key features of the IBM TotalStorage Expert.

25.2.1 Performance management

In your role as a system administrator, you may need to analyze performance information — for applications you maintain, or when an end user feels that something has gone wrong with an application's performance. In this case, you will need to determine which parts of your installation may be degrading performance; for example: application programs, database management systems, host processors, or I/O subsystems.

Enterprise Tape Libraries are much more than a collection of physical tape drives. This is particularly true with the IBM TotalStorage 3494 Virtual Tape Server, and IBM TotalStorage 3494 Peer-to-Peer Virtual Tape Server. These are complexes of intelligent controller, disk cache, and physical tape library to emulate tape drives at the front end, and use reliable, high performance tape drives along with high capacity tape media at the back end. So you need different performance indexes from what you used to have for native tape drives. Just taking the physical specification of tape drives into account will not help monitor and analyze performance.

The TotalStorage Expert can gather performance information for all kinds of ETLs — the IBM TotalStorage 3494 Tape Library, IBM TotalStorage 3494 Virtual Tape Server, and IBM TotalStorage 3494 Peer-to-Peer Virtual Tape Server. This is especially useful for the VTS. Since the TotalStorage Expert shows you the VTS activity along with the back end native

drive activity, you can correlate this information to determine how well your VTS is working and to help perform capacity planning.

These performance statistics have been available for OS/390 or z/OS as System Management Facility (SMF) records. The ETL generates performance statistics every hour and OS/390 or z/OS can store the information.

The TotalStorage Expert allows you to manage an ETL's performance no matter what kind of host platform is used. Furthermore, it can provide more up-to-date status information and statistics than SMF records.

25.2.2 Asset management

A storage server environment is dynamic in regard to the hardware and software incorporated in it. To ensure the administrator has a comprehensive understanding of all components in the enterprise, TotalStorage Expert has asset tools to provide a comprehensive view. This information can be used for resource planning, budgetary considerations, and maintenance.

The Enterprise Tape Library (ETL) consists of three kinds of libraries, VTS, Library, or Composite. From a host system's viewpoint, each of these can be seen as a tape library, but its physical configuration is very different from the others. For example, the VTS is comprised of the VTS controller and its back-end native tape library. You may want to review which VTS uses which native tape library. Or, you may want to review how many physical tape drives in a native library are allocated for a VTS's use.

The TotalStorage Expert shows you the relationship between libraries, and you can manage the configuration while reducing the overhead of maintenance.

25.2.3 Capacity management

The TotalStorage Expert shows capacity information for the VTS, such as, how much data is active in written tape cartridges, and how many empty tape cartridges are available for backing logical volumes. This number might reflect multiple versions of logical volumes, therefore it might exceed the number of defined logical volumes. In this comprehensive reporting feature, you can determine the average, total, and maximum logical volume sizes. This information is important not only from a capacity planning perspective, but also from the VTS performance viewpoint, since the tape cartridge utilization has a tight relationship with the VTS's back-end operation.

To enable you to manage your VTS free storage capacity on the subsystem, a throttling value can be custom defined. The alarm level can be set at the Library Manager console.

Other valuable ETL-VTS capacity management report features include; stacked volumes managed by the VTS, reclaim threshold percentage currently set for the VTS, and backstore compression ratio which measures the data moved from the Tape Volume Cache to the drive. The ETL-VTS capacity management reporting allows for customization and provides you with necessary tape library and server details for current resource utilization in addition to future growth and usage considerations.

25.3 How IBM TotalStorage Expert works

Figure 25-1 on page 478 shows how the TotalStorage Expert works with your ETL. The TotalStorage Expert communicates with your ETL through the TCP/IP network. Therefore, you can gather information on any ETL around the world as long as you have a

communication path between the TotalStorage Expert and the ETL through your intranet or the World Wide Web (Internet).

Note: Due to built-in and customized constraints, TotalStorage Expert implementations across a firewall between the Expert host(s) and the ESS and ETL is not recommended.

The TotalStorage Expert itself is installed in a Windows 2000 Server or AIX V4.3.3 or 5.1 operating system environment. However, you do not have to operate the ESS Expert right where it runs, since the user interface of the TotalStorage Expert is the Web browser. In other words, you can operate the Expert through Netscape Navigator or Microsoft Internet Explorer from any machine which has network access to the machine where the Expert is installed.

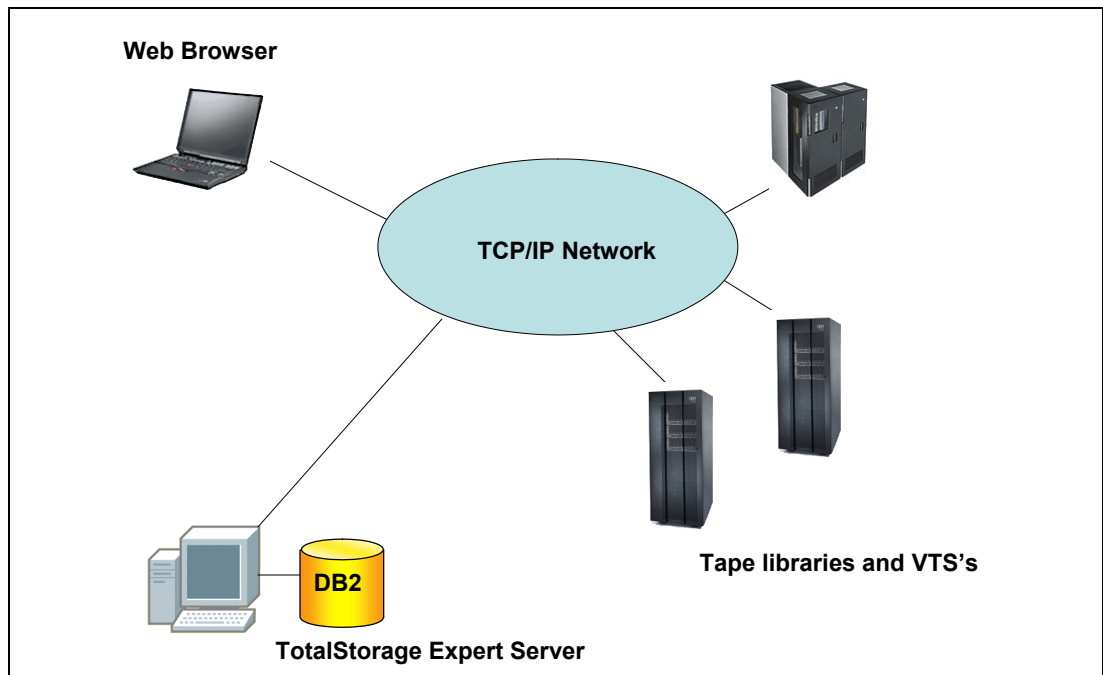


Figure 25-1 IBM TotalStorage Expert sample configuration

The TotalStorage Expert solicits your ETL's to send information about their capacity or performance. When the TotalStorage Expert receives this information, it is stored into tables within a DB2 database. Therefore, you can prepare and produce customized reports, in addition to the built-in Expert reports, containing just the information you need using traditional DB2 commands.

IBM TotalStorage Expert uses the security and auditability features of the host operating system. The host administrator is responsible for evaluation, selection, and implementation of security features, administrative procedures, and appropriate controls in application systems and communication facilities.

25.4 Planning and implementation considerations

This section provides information about planning and implementing your TotalStorage Expert environment including network, operating system co-requisite software requirements.

Operating system requirements

The IBM TotalStorage Expert can run under AIX or Windows 2000 Server, and will typically be run on an AIX or Windows 2000 Server that you already have; the Expert does not have to be on a dedicated server. Both AIX and Windows 2000 require TCP/IP.

Product packaging

The following materials are shipped with IBM TotalStorage Expert:

- ▶ IBM TotalStorage Expert Installation CD for AIX and Windows NT
- ▶ IBM TotalStorage Expert Installation Guide
- ▶ IBM TotalStorage Expert License Information

Co-requisite products

The TotalStorage Expert requires and installs five other software products as its co-requisites.

- ▶ IBM WebSphere Application Server
- ▶ IBM HTTP Server
- ▶ Viador Sage
- ▶ IBM DB2 UDB
- ▶ IBM Developer Kit for the Java Platform

All the software installs at one time, with a “one-button install” process which is designed for systems that do not already have DB2 or a Web-server installed.

25.5 More information

For more information about the IBM TotalStorage Expert refer to this Web site:

<http://www.ibm.com/servers/storage/software/expert/index.html>

These Redbooks provide detailed information about using the IBM TotalStorage Expert:

- ▶ *IBM TotalStorage Expert Hands-On Usage Guide*, SG24-6102
- ▶ *IBM TotalStorage Expert Reporting: How to Produce Built-In and Customized Reports*, SG24-7016



Part 5

z/OS storage management

Part 5 describes these products:

- ▶ DFSMS
- ▶ DFSORT



Data Management with DFSMS

DFSMS is not a single entity, it consists of one base and four optional products, which together can automatically manage data from creation to expiration. This chapter provides you with a brief overview of the DFSMS family which provide storage management for z/OS.

26.1 Managing Enterprise data

All data has a lifecycle, whether it is a temporary work file or customer information that regulatory laws require to be kept for several years. There is more to managing data than simply creating and deleting a data set and the DFSMS group of products support all the stages involved.

26.2 The components of DFSMS

DFSMS is made up of the following components:

- ▶ **DFSMSdfp** (Data Facility Product), a base element of z/OS: Provides storage, data, program, and device management functions and DFSMS Copy Services capabilities.
- ▶ **DFSMSdss™** (Data Set Services): An optional feature of z/OS providing data movement, copy, backup, and space management functions.
- ▶ **DFSMShsm™** (Hierarchical Storage Manager): An optional feature of z/OS providing backup, recovery, migration, and space management functions.
- ▶ **DFSMSrmm** (Removable Media Manager): An optional feature of z/OS providing management functions for removable media.
- ▶ **DFSMSstvs** (Transactional VSAM Services): An optional feature of z/OS that enables batch jobs and CICS® online transactions to update shared VSAM data sets concurrently.

26.3 DFSMSdfp (Data Facility Product)

DSDFSMSdfp performs the essential data, storage, and device management functions of the system, providing the methods that allow data to be read from and written to storage devices. As well as providing data access to devices in the native z/OS applications, it also supports data on other platforms such as AIX/UNIX, Windows, and OS/400.

26.3.1 Organizing and accessing data with DFSMSdfp

Data is stored in the form of a record, the size of which will depend on the type of data to be processed. For example records containing customer names and addresses will vary in size, but records containing the source code of a program will be of a fixed size.

Space on DASD volumes is used most efficiently in large sizes, which also improves the performance of the transfer of data to and from the volume. Therefore small data records are generally grouped together to form a block, the optimum size of which will depend on whether the data is to be processed randomly or sequentially. Data sets on tape can only be processed efficiently sequentially and large block sizes are the most efficient.

In common with the rest of z/OS, DFSMSdfp has evolved over some 40 years and is a mix of old and new data set organizations and techniques (access methods) for processing them.

Traditional data set organizations:

- ▶ **Sequential data set:** the simplest form where data is written in the order in which the program supplies it.

- ▶ **Partitioned data set (PDS):** this consists of a collection of, usually, small sequential elements (members) which have a common characteristic, such as being program source data. The individual elements are stored in a single data set and are located by a directory.
- ▶ **Direct access data set (BDAM):** the data is stored so that the contents of a block decide its position in the data set, which allows any block to be located and retrieved quickly.

Traditional access methods

- ▶ **Basic Sequential Access Method (BSAM)**

This method requires the user to group records into blocks for writing and to extract records from blocks when reading. The program controls when data is read or written and gives fine control when several data sets are being read concurrently.

- ▶ **Queued Sequential Access Method (QSAM)**

This method takes the records provided by the program and automatically puts them into blocks for writing and supplies records when reading. When reading QSAM will read blocks from the data set in advance of the program requesting data, which means that the program does not have to wait and elapsed times are reduced.

- ▶ **Basic Partitioned Access Method (BPAM)**

This method provides the support for a PDS. Once a member is located it can be processed by BSAM or QSAM

Modern data set organizations and access methods

- ▶ **Virtual Storage Access Method (VSAM)**

VSAM makes use of newer techniques to improve the performance of reading and writing data and provides support for the following four data set organizations. In the first three, the data records are stored in an internal structure with control information which is not known to the program.

- **Key-sequenced data set (KSDS)**

A piece of data that can be used to identify a specific record is known as a key, e.g. an employee personnel number which is unique to each record. Data that contains keys and is required to be processed randomly, such as by a Human Resources department, can be stored in a key-sequenced data set (KSDS). The data set is divided into two parts - index and data. The data part contains the complete records and the key part contains the keys and pointers to the records containing the keys.

- **Entry-sequenced data set (ESDS)** is a sequential data set.

- **Relative-record data set (RRDS)** is the equivalent of a BDAM (basic direct access method) data set.

- **Linear data set (LDS)** - this format does not contain any control information and it is up to the program to keep track of where records are stored.

- ▶ **Object Access Method (OAM)**

OAM processes named byte streams (objects) that have no record boundary or other internal orientation that the system maintains. These objects are recorded either on DASD in a DB2 database, or on an optical or tape storage volume.

Hierarchical File System

DFSMS provides access to enterprise data in an open system environment. You can use the standard BSAM, QSAM, and VSAM access methods to access data in hierarchical file system (HFS) files. The HFS data set contains the HFS file structure. This structure is a framework of directories and HFS files called a file system. The structure resembles a tree

with subtrees, each consisting of a directory and all its related files. The HFS files are identified and accessed by specifying the path leading to them.

zFS

The z/OS Distributed File Service zSeries File System (zFS) is a z/OS UNIX file system that can be used in addition to the Hierarchical File System (HFS). The zFS file systems contain files and directories, including Access Control Lists (ACLs), that can be accessed with the z/OS HFS Application Programming Interfaces (API's). zFS file systems can be mounted into the z/OS UNIX hierarchy along with other local (or remote) file system types (for example, HFS, bTFS, AUTOMNT, NFS, etc.). zFS does not replace HFS, it is complementary to HFS. HFS is still required for z/OS installation and the root file system must be HFS.

26.3.2 Locating data sets

The catalog

Manual tracking of data sets in a large data centre, which may have thousands of volumes containing hundreds of thousands of data sets, is not an option. When a data set is created an entry is made in a system owned data set, known as a catalog.

Catalogs are in KSDS format; the data set name is the key pointing to a record which contains the identification (volume serial number) of the volume, or volumes on which the data set has data. Once the volume has been identified, details of the data set can be found in the Volume Table of Contents (VTOC), on the volume. There are two types of catalogs, *Master*, of which there is only one in a z/OS system and *user*, of which there will be several. User catalogs are located via pointers in the Master catalog.

The VTOC

Each DASD volume contains a VTOC, this data set contains the names, characteristics and details of the space used by all of the data sets on the volume.

Tape libraries and volumes

Tape libraries contain tape drives and thousands of tapes, which can only be mounted on those tape drives. Therefore the catalog information for tape data sets also has to indicate the library that the tape volume is stored in. To achieve this a special catalog is created which can only contain entries for tape volumes and tape libraries.

26.4 System Managed Storage

The historical method of allocating data sets requires the user to make decisions about their placement and physical characteristics. The relevant parameters then had to be coded in Job Control Language (JCL). This is an inefficient use of a programmer's time and does not necessarily lead to the best positioning of data sets and the optimum use of DASD resources. The Storage Management Subsystem (SMS) was therefore introduced as the basis for an optional method of imposing installation standards to data sets.

The installation standards are defined in definitions, known as constructs, which are used as input to user written Automatic Class Selection (ACS) routines. The constructs are held in a Control Data Set (CDS).

The data set name is the key element to the processing performed by an ACS routine and the name alone can be sufficient to determine all of the requirements for a data set. The ACS

routines allow the installation to place data sets with specific requirements, such as space and performance, on the most appropriate volumes.

26.4.1 Parallel Sysplex

In a Parallel Sysplex environment the SMS functions on each system share the CDS and form an SMSplex.

26.4.2 Naming conventions

As stated previously, the processing in a SMS environment is driven by the name of the data set, therefore data set names have to be consistent. A data set owned by an online system, such as DB2, has different requirements to one being used by a program developer and the ACS routines must be able to identify which is which.

26.4.3 Extended Format data sets

There are options which apply only to data sets which are created by SMS: Extended Addressability, compression and striping. Data sets which have any of these attributes are referred to as being in Extended Format (EF). The benefits of these attributes are as follows:

- ▶ Extended Addressability removed architectural limitations on the size of VSAM and sequential data sets
- ▶ Compression can significantly reduce the data stored on DASD and therefore increase utilization
- ▶ Striping is a performance option for large sequential data sets. Tracks of data are split across several volumes (stripes) and as many I/O operations can run concurrently as there are stripes.

26.4.4 Constructs

There are four constructs: data classes, storage classes, management classes and storage groups, which are defined using the Interactive Storage Management Facility (ISMF).

Data Classes

A data class simplifies JCL coding as it can supply all of the space allocation and data set characteristics. It is possible to define a dataclass such that a user only has to specify a data set name. This construct also specifies whether one or both of Extended Addressability and compression apply to the data set. Some other benefits are:

- ▶ Use allocation defaults, specifying allocation values that make efficient use of storage
- ▶ Retry data set allocations on new volumes when allocation fails due to space constraints. Retry allocation requests that fail, because there is not enough space available, with a smaller request

Storage classes

A storage class defines the performance and availability goals required by the data set. SMS uses it to select a device that can best meet those goals and requirements in terms of the performance of the device, and how available a data set or object can be on that device

The availability options are dual copy, and RAID versus non-RAID. The second option does not really apply these days as all modern DASD subsystems are RAID. The dual copy option

however is becoming more and more common as companies need rapid recovery of systems in 24 x 7 environments

The class can specify the response times required for sequential or direct processing. This will direct the data set to a subsystem with sufficient caching capability. The class can also specify that PAVs (parallel access volumes) are required for a data set. Striping is specified in this class as it is performance related.

