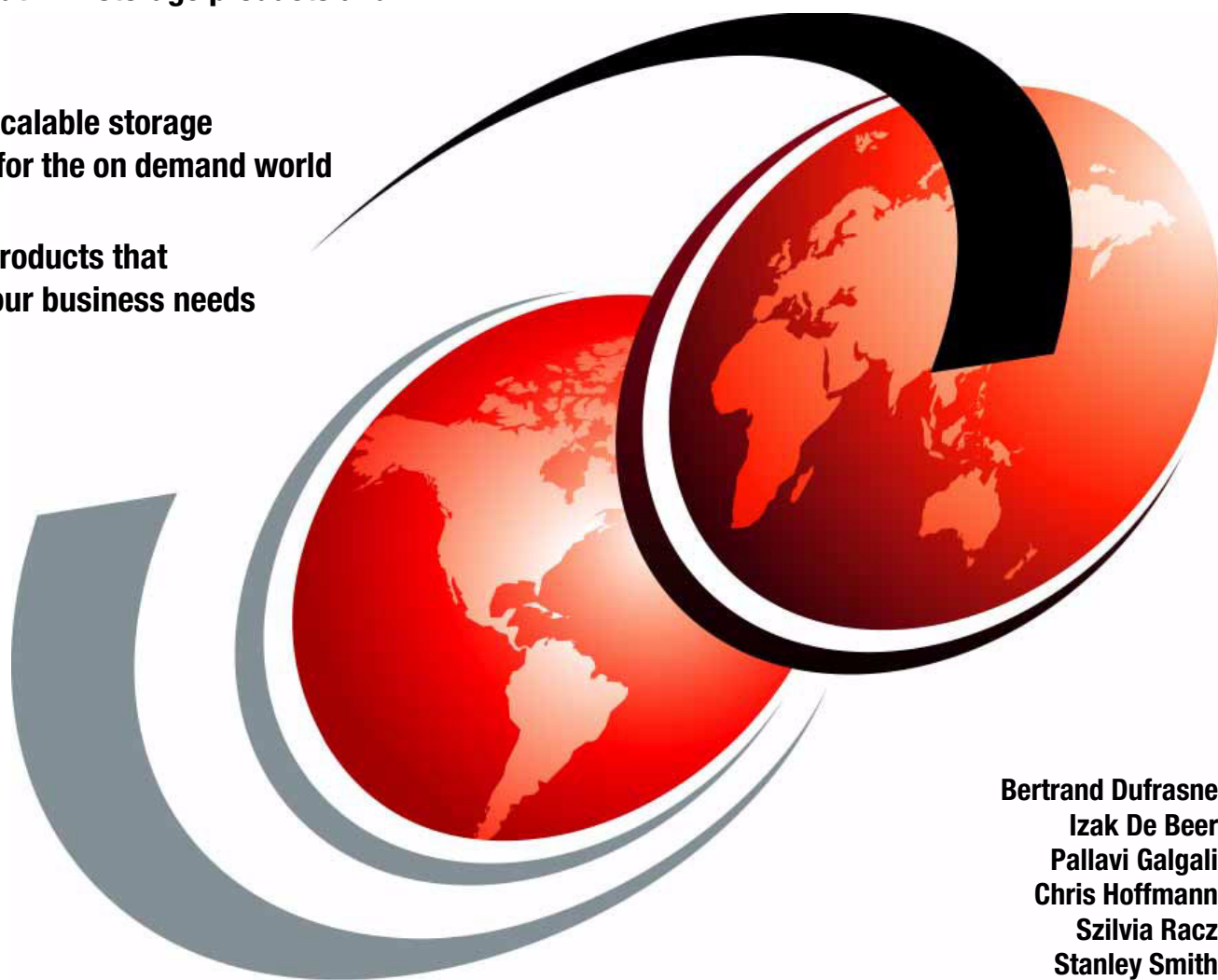


IBM TotalStorage Solutions Handbook

Learn about IBM storage products and solutions

Flexible, scalable storage solutions for the on demand world

Find the products that answer your business needs



Bertrand Dufrasne
Izak De Beer
Pallavi Galgali
Chris Hoffmann
Szilvia Racz
Stanley Smith

Redbooks



International Technical Support Organization

IBM TotalStorage Solutions Handbook

July 2005

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

Sixth Edition (July 2005)

This edition applies to the IBM TotalStorage products portfolio as of March 2005.

© Copyright International Business Machines Corporation 1999, 2003, 2004, 2005. 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	xiii
Notices	xix
Trademarks	xx
Preface	xxi
The team that wrote this redbook.	xxi
Become a published author	xxiii
Comments welcome.	xxiii
Chapter 1. Introduction: Information on demand with IBM TotalStorage	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	8
Part 1. Disk storage solutions	9
IBM TotalStorage point-in-time copy comparison	12
FlashCopy function definitions	12
IBM TotalStorage disk mirroring comparison	13
Disk mirroring function definitions	14
Chapter 2. Entry-level disk storage	17
2.1 Overview	18
2.2 Specifications	20
2.3 IBM TotalStorage DS300	20
2.4 IBM TotalStorage DS400	22
2.5 ServeRAID Manager	22
2.6 Premium features	24
2.6.1 EXP400 expansion units.	24
2.6.2 FlashCopy	24
2.6.3 Metro Mirror	24
2.7 Advanced software features	25
2.8 More information	26
Chapter 3. Mid-range disk systems	27
3.1 Highlights: What's new	28
3.2 Overview	29
3.3 DS4100 Storage Server	30
3.3.1 Features	30
3.3.2 Specifications	32
3.4 DS4300 Storage Server	32
3.4.1 Features	33
3.4.2 Specifications	34
3.5 DS4300 Turbo Storage Server	34
3.5.1 Features	35
3.5.2 Specifications	36

3.6 DS4400 Storage Server	37
3.6.1 Features	37
3.6.2 Specifications	38
3.7 DS4500 Storage Server	39
3.7.1 Features	39
3.7.2 Specifications	39
3.8 EXP100 expansion unit	40
3.8.1 Features	40
3.8.2 Hardware description	41
3.8.3 Specifications	41
3.8.4 SATA versus FC disks	42
3.9 EXP700 expansion unit	43
3.9.1 Features	43
3.9.2 Specifications	43
3.10 EXP710 expansion unit	44
3.10.1 Features	44
3.10.2 Specifications	44
3.11 DS4000 Storage Manager	45
3.12 Copy services	47
3.12.1 FlashCopy	47
3.12.2 VolumeCopy	48
3.12.3 Enhanced Remote Mirroring	50
3.13 Integrated Backup for Databases	53
3.14 DS4000 Service Alert	53
3.15 DS4000 Alert Manager	54
3.16 IBM Management Suite Java (FASTT MSJ)	54
3.17 More information	54
Chapter 4. Enterprise Disk Systems	57
4.1 Introduction	58
4.1.1 Storage consolidation	59
4.1.2 Performance	59
4.1.3 Data protection	61
4.1.4 Common set of functions	61
4.2 Overview of the DS6000 and DS8000 series	61
4.2.1 Enhancements of the DS6000 and DS8000 series	62
4.3 The DS6000 series	65
4.3.1 The DS6800 Controller Enclosure	66
4.3.2 The DS6000 Expansion Enclosure	68
4.3.3 DS6800 Model 1750-511	69
4.3.4 DS6000 Model 1750-EX1	70
4.3.5 More information	71
4.4 The DS8000 series	72
4.4.1 The DS8000 design enhancements	72
4.4.2 The DS8000 architecture	74
4.4.3 DS8000 Models	79
4.4.4 IBM Standby Capacity on Demand offering for the DS8000	80
4.4.5 More information	81
4.5 The Enterprise Storage Server Family	81
4.6 Storage management	83
4.6.1 ESS Specialist	83
4.6.2 DS Management Console	84
4.6.3 IBM TotalStorage Expert	86

4.6.4 IBM TotalStorage Productivity Center	86
4.7 Copy Services	86
4.7.1 IBM TotalStorage FlashCopy (Point-in-time Copy feature)	86
4.7.2 Remote Mirror and Copy feature	88
4.8 Comparison of products	92
4.9 Additional information	94
Chapter 5. 7133 serial disk system	97
5.1 Overview	98
5.2 7133 key features	98
5.3 More information	100
Chapter 6. Expandable Storage Plus	101
6.1 Overview and key features of the 320 models	102
6.1.1 High availability	102
6.1.2 Flexible and scalable storage	103
6.2 More information	103
Part 2. Tape products	105
Chapter 7. IBM TotalStorage LTO products	107
7.1 Overview	108
7.2 The IBM LTO Ultrium 3 Technology	109
7.2.1 Description	109
7.2.2 Enhanced features	110
7.3 IBM LTO tape family product overview	112
7.3.1 IBM TotalStorage Ultrium External Tape Drive 3580	113
7.3.2 IBM TotalStorage LTO Ultrium Tape Autoloader 3581	115
7.3.3 IBM TotalStorage LTO Ultrium Tape Library 3582	117
7.3.4 IBM TotalStorage LTO Ultrium Tape Library 3583	118
7.3.5 IBM TotalStorage Ultrium Tape Library 3584	123
7.4 More information	129
Chapter 8. IBM Enterprise Tape	131
8.1 Overview	132
8.2 Enterprise Tape System 359x family	133
8.2.1 IBM 3590 Enterprise Tape Drive	133
8.2.2 IBM 3592 Enterprise Tape Drive	135
8.2.3 IBM TotalStorage Tape Controller 3590 Model A60	137
8.2.4 IBM TotalStorage Tape Controller 3592 Model J70	138
8.2.5 IBM 3590 cartridges	139
8.2.6 IBM 3592 cartridges	139
8.3 IBM 3494 Enterprise Tape Library	139
8.3.1 Tape Library Base Frame	140
8.3.2 Tape Drive Expansion Frame	141
8.3.3 Tape Storage Frame	142
8.3.4 High Availability Frame	142
8.4 3494 Virtual Tape Server	143
8.4.1 VTS models	146
8.5 More information	147
Chapter 9. Complementary storage products	149
9.1 Overview	150
9.2 IBM 7204 external disk drive	152

9.3 IBM 7205 SDLT tape drive	152
9.4 IBM 7206 4mm tape drive	153
9.4.1 Model 110	153
9.4.2 Model 220	153
9.4.3 Model 336	154
9.4.4 Model VX2	154
9.5 IBM 7207 SLR tape drive	154
9.5.1 Model 122	154
9.5.2 Model 330	154
9.6 IBM 7208 8mm tape drive	155
9.7 IBM 7210 DVD-RAM drive	155
9.7.1 Model 025	156
9.7.2 Model 030	156
9.8 IBM 7212 storage device enclosure	156
9.9 IBM 7212-312 VXA-2 autoloader	157
9.10 IBM 7332 4mm autoloader	157
9.10.1 Model 110	157
9.10.2 Model 220	158
9.11 More information	158
Part 3. Storage networking	159
Chapter 10. Introduction to storage networking	161
10.1 Overview	162
10.2 Storage Area Network (SAN)	162
10.2.1 Fibre Channel	163
10.2.2 SAN topologies	164
10.2.3 Physical components of the SAN infrastructure	168
10.2.4 Naming and addressing	169
10.2.5 FICON	170
10.3 Extended distance fabrics	170
10.4 IP Storage Networking technologies	171
10.4.1 iSCSI	172
10.4.2 FCIP	173
10.4.3 iFCP	174
10.4.4 Comparison iSCSI - FCIP - iFCP	175
10.5 Network Attached Storage (NAS)	176
10.5.1 NAS gateways	178
10.6 Wavelength Division Multiplexing	178
10.7 Selecting the best alternative	179
10.8 More information	179
Chapter 11. Storage Area Network products	181
11.1 SAN switches and connectivity products	182
11.1.1 Common characteristics	182
11.1.2 Other switch features	182
11.1.3 More information	184
11.2 Entry-level SAN switches	184
11.2.1 IBM TotalStorage Switch L10	184
11.2.2 IBM TotalStorage SAN Switch H08	185
11.2.3 IBM TotalStorage SAN Switch SAN12M-1	186
11.3 Mid-range SAN switches	187
11.3.1 IBM TotalStorage SAN Switch H16	188
11.3.2 IBM TotalStorage SAN Switch F32	188