Some other benefits are:

- ▶ Reduce necessity for user awareness of physical device characteristics.
- ▶ Provide for the separation of physical device characteristics from a data set's logical requirements for performance, availability, and space.
- ▶ Designate tape data sets for placement in tape libraries.

Management classes

Data sets need to be managed for several reasons, including:

- ▶ Deletion when they are no longer required (retention period)
- ▶ Transfer to a less expensive media if data is not referenced, but still retain the data (migration)
- ▶ Copy the data for recovery purposes (backup)

A management class defines the time periods for these actions to be taken and the method of performing them. For data sets, these attributes control retention, migration, backup, and release of allocated but unused space. For objects, the attributes control retention, backup, and class transition. The processing associated with many of the attributes in this class is performed by DFSMSHsm.

As these functions are performed automatically, the effort users and the storage administrator require for storage management is reduced. It also means that more effective use can be made of storage devices.

Storage groups

A storage group is a collection of volumes to meet a specific service or business strategy. If a department has to pay for its use of storage then a storage group can limit the number of volumes that the users in the department have access to. A storage group used by online services might have a large amount of free space to make sure that data sets can extend without running into problems.

Storage groups are also used to specify which SMS-managed volumes should be processed by the DFSMSHsm functional component of DFSMS and the systems on which they should be processed.

Some of the benefits of a storage group include:

- ▶ Provide extended storage groups that can be used during extend processing if the currently allocated storage group does not have sufficient space.
- ▶ Provide overflow storage groups that can be used during primary space allocation if the non-overflow storage groups are getting too full.
- ▶ Allow data set separation so that critical data is spread over different physical control units (PCU), thereby reducing the impact of PCU failure.
- ▶ Simplify device installation and hardware configuration management. An administrator can add or delete devices through the ISMF Storage Group Application.

- ▶ Simplify data set allocation for users, because they do not have to provide unit and volume information during allocation.
- ▶ Improve the balance of I/O activity across devices.

26.4.5 ISMF

The user interface to setup, display and control the actions of DFSMS is provided by the Interactive Storage Management Facility (ISMF). ISMF is an Interactive System Productivity Facility (ISPF) application and provides access to the functions of the DFSMS family of products, as represented in Figure 26-1.

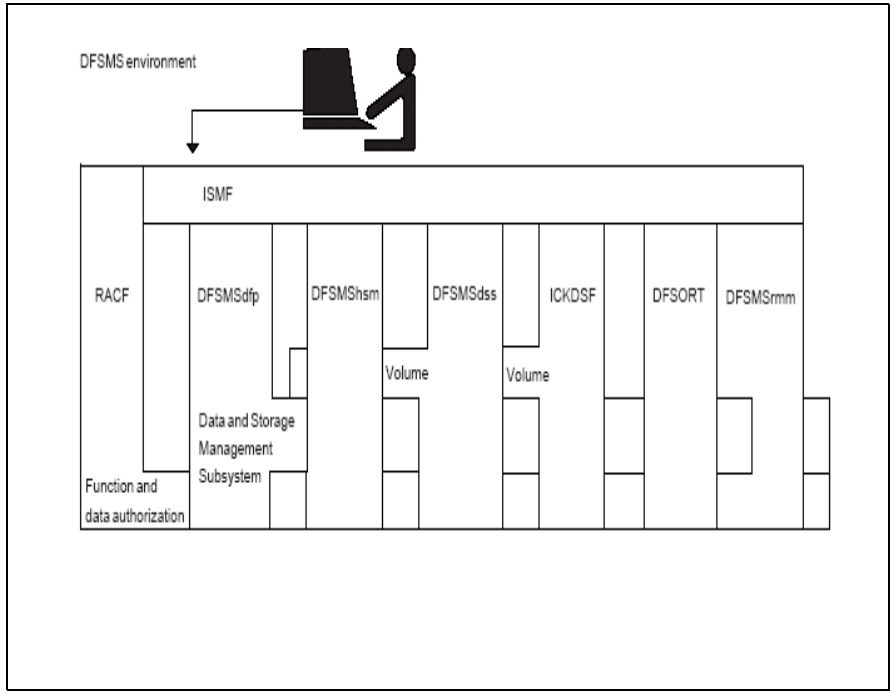


Figure 26-1 DFSMS environment

ISMF is a panel-driven interface to the following tasks:

- ▶ Display and print lists of information about specific data sets, DASD volumes, mountable optical volumes, and mountable tape volumes
- ▶ Generate lists of data, storage, and management classes to find out how data sets are being managed
- ▶ Display and manage lists saved from various ISMF applications

To determine which data sets will appear in a data set list or which volumes will appear in a volume list, you complete selection entry panels. Figure 26-2 on page 490 shows an example of the Data Entry Panel. ISMF generates a list based on your selection criteria. Once the list is built, you can use ISMF entry panels to perform space management or backup and recovery tasks against the entries in the list.

```

Panel Defaults Utilities Scroll Help
-----
                        DATA SET SELECTION ENTRY PANEL                Page 1 of 5
Command ==>>>

For a Data Set List, Select Source of Generated List . . 2 (1 or 2)

 1 Generate from a Saved List          Query Name To
   List Name . .                      Save or Retrieve

 2 Generate a new list from criteria below
   Data Set Name . . . **
   Enter "/" to select option          Generate Exclusive list
   Specify Source of the new list . . 2 (1 - VTOC, 2 - Catalog)
 1 Generate list from VTOC
   Volume Serial Number . . .         (fully or partially specified)
 2 Generate list from Catalog
   Catalog Name . . .
   Catalog Password . . . .          (if password protected)
   Volume Serial Number . . .         (fully or partially specified)
   Acquire Data from Volume . . . . . N (Y or N)
   Acquire Data if DFSMSHsm Migrated . . N (Y or N)
Use ENTER to Perform Selection; Use DOWN Command to View next Selection Panel;
Use HELP Command for Help; Use END Command to Exit.

```

Figure 26-2 Example of an ISMF Data Entry Panel

As a user performing data management tasks against individual data sets or against lists of data sets or volumes, you can use ISMF to:

- ▶ Edit, browse, and sort data set records
- ▶ Delete data sets and backup copies
- ▶ Protect data sets by limiting their access
- ▶ Recover unused space from data sets and consolidate free space on DASD volumes
- ▶ Copy data sets or DASD volumes to the same device or another device
- ▶ Migrate data sets to another migration level
- ▶ Recall data sets that have been migrated so that they can be used
- ▶ Back up data sets and copy entire volumes for availability purposes
- ▶ Recover data sets and restore DASD volumes, mountable optical volumes, or mountable tape volumes

ISMF also works with NaviQuest, allowing more automation of storage management tasks. NaviQuest is a testing and reporting tool that speeds and simplifies the tasks that are associated with DFSMS initial implementation and ongoing ACS routine and configuration maintenance.

26.4.6 Data management utilities

DFSMSdftp offers the following data management utility programs to perform a variety of tasks, such as moving or copying data.

- ▶ **IDCAMS** creates and maintains VSAM data sets.
- ▶ **IEBCOMPR** compares logical records in sequential or members of partitioned data sets.
- ▶ **IEBCOPY** copies and merges partitioned data sets.
- ▶ **IEBDG** creates a pattern of test data for aid in debugging programs.
- ▶ **IEBEDIT** edits jobs and job steps into a single output data set.
- ▶ **IEBGENER** copies records from a sequential data set or converts sequential data sets into members of partitioned data sets.

- ▶ **IEBTPCH** punches or prints records from sequential or partitioned data sets.
- ▶ **IEBUPDTE** changes source language statements in sequential or partitioned data sets.
- ▶ **IEHINITT** labels tapes in stand-alone devices and automated tape libraries.
- ▶ **IEHLIST** lists entries in catalogs, partitioned data set or PDSE directories, or volume tables of contents (

26.5 DFSMSdss (Data Set Services)

DFSMSdss is a utility which can be used to quickly move, copy and backup data.

26.5.1 Overview

DFSMSdss is an optional component of DFSMS, the primary function of which is to move and copy data. It can operate at both the logical and physical level and can move or copy data between volumes of like and unlike device types. Its other features include:

- ▶ **DEFRAG:** The space occupied by a data set is made up by one or more extents, each extent being contiguous space on the volume. There is a limit to the number of extents that an individual data set is allowed to have the space allocation, therefore it is best if a space request is satisfied by a single extent. As data sets are created and deleted on volumes then small areas can become isolated and requests for more space receive more than one extent. The DEFRAG function moves data sets, that are not in use, and thus can reduce or eliminate the free-space fragmentation.
- ▶ **Standalone program:** A standalone Restore Program that enables vital system packs to be restored, during disaster recovery - without a host operating system
- ▶ **Convert data sets and volumes:** DFSMSdss can convert data sets and volumes to be SMS managed, or return them to a non SMS managed state, as part of a recovery procedure - all without data movement

26.5.2 Interface to other services

DFSMSdss can make use of two features of the IBM Enterprise Storage Server (ESS), including DS6000 and DS8000, which are:

- ▶ **FlashCopy:** which is a point-in-time copy function that can quickly copy data from a source location to a target location.
- ▶ **Concurrent Copy (CC):** which is a copy function that generates a copy of data while applications are updating that data.

DFSMSdss does not communicate directly with the ESS to use these features, this is performed by a component of DFSMSdss, the System Data Mover (SDM).

26.6 DFSMShsm (Hierarchical Storage Manager)

DFSMShsm is an optional component of DFSMS that provides facilities for the management of data sets.

26.6.1 Overview

In a large data centre with many thousands of data sets it is not possible to manually identify data sets which little or no activity. DFSMShsm provides automatic management of such data

sets and also provides automatic backup and recovery of active data in both SMS and non-SMS environments. DFSMSShsm accomplishes this by providing three types of management, Space, Automatic Storage and Availability

In a Parallel Sysplex environment there can be an instance of DFSMSShsm running on each system. These share the control data sets and form an Hsmplex.

26.6.2 Space management

DFSMSShsm helps manage storage space utilization using the following methods.

Data Hierarchy

DFSMSShsm has three levels at which data may exist, 0, 1, and 2. A user can only directly access data at level 0. To save on DASD space, and therefore cost, data sets that have low activity, or are inactive can be moved (migrated) to DFSMSShsm owned DASD or tape volumes. DFSMSShsm owned DASD volumes are known as migration level 1 (ML1) and tape volumes are known as migration level 2 (ML2). Although ML1 data is on DASD it is stored in a space-saving format and is therefore not directly accessible by users - it must first be returned to a level 0 volume for use.

Migration

At a time specified by the installation, DFSMSShsm automatically manages DASD space by checking when data sets were last referenced. If this is within a user defined time period the data set is considered active and remains on level 0. If the data set is found to be inactive it is moved (migrated) to ML1. Although this is also DASD, the technique used for storing data is very efficient and uses less space than the original data set.

Data sets on ML1 are checked to see how long they have been there and if this meets an installation defined value the data set is migrated to ML2.

Data sets on ML2 are checked to see how long they have been there and if this meets an installation defined value the data set is deleted.

Although the usual process is to move first to ML1 and then to ML2, it is possible to set the parameters such that an inactive data set may be migrated directly to ML2.

Recall

Data sets that have been migrated have a catalog entry indicating that they are on a DASD volume with a serial number of MIGRAT. If a user or program attempts to reference a data set that has been migrated the system knows, from the catalog entry, the data set has to be returned from migration (recalled). Control is passed to DFSMSShsm and the data set is recalled. ML2 recall can take quite a long time, as a tape mount is involved (many installations use a VTS for ML2 and this makes the recall faster).

Control Data Set

The information about data that has been migrated is stored in the Migration Control Data Set (MCDS).

26.6.3 Availability management

DFSMSShsm also helps keep data available.

Backup

At a time specified by the installation DFSMSHsm checks to see whether data sets have been updated. If a data set has been updated then it can have a backup taken.

If a data sets are damaged or accidentally deleted, then it can be recovered from a backup copy. There can be more than one backup version, which assists in the recovery of a data set which has been damaged for some time, but this has only recently been detected.

Duplication of data

ML2 and backup tapes can have alternate tapes. The alternate tapes are produced by either duplicating the original tapes, or by using the duplex tape function to concurrently create a copy when the original ML2 or backup tapes are created. Alternate tapes give protection against media damage to the primary tape and can be stored off-site, to provide site disaster protection.

Control Data Set

The information about data that has been backed up is stored in the Backup Control Data Set (BCDS).

26.6.4 Tape environment

DFSMSHsm works with all major tape management systems. DFSMSHsm provides an internal interface to an IBM tape management system, DFSMSrmm, and provides a general exit as an interface to other tape management systems.

DFSMSHsm is generally used with a global or public scratch pool. That is, DFSMSHsm makes a request for the next available scratch tape, which it uses as ML2 or backup until all of the data sets on it are no longer required, then the tape is returned to the scratch pool. DFSMSHsm is also able to work with a private scratch pool when tapes are specifically identified for the private use of DFSMSHsm.

As data sets are recalled or deleted, tapes can have large gaps with no valid data. DFSMSHsm monitors the percentage of a tape in use and when it drops below a preset value the data is recycled. That is the data is copied to a new tape and the old tape is returned to the scratch pool.

DFSMSHsm takes advantage of the single-file format and the compaction algorithms available on cartridge-type devices, such as the IBM 3480, 3490, 3490E, and 3590-1. The single-file format allows for better performance and utilization of cartridges since they can contain hundreds or thousands of data sets on a single tape cartridge.

Because DFSMSHsm stores many data sets on a single cartridges, contention can arise for the use of the cartridge. Requests for migration and/or recall from one or more systems could require the same cartridge. DFSMSHsm gives priority to Recall, as active work is being delayed and will temporarily stop a migrate action in order to reposition the tape and satisfy the recall request.

Physical volume dumps can have up to five copies made concurrently and the dumps from multiple DASDs can be written to a single tape cartridge.

26.7 DFSMSrmm (Removable Media Manager)

DFSMSrmm is an optional component of DFSMS for managing removable media resources, including automatic libraries.

Introduction

In a large z/OS environment the control of removable tape media using manual methods is not a possibility. A pool of tapes which are available for general use (scratch) has to be maintained and tapes which are in use (private) have to be protected from being overwritten.

The traditional tape library, with tapes mounted on drives by operators, has essentially been replaced by automated and virtual tape libraries.

26.7.1 Overview

DFSMSrmm is primarily used for tape management, but it can manage any removable media, such as optical disk. DFSMSrmm can manage all of the tape volumes in an enterprise and the data sets on those volumes. It maintains a scratch pool and protects tape data sets from being accidentally overwritten. It also manages the movement of tape volumes between libraries and vaults over the life of the tape data sets, all according to policies that the installation defines. For optical disks the shelf location can be recorded but the data objects on the disks are not managed.

In addition to on-site libraries, off-site storage locations (for DR purposes), also known as vaults or stores are supported by DFSMSrmm.

DFSMSrmm has retention policies, which apply to both volumes and data sets. When a private volume reaches its expiration date the data sets on it are checked and if they have all expired the volume is returned to scratch status.

Volume movement and retention are fully automated interfaces exist for manual intervention if required, such as removing a volume from a library. This type of activity is protected by security controls to make sure that they are only carried out by authorized staff.

Volume management

Many of the volumes that DFSMSrmm manages in an installation will be scratch volumes. Each time a volume is used it is retained and managed by policies that the installation defines to DFSMSrmm. When the data is no longer required, the volume is returned to scratch status and is ready for use by another user.

DFSMSrmm records information for data sets on all files of a tape volume.

Validation

DFSMSrmm automatically validates volumes, ensuring that only valid scratch volumes are mounted and that the right volume is mounted for a private mount request. This validation eliminates the unintentional overwriting of data.

Location management

Tape media can exist in libraries, or in racks for manual mounting on drives by operators. In a VTS the media is virtual, with the real (back-end) tapes managed by the library code. DFSMSrmm keeps track of the location of real volumes, whether they are in the data centre or have been moved to another location, for example to a vault for disaster recovery use.

26.7.2 Policies for retention and movement

DFSMSrmm provides policy management for movement and retention at the data set level. Every tape data set can have a policy, and each policy can specify movement as well as retention. An example of this would be data that is required to be active for a period of time and then has to be archived. The policy would cause the movement of the data to a vault for long term storage.

26.7.3 Control Data Set

All data set, volume, and policy information is kept in the control data set (CDS), which is distinct from the one used by either DFSMS or DFSMSHsm. DFSMSrmm provides utilities to manage your inventory and create reports.

26.8 More information

For more information about DFSMS, refer to this Web site:

<http://www.storage.ibm.com/software/sms/index.html>

These Redbooks also provide more information:

- ▶ *DFSMSHsm ABARS and Mainstar Solutions*, SG24-5089
- ▶ *IBM TotalStorage Virtual Tape Server: Planning, Implementing, and Monitoring*, SG24-2229
- ▶ *z/OS V1R3 DFSMS Technical Guide*, SG24-6569
- ▶ *Z/OS V1R3 and V1R5 DFSMS Technical Guide*, SG24-6979



SORT products for the System z

IBM offers two optional, licensed products for the System z: DFSORT for z/OS and DFSORT/VSE for z/VSE™. Both provide high-performance sort, merge, copy, analysis, and reporting.

27.1 Overview

The two products, DFSORT for z/OS and DFSORT/VSE for z/VS, have many similarities and the following topics can be considered to apply to both unless otherwise stated.

Although these products are marketed as sort programs, they have far greater functionality than simply arranging data into a sequence. As well as simple tasks such as alphabetizing a list of names, complex tasks such as taking inventory or running a billing system can be performed. Sort programs manipulate data stored in data sets (files) and can perform various types of processing for all of the records of a data set, for a subset of the records of a data set, and for specific fields, bytes, and bits.