11.3.3	IBM TotalStorage SAN Switch SAN32B-2 Fabric Switch	190
11.3.4	IBM TotalStorage SAN Switch SAN24M-1	190
11.3.5	IBM TotalStorage SAN32M-1	192
11.3.6	Cisco MDS 9120 and 9140 Multilayer Fabric Switch	193
11.3.7	Cisco MDS 9216 Multilayer Fabric Switch	194
11.4	Enterprise SAN switches	196
11.4.1	IBM TotalStorage SAN Director M14	196
11.4.2	IBM TotalStorage SAN16B-R multiprotocol router	197
11.4.3	IBM TotalStorage SAN140M	198
11.4.4	IBM TotalStorage SAN16M-R multiprotocol SAN router	199
11.4.5	IBM TotalStorage SANC40M Cabinet	200
11.4.6	IBM TotalStorage SAN256M	201
11.4.7	Cisco MDS 9506 Multilayer Director	202
11.4.8	Cisco MDS 9509 Multilayer Director	203
11.4.9	IBM TotalStorage SAN256N	204
11.4.10	CNT UltraNet Multi-service Director	205
11.5	SAN solutions	206
11.5.1	IBM TotalStorage entry-level SAN solutions	207
11.5.2	IBM TotalStorage mid-range SAN solutions	208
11.5.3	IBM TotalStorage enterprise SAN solutions	209
Part 4.	IBM TotalStorage open software	211
Chapter 12.	IBM TotalStorage Virtualization	213
12.1	What storage “virtualization” is	214
12.1.1	Levels of storage virtualization	215
12.2	IBM TotalStorage approach to virtualization	217
12.3	IBM TotalStorage virtualization family	219
12.4	More information	219
Chapter 13.	SAN Volume Controller	221
13.1	IBM TotalStorage SAN Volume Controller (SVC)	222
13.1.1	Overview	222
13.2	Glossary of commonly used terms	223
13.2.1	SVC architecture	225
13.2.2	SVC virtualization	226
13.2.3	SVC software licensing	228
13.2.4	SVC usage	228
13.2.5	SVC compatibility	230
13.3	More information	231
Chapter 14.	SAN File System	233
14.1	Overview	234
14.2	What a SAN File System is	235
14.3	SAN File System architecture	236
14.4	SAN File system terminology	238
14.5	Supported systems and planning considerations	241
14.6	Key features	243
14.7	More information	245
Chapter 15.	IBM TotalStorage Productivity Center	247
15.1	Overview	248
15.2	Launchpad	248
15.3	Rebranding to IBM TotalStorage Productivity Center	249

15.4	Productivity Center with Advanced Provisioning	250
15.5	Architecture	251
15.6	Key benefits and highlights	253
15.7	More Information	253
Chapter 16. IBM TotalStorage Productivity Center for data		255
16.1	Overview	256
16.2	Product highlights	257
16.3	Architecture	258
16.4	Planning considerations and prerequisite	260
16.4.1	Server recommendations	260
16.4.2	Supported systems and databases	261
16.4.3	Security considerations	261
16.5	Key features	262
16.5.1	Basic menu displays	262
16.5.2	Discover and monitor information	263
16.5.3	Reporting	265
16.5.4	Alerts	267
16.5.5	Chargeback: Charging for storage usage	268
16.6	More information	270
Chapter 17. IBM TotalStorage Productivity Center for Fabric		271
17.1	SAN management	272
17.1.1	Standards and SAN management tools	272
17.1.2	Discovery	272
17.1.3	Zones and zone control	274
17.1.4	Topology views	275
17.1.5	SNMP events	275
17.1.6	Fault management	275
17.1.7	iSCSI support	276
17.1.8	Tivoli Common Agent Services	277
17.1.9	Interfaces to Fabric Manager	278
17.1.10	Integrating Fabric Manager into an SNMP management solution	278
17.2	TPC for Fabric - overview	278
17.3	TPC for Fabric - architecture and components	279
17.3.1	Manager	279
17.3.2	Agents	280
17.3.3	DB2 and WebSphere®	280
17.3.4	NetView	280
17.3.5	Optional components	280
17.4	TPC for Fabric V2.2 - New features	281
17.4.1	Zone control	281
17.4.2	IBM TotalStorage Productivity Center for Fabric API	282
17.4.3	IBM Cloudscape database	282
17.4.4	Database snapshot of topology	283
17.4.5	Selectable launch	283
17.5	Functions	284
17.5.1	Display a topology map of the SAN in physical and logical views	284
17.5.2	Real-time monitoring and alerts	291
17.5.3	Error Detection and Fault Isolation	293
17.5.4	Launch vendor-provided applications to manage components	295
17.5.5	System management	296
17.6	More information	296

Chapter 18. TotalStorage Productivity Center for Disk	297
18.1 Overview	298
18.2 TotalStorage Productivity Center Common Base	298
18.3 Performance monitoring and management	300
18.3.1 IBM TotalStorage Expert	306
18.4 More information	306
Chapter 19. TotalStorage Productivity Center for Replication	307
19.1 Overview	308
19.2 TotalStorage Productivity Center Common Base	308
19.3 Replication Manager functionality	309
19.3.1 Replication session	311
19.3.2 Storage group	312
19.3.3 Storage pools	312
19.3.4 Relationship of group, pool, and session	312
19.3.5 Copyset and sequence concepts	313
19.4 More Information	314
Chapter 20. IBM Tivoli Storage Manager	315
20.1 Introduction to IBM Tivoli Storage Manager	316
20.1.1 IBM Tivoli Storage Manager family of products	316
20.1.2 Licensing	316
20.1.3 IBM Tivoli Storage Manager core functionality	317
20.1.4 IBM Tivoli Storage Manager key attributes	319
20.2 How Tivoli Storage Manager works	323
20.2.1 Tivoli Storage Manager backup and archive concepts	324
20.2.2 Tivoli Storage Manager Administration Center	324
20.2.3 IBM Tivoli Storage Manager backup/archive client	329
20.2.4 IBM Tivoli Storage Manager storage agent	330
20.2.5 Tivoli Storage Manager security	331
20.2.6 Tivoli Storage Manager API	331
20.3 Client backup and restore operations	332
20.3.1 Traditional LAN and WAN backup topology	335
20.3.2 SAN (LAN-free) backup topology	335
20.3.3 Server-free backup	336
20.3.4 Split-mirror/point-in-time copy backup using SAN	337
20.3.5 NAS backup and restore	338
20.3.6 Image backup	339
20.3.7 Progressive incremental backups	340
20.4 IBM Tivoli Storage Manager for Data Retention	341
20.4.1 Overview	341
20.5 IBM Tivoli Storage Manager Disaster Recovery	342
20.5.1 DRM and Tivoli Storage Manager clients	344
20.6 Tivoli Storage Manager server-to-server communications	345
20.6.1 Server-to-server communication	345
20.6.2 Server-to-server virtual volumes	346
20.6.3 Using server-to-server virtual volumes for disaster recovery	348
20.6.4 Considerations for server-to-server virtual volumes	348
20.7 Tivoli Storage Manager and high availability clustering	349
20.7.1 IBM High Availability Cluster Multi-Processing (HACMP) and TSM	349
20.7.2 TSM backup/archive and HSM client support with HACMP	351
20.7.3 Tivoli Storage Manager and Microsoft Cluster Server (MSCS)	352
20.7.4 Tivoli Storage Manager backup/archive client support with MSCS	353
20.8 Tivoli Storage Manager and tape vaulting	354

20.8.1 Electronic tape vaulting	355
20.9 Remote disk mirroring and tape vaulting solutions	356
20.9.1 Collocation considerations for off-site vaulting	357
20.9.2 Reclamation considerations for off-site vaulting	357
20.10 More information	359
Part 5. z/OS storage management	361
Chapter 21. DFSMS	363
21.1 DFSMSdfp (Data Facility Product)	364
21.1.1 Overview	364
21.1.2 ISMF	365
21.1.3 System Managed Storage	367
21.1.4 Managing data with DFSMSdfp	369
21.2 DFSMSdss (Data Set Services)	371
21.2.1 Overview	371
21.2.2 Interface to other services	371
21.3 DFSMSHsm (Hierarchical Storage Manager)	372
21.3.1 Overview	372
21.3.2 Data Hierarchy	373
21.4 DFSMSRmm (Removable Media Manager)	375
21.4.1 Overview	376
21.4.2 Volume management	376
21.4.3 Validation	377
21.4.4 Library and storage location management	377
21.4.5 Policies for retention and movement	379
21.4.6 Environment	379
21.5 More information	380
Chapter 22. DFSORT	381
22.1 Overview	382
22.1.1 Other DFSORT utilities and features	382
22.2 More information	385
Chapter 23. TotalStorage Expert	387
23.1 Overview	388
23.2 Key features	390
23.2.1 Performance management for ESS and ETL	390
23.2.2 Asset management for your ESS and ETL	392
23.2.3 Capacity management for ESS and ETL	393
23.3 How IBM TotalStorage Expert works	394
23.4 Planning and implementation considerations	395
23.5 More information	397
Part 6. Storage solutions	399
Chapter 24. IBM TotalStorage DR550	401
24.1 Data retention and compliance	402
24.1.1 Characteristics of retention managed data	402
24.1.2 IBM strategy and positioning	402
24.2 Positioning of the DR550	403
24.3 DR550 overview	404
24.3.1 Hardware	404
24.3.2 Software	405

24.4	Data retention features	406
24.4.1	Methods of archive retention	406
24.5	Document Management Systems	409
24.6	More information	410
Chapter 25.	Business Continuity solutions	411
25.1	Evolving definitions of disaster and recovery	412
25.2	Types of contingency plans	413
25.2.1	Business Continuity Plan (BCP)	413
25.2.2	Business Recovery Plan (BRP)	413
25.2.3	Continuity of Operations Plan (COOP)	413
25.2.4	Incident Response Plan (IRP)	413
25.2.5	Occupant Emergency Plan (OEP)	413
25.2.6	Disaster recovery plan (DRP)	414
25.3	Disaster recovery planning phases	414
25.4	Project phases in detail	416
25.4.1	Phase 1 - Project initiation and team selection	416
25.4.2	Phase 2 - Data collection and critical needs	416
25.4.3	Phase 3 - Risk analysis	417
25.4.4	Phase 4 - Data protection	420
25.4.5	Phase 5 - Recovery Plan	421
25.4.6	Phase 7 - Maintenance and changes	424
25.5	Notification and activation procedures	424
25.5.1	Notification procedures	424
25.5.2	Activation	426
25.6	Recovery procedures	427
25.6.1	Sequence of recovery activities	427
25.6.2	Recovery procedure highlights	427
25.7	Reconstitution phase	428
25.8	Tiers of Business Continuity	428
25.8.1	Tier 0: No off-site data	429
25.8.2	Tier 1: Data backup with no Hot Site	429
25.8.3	Tier 2: Data backup with a Hot Site	429
25.8.4	Tier 3: Electronic vaulting	430
25.8.5	Tier 4: Point-in-time copies	430
25.8.6	Tier 5: Transaction integrity	430
25.8.7	Tier 6: Zero or little data loss	430
25.8.8	Tier 7: Highly automated, business integrated solution	431
25.9	Blending tiers into an optimized solution	431
25.9.1	Three Business Continuity solution segments	431
25.10	Select a solution by segment	432
25.11	Continuous Availability solutions	433
25.11.1	Geographically Dispersed Parallel Sysplex (GDPS)	433
25.11.2	Geographically Dispersed Open Clusters	437
25.11.3	TotalStorage Continuous Availability for AIX	439
25.11.4	TotalStorage Continuous Availability for Windows	441
25.12	Rapid Data Recovery	442
25.12.1	zSeries and mixed z+Open platforms	442
25.12.2	TotalStorage Rapid Data Recovery for UNIX and Windows	442
25.12.3	IBM TotalStorage SAN Volume Controller	443
25.12.4	Global Mirror Utility	443
25.13	Backup and Restore	443
25.13.1	Microsoft Volume Shadow Copy Service (VSS)	444

25.13.2 SAN File System FlashCopy	444
25.13.3 IBM Tivoli Storage Manager	444
25.13.4 IBM TotalStorage DS4000 Integrated Backup for Data	445
25.13.5 TotalStorage Data Manager for SQL 2000	445
25.14 More information	446
Chapter 26. Digital media solutions	447
26.1 The magic of digital	448
26.1.1 Digital Media Center	448
26.1.2 Digital content creation	450
26.1.3 Digital Video Surveillance and Security	450
26.2 Support	451
26.3 More information	451
Chapter 27. Solutions for the life sciences	453
27.1 Life sciences solutions and services	454
27.2 Storage for medical imaging	454
27.2.1 PACS overview	455
27.3 IBM Medical Archive Solution	456
27.4 More information	456
Part 7. Appendixes	457
Appendix A. Introduction to SMIS and CIM/WBEM	459
Appendix B. Redundant Arrays of Independent Disks (RAID)	463
Abbreviations and acronyms	469
Glossary	473
Related publications	479
IBM Redbooks	479
Other publications	480
Online resources	480
How to get IBM Redbooks	481
Help from IBM	481
Index	483

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	DS300	18
2-2	Cost-effective iSCSI SAN using DS300.	18
2-3	SAN connection options for DS400	19
2-4	DS300 (1701-1RL) storage subsystem rear view	21
2-5	DS300 (1701-1RS) storage subsystem rear view	21
2-6	DS300 storage subsystem model 1701-2RD rear view.	22
2-7	DS400 (1700-2RD) storage subsystem rear view	22
2-8	ServeRAID Manager	23
2-9	Backup with Metro Mirror and FlashCopy	25
3-1	DS4100 Storage Server.	31
3-2	DS4300 storage server	33
3-3	DS4300 turbo storage server.	35
3-4	DS4400 storage server (left) and rack-mounted storage servers (right)	37
3-5	EXP100 storage expansion unit.	41
3-6	Storage data in a networks storage environment	42
3-7	EXP700 expansion unit	43
3-8	EXP710 expansion unit	44
3-9	IBM DS4000 Storage Manager evolution	45
3-10	VolumeCopy	49
3-11	VolumeCopy integration with FlashCopy.	50
3-12	Enhanced Remote Mirroring	51
3-13	Integrated Backup for Databases	53
4-1	Enterprise Disk Systems	58
4-2	The server based architecture brings continuous performance improvements	60
4-3	Scaling options.	62
4-4	Disk enclosure switched connections	63
4-5	Four path to each disk drive.	63
4-6	Grouping of volumes in LSSs	65
4-7	The DS6000 series	66
4-8	DS6800 controller card	66
4-9	DS6000 disk architecture.	67
4-10	DS6800 controller architecture	68
4-11	SBOD controller card.	69
4-12	Example of expansion enclosure cabling.	69
4-13	DS6800 Model 1750-511 rear view	70
4-14	DS6000 with 2 expansion enclosures	71
4-15	The DS8000 series	72
4-16	Logical partitions (LPARs).	73
4-17	DS8000 - Primary Frame.	74
4-18	Array across loops	75
4-19	The DS8000 architecture.	76
4-20	RIO-G design.	77
4-21	DS8000 Frames.	78
4-22	DS8100	79

4-23	Model 922 and 9A2	80
4-24	ESS Family	82
4-25	ESS Launch Panel	84
4-26	DS6000 Storage Manager GUI	85
4-27	FlashCopy	87
4-28	Establish FlashCopy on existing Remote Mirror and Copy Primary	88
4-29	Metro Mirror	89
4-30	Global Copy	89
4-31	Global Mirror	90
4-32	z/OS Global Mirror	91
4-33	z/OS Metro/Global Mirror	92
4-34	DS Storage Manager feature activation	95
5-1	7133 Serial Advanced Models D40 and T40	99
6-1	IBM 2104 expandable storage plus 320 models DS4 and TS4	102
7-1	Ultrium 1, 2 and 3 compatibility	112
7-2	IBM LTO Ultrium family of tapes and libraries	113
7-3	The IBM TotalStorage Ultrium External Tape Drive 3580 Model L33	113
7-4	BM TotalStorage LTO Ultrium Tape Autoloader 3581	115
7-5	IBM TotalStorage LTO Ultrium Tape Library 3582	117
7-6	IBM 3583 Ultrium Scalable Library	119
7-7	IBM TotalStorage LTO Ultrium Tape Library 3584	123
8-1	IBM TotalStorage Enterprise Tape Drive 3590 in a rack	134
8-2	3592 model compared to 3590 model	136
8-3	IBM 3592 model J70 tape controller	138
8-4	3494 Tape Library	140
8-5	3494 Virtual Tape Server	144
8-6	VTS Virtual Drives and Hard disk caching	145
9-1	IBM 7204 external disk drive	152
9-2	IBM 7205 SDLT tape drive	153
9-3	IBM 7206 4mm tape drive	153
9-4	IBM 7207 SLR tape drive	154
9-5	IBM 7208 8mm tape drive	155
9-6	IBM 7210 DVD-RAM drive	155
9-7	IBM 7212-102 storage device enclosure	156
9-8	IBM 7212-312 VXA-2 autoloader	157
9-9	IBM 7332 4mm autoloader	158
10-1	A picture of SAN	162
10-2	Fibre Channel protocol layers	164
10-3	Fibre Channel point-to-point topology	164
10-4	Fibre Channel loop topology	166
10-5	Fibre Channel switched topology	166
10-6	Fibre Channel switched topology (non-cascaded switches)	167
10-7	Fibre Channel switched topology (cascaded switches)	168
10-8	LC and SC connectors	168
10-9	SFP (Small Form Factor Pluggable Media) and GBIC (Gigabit Interface Converter)	169
10-10	Host Bus Adapter	169
10-11	Extending the distance of fabric using extender	171
10-12	Extended distance with a WAN	171
10-13	IP Storage networking technology	172
10-14	iSCSI protocol layering model	173
10-15	FCIP with IP tunneling	173
10-16	FCIP Encapsulation model	174
10-17	iFCP Gateway architecture	175

10-18 iSCSI, FCIP and iFCP protocol stacks	176
10-19 A picture of NAS appliance	177
10-20 Wave® Division Multiplexer concept	178
11-1 IBM TotalStorage Switch L10	185
11-2 IBM TotalStorage SAN Switch H08	186
11-3 IBM TotalStorage SAN Switch SAN12M-1	187
11-4 IBM TotalStorage SAN Switch H16	188
11-5 IBM TotalStorage SAN Switch F32	189
11-6 IBM TotalStorage SAN32B-2 fabric switch	190
11-7 IBM TotalStorage SAN24M-1 SAN Switch	191
11-8 IBM TotalStorage SAN32M-1 SAN Switch	193
11-9 Cisco MDS 9120 and 9140 Multilayer Fabric Switch	194
11-10 Cisco MDS 9216 Multilayer Fabric Switch	195
11-11 IBM TotalStorage SAN Director M14	197
11-12 IBM TotalStorage SAN16B-R multiprotocol router	198
11-13 IBM TotalStorage SAN 140M	199
11-14 IBM TotalStorage SAN16M-R multiprotocol SAN router	200
11-15 IBM TotalStorage SAN256M	201
11-16 Cisco MDS 9506 Multilayer Director	202
11-17 Cisco MDS 9509 Multilayer Director	204
11-18 IBM TotalStorage SAN256N	205
11-19 CNT UltraNet Multi-service Director	206
11-20 Storage consolidation and high availability server clustering solution	208
11-21 Disaster tolerant metro mirror and data protection solutions	209
11-22 Storage and SAN island consolidation and high availability server clustering	210
12-1 SAN with virtualization	214
12-2 Levels of application of storage virtualization	215
12-3 In-band and out-of-band virtualization models	216
12-4 SNIA storage model	218
13-1 SVC components	225
13-2 Extents being used to create a vDisk	226
13-3 Relationship between physical and virtual disks	227
13-4 SVC logical view	227
13-5 Data migration using SAN	229
13-6 Data migration using SANVC	230
14-1 File level virtualization	234
14-2 File system types	235
14-3 SAN File System architecture	236
14-4 SAN File System administrative structure	237
14-5 GUI welcome screen	238
14-6 Global namespace	239
14-7 Policy execution for file placement	240
14-8 FlashCopy Image viewable on Windows client	241
15-1 IBM TotalStorage Productivity Center Launchpad	249
15-2 A typical IBM TotalStorage Productivity Center configuration	251
15-3 IBM TotalStorage Productivity Center architecture overview diagram	252
15-4 Architectural layers	253
16-1 Storage life cycle resource management	256
16-2 IBM TotalStorage Productivity Center for Data high level architecture	259
16-3 IBM TotalStorage Productivity Center for Data - dashboard	262
16-4 Availability Report - ping	264
16-5 Asset report of discovered disks	264
16-6 Summary view by file system, disk space used and disk space free	265