There are three basic functions - sorting, merging, and copying, which provide the basis for a host of features, some of which will be outlined in the following sections.

27.2 Performance considerations

Many organizations have a 'batch window', the time during which the overnight batch jobs have to perform processing related to the previous day's activity. As system capabilities and usage have increased, so too, the number of reports required and the amount of data to be processed by the batch jobs has increased. As the batch window cannot be extended, ways have to be found to process the data as quickly and efficiently as possible. As many of the batch jobs require data to be sorted, particularly those producing reports, improved sorting can make a contribution to the goal of processing more data in the same period of time.

The initial processing by the sort program is to read the data records to be sorted - these may be from a data set or obtained directly from a program. The final processing will be to write the (now sorted) data either to a data set or to a program interface. Between these two phases, the data records have to be stored, either in memory or in work files. From a performance perspective, the optimum number of I/O operations to work files is zero, that is all data records should be kept in memory.

27.3 DFSORT performance

The performance of DFSORT can be enhanced, where appropriate, by the use of the following functions:

- ▶ Extended Format data sets
- ▶ The ICETPEX interface
- ▶ Sorting in memory

27.3.1 Extended Format data sets

For a standard sequential file, only one I/O operation can be active at any time and the data is not compressed. Extended Format data sets have an internal structure which is optimized for I/O read and write performance and they can also be defined such that the data is compressed and striped. Compression replaces repeating character strings which means that more data can be fitted onto a disk track. Striping spreads the data records across more than one volume and data is read or written by as many concurrent I/O operations as there are stripes. The use of Extended Format can:

- ▶ Significantly reduce the elapsed time DFSORT spends reading and writing data
- ▶ Reduce the space requirements of very large data sets.

27.3.2 ICETPEX

In order to make decisions about the amount of space required to perform a sort, it is necessary to know how many records are to be sorted. With data sets on disk, this can be calculated, as the size of the data set can be obtained. The use of tapes managed by a tape management system that uses ICETPEX, such as DFSMSrmm, allows DFSORT to obtain accurate information about input file size and data set characteristics. This can result in improved performance and more efficient use of both main storage and intermediate work storage.

27.3.3 Memory sorting

While the performance of DFSORT, for very large sorts, will improve with the use of memory, remember that a z/OS environment will be running numerous tasks, possibly including other sorts, concurrently. Therefore, it is part of the design criteria that DFSORT should not use system resources to the extent where it affects the performance of other workloads.

All jobs in a z/OS environment run in an address space which, until the introduction of z/OS (e.g. with OS/390), had a maximum virtual storage size of 2^{31} (2 Gigabytes). If more storage was required, one or more Hiperspaces or Dataspaces could be used, both of which made another 2 GB available. The size of an address space was increased to 2^{64} by z/OS with the new memory addresses having the potential to be used as Memory Objects.

DFSORT can make use of Hiperspaces, Dataspaces or Memory Objects - these can be used in conjunction with work data sets on DASD, if necessary, but not with one another. All of these options will increase the amount of real storage being used by the job and DFSORT uses an interface to the Real Storage Manager (RSM™) component of z/OS to determine how much storage can be used without degrading the performance of the other workloads currently active.

DFSORT has more than one sorting technique; the preferred and most efficient is BLOCKSET. The following sections, on memory sorting, assume that this technique is being used.

Hiperspace

A Hiperspace™ dates back to when the maximum amount of central storage that the predecessors to z/OS supported was 2 GB. Additional storage (known as expanded), which was of slower performance than central, could be added and used as a work area for system functions. This gave better performance than DASD and programs could request that their data be stored in a Hiperspace, which has a maximum size of 2 GB. However all requests to store to or retrieve data from a Hiperspace are performed by a z/OS service routine. All of the currently marketed System z software and hardware supports 64-bit addressing of central storage and expanded storage no longer exists. It is still possible to use the facility for compatibility reasons, but it is emulated and, although it reduces I/O processing, which in turn reduces elapsed time and channel usage, it is not the best of the three options.

Dataspace

An Address Space is a maximum size of 2 GB but z/OS usage reduces the amount of this available to the program. A Dataspace can have a maximum size of 2 GB, but all of it is available for storing data. The movement of data to and from a Dataspace is under direct program control and therefore has less overhead than a Hiperspace. The major drawback to the Dataspace is that to perform I/O operations on the data it must first be moved into storage in the address space.

Memory object

When the Address Space size was increased to 2^{64} , the old upper limit of 2^{31} remained as a limit for programs to execute in and is referred to as the 'bar'. A memory object is a data area in virtual storage that is allocated above the bar and backed by central storage. With memory object sorting, a memory object is used in place of and along with disk for temporary storage of records during a sort. Memory object sorting reduces I/O processing, which in turn reduces elapsed time, EXCPs, and channel usage. Memory object sorting is recommended for large input data sets when a sufficient amount of central storage is available.

You can control the maximum size of a memory object for a memory object sorting application with the MOSIZE parameter.

Setting limits on memory usage

In an environment where several sorts are running concurrently, it can give the best overall performance if the resources are shared between them. To assist in this, DFSORT has parameters, for all of the three methods described previously, which will limit the maximum storage to be used by any individual sort. Without these predefined limits, it is possible that one very large sort could consume all of the available resources.

27.3.4 ICEGENER utility

The standard data mover utility supplied with z/OS is IEBGENER. This is a very old program and does not exploit system facilities to optimize performance. Part of the DFSORT product is ICEGENER, a replacement for IEBGENER that can provide excellent performance improvements because it uses DFSORT copy to process the data. ICEGENER can be installed so that no changes are required to existing IEBGENER jobs. There are a few functions that ICEGENER does not support and in these cases control is passed to IEBGENER.

27.4 Symbols

A symbol is a name (preferably something meaningful like Price-of-Item) that can be used to represent a field or a constant. Sets of symbols, also called mappings, can be used to describe a group of related fields and constants such as the information in a particular type of record. Such mappings can be used to refer to fields and constants by their symbols without knowing the position, length and format of a field, the value of a constant, or the position of an output column. In addition, system symbols (for example, &SYSPLEX and &JOBNAME) can be used in symbol constants.

DFSORT's symbols processing feature is a powerful, simple, and flexible way for a programmer or site to create symbols for the fields in commonly used record types and for constants associated with those fields. Once the symbols are defined, they can be used in DFSORT control statements and ICETOOL operators instead of the equivalent hard-coded fields, constants, and output columns.

Symbols turn DFSORT's syntax into a high-level language. Symbols can help to standardize DFSORT applications and increase productivity. DFSORT symbols can be up to 50 characters, are case-sensitive and can include underscore and hyphen characters. Therefore, they can be meaningful and descriptive.

27.5 DFSORT/VSE memory sorting

There are two options available to DFSORT/VSE for using memory in a z/VSE environment:

- ▶ Dataspace
- ▶ GETVIS

Dataspace

Dataspace sorting is a DFSORT/VSE capability that uses data space available with z/VSE systems in place of the intermediate work space to improve the performance of sort applications. A data space is an area of contiguous storage backed by processor or auxiliary storage, whichever is necessary as determined by the system.

With data space sorting, the data space is used as an extension of main storage for storing records during a sort. Data space sorting reduces CPU time and elapsed time by using data space in place of disk work files.

The maximum amount of data space used for data space sorting can be controlled with the DSPSIZE option.

GETVIS

GETVIS sorting is a DFSORT/VSE capability that uses partition virtual storage for sort applications. GETVIS sorting is similar to data space sorting except that it uses the partition GETVIS area for virtual storage instead of data space.

The maximum amount of GETVIS area used for getvis sorting can be controlled with the GVSIZE option.

27.6 Sort features

DFSORT includes some built-in data processing features which can be used instead of having to write a separate program to process data.

Include and omit

Select records to be included in or omitted from the output data set using five different types of tests optionally combined with OR and AND to form complex logical expressions.

Reformat

Records can be reformatted in one of three ways.

- ▶ **BUILD:** Reformat each record by specifying all of its items one by one. Build gives you complete control over the items you want in your reformatted records and the order in which they appear. You can delete, rearrange and insert fields and constants.
- ▶ **OVERLAY:** Reformat each record by specifying just the items that overlay specific columns. Overlay lets you change specific existing columns without affecting the entire record.
- ▶ **IFTHEN clauses:** Reformat different records in different ways by specifying how build or overlay items are applied to records that meet given criteria. IFTHEN clauses let you use sophisticated conditional logic to choose how different record types are reformatted.

Multiple output and reports

One or more output data sets or reports can be created for a sort, copy, or merge application from a single pass over one or more input data sets.

Locales

You can modify collating behavior for different cultural environments by selecting an active locale. For example a Locale of FR_CA would indicate rules associated with French Canada

27.7 User exits

You can direct DFSORT to pass control during run time to routines you design and write yourself, such as user exit routines to:

- ▶ Summarize, insert, delete, shorten, or otherwise alter records during processing
- ▶ Correct I/O errors that DFSORT does not handle
- ▶ Perform any necessary abnormal end-of-task operation before DFSORT terminates.

However, keep in mind that the extensive editing capabilities provided by DFSORT can eliminate the need to write some types of user exit routines.

27.8 ICETOOL utility

ICETOOL is supplied with both DFSORT and DFSORT/VSE. It is a versatile data set processing and reporting utility that provides an easy-to-use batch front-end. The user can perform complex sorting, copying, reporting, matching, joining, and analytical tasks using multiple inputs in a single job step. ICETOOL works by building the control statements and options required to perform the requested operations and then passing control to the relevant DFSORT program.

There are 13 ICETOOL operators which can be used to perform a variety of functions. By using various combinations of these operators, applications can easily be created that perform many complex tasks.

For more information about the use of ICETOOL, refer to the Web site shown in 27.9, "More information" on page 502.

27.9 More information

For papers, online books, news, tips, examples and more, visit the DFSORT and DFSORT/VSE Web sites at:

<http://www.ibm.com/storage/dfsor/>
<http://www.ibm.com/storage/dfsorvse/>

If you're not familiar with DFSORT and DFSORT's ICETOOL, we recommend reading through *z/OS DFSORT: Getting Started*, SC26-7527. It's an excellent tutorial, with lots of examples, that will show you how to use DFSORT, DFSORT's ICETOOL and DFSORT Symbols. You can access it online, along with all of the other DFSORT books, from:

<http://www.ibm.com/servers/storage/support/software/sort/mvs/srtmpub.html>



Part 6

Appendixes



A

7133 serial disk system

This appendix describes the IBM TotalStorage 7133 serial disk system models D40 and T40.

Note: The 7133 models D40 and T40 were withdrawn from marketing on June 3, 2005.

The IBM TotalStorage 7133 serial disk system can connect and share storage with multiple UNIX, Microsoft Windows, and Novell NetWare hosts.

Overview

The IBM TotalStorage 7133 Serial Disk System Models D40 and T40 provide highly available storage for UNIX, Windows NT, and Novell NetWare servers. Through its implementation of Serial Storage Architecture (SSA), the 7133 Models provide outstanding performance, availability, multihost attachment, and economical RAID 5 data-protected storage.

Serial Storage Architecture

SSA is an open storage interface developed to overcome many of the SCSI limitations. SSA technology enables simultaneous communication between multiple devices, subsystems, and local host processors throughout an open systems environment.

Increased throughput for higher performance

Compared to SCSI, SSA offers higher bandwidth that yields excellent performance, even under heavy workloads. SSA addresses SCSI's throughput bottleneck with its fundamental building block—the SSA connection, or node. Each SSA node has two ports that can each carry on two 40 MBps conversations at once (one inbound and one outbound), to enable a total of 160 MBps of throughput. A single SCSI bus, on the other hand, can easily be saturated. Enterprises can take advantage of SSA's high performance to develop the types of applications (such as data mining and multimedia) that are currently beyond the capabilities of today's systems.

High system availability

Overall system availability is enhanced by SSA devices' ability to act independently, and peripherals can be attached in multiple paths, making a fault-tolerant system much easier to implement. In addition to enhancing fault tolerance, avoiding single points of failure in the path design also simplifies network maintenance. SSA supports flexible configurations, enabling an assortment of system connection options that help improving system throughput by isolating sub-tasks from production. One of the key benefits of the 7133 is its ability to fit into the enterprise environment with minimal disruption, even during the actual integration. *Hot-swapping* and automatic configuration of new devices support greater system availability while freeing up technical staff to concentrate on more strategic tasks.

Figure A-1 shows the 7133 models.



Figure A-1 7133 Models D40 and T40

The 7133 subsystem can contain four to sixteen drives per 7133 and one to six 7133 enclosures per 19 inch rack. Each 7133 has a Storage Serial Architecture (SSA) four-port adapter that allows two loops and delivers up to 40 MBps per adapter (therefore, up to 160 MBps per subsystem).

The 7133 can also attach up to 96 SSA disk drives, 48 per loop and can be configured with up to 8 adapters in up to 8 host processors within each SSA loop. The SSA RAID adapters can be configured with multiple RAID 5 arrays.

SSA is a serial link with a two-signal connection (transmit and receive) providing full duplex communication.

There are significant availability benefits inherent in the SSA technology that enhance the IBM 7133 solution. With the SSA loop design, a single cable failure will not necessarily cause loss of access to data. If there is a failure on a loop, the SSA adapter automatically reconfigures itself and continues accessing the devices without the loop configuration. After the path is restored, the adapter automatically reconfigures to resume normal operation. If there is a disk failure, the hot-swappable disks can be replaced without the loss of communication between the adapter and the other disks on the loop. The 7133 is also designed with redundant power supplies and fans that can be replaced without degrading system operations.

The rack-mountable 7133 Model D40 fits into an industry standard 19 inch rack, while the 7133 Advanced Model T40 is a free-standing deskside tower unit.

Both 7133 Models can be populated with 145.6 GB, 72.8 GB, 36.4 GB and 18.2 GB 10,000 RPM and 72.8 GB and 36.4 GB 15,000 RPM disk drives. Drive capacities can be intermixed. Each 7133 includes three auto-docking power/cooling units, which provide redundant power to all sixteen disk drive module positions.

An Advanced SSA Optical Extender (Feature 8851) is available on all models. This feature supplies a pair of optical converters that convert electrical SSA signals into optical signals for transmission using optical fibers. This feature allows the maximum distance between SSA nodes to be increased from 25 meters up to 10 kilometers.

More information

7133 products are supported in the environments shown at this Web site:

<http://www.ibm.com/servers/storage/support/disk/7133/index.html>

For more information about the 7133, also see the redbook *Understanding SSA Subsystems in Your Environment*, SG24-5750.



Introduction to SMI-S and CIM/WBEM

SMI-S

Efficiently managing multivendor Storage Area Networks (SANs) is a key concern for end-users and integrators alike. In mid-2002 the Storage Networking Industry Association (SNIA) launched the Storage Management Initiative (SMI) to create and foster the universal adoption of a highly functional open interface for the management of storage networks. The SMI's goal is to deliver open storage network management interface technology in the form of an *SMI Specification* (SMI-S).

SMI-S is intended to be the unifying factor between the vast objects that must be managed in a storage network and the tools used to manage them. SMI-S is based on the *Common Information Model* (CIM) and *Web Based Enterprise Management* (WBEM) standards developed by the Distributed Management Task Force (DMTF). As an industry standard SAN management interface, SMI-S provides new features that extend CIM/WBEM standards to create interoperable storage networking product implementations.

CIM/WBEM

Web Based Enterprise Management (WBEM) is a set of management and Internet standard architectures developed by the Distributed Management Task Force (DMTF) to unify the management of enterprise computing environments traditionally administered through traditional management stacks like SNMP or CMIP. WBEM provides the ability for the industry to deliver a well-integrated set of standard-based management tools leveraging emerging Web technologies. The DMTF has developed a core set of standards that make up WBEM, which includes a data model, the Common Information Model (CIM) standard; an encoding specification, xmlCIM Encoding Specification; and a transport mechanism, CIM Operations over HTTP.

- ▶ **Common Interface Model (CIM) Standard:** CIM is a formal object oriented modeling language that is used to describe the management aspects of systems.
- ▶ **xmlCIM Encoding Specification:** This is a grammar to describe CIM declarations and messages used by the CIM protocol

- ▶ **CIM Operations over Hypertext Transfer Protocol (HTTP):** HTTP is used as a way to enable communication between a management application and a device that both use CIM.

The WBEM architecture defines the following elements:

- ▶ **CIM Client:** The CIM Client is a management application like TotalStorage Productivity Center that uses CIM to manage devices. A CIM Client can reside anywhere in the network, because it uses HTTP to talk to CIM Object Managers and Agents.
- ▶ **CIM Managed Object:** A Managed Object is a hardware or software component that can be managed by a management application by using CIM. e.g. IBM SAN Volume Controller.
- ▶ **CIM Agent:** The CIM Agent is embedded into a device and interfaces with the management application. It interprets CIM requests and responses to it.
- ▶ **CIM Object Manager (CIMOM):** A common conceptual framework for data management that receives, validates and authenticates the CIM requests from client application. It then directs the requests to appropriate component or CIM provider.

Note: The terms CIM Agent and CIM Object Manager (CIMOM) are often used interchangeably. At this time only few devices come with an integrated CIM Agent, most devices need a external CIMOM for CIM enable management applications (CIM Clients) to be able to talk to the device.

- ▶ **Device Provider:** A CIM Provider is the element that translates CIM calls to the device specific commands, it is like a device driver. A provider is always closely linked to a CIM Object Manager.
- ▶ **CIM Server:** A CIM Server is the software that runs the CIMOM and the CIM provider for a set of devices. This approach is used when the devices do not have an embedded CIM Agent. This term is often not used, instead people often use the term CIMOM but they mean the CIM Server.