16-7	Asset reporting	266
16-8	Reports on the Web	267
16-9	Alert Log and details	268
17-1	Components of TotalStorage Productivity Center for Fabric	279
17-2	SAN management across zones	282
17-3	Root menu	284
17-4	Explorer display	284
17-5	SAN submap	285
17-6	NetView physical topology display	285
17-7	Map showing host connection lost	286
17-8	Zone view submap	286
17-9	Zone members	287
17-10	Device Centric View	287
17-11	Device Centric View - explorer display	288
17-12	Host Centric View	289
17-13	Host Centric View - logical volumes and LUN	289
17-14	Navigation tree for IBM TotalStorage Productivity Center for Fabric	290
17-15	Graph of number frames transmitted over 8 ports in a 2-minute interval	291
17-16	Number of frames transmitted over time	292
17-17	Switch events	293
17-18	EDFI configuration	293
17-19	Adornment of icons in NetView	294
17-20	Map showing effects of switch losing power	295
17-21	Vendor application launch	295
18-1	IBM Director Console	299
18-2	Performance management functions	301
18-3	Performance threshold properties	302
18-4	Gauges	303
18-5	TotalStorage Productivity Center for Disk database purge task	304
18-6	Volume Performance Advisor	305
19-1	IBM Director Console	309
19-2	TotalStorage Productivity Center for Replication	311
19-3	Relationships of groups, pools and sessions	313
19-4	RM PPRC sequence relationship example	314
20-1	Tivoli Storage Manager client, server, application, and database support	320
20-2	Data movement with Tivoli Storage Manager and the storage hierarchy	324
20-3	Integrated Solutions Console Administration Center	325
20-4	Enterprise management	326
20-5	Storage devices	327
20-6	Policy Domains and client nodes	328
20-7	Server maintenance	329
20-8	The Tivoli Storage Manager Version 5 client interface	330
20-9	How the API fits into the DR550 solution	332
20-10	Tivoli Storage Manager LAN and WAN backup	335
20-11	Tivoli Storage Manager LAN-free backup	336
20-12	Tivoli Storage Manager server-free backup	337
20-13	Tivoli Storage Manager split-mirror/point-in-time copy backup	338
20-14	IBM TotalStorage NAS backup	338
20-15	Tivoli Storage Manager and NDMP backup	339
20-16	Tivoli Storage Manager image backup	340
20-17	Tivoli Disaster Recovery Management (DRM) functions	343
20-18	Tivoli Storage Manager server-to-server communications	345
20-19	Server-to-server virtual volumes	347

20-20	HACMP and Tivoli Storage Manager server high availability configuration	350
20-21	HACMP and Tivoli Storage Manager client high availability configuration	351
20-22	MSCS and Tivoli Storage Manager server high availability configuration	352
20-23	MSCS and Tivoli Storage Manager client high availability configuration	354
20-24	Tivoli Storage Manager and electronic tape vaulting	355
21-1	DFSMS environment	365
21-2	Example of an ISMF Data Entry Panel	366
23-1	IBM TotalStorage Expert	389
23-2	IBM TotalStorage Disk Utilization Report.	391
23-3	TIBM TotalStorage Expert sample configuration	395
23-4	TotalStorage Expert with ESS or ETL on a private network	396
24-1	Representation of components of the DR550	403
24-2	DR550 with management console	405
24-3	Chronological retention policy	408
24-4	Event based retention policy	409
24-5	Deletion hold and release	409
25-1	Interrelationship of emergency preparedness plans	414
25-2	Diagram of disaster recovery planning phases	415
25-3	Disaster recovery plan as an element of risk management implementation	418
25-4	Critical risk assessment - disaster recovery plan relationship	418
25-5	Example of a business impact analysis process	419
25-6	Hot site alternative	422
25-7	Sample call tree	426
25-8	Seven tiers of Business Continuity	429
25-9	Three Business Continuity solution segments.	432
25-10	GDPS/PPRC topology	434
25-11	GDPS/XRC topology	436
25-12	GDOC functional overview.	438
26-1	Digital Media Center components	449
27-1	PACS storage solution.	455
A-1	Component Interaction in CIM Environment	461
B-1	RAID 0	464
B-2	RAID 1	465
B-3	RAID 5	466
B-4	RAID-00	467
B-5	RAID-10	468
B-6	RAID-50	468

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:

@server™	Enterprise Storage Server®	OS/390®
ibm.com®	Enterprise Systems Connection	OS/400®
iSeries™	Architecture®	Parallel Sysplex®
i5/OS™	ESCON®	Power PC®
pSeries®	FlashCopy®	PowerPC®
xSeries®	FICON®	Predictive Failure Analysis®
z/OS®	Geographically Dispersed Parallel	POWER5™
zSeries®	Sysplex™	Redbooks™
AFS®	GDPS®	Redbooks (logo)  ™
AIX 5L™	Hiperspace™	RACF®
AIX®	Hummingbird®	RMF™
AS/400®	Hypervisor™	RS/6000®
BladeCenter™	HyperSwap™	S/390®
Cloudscape™	HACMP™	ServeRAID™
CICS®	Informix®	System/390®
Domino®	IntelliStation®	Tivoli Enterprise™
DB2 Universal Database™	IBM®	Tivoli Enterprise Console®
DB2®	IMS™	Tivoli®
DFS™	Lotus Notes®	TotalStorage®
DFSMS/MVS®	Lotus®	Virtualization Engine™
DFSMSdfp™	Micro Channel®	VSE/ESA™
DFSMSdss™	MVS™	Wave®
DFSMSHsm™	Netfinity®	WebSphere®
DFSMSrmm™	NetView®	
DFSORT™	Notes®	

The following terms are trademarks of other companies:

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel Inside (logos), MMX, and Pentium are trademarks of Intel Corporation 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, and 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 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 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.

We cover these topics:

- ▶ 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.
- ▶ Part 2 is an overview of tape products, reviewing the LTO family with recently announced LTO3 support, and enterprise system like the 3494 library and Virtual Tape Server (VTS).
- ▶ Part 3 starts by describing storage networking infrastructure and protocols, then presents the switches and directors to form SAN solutions.
- ▶ Part 4 discusses the IBM TotalStorage software portfolio for open systems and includes Storage Virtualization products such as the SAN Volume Controller and SAN File System, TotalStorage Productivity Center, and Tivoli® Storage Manager.
- ▶ Part 5 describes the z/OS® storage management software: DFSMS, DFSORT™, ESS and ETL Expert.
- ▶ Part 6 presents some of the solutions and services offered by IBM, like Disaster Recovery and Business Continuity, Information Lifecycle Management, life sciences, and digital media.
- ▶ Part 7 is the appendixes, which cover standards, such as Storage Management Initiative Specification (SMIS), Common Information Model (CIM), and Web Based Enterprise Management (WBEM); also discussed are redundant array of independent disks (RAID).

This redbook is intended as a reference for basic and comprehensive information about the IBM 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, San Jose Center.

Bertrand DufRASne is a Certified Consulting I/T Specialist and Project Leader for Disk Storage Systems at the International Technical Support Organization, San Jose Center. He has worked for IBM in many I/T areas. Before joining the ITSO, he worked for IBM Global Services as an I/T Architect. He holds a degree in Electrical Engineering.

Izak De Beer is an IBM Tivoli Storage Manager Administrator from the Integrated Data Centre in South Africa. He has been with IBM for six years and has experience on many different platforms including Mainframe, AIX®, and Windows®. His responsibility involves

maintaining and implementing backup solutions at international companies. His areas of expertise include IBM Tivoli Storage Manager planning and implementation, Tivoli Data Protection products and other mainframe products such as DFSMS.

Pallavi Galgali is a Software Engineer at IBM India, Software Lab, Pune Center. She has about three years of industry experience. She has been involved in development projects with products such as SAN File System, Advanced Distributed File System and Distributed Computing Environment. She holds a degree in Computer Engineering from Pune Institute of Computer Technology, Pune (India). Her areas of expertise include storage networking, file systems and device drivers.

Chris Hoffmann is a Technical Consultant for Tardis Services, an IBM Business Partner in Sydney, Australia. He is an IBM Certified Specialist - Open Systems Storage Solutions and TotalStorage Networking and Virtualization Architecture and also holds an MCSE and CNE. Chris has over 15 years of experience in the I/T industry and his areas of expertise include IBM TotalStorage disk and tape products, Storage Area Networks, and Microsoft® products. Chris has a Ph.D in Chemistry from the University of Bristol, England.

Szilvia Racz has worked at IBM for five years, providing I/T support for storage manufacturing in Hungary. She currently works for IBM Data Storage Systems Ltd. in Vac, Hungary, where the ESS and DS8000 storage systems are built. She holds AIX Advanced Technical Expert, TSM and Cisco CCNA certifications. Prior to joining IBM she worked as an AIX system administrator.

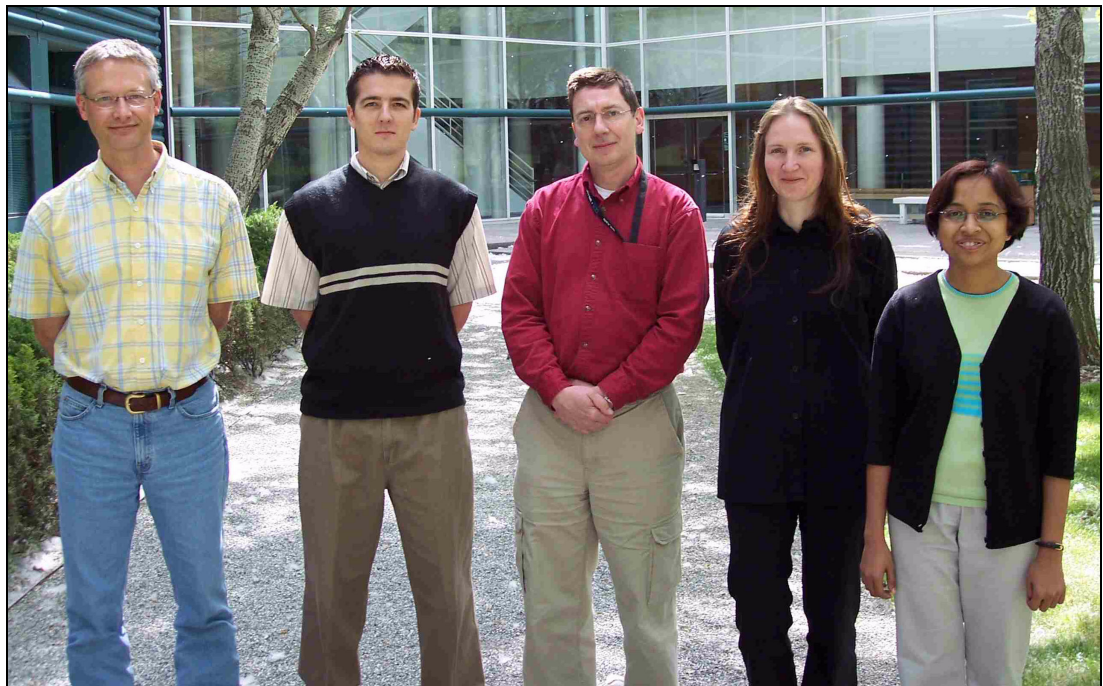


Figure 1 Chris, Izak, Bertrand, Szilvia, Pallavi



Stanley Smith is a Competitive Sales Specialist responsible for storage network and management competition at the IBM EMEA Compete Center in Copenhagen, Denmark. He has several years of experience working with storage solutions, computer networks, and telecommunications.

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

Cathy Warrick
Charlotte Brooks
Roland Tretau
Jon Tate
Mary Lovelace
Emma Jacobs
Gabrielle Velez
International Technical Support Organization

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'll team with IBM technical professionals, Business Partners and/or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll 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. QXXE Building 80-E2
650 Harry Road
San Jose, California 95120-6099



Introduction: Information on demand with IBM TotalStorage

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's 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 TotalStorage 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 security and durability of information.
3. Information Lifecycle Management (ILM): Efficiently managing information over its lifecycle.

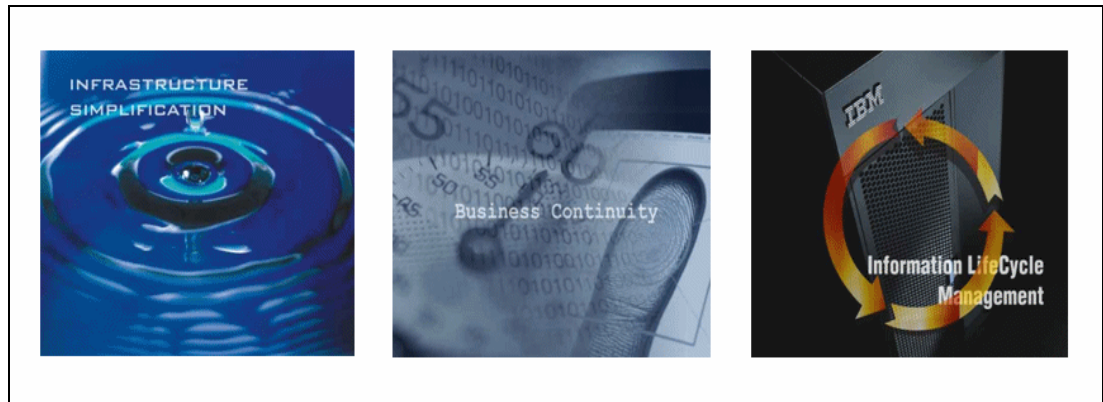


Figure 1-1 IS, BC, and ILM

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 TotalStorage 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 TotalStorage 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's 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 using our advanced automated management software.

Note: For more information about Infrastructure Simplification, refer to:

<http://www-1.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 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.

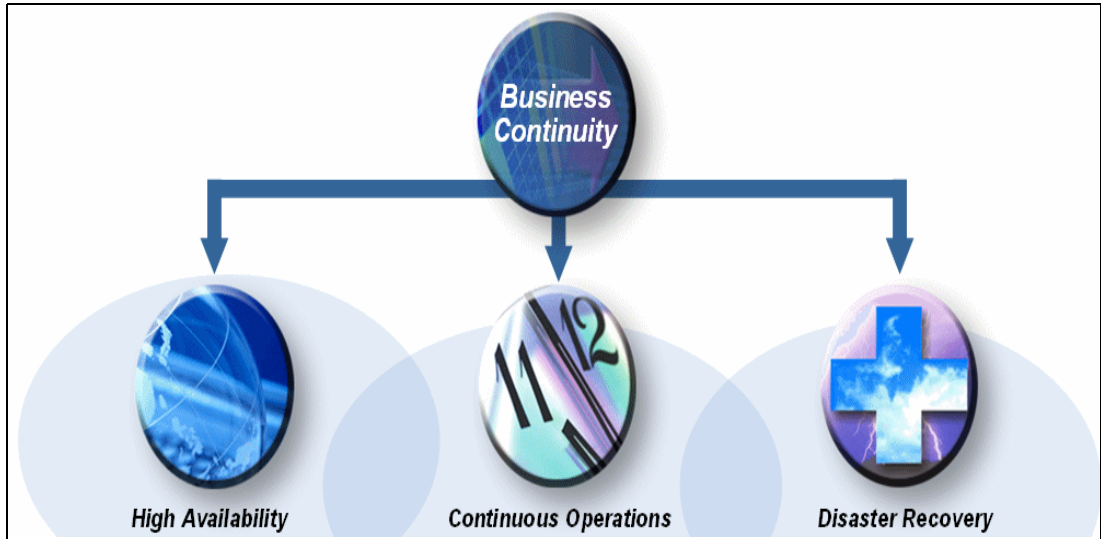


Figure 1-2 Business Continuity

Note: For more information about Business Continuity, refer to:

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.

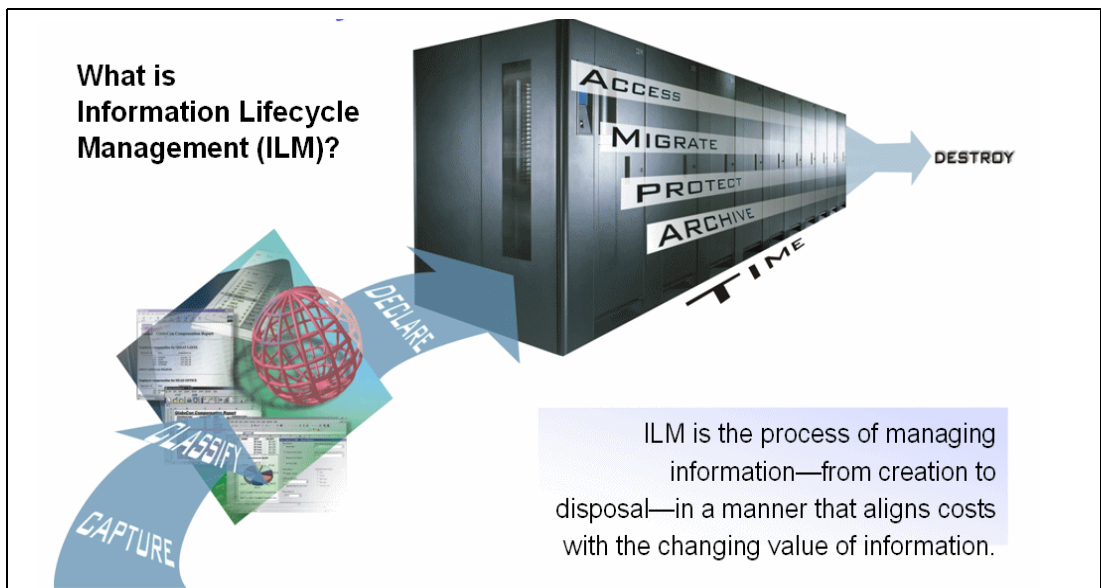


Figure 1-3 Information Lifecycle Management

IBM's unique experience with the broad range of ILM technologies, and its broad portfolio of offerings and solutions, including offerings in TotalStorage hardware, TotalStorage 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, refer to:

<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 TotalStorage 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 TotalStorage DS family combines the high-performance of the IBM TotalStorage DS6000 and DS8000 series enterprise servers with IBM TotalStorage DS4000 series of mid-range systems (formerly called the FASiT family), with low priced entry systems. This family is complemented by a full range of IBM TotalStorage capabilities like advanced copy services, management tools and virtualization services to help protect your data.

Note: For a summary of IBM's offerings in the area of disk systems, refer to Part 1, "Disk storage solutions" on page 9.

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

Note: For a summary of IBM's offerings in the area of tape systems, refer to Part 2, "Tape products" on page 105.

IBM TotalStorage 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, CNT and McDATA switches and directors. Section 11.5, "SAN solutions" on page 206 gives more information about IBM's SAN solutions.

Note: For a summary of IBM's offerings in the area of SAN solutions, refer to Part 3, "Storage networking" on page 159.

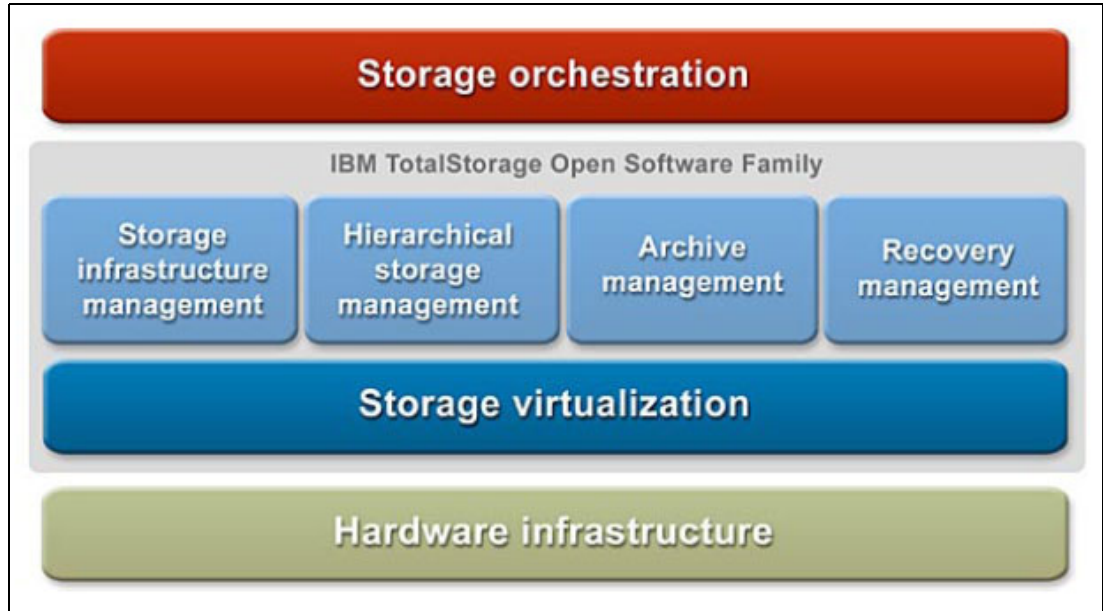


Figure 1-4 On demand storage environment

IBM TotalStorage open software family

IBM TotalStorage 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 as shown in Figure 1-4.

Storage virtualization

Storage virtualization products are designed to improve the flexibility and utilization of your storage. Virtualization works by pooling your storage volumes, files and file systems 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 is also designed to reduce the effects of reconfigurations at the hardware level, to help support Business Continuity.

The Virtualization family includes the IBM TotalStorage SAN Volume Controller to achieve virtualization at the block level and IBM TotalStorage SAN File System to achieve virtualization at the file level.

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

<http://www-1.ibm.com/servers/storage/software/virtualization/sfs/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 Open Software Family HSM 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 Tivoli Storage Manager for Space Management is a significant product for HSM.

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

Archive management

Archive management gives complete solutions designed to help enterprises archive, retain, and manage data, including email, 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 once read many (WORM) tape and magnetic disks.

The IBM TotalStorage DR 550 is designed as a preconfigured, integrated solution to help store, retrieve, manage, share and secure regulated and non-regulated 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-306.ibm.com/software/tivoli/products/storage-mgr/>

Note: To know more about the products under IBM TotalStorage open software family mentioned above, refer to the Part 4, “IBM TotalStorage open software” on page 211.

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 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-306.ibm.com/software/tivoli/products/prov-mgr/>

1.3 More information

- ▶ For the latest information about IBM's TotalStorage offerings, refer to:
<http://www-1.ibm.com/servers/storage/index.html>
- ▶ The following link gives an A-Z listing of all IBM TotalStorage products:
<http://www-1.ibm.com/servers/storage/product/index.html>
- ▶ For more information about virtualization and IBM TotalStorage virtualization offerings, refer to:
<http://www-1.ibm.com/servers/storage/software/virtualization/>



Part 1

Disk storage solutions

Part 1 covers disk products that either can be directly attached to a server (Direct Attached Storage or DAS), or used in a Storage Area Network (SAN) infrastructure, which is generally the case in today's enterprise storage implementations.

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

To support the increasing demands for disk storage, IBM has brought together into one family a broad range of disk systems to help small, medium and large-size enterprises select the right solutions for their needs.

The IBM TotalStorage DS family combines the high-performance of the IBM TotalStorage DS6000 and DS8000 series enterprise servers with the IBM TotalStorage DS4000 series of mid-range systems (formerly called the FASiT family), 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 TotalStorage capabilities such as advanced copy services, management tools, and virtualization services are available to help protect data.

Figure 5 shows the extensive range of IBM disk storage systems that are discussed in more detail in subsequent chapters.