Typical component Interaction

Figure B-1 on page 511 shows a typical interactions between the components in the CIM/WBEM environment. The following description explains the steps involved in the interaction.

(1) The client application locates the CIMOM by calling a Service Location Protocol (SLP) directory service.

Note: The SMI-S specification introduces (Service Location Protocol) SLP as the method for the management applications (the CIM clients) to locate managed objects.

For more information about SLP, refer the IBM technote TIPS0523 at following link:

<http://www.redbooks.ibm.com/abstracts/tips0523.html?Open>

(2) When the CIMOM is first invoked, (3) it registers itself to the SLP and supplies its location, IP address, port number, and the type of service it provides. (4) With this information, the client application starts to directly communicate with the CIMOM.

The client application then (5) sends CIM requests to the CIMOM. As requests arrive, the CIMOM validates and authenticates each request. (6) It then directs the requests to the appropriate functional component of the CIMOM or to a device provider. (7) The provider

makes calls to a device-unique programming interface on behalf of the CIMOM to satisfy (8)-(9)-(10) client application requests.

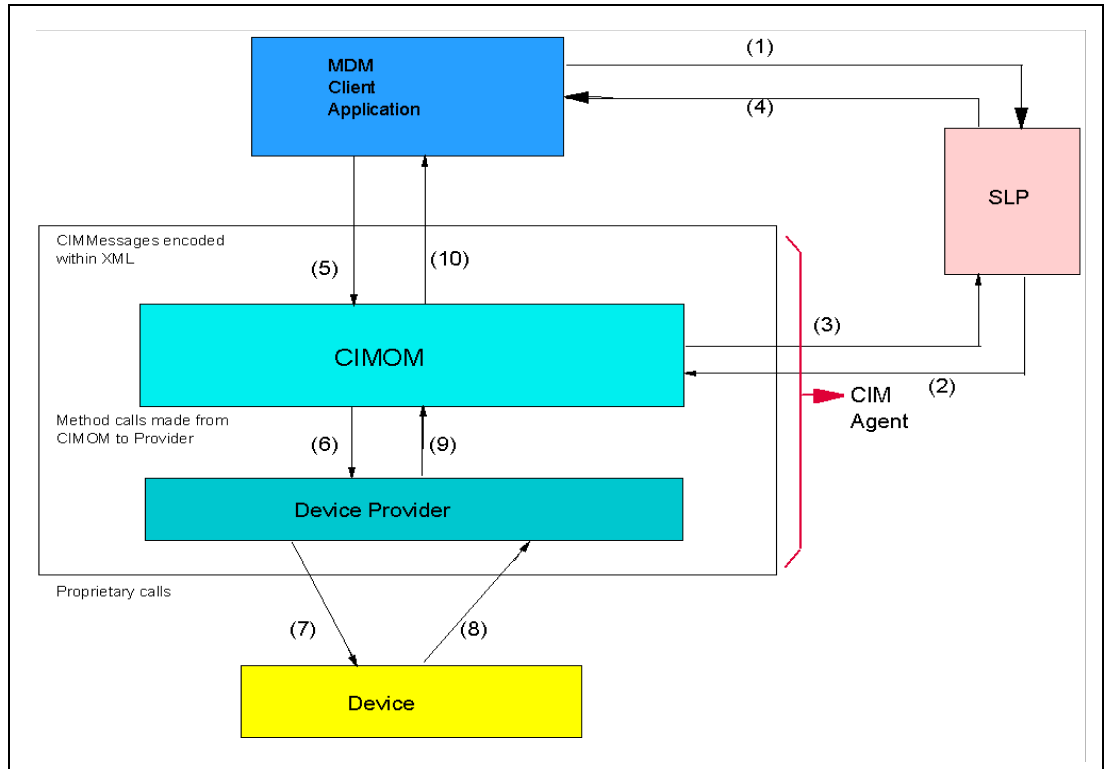


Figure B-1 Component interaction in CIM environment

More Information

For more information, refer to following links:

<http://www.snia.org/smi/home>

http://www.snia.org/tech_activities/SMI/cim



Redundant Arrays of Independent Disks (RAID)

RAID is a collection of techniques that treats multiple disks as combined units (arrays) and allows various levels of fault tolerance to be implemented in the disk subsystem.

RAID stands for *redundant array of independent disks*. It is a strategy created to bridge the gap between computer I/O requirements and the latency and throughput restrictions of single disk drives, while also allowing for greater degrees of fault tolerance.

RAID is the technology of grouping several physical drives in a computer into an array that can be defined as one or more logical drives. Each logical drive appears to the operating system as a single drive. This grouping technique greatly enhances logical drive capacity and performance.

SCSI has the ability to process multiple, simultaneous I/O requests. With RAID, I/O performance is improved because all drives can contribute to system I/O throughput.

In order to know RAID strengths and limitations, a clear understanding of the different RAID architectures is required.

RAID levels

Current RAID levels most commonly used are:

- ▶ 0, 1, 3, 5
- ▶ Enhanced 1, Enhanced 5
- ▶ 00, 10, 50

Data Striping

Data striping is the process of storing data across all the disks in an array.

- ▶ A data stripe is the collection of stripe units across the array.
 - ▶ A stripe unit is the amount of data written on a disk before writing on the next disk.
- ▶ Choosing the correct stripe unit size will maximize the performance of the ServeRAID adapter, improving overall server performance.

- ▶ Common RAID levels that use striping are 0, 1 Enhanced, 5, and 5 Enhanced.

Parity is defined as redundant information about user data, which allows it to be regenerated in the event of a disk failure. In the following illustrations, data can mean a byte or block, not necessarily an entire file.

RAID 0: Data striping, no parity

- ▶ Distributes (stripes) the data evenly across all disks in the array.
- ▶ Offers the fastest performance for multiple concurrent requests.
- ▶ Does not provide any level of fault tolerance.
- ▶ Offers most storage capacity of all the RAID levels.
- ▶ Total disk capacity of n (where n is the number of drives).

As shown in Figure C-1, in the event of a single disk failure, the data residing on the disk cannot be regenerated, and because of data striping, all of the striped data becomes unavailable.

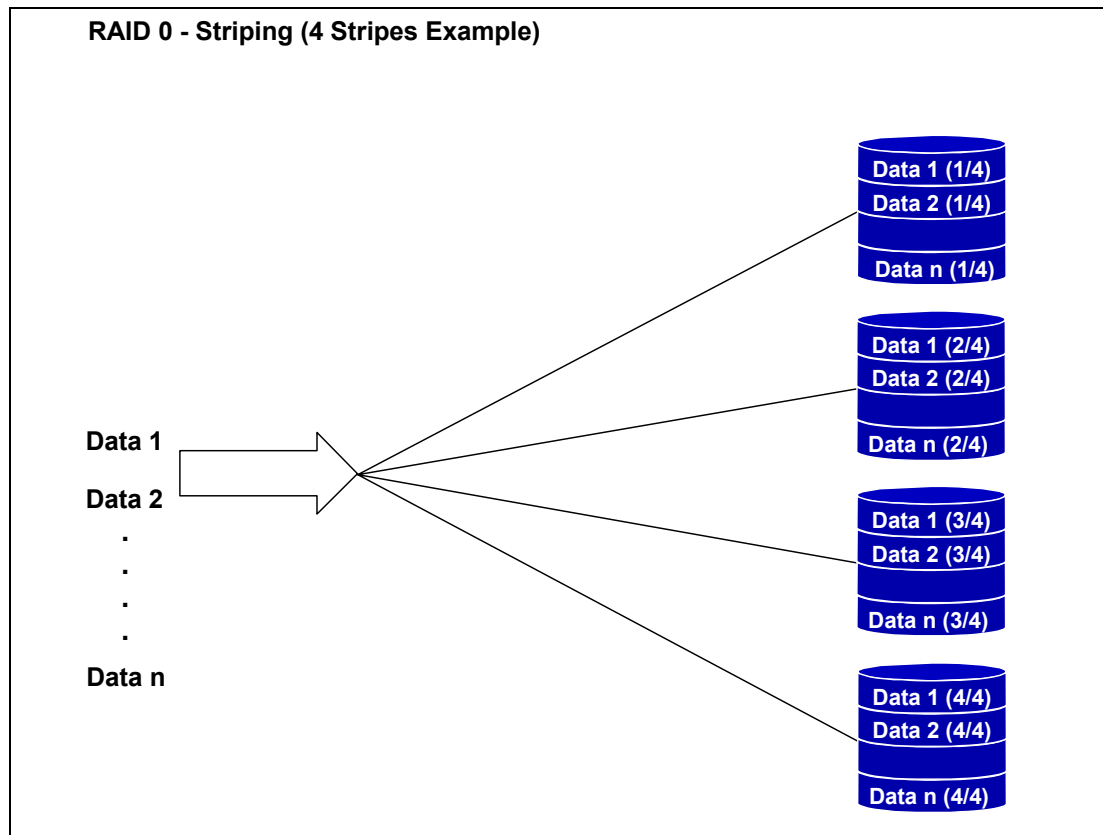


Figure C-1 RAID 0

Tip: Never use RAID level 0 for critical applications requiring high data availability. Consider it only for applications that would benefit from the performance capabilities of this level.

RAID 1: Data Mirroring

- ▶ Data is written simultaneously to two drives.
- ▶ RAID-1 can be faster than a single drive.
- ▶ High cost of implementation.
- ▶ Allows for the loss of one disk in the array.
- ▶ Total disk capacity is $n/2$ (n is the number of disks).

The RAID 1 implementation (Figure C-2) employs data mirroring to achieve redundancy.

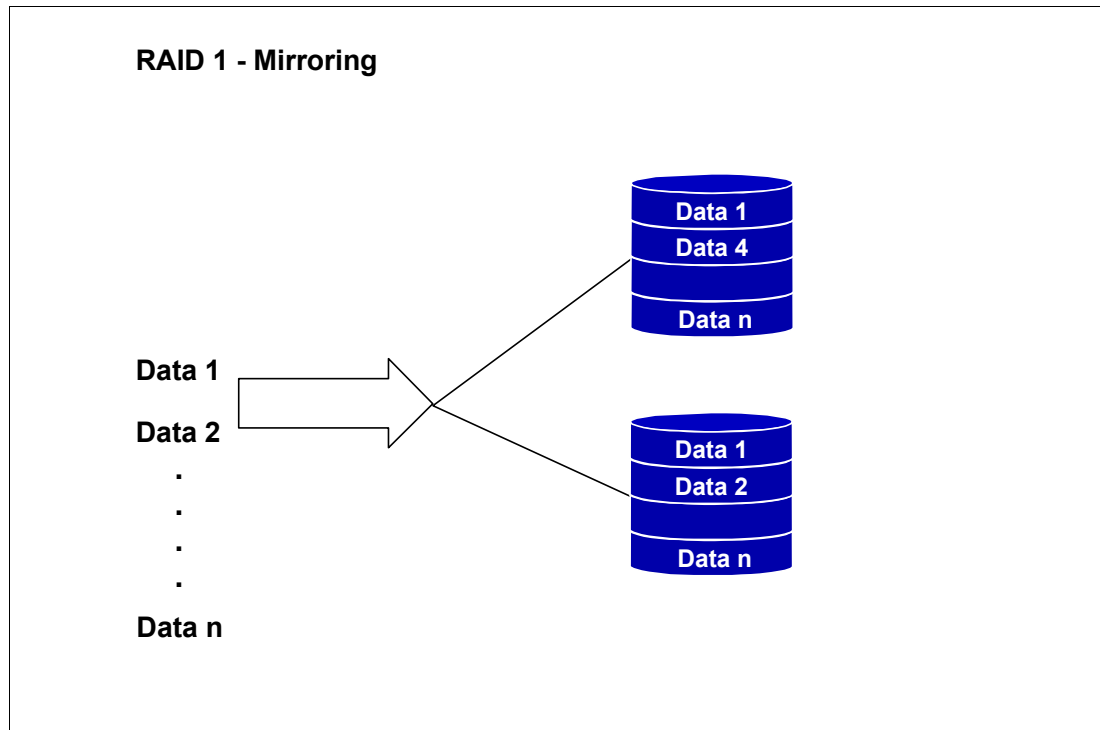


Figure C-2 RAID 1

Tip: Consider RAID 1 when cost is not a factor for applications requiring high data availability and high performance.

IBM Enhanced RAID-1

Raid-1 Enhanced (RAID-1E) is an IBM exclusive that combines data mirroring and data striping.

- ▶ Mirrors and stripes data across all disks (mirrored stripe)
- ▶ Allows disk mirroring with 3 or more disk drives
- ▶ Approximates RAID-0 performance for reads
- ▶ Allows for the loss of 1 disk in the array
- ▶ Total disk capacity of $n/2$ (n is the number of disks)

RAID 3: Sequential access to large files

RAID-3 is a parallel process array mechanism, where all drives in the array operate in unison. Similar to data striping, information to be written to disk is split into chunks (a fixed amount of

data), and each chunk is written out to the same physical position on separate disks (in parallel). This architecture requires parity information to be written for each stripe of data.

Performance is very good for large amounts of data, but poor for small requests because every drive is always involved, and there can be no overlapped or independent operation. It is well suited for large data objects such as CAD/CAM or image files, or applications requiring sequential access to large data files. Select RAID-3 for applications that process large blocks of data. It provides redundancy without the high overhead incurred by mirroring in RAID-1

RAID 5: Independent access, data striping with distributed parity

- ▶ Offers optimal balance between price and performance
- ▶ Cost of RAID-5 is the capacity of 1 disk
- ▶ Allows for the loss of 1 disk in the array
- ▶ Loss of a drive will cause system degradation
- ▶ Read performance approximates RAID-0
- ▶ Data and checksum (parity) are evenly spread across drives
- ▶ Requires a minimum of 3 disks
- ▶ Total capacity of $n-1$ (n is number of disks)

In RAID 5 the access arms can move independently of one another (Figure C-3). This enables multiple concurrent accesses to the array devices, thereby satisfying multiple concurrent I/O requests and providing higher transaction throughput. RAID 5 is best suited for random access data in small blocks.

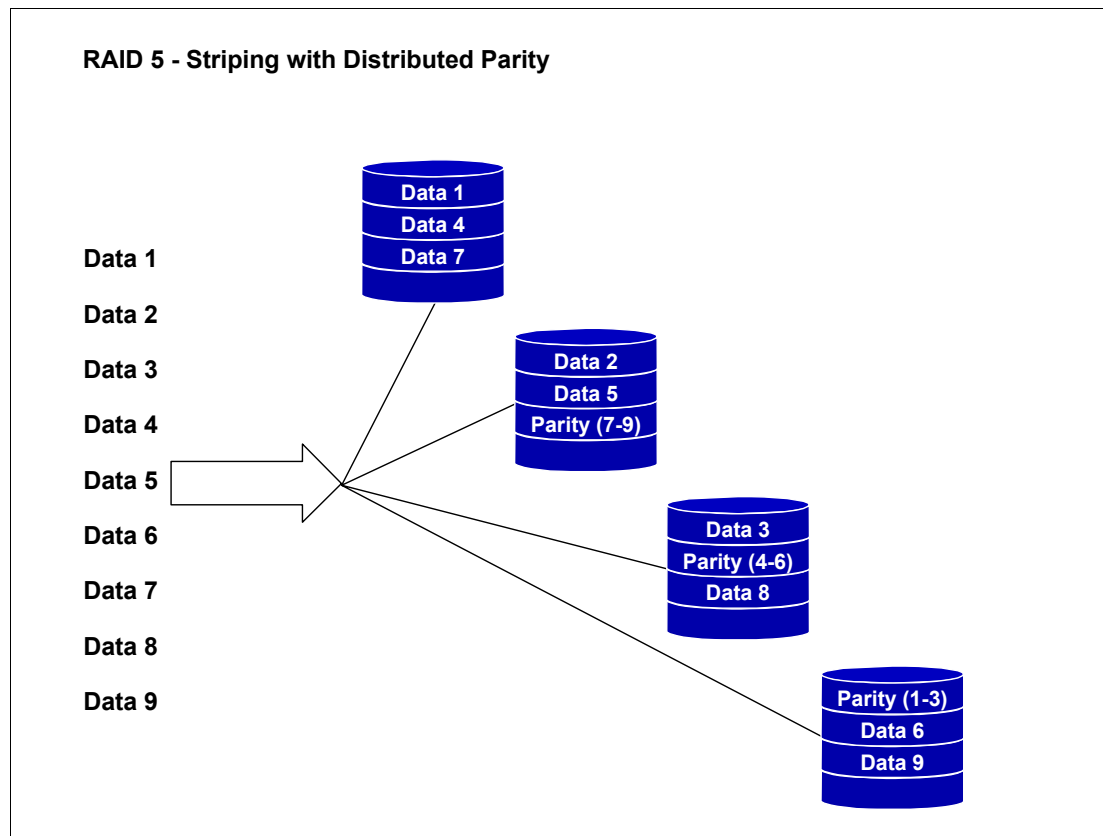


Figure C-3 RAID 5

Tip: RAID 5 should be considered for environments requiring high data availability and with applications that process relatively short data records or a mixture of large sequential records and short random blocks.

RAID 5 Enhanced

- ▶ RAID -5E is standard RAID-5 with a hot spare built into the array (distributed hot spare)
 - RAID-5 uses a physical dedicated hot-spare disk
 - RAID-5E uses a logical spare
 - The once-dedicated hot spare is now distributed across the array
- ▶ RAID-5E works with a minimum of 4 disks
- ▶ Performance is good (better than RAID-5 with hot spare defined when compared with same number of total disks)
 - No idle hot-spare disk needed; provides an additional drive to stripe data across
 - up to 20% faster in small RAID-5 configurations.