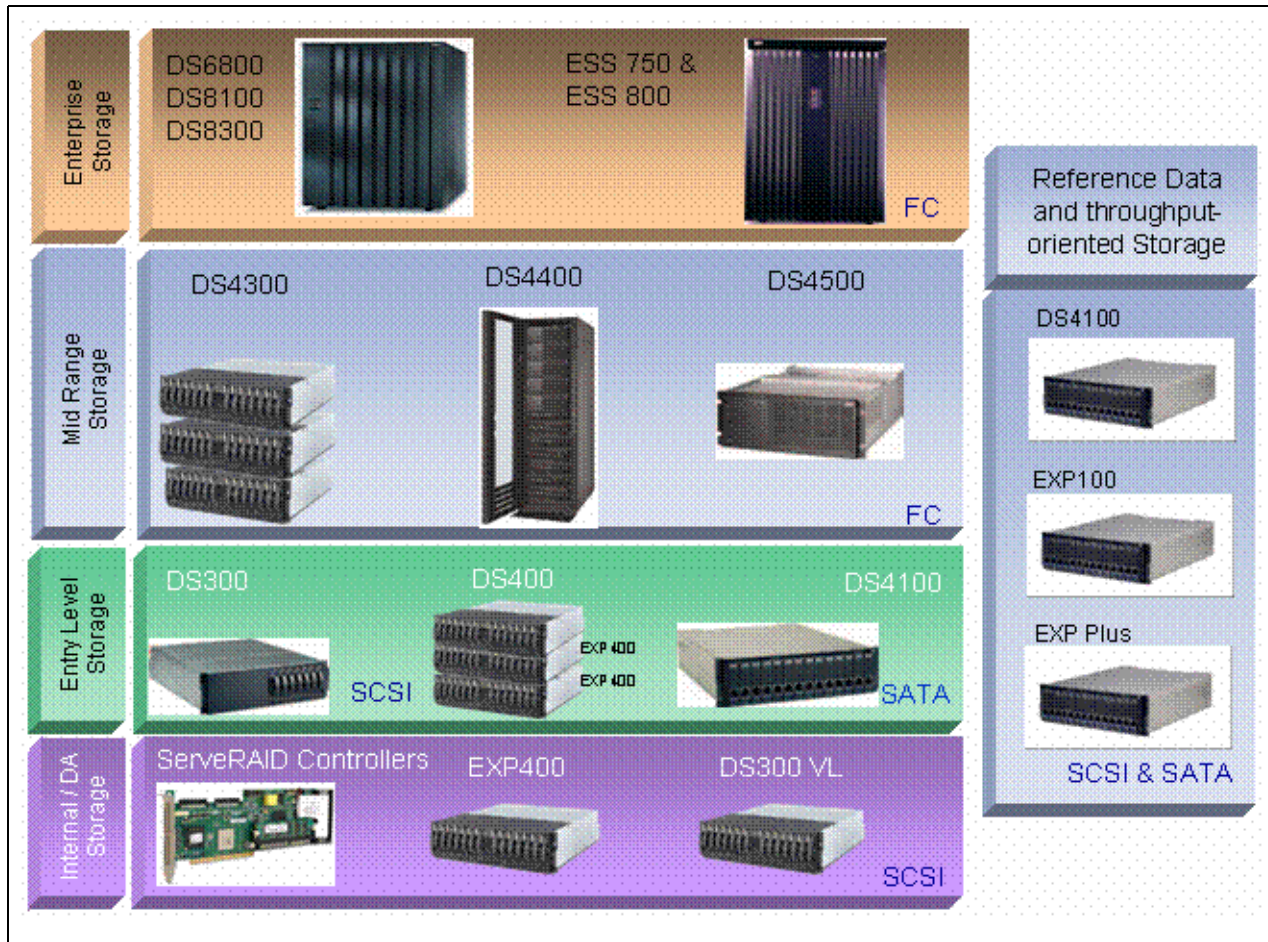


Figure 5 IBM Disk Storage systems

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
- ▶ Designed to offer high performance, highly scalable, very high function storage sub systems
- ▶ Designed to avoid any single points of failure for high availability. Highly reliable architecture with hot swap redundant components.
- ▶ Designed to be nondisruptive for 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 TotalStorage mid-level storage disks systems are characterized as:

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

In particular the DS4000 series is:

- ▶ Designed to offer high reliability with hot swap redundant power supplies, controllers and drives
- ▶ Designed to offer high performance dual ported 2 Gb FC drives
- ▶ Easy to use, common DS Storage Manager

Near-line storage and reference data

For near-line storage and reference data, IBM TotalStorage 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 TotalStorage 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

IBM TotalStorage point-in-time copy comparison

The following table provides a quick reference to the point-in-time copy (FlashCopy®) functions available on the IBM TotalStorage 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	SAN File System	IBM DS6000, DS8000, ESS	IBM DS4000
File level	N	Y	Y- z/OS only	N
Block level	Y	N	Y	Y
Physical Copy	Y	N	Y	Y
Logical Copy (no copy)	Y	Y	Y	Y
Transition NoCopy -> Copy	Y	N	Y	N
Logical Copy (space efficient)	N	Y	N	Y
Persistent Copies	Y	N	Y	N
Multiple Concurrent Copies	N	N	Y (12)	Y (4)
Source / Target Read / Write Capable	Y	N	Y	Y
Incremental Copies	N	N	Y	N
Consistency Groups	Y	Y	Y	N
Across heterogeneous disk vendors	Y	Y	N	N

Figure 6 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, for example, ESS/DS6000/DS8000 NOCOPY, or SVC 0 background copy rate parameter?

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

Figure 7 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 – 10Km Fibre Channel - channel extender IP/DWDM based solutions for longer distances.
- ▶ ESS / DS6000 / DS8000 – 103Km over ESCON®, 300 KM over Fibre Channel or IP/DWDM based network; unlimited distance supported for asynchronous mirroring.
- ▶ DS4000 – 10Km 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, does I/O stop (freeze) to the consistency group, and an error posted to the host for automation to act upon and then host I/O is resumed (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 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 @server xSeries® 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 and Metro Mirroring.

2.1 Overview

The IBM TotalStorage DS300 and DS400 are entry-level, low cost workgroup storage subsystems for xSeries 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 2-1 DS300

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

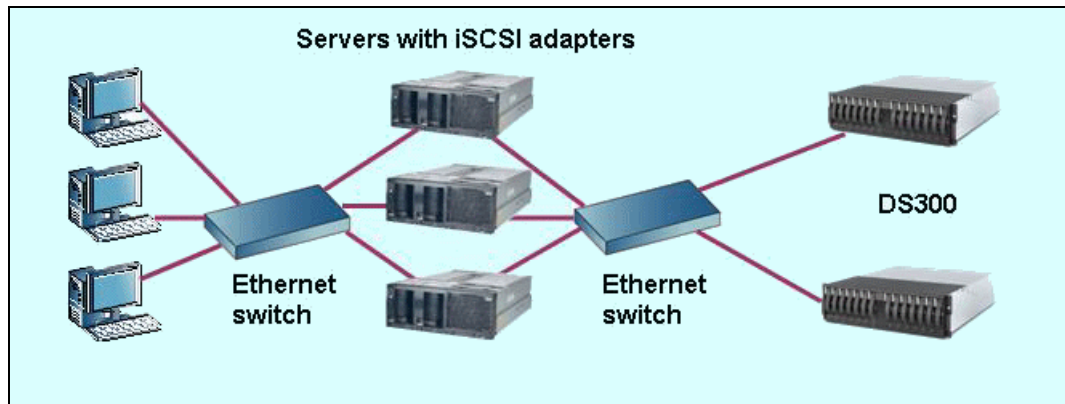


Figure 2-2 Cost-effective iSCSI SAN using DS300

As an alternative to the iSCSI adapters in the servers is to use iSCSI software initiators instead. These have the disadvantage of less security and increased CPU utilization. Suggested software initiators include:

- ▶ Microsoft iSCSI software initiator
- ▶ Cisco iSCSI initiator for Linux® and Windows
- ▶ Novell iSCSI initiator for NetWare

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 2-3.

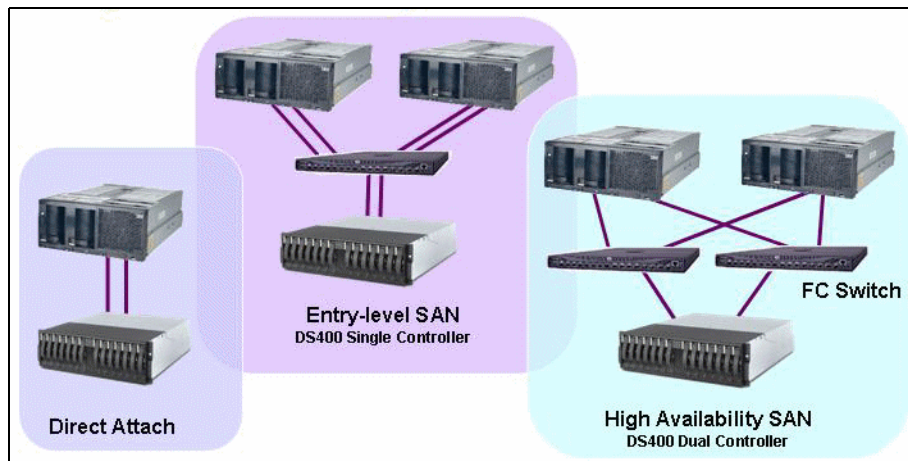


Figure 2-3 SAN connection options for DS400

The modular design of the DS300 and DS400 allows you to easily perform upgrades and additions to help meet the growing needs of your application environment—scaling to support increasing volumes of data flexibly and affordably. In addition, advanced software features on the DS300 and DS400—including access control lists and online array expansion—are designed to give administrators the power to configure storage according to changing usage needs. These features also help enable you to share storage across multiple application servers.

Both units are designed to deliver advanced functionality for business continuance and disaster recovery at an entry-level price. 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 for management and redundant data paths so that—even if a line or a controller fails—multiple paths between servers and storage can help maintain a connection.

Metro mirroring on the DS300/400 is designed to enable disaster recovery implementations to help protect critical data. 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 xSeries 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. Configuration wizards simplify initial storage setup. IBM ServeRAID Manager is designed to support most common operating environments for workgroup applications, including Microsoft Windows, Novell NetWare and Linux platforms.

The management software of the DS300/400 is built directly into the client/server workgroup manager IBM Director. IBM Director is designed to allow IT administrators to view and track the hardware configuration of remote systems in detail and monitor the usage and performance of critical components, such as processors, disks and memory.

2.2 Specifications

Table 2-1 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, NetWare	Windows 2000, Windows 2003, Linux, NetWare, VMware
Host Connectivity	iSCSI	Fibre Channel
SAN Support	Direct, Switched Ethernet	Direct, FC-AL, Switched Fabric
Copy Services	FlashCopy, Metro Mirror	FlashCopy, Metro Mirror
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)	None (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, 2 TB	36 GB, 5.8 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	36 GB, 73 GB, 146 GB, 300 GB 10,000 RPM Disk Drives; 36 GB, 73 GB, 146 GB 15,000 RPM Disk Drives
Certifications	Microsoft Windows MSCS	Microsoft Windows MSCS

2.3 IBM TotalStorage DS300

The IBM TotalStorage DS300 comes in a variety of models.

- The DS300 storage subsystem 1701-1RL comprises 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).

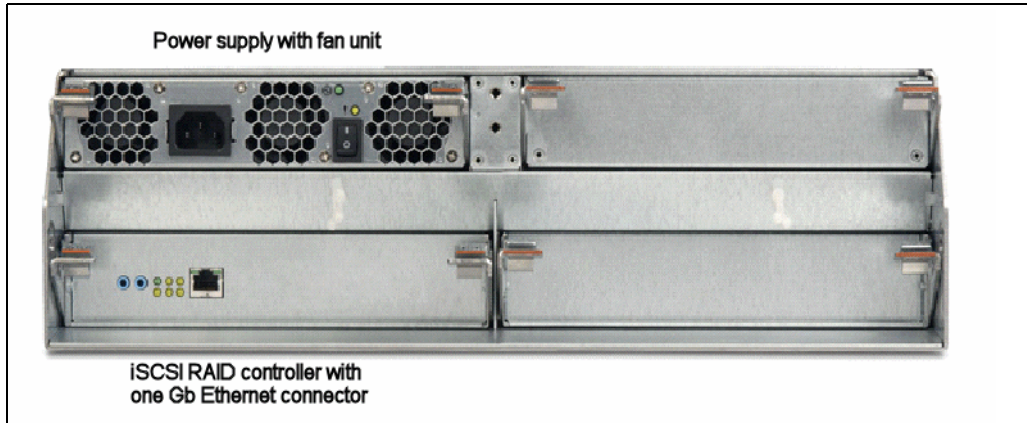


Figure 2-4 DS300 (1701-1RL) storage subsystem rear view

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

- One iSCSI RAID controller. 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).

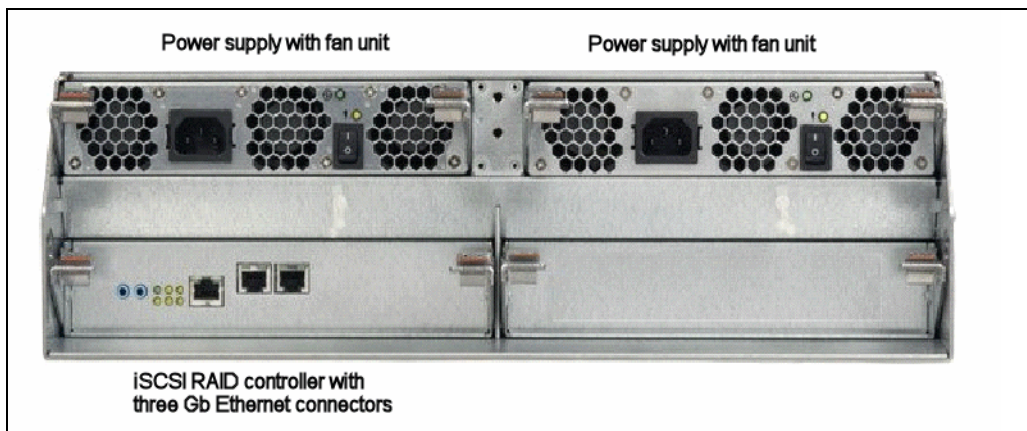


Figure 2-5 DS300 (1701-1RS) storage subsystem rear view

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

- Two iSCSI RAID controller. 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).

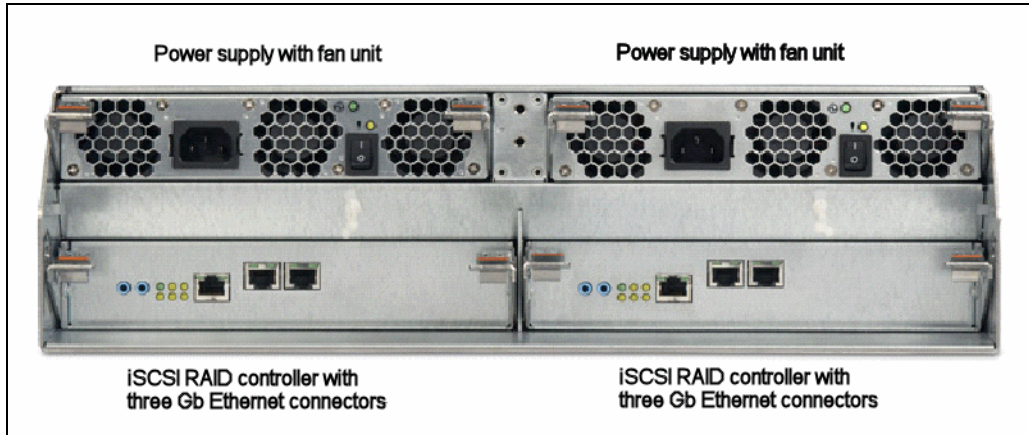


Figure 2-6 DS300 storage subsystem model 1701-2RD rear view

2.4 IBM TotalStorage DS400

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 (gigabit per second) Ethernet connector for management,
- ▶ Two 2 Gbps Fibre Channel host ports
- ▶ One Fibre Channel RAID controller
- ▶ Two power-supply-with-fan units
- ▶ 14 hard disk drive filler panels

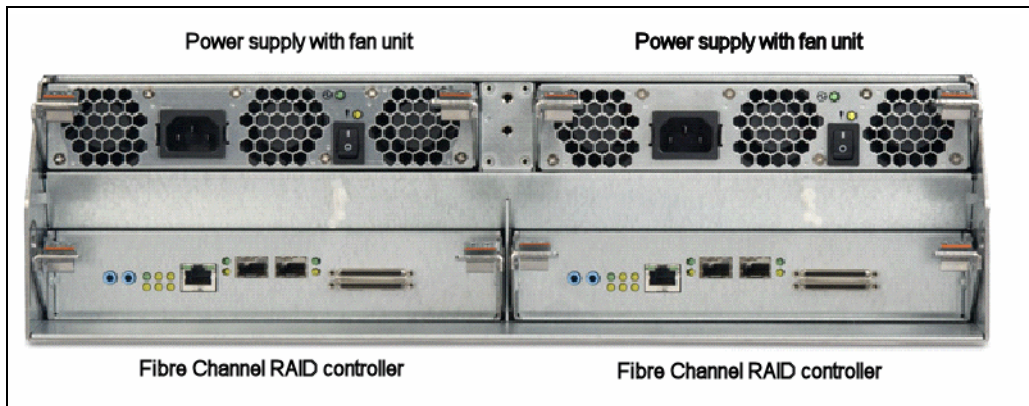


Figure 2-7 DS400 (1700-2RD) storage subsystem rear view

2.5 ServeRAID Manager

The ServeRAID Manager is used to configure the storage subsystem and to monitor its status. The steps to follow to configure the DS300/400 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

When you log in for the first time, you must add the management station before you can begin to set up your network storage. The management station agent is a monitoring agent for network storage. After you add an agent to a management station, you can monitor and configure the attached storage system enclosures from a ServeRAID Manager console.

Figure 2-8 shows how IBM ServeRAID Manager displays the Physical and Logical device views, which show the physical devices and logical devices connected to the controller or enclosure.

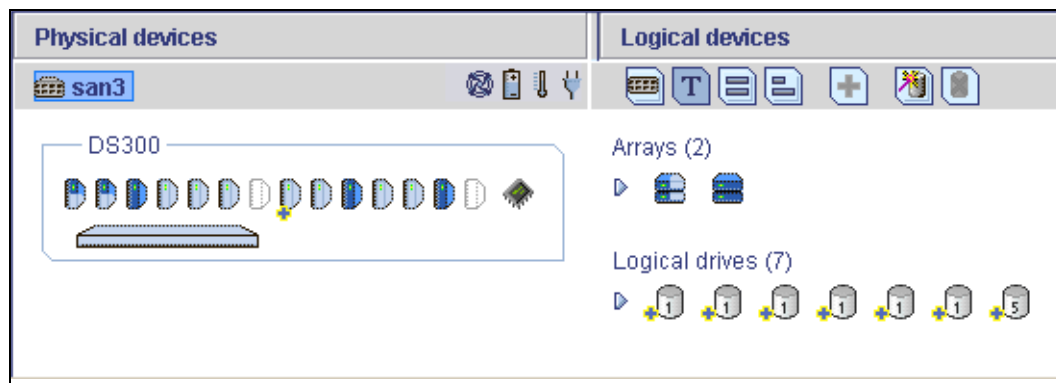


Figure 2-8 ServeRAID Manager

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 DS400 Problem Determination Guide provides examples of typical configuration and problem determination flowcharts and is available at:

ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers/ds400pd.zip

2.6 Premium features

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

2.6.1 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 4 TB of additional storage.

The EXP400 delivers fast, high-volume data transfer, retrieval, and storage functions across multiple drives, to multiple hosts. The expansion unit is designed for continuous, reliable service; the modular, redundant disk drives, power supply with fan units, and SCSI Bus Expander and Enclosure Services Module (ESM) use 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 expansion units:

- ▶ *Hardware Maintenance Manual and Troubleshooting Guide*
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/90p3548.pdf)
- ▶ *Installation Guide*
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/88p9249.pdf)
- ▶ *User's Guide* (ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/88p9250.pdf)

2.6.2 FlashCopy

FlashCopy provides the ability to make an instantaneous point-in-time copy of a volume. For the DS300/DS400 it is 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 drives. You can only use the FlashCopy Management Command-Line Tool to create and manage FlashCopies, although 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 application development or short term backup and recovery scenarios.

2.6.3 Metro Mirror

Metro Mirror is an option, available for the IBM TotalStorage DS300 and DS400 single controller models, that provides a license activation key enabling you to maintain a remote mirror copy of a logical volume on another DS300 or DS400 Storage Server at a distant site. Mirroring between dissimilar models is not supported. This feature supports a synchronous copy of your primary data across an extended distance to a secondary site using an Ethernet Fabric. The Metro Mirror feature can help protect your critical data from catastrophic loss.

Note: At the time of writing Metro Mirror had a planned availability date of late 2005.

A combination of FlashCopy and Metro Mirror can be used for improved backup scenarios as shown in Figure 2-9. A FlashCopy volume on the mirrored storage (e.g. FlashCopy A) can be mounted and backed up to tape without affecting the performance of the primary storage (e.g. Volume A).

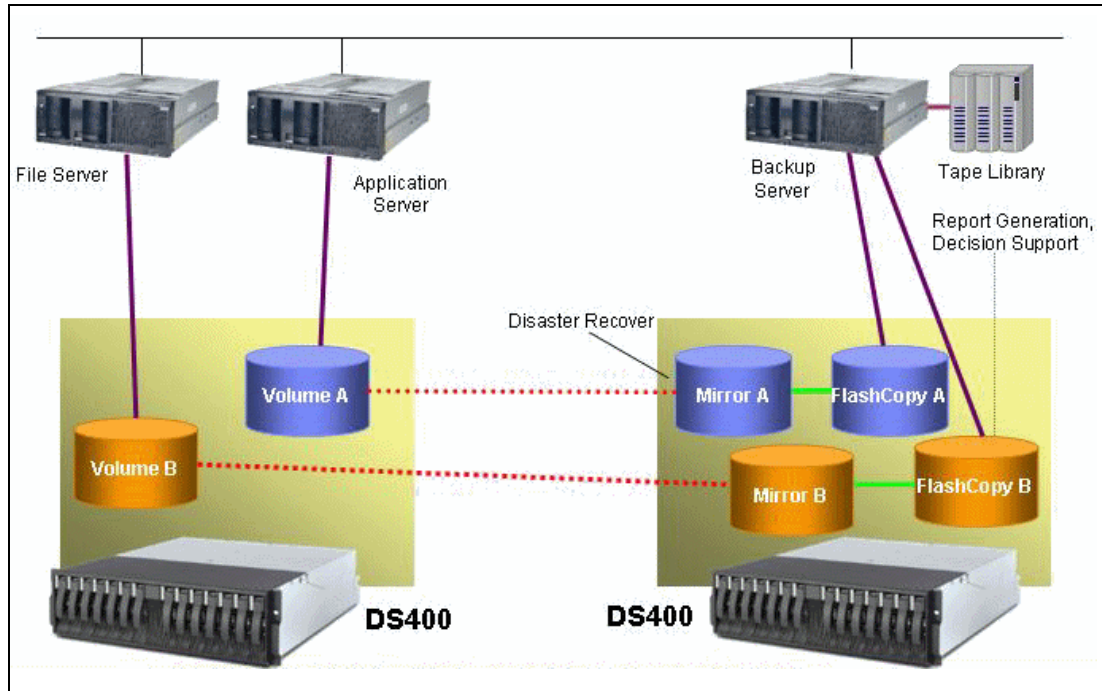


Figure 2-9 Backup with Metro Mirror and FlashCopy

2.7 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

Note: At the time of writing Online Capacity Expansion and RAID level migration had a planned availability date of mid 2005.