RAID -00

RAID-00 is RAID-0 plus RAID-0

- ▶ RAID-00 works with a minimum of two disks
- ▶ Performance: excellent
- ▶ Protection: none
- ▶ Capacity: all
- ▶ Maximum configuration is 60 disks (ServeRAID-4H)

RAID-00 Figure C-4 is a combination of multiple, local RAID-0 arrays.

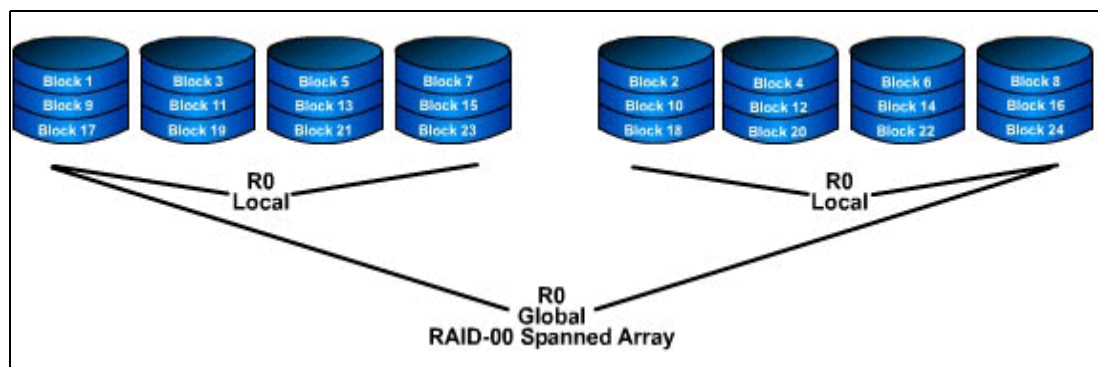


Figure C-4 RAID-00

RAID-10

RAID-10 is RAID-1 plus RAID-0

- ▶ RAID-10 works with a minimum of four disks
- ▶ Performance: very good
- ▶ Protection: one drive in each local array may fail
- ▶ Capacity: $n/2$
- ▶ Maximum configuration is 16 disks (ServeRAID-4H)

The RAID 10 Figure C-5 consists of a set of disks for user data plus their mirrored disks counterparts. There is no parity disk to rebuild a failed disk. In case one disk becomes unusable, then its mirror will be used to access the data and also to build the spare.

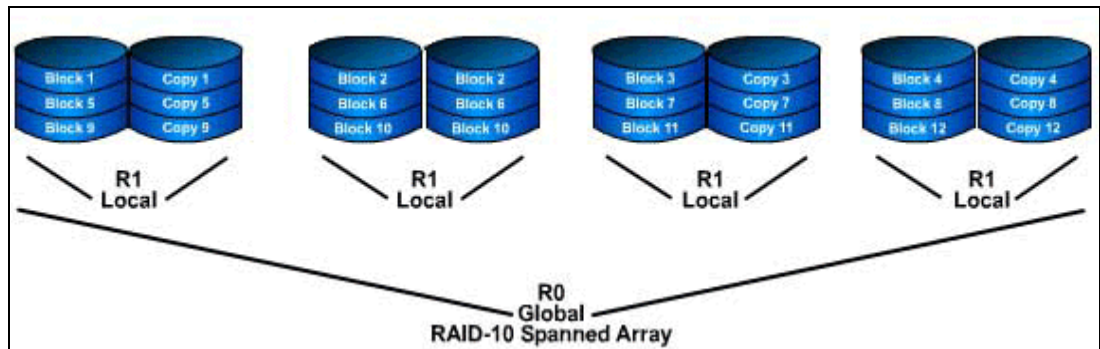


Figure C-5 RAID-10

RAID 10 is also known as RAID 0+1, because it is a combination of RAID 0 (striping) and RAID 1 (mirroring). The striping optimizes the performance by striping volumes across several disk drives (for example, in the ESS Model 800 implementation, three or four DDMs). RAID 1 is the protection against a disk failure by having a mirrored copy of each disk. By combining the two, RAID 10 provides data protection with I/O performance.

RAID-50

RAID-50 Figure C-6 is RAID-5 plus RAID-0

- ▶ RAID-50 works with a minimum of 6 disks
- ▶ Performance: good (improves with size of local array)
- ▶ Protection: one drive in each local array may fail
- ▶ Capacity: $n-x$ (where x = number of local arrays in RAID-0 stripe)
- ▶ Maximum configuration is 60 disks (SERVERAID-4H)

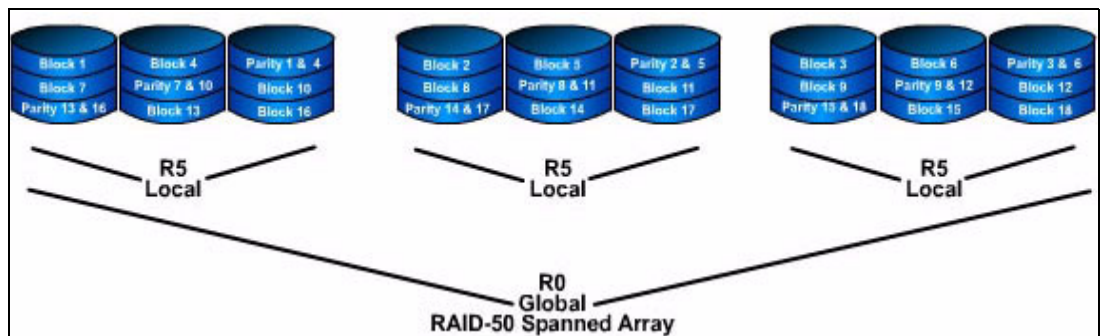


Figure C-6 RAID-50

For more information about RAID Levels see this Web site:

<http://www.acnc.com>

Glossary

A

Agent A software entity that runs on endpoints and provides management capability for other hardware or software. An example is an SNMP agent. An agent has the ability to spawn other processes.

AL See arbitrated loop.

Allocated storage The space that is allocated to volumes, but not assigned.

Allocation The entire process of obtaining a volume and unit of external storage, and setting aside space on that storage for a data set.

Arbitrated loop A Fibre Channel interconnection technology that allows up to 126 participating node ports and one participating fabric port to communicate. See also Fibre Channel Arbitrated Loop and loop topology.

Array An arrangement of related disk drive modules that have been assigned to a group.

B

Backup A copy of computer data that is used to recreate data that has been lost, mislaid, corrupted, or erased. The act of creating a copy of computer data that can be used to recreate data that has been lost, mislaid, corrupted or erased.

Bandwidth A measure of the data transfer rate of a transmission channel.

BCP See Business Continuity planning

Bridge Facilitates communication with LANs, SANs, and networks with dissimilar protocols.

Business Continuity Planning An enterprise-wide planning process which creates detailed procedures to be used in the case of a disaster. Business Continuity Plans take into account processes, people, facilities, systems, and external elements.

C

Channel (1) A path along which signals can be sent; for example, data channel and output channel. (2) A functional unit, controlled by the processor, that handles the transfer of data between processor storage and local peripheral equipment.

CIFS See Common Internet File system.

CIM See Common Information Model.

CIM Object Manager The CIMOM is the core component to the implementation of the CIM specification. The CIMOM manages the CIM schema, instantiation, communication, and operation of the physical Providers that represent the CIM classes stored within the namespace of the local host.

CIMOM See CIM Object Manager.

Client A function that requests services from a server, and makes them available to the user. A term used in an environment to identify a machine that uses the resources of the network.

Client authentication The verification of a client in secure communications where the identity of a server or browser (client) with whom you want to communicate is discovered. A sender's authenticity is demonstrated by the digital certificate issued to the sender.

Client-server relationship Any process that provides resources to other processes on a network is a server. Any process that employs these resources is a client. A machine can run client and server processes at the same time.

Common Information Model An object oriented description of the entities and relationships in a business' management environment maintained by the Distributed Management Task Force.

Common Internet File System CIFS provides an open cross-platform mechanism for client systems to request file services from server systems over a network. It is based on the SMB protocol widely used by PCs and workstations running a wide variety of operating systems.

Console A user interface to a server.

CWDM See WDM.

D

DASD Acronym for Direct Access Storage Device. This term is common in the z/OS environment to designate a disk or z/OS volume.

DATABASE 2 (DB2) A relational database management system. DB2 Universal Database is the relational database management system that is Web-enabled with Java support.

Device driver A program that enables a computer to communicate with a specific device, for example, a disk drive.

Disaster Recovery Planning Functions as a logical subset to the BCP process. IT operation's teams manage the DRP process to ensure continuity of operations in the event of a wide variety of disaster scenarios

Disk group A set of disk drives that have been configured into one or more logical unit numbers. This term is used with RAID devices.

DRP See Disaster Recovery Planning

DWDM See WDM.

E

E_Port An E_Port is used as an inter-switch expansion port to connect to the E_Port of another Fibre Channel switch, in order to build a larger switched fabric.

Enterprise network A geographically dispersed network under the backing of one organization.

ESCON Enterprise Systems Connection Architecture®. An zSeries 900 and S/390 computer peripheral interface. The I/O interface utilizes S/390 logical protocols over a serial interface that configures attached units to a communication fabric.

Event In the Tivoli environment, any significant change in the state of a system resource, network resource, or network application. An event can be generated for a problem, for the resolution of a problem, or for the successful completion of a task. Examples of events are: the normal starting and stopping of a process, the abnormal termination of a process, and the malfunctioning of a server.

F

F_Port An F_Port is a fabric port that is not loop capable. It is used to connect a N_Port to a switch.

Fabric The Fibre Channel employs a fabric to connect devices. A fabric can be as simple as a single cable connecting two devices. The term is often used to describe a more complex network utilizing hubs, switches, and gateways.

FC See Fibre Channel.

FCIP See Fibre Channel over IP.

FCP See Fibre Channel protocol.

FCS See Fibre Channel standard.

Fiber optic The medium and the technology associated with the transmission of information along a glass or plastic wire or fiber.

Fibre Channel A technology for transmitting data between computer devices at a data rate of up to 2 Gbps. It is especially suited for connecting computer servers to shared storage devices and for interconnecting storage controllers and drives.

Fibre Channel Arbitrated Loop A reference to the FC-AL standard, a shared gigabit media for up to 127 nodes, one of which can be attached to a switch fabric. See also arbitrated loop and loop topology.

Fibre Channel over IP Fibre Channel over IP is defined as a tunneling protocol for connecting geographically distributed Fibre Channel SANs transparently over IP networks.

Fibre Channel Protocol The serial SCSI command protocol used on Fibre Channel networks.

Fibre Channel Standard An ANSI standard for a computer peripheral interface. The I/O interface defines a protocol for communication over a serial interface that configures attached units to a communication fabric. Refer to ANSI X3.230-199x.

FICON An I/O interface based on the Fibre Channel architecture. In this new interface, the ESCON protocols have been mapped to the FC-4 layer, that is, the Upper Level Protocol layer, of the Fibre Channel Architecture. It is used in the S/390 and zSeries environments.

File system An individual file system on a host. This is the smallest unit that can monitor and extend. Policy values defined at this level override those that might be defined at higher levels.

FL_Port An F_Port is a fabric port that is loop capable. It is used to connect NL_Ports to the switch in a loop configuration.

FlashCopy A hardware based local copy option that provides an online point-in-time copy of data.

G

Gateway In the SAN environment, a gateway connects two or more different remote SANs with each other. A gateway can also be a server on which a gateway component runs.

Global Mirror A hardware based remote copy option that provides asynchronous volume copy across storage subsystems for disaster recovery, device migration, and workload migration.

H

Hardware zoning Hardware zoning is based on physical ports. The members of a zone are physical ports on the fabric switch. It can be implemented in the following configurations: one to one, one to many, and many to many.

HBA See host bus adapter.

Host Any system that has at least one Internet address associated with it. A host with multiple network interfaces can have multiple Internet addresses associated with it. This is also referred to as a server.

Host bus adapter A Fibre Channel HBA connection that allows a workstation to attach to the SAN network.

Hub A Fibre Channel device that connects up to 126 nodes into a logical loop. All connected nodes share the bandwidth of this one logical loop. Hubs automatically recognize an active node and insert the node into the loop. A node that fails or is powered off is automatically removed from the loop.

I

I/O group A group containing two SVC nodes defined by the configuration process. The nodes in the I/O group provide access to the vDisks in the I/O group.

ICAT IBM Common Information Model [CIM] Agent Technology.

iFCP See Internet Fibre Channel Protocol.

Internet Fibre Channel Protocol The Internet Fibre Channel Protocol specification defines iFCP as a gateway-to-gateway protocol for the implementation of a Fibre Channel fabric in which TCP/IP switching and routing elements replace Fibre Channel components.

Internet Protocol A protocol used to route data from its source to its destination in an Internet environment.

Internet SCSI Internet SCSI encapsulates SCSI commands into TCP packets; therefore enabling the transport of I/O block data over IP networks.

IP Internet protocol.

iSCSI See Internet SCSI.

J

Java A programming language that enables application developers to create object-oriented programs that are very secure, portable across different machine and operating system platforms, and dynamic enough to allow expendability.

Java runtime environment The underlying, invisible system on your computer that runs applets the browser passes to it.

Java Virtual Machine The execution environment within which Java programs run. The Java virtual machine is described by the Java Machine Specification which is published by Sun Microsystems™. Because the Tivoli Kernel Services is based on Java, nearly all ORB and component functions execute in a Java virtual machine.

JBOD Just a Bunch Of Disks. A disk group configured without the disk redundancy of the RAID arrangement. When configured as JBOD, each disk in the disk group is a rank in itself.

JRE See Java runtime environment.

JVM See Java Virtual Machine.

L

Logical unit number The LUNs are provided by the storage devices attached to the SAN. This number provides you with a volume identifier that is unique among all storage servers. The LUN is synonymous with a physical disk drive or a SCSI device. For disk subsystems such as the IBM Enterprise Storage Server, a LUN is a logical disk drive. This is a unit of storage on the SAN which is available for assignment or unassignment to a host server.

Loop topology In a loop topology, the available bandwidth is shared with all the nodes connected to the loop. If a node fails or is not powered on, the loop is out of operation. This can be corrected using a hub. A hub opens the loop when a new node is connected and closes it when a node disconnects. See also Fibre Channel Arbitrated Loop and arbitrated loop.

LUN See logical unit number.

LUN assignment criteria The combination of a set of LUN types, a minimum size, and a maximum size used for selecting a LUN for automatic assignment.

LUN masking This allows or blocks access to the storage devices on the SAN. Intelligent disk subsystems like the IBM Enterprise Storage Server provide this kind of masking.

M

MAN See Metropolitan Area Network.

Managed object A managed resource.

Managed resource A physical element to be managed.

Management Information Base A logical database residing in the managed system which defines a set of MIB objects. A MIB is considered a logical database because actual data is not stored in it, but rather provides a view of the data that can be accessed on a managed system.

Metro Mirror A hardware based remote copy option that provides a synchronous volume copy across storage subsystems for disaster recovery, device migration, and workload migration.

Metropolitan Area Network A network that connects nodes distributed over a metropolitan (city-wide) area as opposed to a local area (campus) or wide area (national or global).

MIB See Management Information Base.

MIB object A MIB object is a unit of managed information that specifically describes an aspect of a system. Examples are CPU utilization, software name, hardware type, and so on. A collection of related MIB objects is defined as a MIB.

Multiple Device Manager The Multiple Device Manager is software that has been designed to allow administrators to manage Storage Area Networks (SANs) and storage from a single console. It is now known as TotalStorage Productivity Center

N

N_Port node port A Fibre Channel-defined hardware entity at the end of a link which provides the mechanisms necessary to transport information units to or from another node.

NAS See Network attached storage.

Network attached storage A NAS device is attached to a TCP/IP- based network (LAN or WAN), and accessed using CIFS and NFS - specialized I/O protocols for file access and file sharing.

Network topology A physical arrangement of nodes and interconnecting communications links in networks based on application requirements and geographical distribution of users.

NL_Port node loop port A node port that supports arbitrated loop devices.

O

Open system A system whose characteristics comply with standards made available throughout the industry, and therefore can be connected to other systems that comply with the same standards.

P

Peer-to-Peer remote copy .See Metro Mirror

Point-to-point topology It consists of a single connection between two nodes. All the bandwidth is dedicated for these two nodes.

Port An endpoint for communication between applications, generally referring to a logical connection. A port provides queues for sending and receiving data. Each port has a port number for identification. When the port number is combined with an Internet address, it is called a socket address.

Port zoning In Fibre Channel environments, port zoning is the grouping together of multiple ports to form a virtual private storage network. Ports that are members of a group or zone can communicate with each other but are isolated from ports in other zones. See also LUN masking and subsystem masking.

PPRC See Metro Mirror.

Protocol The set of rules governing the operation of functional units of a communication system if communication is to take place. Protocols can determine low-level details of machine-to-machine interfaces, such as the order in which bits from a byte are sent. They can also determine high-level exchanges between application programs, such as file transfer.

R

RAID Redundant array of inexpensive or independent disks. A method of configuring multiple disk drives in a storage subsystem for high availability and high performance.

RDAC See Redundant Disk Array Controller.

Redundant Disk Array Controller Controller failover facility provided for some operating systems with the FAST product line.

Remote volume mirroring A hardware based remote copy option that provides a synchronous volume copy with the FAST product line.

RVM See Remote volume mirroring.

S

SAN See Storage Area Network.

SAN agent A software program that communicates with the manager and controls the subagents. This component is largely platform independent. See also subagent.

SAN File System SAN file systems allow computers attached via a SAN to share data. They typically separate the actual file data from the metadata, using the LAN path to serve the metadata, and the SAN path for the file data.

SAN Integration Server The SIS is a pre-packaged system comprising an SVC, backend storage, SAN and Ethernet Switches, and a master controller assembled and pre-configured in a rack.

SAN Volume Controller The SVC is a SAN appliance designed for attachment to a variety of host computer systems, which carries out block level virtualization of disk storage.

SCSI Small Computer System Interface. An ANSI standard for a logical interface to computer peripherals and for a computer peripheral interface. The interface utilizes a SCSI logical protocol over an I/O interface that configures attached targets and initiators in a multi-drop bus topology.

SDD See subsystem device driver.

Serial Storage Architecture An IBM standard for a computer peripheral interface. The interface uses a SCSI logical protocol over a serial interface that configures attached targets and initiators in a ring topology

Server A program running on a mainframe, workstation, or file server that provides shared services. This is also referred to as a host.

Shared storage Storage within a storage facility that is configured such that multiple homogeneous or divergent hosts can concurrently access the storage. The storage has a uniform appearance to all hosts. The host programs that access the storage must have a common model for the information on a storage device. You need to design the programs to handle the effects of concurrent access.

Simple Network Management Protocol A protocol designed to give a user the capability to remotely manage a computer network by polling and setting terminal values and monitoring network events.

SIS See SAN Integration Server.

SNMP See Simple Network Management Protocol.

SNMP agent An implementation of a network management application which is resident on a managed system. Each node that is to be monitored or managed by an SNMP manager in a TCP/IP network, must have an SNMP agent resident. The agent receives requests to either retrieve or modify management information by referencing MIB objects. MIB objects are referenced by the agent whenever a valid request from an SNMP manager is received.

SNMP manager A managing system that executes a managing application or suite of applications. These applications depend on MIB objects for information that resides on the managed system.

SNMP trap A message that is originated by an agent application to alert a managing application of the occurrence of an event.

Software zoning Is implemented within the Simple Name Server (SNS) running inside the fabric switch. When using software zoning, the members of the zone can be defined with: node WWN, port WWN, or physical port number. Usually the zoning software also allows you to create symbolic names for the zone members and for the zones themselves.

SQL Structured Query Language.

SSA See Serial Storage Architecture.

Storage administrator A person in the data processing center who is responsible for defining, implementing, and maintaining storage management policies.

Storage area network A managed, high-speed network that enables any-to-any interconnection of heterogeneous servers and storage systems.

Subagent A software component of SAN products which provides the actual remote query and control function, such as gathering host information and communicating with other components. This component is platform dependent. See also SAN agent.

Subsystem Device Driver The Subsystem Device Driver is a multipathing software designed especially to use with the IBM Enterprise Storage Server 2105. It cannot be used with any other storage servers or storage devices.

Subsystem masking The support provided by intelligent disk storage subsystems like the Enterprise Storage Server. See also LUN masking and port zoning.

SVC See SAN Volume Controller.

Switch A component with multiple entry and exit points or ports that provide dynamic connection between any two of these points.

Switch topology A switch allows multiple concurrent connections between nodes. There can be two types of switches, circuit switches and frame switches. Circuit switches establish a dedicated connection between two nodes. Frame switches route frames between nodes and establish the connection only when needed. A switch can handle all protocols.

T

TCP See Transmission Control Protocol.

TCP/IP Transmission Control Protocol/Internet Protocol.

Topology An interconnection scheme that allows multiple Fibre Channel ports to communicate. For example, point-to-point, arbitrated loop, and switched fabric are all Fibre Channel topologies.

Transmission Control Protocol A communications protocol used in the Internet and in any network that follows the Internet Engineering Task Force (IETF) standards for Internetwork protocol. TCP provides a reliable host-to-host protocol between hosts in packet-switched communications networks and in interconnected systems of such networks. It uses the Internet Protocol (IP) as the underlying protocol.

V

vDisk See Virtual Disk.

Virtual Disk An SVC device that appears to host systems attached to the SAN as a SCSI disk. Each vDisk is associated with exactly one I/O group.

W

WAN Wide Area Network.

Wave Division Multiplexing WDM allows the simultaneous transmission of a number of data streams over the same physical fiber, each using a different optical wavelength. WDM receives incoming optical signals from many sources (Fibre Channel, IP, ESCON, FICON) which it converts to electrical signals, it then assigns them a specific wavelength (or lambdas) of light and retransmits them on that wavelength. This method relies on the large number of wavelengths available within the light spectrum. Coarse WDM (CWDM) and Dense WDM (DWDM) are based on the same methodology as WDM enabling more data streams over the same physical fiber.

WDM See Wave Division Multiplexing.

World Wide Name A unique number assigned to Fibre Channel devices (including hosts and adapter ports) - analogous to a MAC address on a network card.

WWN See World Wide Name.

Z

Zoning In Fibre Channel environments, zoning allows for finer segmentation of the switched fabric. Zoning can be used to instigate a barrier between different environments. Ports that are members of a zone can communicate with each other but are isolated from ports in other zones. Zoning can be implemented in two ways: hardware zoning and software zoning.

Abbreviations and acronyms

ACL	Access Control List	DHCP	Dynamic Host Configuration Protocol
AD	Microsoft Active Directory®	DLC	Data Link Control
AFS®	Andrew File System	DLL	Dynamic Load Library
AIX	Advanced Interactive eXecutive	DRP	Disaster Recovery Planning
ANSI	American National Standards Institute	DWDM	Dense Wave Division Multiplexing
API	Application Programming Interface	EFS	Encrypting File Systems
APPC	Advanced Program to Program Communication	EMS	Event Management Services
ASCII	American National Standard Code for Information Interchange	ERD	Emergency Repair Disk
ATM	Asynchronous Transfer Mode	ERP	Enterprise Resources Planning
BCP	Business Continuity Planning	ESCON	Enterprise System Connection
BDC	Backup Domain Controller	ESS	Enterprise Storage Server
BIND	Berkeley Internet Name Domain	ETL	Enterprise Tape Library
BOS	Base Operating System	FAT	File Allocation Table
BSD	Berkeley Software Distribution	FC	Fibre Channel
CA	Certification Authorities	FCIP	Fibre Channel over IP
CDE	Common Desktop Environment	FCP	Fibre Channel Protocol
CDMF	Commercial Data Masking Facility	FCS	Fibre Channel Standard
CDS	Cell Directory Service	FDDI	Fiber Distributed Data Interface
CERT	Computer Emergency Response Team	FEC	Fast EtherChannel technology
CGI	Common Gateway Interface	FIFO	First In/First Out
CIFS	Common Internet File System	FQDN	Fully Qualified Domain Name
CIM	Common Information Model	FSF	File Storage Facility
CIMOM	CIM Object Manager	FTD	Fault-Tolerant Disk
C-SPOC	Cluster Single Point Of Control	FTP	File Transfer Protocol
CWDM	Coarse Wave Division Multiplexing	GDA	Global Directory Agent
DASD	Direct Access Storage Device	GDS	Global Directory Service
DBM	DataBase Management	GID	Group Identifier
DCE	Distributed Computing Environment	GL	Graphics Library
DCOM	Distributed Component Object Model	GUI	Graphical User Interface
DDE	Dynamic Data Exchange	HA	High Availability
DDNS	Dynamic Domain Name System	HACMP	High Availability Cluster Multiprocessing
DEN	Directory Enabled Network	HAL	Hardware Abstraction Layer
DES	Data Encryption Standard	HBA	Host Bus Adapter
DFS™	Distributed File System	HCL	Hardware Compatibility List
		HSM	Hierarchical Storage Management
		HTTP	Hypertext Transfer Protocol
		IBM	International Business Machines Corporation

ICAT	IBM Common Information Model (CIM) Agent Technology	LVM	Logical Volume Manager
IDE	Integrated Drive Electronics	MAN	Metropolitan Area Network
IEEE	Institute of Electrical and Electronic Engineers	MBR	Master Boot Record
IETF	Internet Engineering Task Force	MBps	Megabytes per second
IFCP	Internet Fibre Channel Protocol	MCA	Micro Channel® Architecture
IGMP	Internet Group Management Protocol	MDC	Meta Data Controller
IIS	Internet Information Server	MDM	Multiple Device Manager
IKE	Internet Key Exchange	MDS	Metadata Server Engine
IMAP	Internet Message Access Protocol	MFT	Master File Table
IP	Internet Protocol	MIB	Management Information Base
IPC	Interprocess Communication	MIPS	Million Instructions Per Second
IPL	Initial Program Load	MMC	Microsoft Management Console
IPsec	Internet Protocol Security	MOCL	Managed Object Class Library
IPX™	Internetwork Packet eXchange™	MPTN	Multiprotocol Transport Network
ISA	Industry Standard Architecture	MSCS	Microsoft Cluster Server
iSCSI	Internet SCSI	MSS	Modular Storage Server
ISDN	Integrated Services Digital Network	MWC	Mirror Write Consistency
ISO	International Standards Organization	NAS	Network Attached Storage
ISS	Interactive Session Support	NBC	Network Buffer Cache
ISV	Independent Software Vendor	NBPI	Number of Bytes per I-node
ITSO	International Technical Support Organization	NCP	NetWare Core Protocol
JBOD	Just a Bunch Of Disks	NCS	Network Computing System
JFS	Journalled File System	NCSC	National Computer Security Center
JRE™	Java Runtime Environment	NDIS	Network Device Interface Specification
JVM™	Java Virtual Machine	NDMP	Network Data Management Protocol
LAN	Local Area Network	NDS	NetWare Directory Service
LCN	Logical Cluster Number	NetBEUI	NetBIOS Extended User Interface
LDAP	Lightweight Directory Access Protocol	NetDDE	Network Dynamic Data Exchange
LFS	Log File Service (Windows NT)	NETID	Network Identifier
LFT	Low Function Terminal	NFS	Network File System
LOS	Layered Operating System	NIM	Network Installation Manager
LPAR	Logical Partition	NIS	Network Information System
LPC	Local Procedure Call	NNS	Novell Network Services
LPP	Licensed Program Product	NTFS	NT File System
LSA	Local Security Authority	NTLDR	NT Loader
LUID	Login User Identifier	NTLM	NT LAN Manager
LUN	Logical Unit Number	NTP	Network Time Protocol
LVCB	Logical Volume Control Block	NVRAM	Non-Volatile Random Access Memory
LVDD	Logical Volume Device Driver	ODBC	Open Database Connectivity
		ODM	Object Data Manager
		OLTP	OnLine Transaction Processing

OMG	Object Management Group	ROS	Read-Only Storage
ONC™	Open Network Computing	RPC	Remote Procedure Call
OS	Operating System	RRIP	Rock Ridge Internet Protocol
OSF	Open Software Foundation	RSCT	Reliable Scalable Cluster Technology
PAP	Password Authentication Protocol	RSM	Removable Storage Management
PB	Petabytes	RSVP	Resource Reservation Protocol
PCI	Peripheral Component Interconnect	RVM	Remote Volume Mirroring
PCMCIA	Personal Computer Memory Card International Association	SAK	Secure Attention Key
PDC	Primary Domain Controller	SAM	Security Account Manager
PDF	Portable Document Format	SAN	Storage Area Network
PDS	Partitioned data set	SASL	Simple Authentication and Security Layer
PDSE	Partitioned data set extended	SCSI	Small Computer System Interface
PDT	Performance Diagnostic Tool	SDD	Subsystem Device Driver
PEX	PHIGS Extension to X	SDK	Software Developer's Kit
PHIGS	Programmer's Hierarchical Interactive Graphics System	SFP	Small form-factor pluggable
PID	Process Identification Number	SID	Security Identifier
PIN	Personal Identification Number	SIS	SAN Integration Server
PMTU	Path Maximum Transfer Unit	SLIP	Serial Line Internet Protocol
POSIX	Portable Operating System Interface for Computer Environment	SMB	Server Message Block
POST	Power-On Self Test	SMIT	System Management Interface Tool
PP	Physical Partition	SMP	Symmetric Multiprocessor
PPP	Point-to-Point Protocol	SMS	Systems Management Server
PPRC	Peer to Peer Remote Copy	SNA	Systems Network Architecture
PPTP	Point-to-Point Tunneling Protocol	SNMP	Simple Network Management Protocol
PSM	Persistent Storage Manager	SP	System Parallel
PSSP	Parallel System Support Program	SPX	Sequenced Packet eXchange
PV	Physical Volume	SQL	Structured Query Language
PVID	Physical Volume Identifier	SRM	Storage Resource Manager
QoS	Quality of Service	SSA	Serial Storage Architecture
RACF®	Resource Access Control Facility	SSL	Secure Sockets Layer
RAID	Redundant Array of Independent Disks	SVC	SAN Volume Controller
RAS	Remote Access Service	TCB	Trusted Computing Base
RDAC	Redundant Disk Array Controller	TCO	Total cost of ownership
RDBMS	Relational Database Management System	TCP/IP	Transmission Control Protocol/Internet Protocol
RFC	Request for Comments	TDI	Transport Data Interface
RGID	Real Group Identifier	TDP	Tivoli Data Protection
RISC	Reduced Instruction Set Computer	TLS	Transport Layer Security
RMC	Resource Monitoring and Control	TOS	Type of Service
		TSM	Tivoli Storage Manager
		UCS	Universal Code Set

UDB	Universal Database
UDF	Universal Disk Format
UDP	User Datagram Protocol
UFS	UNIX File System
UID	User Identifier
UMS	Ultimedia Services
UNC	Universal Naming Convention
UPS	Uninterruptible Power Supply
URL	Universal Resource Locator
USB	Universal Serial Bus
UTC	Universal Time Coordinated
UUCP	UNIX to UNIX Communication Protocol
UUID	Universally Unique Identifier
VAX	Virtual Address eXtension
VCN	Virtual Cluster Name
VFS	Virtual File System
VG	Volume Group
VGDA	Volume Group Descriptor Area
VGID	Volume Group Identifier
VGSA	Volume Group Status Area
VIPA	Virtual IP Address
VMM	Virtual Memory Manager
VPD	Vital Product Data
VPN	Virtual Private Network
VSM™	Virtual System Management
VTC	Virtual Tape Controller
VTS	Virtual Tape Server
W3C	World Wide Web Consortium
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WINS	Windows Internet Name Service
WLM	Workload Manager
WWN	World Wide Name
WWW	World Wide Web
XCMF	X/Open Common Management Framework
XDM	X Display Manager
XNS	XEROX Network Systems

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

IBM Redbooks

For information on ordering these publications, see “How to get IBM Redbooks” on page 530. Note that some of the documents referenced here may be available in softcopy only.

- ▶ *IBM TotalStorage DS300 and DS400 Best Practices Guide*, SG24-7121
- ▶ *The IBM TotalStorage DS6000 Series: Concepts and Architecture*, SG24-6471
- ▶ *The IBM TotalStorage DS8000 Series: Concepts and Architecture*, SG24-6452
- ▶ *FICON Native Implementation and Reference Guide*, SG24-6266
- ▶ *iSeries in Storage Area Networks A Guide to Implementing FC Disk and Tape with iSeries*, SG24-6220
- ▶ *IBM System Storage DS4000 Series, Storage Manager and Copy Services*, SG24-7010
- ▶ *DS4000 Best Practices and Performance Tuning Guide*, SG24-6363
- ▶ *Introduction to Storage Area Networks*, SG24-5470
- ▶ *Designing an IBM Storage Area Network*, SG24-5758
- ▶ *IBM SAN Survival Guide*, SG24-6143
- ▶ *IBM SAN Survival Guide Featuring the Cisco Portfolio*, SG24-9000
- ▶ *IBM SAN Survival Guide Featuring the McDATA Portfolio*, SG24-6149
- ▶ *IBM SAN Survival Guide Featuring the CNT Portfolio*, SG24-6150
- ▶ *FICON (FCV Mode) Planning Guide*, SG24-5445
- ▶ *Introduction to IBM S/390 FICON*, SG24-5176
- ▶ *IBM TotalStorage Peer-to-Peer Virtual Tape Server Planning and Implementation Guide*, SG24-6115
- ▶ *IBM TotalStorage Enterprise Tape: A Practical Guide*, SG24-4632
- ▶ *IBM TotalStorage Tape Selection and Differentiation Guide*, SG24-6946
- ▶ *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946
- ▶ *Implementing IBM Tape in UNIX Systems*, SG24-6502
- ▶ *Implementing IBM Tape in Linux and Windows*, SG24-6268
- ▶ *IBM TotalStorage NAS Backup and Recovery Solutions*, SG24-6831
- ▶ *IBM Tivoli Storage Management Concepts*, SG24-4877
- ▶ *Disaster Recovery Strategies with Tivoli Storage Management*, SG24-6844
- ▶ *IBM Tivoli Storage Resource Manager: A Practical Introduction*, SG24-6886
- ▶ *DFSMSHsm ABARS and Mainstar Solutions*, SG24-5089
- ▶ *IBM TotalStorage Business Continuity Solutions Guide*, SG24-6547
- ▶ *IBM System Storage SAN Volume Controller*, SG24-6423

- ▶ *IBM TotalStorage Productivity Center V2.3: Getting Started*, SG24-6490
- ▶ *IBM TotalStorage: Implementing an Open IBM SAN*, SG24-6116
- ▶ *IBM Tivoli Storage Manager Implementation Guide*, SG24-5416
- ▶ *Understanding the IBM System Storage DR550*, SG24-7091

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM Tivoli Storage Manager for Windows Administrator's Guide*, GC32-0782
- ▶ *IBM Tivoli Storage Manager for Windows Installation Guide*, GC32-1602

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

Numerics

2027-256 338
2027-R16 346–347
3494
 Base Frame 247
 High Availability Frame 248
 S10 tape storage 248
 Tape Drive Expansion Frame 248
3580 model L33 216
3588 Model F3A 207
3590 250
 cartridge 224
 H1A 220
 See IBM 3590
3592
 See IBM 3592
7133 506
7204 External Disk Drive 296
7205 296
7207 298

A

Accelis 202
Access Control List 71
accessor 241
ACS 486
adaptive subfile backup 452
address space 499
Administration Center 446–447
 wizards 447
advanced copy services 430
Advanced Library Management System (ALMS) 242
Advanced Performance Monitoring 325
Advanced Security (AS) 325
Advanced Web Tools 324
Advanced Zoning 325
AIX 465
alert 397
Alert Manager 111
ALMS 241
ANSI 36, 305
any-to-any 310
API 439, 453
APM 272
appliance 319
Application Programming Interface (API) 20
application/database replication 22
arbitrated loop 121, 307
archive 439
archive data 41, 188
archive management 7, 44
archive management system 44
archive retention
 chronological 194–195

 event based 194
Array Across Loops (AAL) 134, 141–142
Array Support Library (ASL) 104
asset reporting 396, 399
asymmetry compensation 208
Asynchronous data transfer 306
Asynchronous Mirroring 109
Asynchronous PPRC 148
Asynchronous PPRC-XD 147
Asynchronous Transfer Mode (ATM) 326
ATM. See Asynchronous Transfer Mode
Auto Archive 265
Auto Replication 266
Automated Cartridge Facility (ACF) 220
automated management 2
Automatic file system extension 391
availability 366
availability reporting 399

B

backhitch 208
backhitch reduction 211
backup and restore of data 438
backup reporting 400
backup topology 454
backup window 38
backwards compatibility 205
bands 203–204
 See also servo tracks
bandwidth 307–308
barcode 265
barcode reader 239, 254
barcode scanner 252
bare machine recovery 438
bare metal recovery 453
base logical drive 105
Basic Partitioned Access Method 485
Basic Sequential Access Method 485
batch window 498
battery 83–84, 86, 137
battery backup unit (BBU) 126
BBU 126
block aggregation 372
BM 7206 Model VX2 298
Boot from SAN 71
boss node 369
BPAM 485
Brocade (b-type) 324
BSAM 485
Business Continuity 1–2, 9, 268, 349, 351–352
 tiers 15
Business Impact Analysis (BIA) 472

C

- cache 122
- Calibrated Vektored Cooling 126
- calibration sensor 239
- canister 217
- capacity 206, 222
- Capacity on Demand (CoD) 242
- capacity reporting 399
- cartridge memory 202, 208, 211–212
- CartridgeGroup 284
- CartridgeGroupApplication 284
- cascade 308–310
- catalog 486
- CDB 314
- CDS 486
- channel calibration 212
- chargeback 400
- CIFS 164, 304, 318–319
- CIM 408
 - agent 408
 - object manager 408
- CIMOM 407, 510
- Cisco Fabric Manager 329
- Cisco MDS 9020 328
- Cisco MDS 9120 333
- Cisco MDS 9140 333
- Cisco MDS 9216A 335
- Cisco MDS 9216i Multilayer Fabric Switch 348
- Cisco MDS 9506 Multilayer Director 339
- Cisco MDS 9509 Multilayer Director 340–341
- Cisco MDS 9513 Multilayer Director 341
- CKD volume 123
- CLI 144
- client options file 469
- cluster 133, 371
 - node 124
 - resources 467
- cluster cache 371
- clustered failover 173
- clustered server resources 371
- clustering 350, 468
- clutch 208
- CMIP 509
- CNT (n-type) 324
- Command Line Interface. See CLI.
- Common Information Model (CIM) 408, 509
- Compass architecture 370
- complexity 366
- Compliance framework 189
- compression technology 207
- Concurrent Copy 491
- Consistency Group 25, 109, 146
- consistency management 431
- consolidation 2, 115, 152
- constructs 487
- content manager 45
- Content Manager Family 42
- continuous operations 3, 11
- control data set 495
- Controller service mode 103

- cooling 137
- Copy Services 123, 145
- copy set 434
- copy storage pool 469
- copy-on-write 71, 105
- cost of downtime 470
- Cost of Service (CoS) 316
- Cristie 453
- CWDM 320, 352

D

- DAA 247
- DASD 489
- data
 - backup and restore 438
 - security 470
- Data Backup with no Hot-Site
 - Tier 1 17
- data bands 204–205
- data buffer 207
- data compression 211
- data lifecycle 38–39
- data lifecycle management 42–43
- data management utility program 490
- Data Manager 390
- Data Manager for Chargeback 400
- data migration 147, 378, 492
- Data ONTAP 156, 159, 164
- data path 121
- data permanence 181
- data protection solution 350
- data rate 212, 219
- data retention 7, 38, 440, 460
 - event based 195
- data set 489
- data sharing 38
- data striping 515
- data value 41
- database
 - asset reporting 396
- DataFabric Manager 184
- dataspace 499
- dataspace sorting 501
- DB2 439, 456
- DB2 Content Manager Family 42
- DEFRAG 491
- deletion
 - hold 196
 - release 196
- Dependent Write 24
- DES 453
- DFSMS 491
 - ACS 273
 - catalog 486
 - overview 491
 - Tape Library Application Interface 279
 - user interface 489
- DFSMSdfp 30, 484
 - data class 487
 - management class 488

- storage class 487
- storage group 488
- System Data Mover 491
- DFSMSdss 491
 - overview 491
- DFSMSshm 488, 491
 - Availability Management 492
 - backup 493
 - data duplication 493
 - migration 492
 - recall 492
 - scratch pool 493
 - Space Management 492
 - tape management 493
- DFSMSrmm 493–494, 499
 - environment 495
 - overview 494
 - policies 495
 - scratch volume 494
 - validation 494
 - volume management 494
- DFSORT 498
 - BLOCKSET 499
 - ICEGENER 500
 - ICETOOL 502
 - MOSIZE parameter 500
 - symbol processing 500
 - user exit 502
- DFSORT/MVS
 - DFSORT symbols 500
- DFSORT/VSE 498
 - dataspace sorting 501
 - DSPSIZE parameter 501
 - getvis 501
 - memory sorting 501
- Disaster Recovery 3, 11, 25, 148, 461
- Disaster Recovery Plan file 461
- discovery 408
 - in-band 408
- disk
 - instant copy 456
 - remote replication 470
- DISK CONTRL 127
- DISK EXP 127
- disk sanitization 186
- Distributed Management Task Force (DMTF) 408, 509
- DLT 296
- DMF 37
- DMTF 408
- double parity RAID 165
- DR550 187
 - Enhanced Remote Mirroring 191
- drive head 203
- DriveGroup 284
- DriveGroupApplication 284
- DS CLI 115
- DS Command-Line Interface (CLI) 115, 144
- DS Management Console 142
- DS Open application programming interface (API) 144
- DS Storage Manager 142

- DS300 64
- DS400 64, 68
- DS4000
 - Alert Manager 111
 - Enhanced Remote Mirroring 108
 - FlashCopy 105
 - service alert 110
 - VolumeCopy 106
 - VSS service 21
- DS4000 EXP100 96
- DS4000 Storage Manager 102
- DS4100 80, 82, 191, 193
 - dual controller 82
 - single controller 84
- DS4300 80, 85
 - dual controller 85
 - service alert 86
 - single controller 86
- DS4300 Turbo 80, 87
- DS4500 80
- DS4700 Express 80, 90
- DS4800 81, 93
- DS6000 114
 - VSS service 21
- DS6000 expansion enclosure 126
- DS6800 124
- DS8000 114, 130
 - VSS service 21
- DS8100 Model 921 138
- DS8300 Model 922 138
- DS8300 Model 9A2 132, 138
- dual accessor 241
- duplicate data 38
- DWDM 146, 320, 352, 470–471
- Dynamic Capacity Addition 86
- Dynamic Capacity Expansion (DCE) 104
- Dynamic LUN 123
- Dynamic Multipathing Driver (DMP) 104
- dynamic speed matching 208
- Dynamic Volume Expansion (DVE) 86, 104

E

- EHPCT 224
- electronic tape vaulting 469–470
- Electronic vaulting
 - Tier 3 17
- emulation
 - tape library 263
- enclosure 137
- Enclosure Services Module 70
- endpoint devices 409
- Enhanced Remote Mirroring (ERM) 108
- Enterprise Fabric Connectivity Manager (EFCM) 325, 337–338
- Enterprise Removable Media Manager see eRMM
- Enterprise Tape products
 - overview 226
- eRMM 280
- ERP 44
- error correction 214

- ESCON 134, 214, 471
- ESCON Director 309
- ESDS 485
- ESH2 164
- ESM 70, 99
- ESS 140, 456
 - VSS service 21
- ESS Model 750 142
- ESS Model 800 141
- Ethernet switch 136
- EXN1000 163
- EXP Plus 76
- EXP Plus 320 76
- EXP100 96
- EXP24 74
- EXP400 70
- EXP700
 - upgrading 99
- EXP710 99
- EXP810 100
- expanded storage 499
- expansion enclosure 126
- expansion frame 137
- expansion unit
 - EXP400 70
- Extended Fabric feature 325
- Extended Format data set 487
- Extended Format data sets 498
- Extended Long Busy (ELB) 25
- Extended Remote Copy XRC 149
- External Disk Drive 296

F

- fabric 306–310
- Fabric L1 414
- Fabric L2 415
- Fabric Manager 325
- Fabric Services Architecture 324
- Fabric Watch 325
- failover 465
- fanout 307
- FAST 151
 - service alert 83
- FAST 200 82, 84–86, 88, 90, 93, 95
 - Specification 84, 86, 90, 92, 95
- FAST600 80
- FAST600 Turbo 80
- FAST900 80
- fast-write 125
- fault-tolerant 115
- FC versus SATA disk 98
- FC-AL 121, 125, 134, 152
- FC-FC Routing 345
- FC-FC Routing Service 342, 345
- FCIP 313, 315, 317, 321
- FCIP Tunneling Service 345
- FC-MI 408
- FCP 125, 164
- feature key 107
- Fiber Channel over IP 315

- Fibre Alliance 407
- Fibre Channel 305–308
- Fibre Channel Arbitrated Loop 125
- Fibre Channel network 409
- fibre extenders 470
- FICON 120, 125, 134, 214, 330, 333, 471
- FICON In-Band Management Server 332
- FICON Management Server 338
- filemark 208, 211
- FilerView 183–184
- FlashCopy 71, 105, 145, 151, 375, 435, 439, 456, 491
 - Data Set 145
 - Incremental 145
 - Multiple Relationship 145
- FlashCopy repository 105
- FlashCopy service 20
- FlexClone 166
- FlexPar 338–339
- FlexPort 325, 328, 331
- FlexVol 166
- FSPF 335, 340
- full duplex 307
- full synchronization 110

G

- gateway 319
- GDPS 270, 279
- GDPS/PPRC 25, 27
 - management for Open Systems LUNs 29
 - Multi-Platform Resiliency for zSeries 29
 - Near Continuous Availability of data with HyperSwap 28
 - overview 27
- GDPS/PPRC HyperSwap Manage
 - overview 29
- GDPS/PPRC HyperSwap Manager 29
- GDPS/XRC 27
 - overview 30
 - topology 30
- getvis 501
- Global Copy 109, 146–147
- Global Mirror 25, 110, 146, 148, 377, 435
- Global Mirroring 110
- global spare 125
- grain 369
- group
 - collocation 444

H

- HACMP 31, 192, 465
- Hardware Management Console 132, 136
- HBA 409
- head actuator 207
- head cleaner 209
- heartbeat 465
- HFS 485
- Hierarchical File System 485
- Hierarchical Space Management (HSM) 438, 441
- Hierarchical Storage Management 6

Hierarchical Storage Manager 491
 high availability 3, 11, 371
 High Voltage Differential (HVD) 251
 Hiperspace 499
 hop 316
 host volume 436
 hot spare 83, 85
 hot site 470
 HPCT 224
 HTML output 398
 hub 307
 HyperSwap 28
 Hypervisor 131

I

I/O group 370
 IBM 81
 IBM 3480 215
 IBM 3590 213, 215–216
 IBM 3592 210, 217–218, 221
 backhitch reduction 211
 capacity scaling 212
 cartridge memory 211
 channel calibration 212
 interleaved recording 210
 nonvolatile caching 211
 segmentation 212
 speed matching 211
 IBM 7206 Model 336 297
 IBM 7207 Model 122 298
 IBM 7207 Model 330 298
 IBM 7210 External DVD-RAM Drive 299
 IBM 7212 Storage Enclosure 300
 IBM 7212-312 VXA-2 autoloader 300
 IBM lossless compression algorithm 209
 IBM Magstar 3590 212
 IBM System Storage 3592 Tape Frame Model F05 261
 IBM System Storage 3952 258
 IBM System Storage SAN18B-R multiprotocol router 343
 IBM System Storage TS3100 Tape Library 227
 IBM System Storage TS3200 Tape Library 229
 IBM System Storage TS3310 Tape Library 231
 IBM System Storage TS3500 Tape Library 234
 IBM Tivoli Continuous Data Protection for Files 440
 IBM Tivoli Storage Manager 21, 251
 API 192–193
 Client API 192–193
 IBM TotalStorage
 SAN Volume Controller 26
 IBM TotalStorage 3494 Virtual Tape Server 269
 Advanced Policy Management 272
 logical volume 272
 recall 272
 stacked volumes 273
 Tape Volume Cache 271
 virtual tape drive 271
 virtual volume 271
 IBM TotalStorage 3581 Tape Autoloader 250
 IBM TotalStorage EXP24 74
 IBM TotalStorage Expert 476
 IBM TotalStorage Productivity Center for Fabric 405
 IBM TotalStorage SAN Volume Controller 366
 IBM TotalStorage SAN140M 337
 IBM TotalStorage SAN16B-2 326
 IBM TotalStorage SAN16B-R multiprotocol router 345
 IBM TotalStorage SAN16M-2 327
 IBM TotalStorage SAN16M-R multiprotocol SAN router 346–347
 IBM TotalStorage SAN256B 336
 IBM TotalStorage SAN256M 338
 IBM TotalStorage SAN32B-2 330
 IBM TotalStorage SAN32M-2 331
 IBM Virtualization Engine TS7510 258, 260
 application software 263
 Auto Archive 265
 Auto Replication 266
 backup compression 264
 compression 264
 failover/failback 268
 import/export 264
 Management Console 262
 Network Copy on Export 266
 network replication 265
 Physical Copy on Export 265
 remote copy 267
 tape drives 263
 tape library emulation 263
 virtual tape volumes 263
 ICEGENER 500
 ICETOOL 500, 502
 ICETPEX 499
 IDCAMS 490
 IEBCOMPR 490
 IEBCOP 490
 IEBDG 490
 IEBEDIT 490
 IEBGENER 490, 500
 IEBPTPCH 491
 IEBUPDTE 491
 IEHINIT 491
 IEHLIST 491
 IETF 36
 iFCP 316, 318, 321
 ILM 1–2, 35–36, 268
 SNIA 37
 ILM elements 39
 image backup 459
 inactive data 38
 in-band 408
 increasing complexity 366
 Information Lifecycle Management 1–2, 188
 Information on demand 2
 information on demand 1
 Infrastructure Simplification 1–2, 49, 349, 351
 instant copy 456
 Integrated Service Console 442
 Integrated Solutions Console 446
 Interactive Storage Management Facility 487, 489
 interleaved recording 203, 210, 213
 intermix 103

- Interoperability Matrix 154
- inter-switch link 310
- ISCSI 313
- iSCSI 64, 67, 157, 313–314, 317, 321, 471
- ISL Trunking 326
- ISM 489
- ISMF 487, 489
- ISO 36
- ISV 453
- ITIL 389

J

- J cartridge 224
- JBOD 99
- JCL 486
- journal volume 436

K

- Kerberos 453
- KSDS 485–486

L

- lambda 320
- LAN-free backup 452, 455
- LDS 485
- leader pin 222
- library partitions, 254
- LIC 152
- Licensed Machine Code 152
- life cycle 188
- line module 339
- linear density 206–207
- Linear Tape-Open (LTO) 201
- Linux 83
- local mirroring
 - N series
 - local mirroring 170
- LockVault 183
- logical drive
 - primary 106
- logical partition 131, 135
- logical subsystem 123
- logical unit (LU) 370
- logical volume 459
- Logical Volume Storage Agent 459
- longitudinal positioning 204
- loop 306–307
- Lotus Domino 439
- Lotus Notes 439
- Low Voltage (LVD) 252
- LPAR 28, 130–131, 135
- LPOS 204
- LTO 202
 - cartridge 222
 - compatibility 205
 - data tracks 203, 205
 - interleaved recording 203
 - linear density 206

- longitudinal positioning 204
- servo tracks 203
- shingling 205
- tape cartridge 222
- WORM 207–208
- LTO Ultrium 3 206
- LTO-CM 209
- LTO-DC 209
- LUN 370
- LUN expansion 71
- LVSA 459

M

- Magstar 212, 214
 - 3590 220
- maintaining availability 366
- MAN 471
- managed disk 370
- managed disk group 373
- managed disk group (MDG) 370
- management class 488
- Management Information Base (MIB) 408
- management station 69
- managing storage growth 366
- master console 370
- McData (m-type) 324
- MDisk 370, 372
- MDS 9120 333
- MDS 9140 333
- MDS 9216 335, 348
- media 207
- Media Manager 285
- memory object 499–500
- metal particle 207
- Metro Mirror 32, 71, 146, 376, 435
 - Tier 6 26
- Metro Mirror (Synchronous PPRC) 109
- Metro Mirroring 109
- MIB 408
- Microsoft Cluster Server. *see* MSCS
- Microsoft Exchange 439
 - N series backup 179
 - Single Mailbox Recovery 180
- Microsoft SQL Server 439
 - N series backup 180
- Microsoft Volume Shadow Copy Service 20
- Microsoft VSS 20
- midmarket tape products
 - overview 202
- migration 488
- Migration Control Data Set 492
- mini-map 412
- mining 188
- mirroring 170, 515
- MMC 179
- monitoring agent 69
- MPP 103
- MSCS 180, 467–468
 - and TSM 467
- multipath architecture 229, 232, 252, 254

- Multi-Path Proxy 103
- multipathing 117
- multi-pathing device driver 374
- Multiple Allegiance 119
- multiprotocol router 342
- MultiStore 168
- multi-tiered storage 120
- mutual suspicion algorithm 453

N

N Series

- SnapRestore 21
- SnapVault 21

N series 21, 156

- appliance model 159
- authentication 185
- backup 177
- CLI 184
- clustered failover 173
- consolidation 168
- data backup 175
- Data ONTAP 156, 159, 164
- data permanence 181
- DataFabric Manager 184
- disk sanitization 186
- dynamic volume expansion 166
- encryption 185
- Exchange single mailbox recovery 180
- EXN1000 163
- expansion unit 163
- FCP 164
- FilerView 183–184
- FlexClone 166
- FlexVol 166
- gateway model 159
- instant replication 166
- LockVault 183
- management 183
- mirroring 170
- MS Exchange 179
- MS SQL Server 180
- MultiStore 168
- N3700 157
- N5000 159
- NDMP 177
- operating system 164
- Oracle integration 181
- OSSV 176
- RAID 165
- rapid backup and restore 179
- records retention 182
- remote mirroring 171
- SecureAdmin 185
- SnapDrive 178
- SnapLock 182
- SnapManager 22
- SnapMirror 171, 182
- SnapRestore 170
- SnapShot 21, 169
- SnapValidator 181

- SnapVault 175
- storage virtualization 178
- synchronous local mirroring 170
- SyncMirror 170
- telnet 184
- tiered storage 168
- WORM 156

N5000 159

NAS

- TSM 457

NAS Gateway 319

NAS pooling 319

NaviQuest 490

NDMP 177, 438

network attached storage (NAS) 318, 471

NFS 164, 304, 318–319

NOLIMIT 194

non-cascaded 309

nonvolatile caching 211

NVC 211

NVS 116, 125

O

- off-site backups 147
- off-site data storage 470
- off-site storage 461, 469, 472
- on demand 1–2, 130, 140
- on demand storage environment 5
- open software 6
- optical media 445
- Oracle 439, 456
 - N series validation 181
- Oracle HARD 181
- OSSV 176
- outband discovery 409
- out-of-band 408

P

- Parallel Access Volume 119
- Parallel Access Volumes (PAV) 118
- Parallel Sysplex 492
- partitioned data set 485
- PAV 118–119
- Pay-as-you-grow 324
- PDS 485
- Peer-to-Peer VTS 270
- PFA 122
- pin retention 207
- ping 395, 399
- point-in-time 106–107
- point-in-time (PIT) copies
 - Tier 4 17
- point-to-point 306–307, 310
 - virtual 307
- policy-based archive management 44
- policy-based management 391
- Ports on Demand 327, 330, 333, 345
- power 137
- power management 208

- power subsystem 126
- power supply 126
- POWER5 130
- PowerPC architecture 124
- PPRC 471
- predictive failure analysis 133
- predictive failure analysis (PFA) 122
- primary 106
- primary storage pool 469
- Priority I/O Queuing 119
- probe 396, 398–399
- processor 124, 130
- processor complex 133
- progressive backup 452
- progressive incremental backup methodology 446
- progressive incremental backups 446
- provisioning 3
- PtP Specialist 280

Q

- QSAM 485
- Quality of Service (QoS) 316
- Queued Sequential Access Method (QSAM) 485

R

- R/3 456
- Rack Mount Option 251
- RAID
 - distributed parity 516
 - independent access 516
 - levels
 - RAID 1 515
 - mirroring 515
- RAID 0 514
- RAID 1 515
- RAID 10
 - implementation 518
- RAID 5 516
- RAID-DP 165
- RAID-like 214
- rapid backup and restore 179
- Rapid Data Recovery 22
- Rapid Data Recovery for zSeries 25
- RCMF/PPRC 27
- RCMF/XRC 27
- RDAC 103
- rebranding 81
- recording format 203
- records retention 182
- Recovery management 7
- Recovery Point Objective 13
- Recovery Time Objective 13
- Redbooks Web site 530
 - Contact us xxvii
- reference data 41, 188
- regulation 188
- regulatory requirements 38
- remote disk replication 470
- remote mirror 146

- remote mirror and copy 146
- remote mirroring 171
- remote switch 326
- Removable Media Manager 494
- rendering 188
- replication 148
- reporting
 - assets 399
 - availability 399
 - by userID 397
 - capacity 399
 - usage 399
 - usage violation 400
- repository logical drive 105
- resource pooling 319
- Resume Mirror 110
- retention managed data 41, 188
- retention period 188
- RETINIT 194
- RETMIN 194–195
- RETVR 194–195
- RIO 131
- RIO-G 131, 135
- RNID 409
- robotic cartridge 239
- robotic system 253
- rolling disaster 23, 431
- Routing Service 345
- RRDS 485
- RSM 499
- RTO 16

S

- SAN 470
 - transmission distance 471
- SAN File System
 - FlashCopy solution 22
- SAN management 407
- SAN solutions 349
- SAN Switch
 - 08 330, 332
 - H08 327
- SAN switch
 - features 324
- SAN Switch 16B-2 326
- SAN Switch 16M-2 326
- SAN switch Cisco MDS 9020 326
- SAN Switch L10 326
- SAN switches
 - b-type 323
 - Cisco 323
 - m-type 323
 - n-type 323
- SAN topologies 306
 - loops and hubs 307
 - point-to-point 306
 - switched fabric without cascading 309
- SAN Volume Controller 26, 361, 363, 365, 367
 - architecture 370
 - block aggregation 372

- cache 371
- capacity utilization 379
- cluster 371
- compatibility 379
- configuration node 369
- copy services 369
- data migration 368
- extent 369
- Global Mirror 376–377
- grain 369
- GUI 368
- I/O group 370
- licensing 375
- managed disk 370
- managed disk group 370, 373
- master console 370
- MDisk 370, 372
- Metro Mirror 376
- microcode 371
- scalability 368
- storage utilization 368
- UPS 372
- usage 377
- vDisk 370
- virtual disk 370
- volume management 372
- SAN16M-2 350
- SAN16M-R 346
- SAN256M 338
- SAN32M-2 350
- SANvergence Manager 346
- SAP R/3 439
- Sarbanes-Oxley 42
- SARC 122
- SATA 42, 82, 103, 190
- SATA Interface Card 96
- SATA versus FC disk 98
- SBOD 99, 126
- scalable performance 308
- scaling 212
- scan 398
- scratch pool 493
- scratch volume 494
- ScratchPools 284
- SCSI 506
- SCSI protocol 409
- SDD 117, 374
- SDM 30
- SEC 190
- SecureAdmin 185
- segmentation 212
- self-discovering 307
- Semantic Zooming 413
- sequential data set 484
- Sequential Prefetching in Adaptive Replacement Cache (SARC) 122
- Serial Storage Architecture 141, 506
- serpentine recording format 213–214
- server consolidation 168
- Server replication 22
- ServeRAID Manager 69
- Service Alert 110
- service alert 83, 86, 110
- Service Location Protocol 510
- service processor 133
- servo 208
- servo bands 203–204
- servo bursts 203–204
- servo tracks 203, 205
- session 435
- SFP 339
- shingling 205
- S-HMC 131
- Single Mailbox Recovery 180
- single parity RAID 165
- SLP 510
- SMBR 22
- SMI 37
- SMI-S 37, 368, 408, 509
- SMS 486
- SnapDrive 178
- SnapLock 182
- SnapManager for Microsoft SQL Server 22
- SnapMirror 171, 182
- SnapRestore 21, 170
- SnapShot 21, 169
- SnapValidator 181
- SnapVault 21, 175
- SNIA 36, 362
- SNIA Storage Model 361
- SNMP 409–410
 - events 410
 - trap 397
- spanning 99
- spare 125
- spare drive 122
- speed matching 208, 210–211
- SSA 121, 141, 506
- SSA Optical Extender 507
- standards 36
- Standby Capacity on Demand (Standby CoD) 140
- Statistical Analysis and Reporting System (SARS) 209, 214
- storage
 - asset discovery 396
 - capacity 399
 - growth 399
 - reporting 398
- storage class 487
- storage consolidation 168, 350
- storage group 488
- storage growth 366
- storage infrastructure 6
- storage LPAR 131
- storage management 38
- Storage Management Initiative Specification (SMI-S) 408
- storage mirroring
 - Tier 6 18
- Storage Networking Industry Association (SNIA) 407

- storage orchestration 7
- storage policies 391
- storage pool 445
- storage replication 22
- storage utilization 407
- storage virtualization 6
- Subsystem Device Driver (SDD) 117, 374
- Surface Control Guiding Mechanism 208
- Suspend Mirror 110
- SVC
 - VSS service 21
- SVC see SAN Volume Controller
- switch 83, 85, 307–308, 310, 409
- switched (SAN topology) 306
- Switched Hub architecture 164
- switched point-to-point configuration 310
- Symantec 453
- symbol 500
- Synchronous data transfer 306
- synchronous local mirroring 170
- synchronous mirroring 109
- Synchronous PPRC 146
- SyncMirror 170
- System Data Mover 491

T

tape

- 3580 216
- 3581 250
- 3590 220
- barcode reader 251
- TS3100 227
- TS3200 229
- TS3310 231
- TS3500 234
- tape cartridge 222
- tape duplication 493
- tape library
 - emulation 263
- tape management 493
- tape vaulting 469
- Tape Volume Cache 271
- TCO 36, 41
- TDP 439
- Tier 23
- Tier 0 16
- Tier 1 17
- Tier 2 17
- Tier 3 17
- Tier 4 17, 20, 22
- Tier 5 17
- Tier 6 18, 22
- Tier 7 18
- tiered storage 39, 98, 168
 - SAN Volume Controller
 - tiered storage 368
- tiers
 - blending 18
 - Business Continuity 15
- Tivoli Data Protection. see TDP

- Tivoli Disaster Recovery Manager 461
- Tivoli Ready certification 324
- Tivoli Storage Manager 21
 - client configuration information 461
 - disaster recovery 461
 - image backup 459
 - Logical Volume Storage Agent 459
 - N series backup 177
 - NDMP 177
 - off-site data copy 461
 - Original Block File 459
 - server recovery 461
- Tivoli Storage Resource Manager
 - chargeback 400
- Tivoli Storage Resource Manager for Databases
 - asset reporting 396
- TOC 177
- topology map 408
- Topology Viewer
 - Fabric view 414
 - Graphical View 412
 - launch point 410
 - mini-map 412
 - Semantic Zooming 413
 - Tabular View 412
 - zones 416
 - zoom levels 413
- TotalStorage Productivity Center 144, 363, 382
- TotalStorage Productivity Center for Data 382
- TotalStorage Productivity Center for Disk 382
- TotalStorage Productivity Center for Fabric 324, 382
- TotalStorage Productivity Center for Replication 383
- TPC for Data 390
 - alert 397
 - alert log 397
 - asset reporting 396, 399
 - backup reporting
 - reporting
 - backup 400
 - capacity reporting 399
 - chargeback 400
 - discovery 395
 - ping 395, 399
 - probe 396, 398–399
 - reporting 398
 - scan 398
 - scheduled jobs 395
 - summary reporting 399
 - usage reporting 399
 - user space summary 394
 - username reporting 397
- TPC for Disk
 - threshold settings 425
- TPC for Fabric
 - database 409
 - topology map 409
- TPC for Replication 26, 430
 - copy set 434
 - session 435
- TPC for Replication Two Site Business Continuity 430

- track following 203–204, 210
- tracks 208
 - positioning 203
- transaction integrity
 - Tier 5 17
- TS7510 Cache Controller 261
- TS7510 Cache Module 261
- TS7510 see IBM Virtualization Engine TS7510
- TSM
 - adaptive subfile backup 452
 - adaptive subfile differencing 443
 - Administration Center 446
 - and HACMP 465
 - and MSCS 467–468
 - API 439, 453
 - archive 439, 446
 - backup 445
 - backup and restore 438
 - backup methods 454
 - backup topology 454
 - backup/archive client 445
 - backup/restore operations 454
 - Bare Machine Recovery 440
 - checkpoint restart 443
 - client backup methods 454
 - client configuration information 463
 - client interface 451
 - client node 449
 - client Web browser interface 451
 - client with MSCS 468
 - CLUSTERNODE 469
 - collocation 444, 472, 474
 - command routing 463
 - COMMRESTARTDURATION 465
 - COMMRESTARTINTERVAL 465
 - copy storage pool 445, 472–473
 - CPU utilization 454
 - cyclical redundancy check 443
 - data movement 445
 - data movements 443
 - Data Retention 440
 - database 445, 471
 - database backup 470
 - database mirroring 470
 - Disaster Recovery plan file 464
 - distributed configuration 463
 - DRM 464, 469
 - electronic tape vaulting 469–470
 - encryption 443
 - enterprise administration 463
 - expiration 469, 472
 - HACMP 466
 - high availability clustering 464
 - HSM 441, 467
 - image backup 459
 - LAN-based clients 454
 - LAN-free backup 455
 - logical volume backup 459
 - management class 459
 - metadata 445
 - migration 472
 - monitoring 463
 - NAS 457
 - off-site data copy 469, 472
 - overview 437
 - parallel backup 452
 - point-in-time recovery 459
 - policy 446
 - policy domain 449
 - primary storage pool 472
 - progressive backup 443, 452, 454
 - progressive incremental backups 446
 - reclamation 472–473
 - scheduled operations 466, 468
 - Scheduler Service 466, 469
 - SCSI reserve 468
 - security 453
 - security concepts 453
 - server 445
 - server administration 446
 - server maintenance 450
 - server-to-server communication 463
 - snapshot image backup 459
 - source server 464
 - split-mirror backup 456
 - Storage Agent 455
 - storage device 449
 - storage pool backup 469
 - storage pool hierarchy 445
 - storage pools 471
 - tape reclamation 444
 - tape vaulting 469
 - target server 464
 - vaulting 469
 - Version 5 452
 - virtual volumes 463–464
 - volume history 469
 - volume tracking 469
- TTPC for Data
 - HTML output 398
- Tunneling Service 345
- Turbo
 - FAST600 80
- Turbo II 141

U

- Ultra320 74–75
- UltraBac 453
- Ultrium 202
 - compatibility 209
- Ultrium 1 209
- Ultrium 2 209
- Ultrium 3 206
 - capacity 222
 - tape length 222
 - WORM 207
- Ultrium cartridge 222
- uninterruptible power supply 372
- usage reporting 399
- user space summary 394

V

- vDisk 370–371
- Virtual disk (vDisk) 370
- Virtual Disk Service (VDS) 103
- virtual I/O slots 243
- virtual pool 366
- Virtual SAN (VSAN) 334
- Virtual Storage Access Method 485
- Virtual Tape Controller 279
- Virtual Tape Library 257, 494
- Virtual Tape Server (VTS) 220
- virtualization 2, 123, 131, 358
- vital record retention, archive and retrieval 439
- voltage regulation 372
- volume control region 214
- volume management 372
- volume serial (VOLSER) 242
- Volume Shadow Copy Service 20
- Volume Shadow Copy Service (VSS) 103
- VolumeCopy 106
- VSAM 485
- VSAN 333, 340, 342
- VSS 20, 103, 180, 440
- VTL 257
- VTOC 486
- VTS 234
- VTS see IBM TotalStorage 3494 Virtual Tape Server

W

- WAFL 21, 169
- Wavelength Division Multiplexing (WDM) 320
- WBEM 408
- Web Based Enterprise Management (WBEM) xxv, 408, 509, 546
- WebSphere Application Server 439
- Windows
 - Logical Volume Storage Agent 459
- Windows 2000
 - snapshot image backup 459
- WORM 156, 182, 192, 207–208, 222, 461
- write process 222

X

- XFP 339

Z

- z/OS Global Mirror 149
- z/OS Workload Manager (WLM) 119
- zone 409
- zone member 409
- zone set 409
- zones 416
- zoning 406
- zSeries File System(zFS) 486



Redbooks

IBM System Storage Solutions Handbook

(1.0" spine)

0.875" x 1.498"

460 <-> 788 pages



IBM System Storage Solutions Handbook



Learn about IBM storage products and solutions

Flexible, scalable storage solutions for the on demand world

Find the products that answer your business needs

This IBM Redbook provides overviews and pointers for information about the most current IBM System Storage and TotalStorage products, showing how IBM delivers the right mix of products for nearly every aspect of business continuance and business efficiency. IBM TotalStorage products can help you store, safeguard, retrieve, and share your data.

- ▶ First we introduce the basic storage solutions areas: information lifecycle management, infrastructure simplification, and business continuity.
- ▶ Part 1 describes disk products, including the entire Disk Storage (DS) Series, from entry-level offerings to mid-range with the DS4000 family and DS6000, to high end with the DS8000. It also includes the DR550.
- ▶ Part 2 is an overview of tape products, covering tape drives, tape libraries, and virtualization products, including LTO and 3590, and 3592 technology.
- ▶ Part 3 starts by describing storage networking infrastructure and protocols, and then presents the switches and directors to form SAN solutions.
- ▶ Part 4 discusses the IBM System Storage software portfolio for open systems and includes Storage Virtualization products, such as the SAN Volume Controller, TotalStorage Productivity Center, Tivoli Storage Manager, and ETL Expert.
- ▶ Part 5 describes the z/OS storage management software: DFSMS and DFSORT.
- ▶ Part 6 is the appendixes, which cover standards, such as Storage Management Initiative Specification (SMI-S), Common Information Model (CIM), and Web Based Enterprise Management (WBEM); also discussed is redundant array of independent disks (RAID).

This redbook is intended as a reference for basic and comprehensive information about the IBM System Storage and TotalStorage products portfolio. This book provides a starting point when establishing your own enterprise storage environment.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

**For more information:
ibm.com/redbooks**