Boot from SAN

DS400 and DS300 allow you to boot from the SAN rather than from local disks on individual servers, that can enable you to maximize consolidation of IT resources and minimize 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 you a rapid Disaster Recovery as all the boot information and production data is stored on a local or remote SAN environment.

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.

You can refer to the Microsoft Web site for more information and implementation on booting from SAN:

<http://www.microsoft.com/windowserversystem/storage/solutions/sanintegration/bootfromsaninwindows.mspx>

2.8 More information

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

- ▶ <http://www-1.ibm.com/servers/storage/disk/ds/ds300/index.html>
- ▶ <http://www-1.ibm.com/servers/storage/disk/ds/ds400/index.html>
- ▶ Software Installation Guide
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/25k8191.pdf)
- ▶ Hardware Maintenance Manual
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/25k8190.pdf)
- ▶ Dual Controller Hardware Installation and User's Guide
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/25k8196r.pdf)
- ▶ Dual Controller Software Installation Guide
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/31r1057r.pdf)
- ▶ TotalStorage DS300 - Problem Determination Guide
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers/ds300pd.zip)
- ▶ TotalStorage DS400 - Problem Determination Guide
(ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers/ds400pd.zip)



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 Storage Servers have hot swap and redundant power supplies and fans. With the exception of the base entry model, all DS4000 Storage Servers 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 of storage servers:

- ▶ IBM TotalStorage DS4100
- ▶ IBM TotalStorage DS4300 and DS4300 turbo
- ▶ IBM TotalStorage DS4400
- ▶ IBM TotalStorage DS4500
- ▶ IBM TotalStorage DS4000 EXP700/EXP710 Expansion Units
- ▶ IBM TotalStorage DS4000 EXP100

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

3.1 Highlights: What's new

In this section we summarize the new functions, supported platforms, and features of the DS4000 family of disk storage products, since the last edition of this book in November 2003.

- ▶ Larger capacity FC hard disk drives

The IBM TotalStorage DS4000 series now offers two new optional 2 Gbps high-performance, hot-swappable, Fibre Channel disk drive modules. The new 300 GB disk drive operates at up to 10,000 rpm, while the new 146 GB drive module operates at up to 15,000 rpm. The 300 GB drive doubles the current maximum storage capacity of the DS4000 series and enables configurations greater than 67 terabytes (TB) of physical capacity. The new 300 GB and 146 GB disk drive modules are available for the TotalStorage DS4300 and the EXP700 and EXP710 Fibre Channel Storage Expansion Units.
- ▶ EXP710 expansion unit
 - The EXP710 is a rack-mountable storage expansion enclosure that is an enhancement to the EXP700. The enhancement is that the EXP710 enclosure is a *switched bunch of disks* (SBOD) version of the EXP700.
- ▶ Rebranding of names from FAStT to DS4000 family
 - The DS4000 Series is a rebranding of the former IBM TotalStorage FAStT Storage Servers and also includes newer models.
- ▶ Enhanced Remote Mirroring
 - Metro Mirroring
 - Global Copy
 - Global Mirroring
- ▶ DS4100 (FAStT 100) model introduced
- ▶ Intermix of SATA and FC expansion units are supported
- ▶ Integrated database backup introduced
- ▶ IBM TotalStorage DS4000 Alert Manager
- ▶ Storage Manager updated to Version 9.12 which provides the following enhancements:
 - Delta Logging: allows the primary storage subsystem to track portions of the primary logical drive that have been changed during a communication interruption between the DS4000 subsystems.
 - Suspend and Resume: based on delta logging a user can manually suspend mirror synchronization to the secondary mirror.
 - Asynchronous Write Mode: allows the primary site controller to acknowledge host initiated write requests before data has been successfully mirrored to the secondary DS4000.
 - Write order consistency for Asynchronous Mirrors: allows the user to specify that write requests to the remote Storage Subsystem are in the same order as completed on the local Storage Subsystem.
 - Allow read access to mirror secondary logical drives
 - Support for Microsoft VDS/VSS
 - New RDAC driver architecture
 - Full Command Line Interface
 - Controller service mode
 - Support for Veritas ASL 4.0

3.2 Overview

Storage Servers within the IBM TotalStorage DS4000 series are storage subsystems 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 IBM TotalStorage DS4000 Series of disk storage systems are IBM's solution for mid-range/departmental storage requirements. The storage servers are:

- ▶ IBM TotalStorage DS4100 Storage Server

The DS4100 Storage Server is an entry-level SATA storage system that is available in a single and dual controller configuration. With the DS4100 single controller you get entry-level storage supporting up to 3.5TB. The dual controller model supports up to 28TB with the DS4000 EXP100 expansion. The single controller model has 2 Fibre Channel host connections and the dual controller provides 4 FC ports.

- ▶ IBM TotalStorage DS4300 and DS4300 Turbo Storage Servers

The DS4300 is a mid-level, highly scalable 2 Gbps Fibre Channel Storage Server which is available in a single and dual controller configuration. It is designed to be a cost-effective, scalable storage server for consolidation and clustering applications.

The DS4300 Turbo is a mid-level storage server that can scale to over 32TB, 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.

- ▶ IBM TotalStorage DS4400 Storage Server

This system offers 2 Gbps technology and scales from 36 GB to greater than 64 TB of FC storage using 16 EXP700/710 expansion enclosures.

- ▶ IBM TotalStorage DS4500 Storage Server

IBM DS4500 Storage Server delivers breakthrough disk performance and outstanding reliability for demanding applications in compute intensive environments. The DS4500 is designed to offer investment protection with advanced functions and flexible features. Designed for today's on demand business needs, the DS4500 easily scales from 36 GB to over 64 TB to support growing storage requirements. The DS4500 offers advanced replication services to support business continuance and disaster recovery.

The DS4000 Series is a rebranding of the former IBM TotalStorage FAStT Storage Servers and also includes newer models. The mapping between former DS4000 models and rebranded DS4000 models is detailed in Table 3-1. A recent addition is the DS4000 EXP710 enclosure which is a *switched bunch of disks* (SBOD) version of the EXP700.

Table 3-1 DS4000 product renaming

Original Name	New Name
IBM TotalStorage FAStT Storage Server	IBM TotalStorage DS4000
FAStT	DS4000
FAStT family	DS4000 Series
FAStT900	DS4500
FAStT700	DS4400
FAStT600 turbo	DS4300 turbo
FAStT600	DS4300

Original Name	New Name
FAST100	DS4100
EXP700	DS4000 EXP700
N/A	DS4000 EXP710
EXP100	DS4000 EXP100

Along with the product rebranding, some of the features available with the DS4000 series products have been renamed as shown in Table 3-2. The renaming also includes a couple of new mirroring features which are described in “Copy services” on page 47.

Table 3-2 DS4000 services renaming

Original Name	New Name
FAST Storage Manager	DS4000 Storage Manager
FAST Service Alert	DS4000 Service Alert
FAST FlashCopy	FlashCopy for DS4000
FAST VolumeCopy	VolumeCopy for DS4000
FAST Remote Volume Mirror (RVM) ▶ FAST Synchronous Mirroring ▶ N/A ▶ N/A	Enhanced Remote Mirroring for DS4000 ▶ Metro Mirroring for DS4000 ▶ Global Copy for DS4000 ▶ Global Mirroring for DS4000

Note: Visit the product interoperability matrix for the latest information at:

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

3.3 DS4100 Storage Server

The DS4100 Storage Server, shown in Figure 3-1, is a 3U rack-mountable Serial ATA (SATA) RAID controller and disk drive enclosure. It is designed as 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 28TB.

3.3.1 Features

There are two models available:

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



Figure 3-1 DS4100 Storage Server

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 eight IBM TotalStorage DS4000 EXP100 Expansion Units (a license key for 4th - 7th enclosure is required); this allows for up to 112 disk drives, offering a potential capacity of 28TB.

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.

The DS4100 Single Controller presents the following differences (or restrictions) with 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 three 250 GB SATA disk modules and supports 11 additional internal SATA disk modules.

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

3.3.2 Specifications

Table 3-3 IBM DS4100 Storage Server

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 7200 rpm
RAID	levels 0, 1, 3, 5, 10
Storage partitions	Standard Host Group, upgradeable to 4, 8 or 16 partitions
Maximum drives supported	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

Table 3-4 Physical characteristics

Dimensions	132.3 mm H x 482 mm W x 597 mm D (5.2 in x 19 in x 24 in)
Weight	30 kg (66 lbs)

For the latest specification information, check:

<http://www-1.ibm.com/servers/storage/disk/ds4000/ds4100/index.html>

3.4 DS4300 Storage Server

The DS4300 Storage Server is designed to be an affordable, scalable storage server 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.

This server is designed to help consolidate direct-attached storage into a centrally managed, shared or Storage Area Network (SAN). The IBM TotalStorage DS4300 has been issued a

Certificate of Conformance by National Technical Systems, Inc., for conformance to Network Equipment-Building System requirements for Level 3 type 2 and 4 equipment (NEBS 3).

3.4.1 Features

The DS4300 Storage Server is a 3 U rack-mountable Fibre Channel RAID storage system and disk drive enclosure housing up to 14 FC drives with capacities from 36.4 GB to 300 GB.

More than three times the I/O performance and sustained throughput of the FASTt200, the previous entry-level offering in the DS4000 family of products. Average customer configuration priced below competitive solutions together with longer and more extensive warranties.

The dual active RAID controllers have 512 MB, battery backed up cache in the base model and 2 GB battery backed up cache in the turbo model. It uses the latest in storage networking technology to provide an end-to-end 2 Gbps Fibre Channel solution. The DS4300 is shown in Figure 3-2 on page 33.



Figure 3-2 DS4300 storage server

- ▶ There are four host ports, two for each controller; this enables a cluster solution without the use of a switch.
- ▶ 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.
- ▶ Supports three IBM TotalStorage DS4000 EXP700/710 Expansion Units (a license key is required); this allows for up to 56 disk drives, offering a potential capacity of 16.8TB Fibre Channel disk.

Note: DS4300 only supports the EXP700. The DS4300 does not support connection to the EXP500 as the DS4300 is 2 Gb system only, and does not support 1 Gb devices. (DS4300 **turbo** auto senses to connect to 1 Gb or 2 Gb.)

- ▶ Supports eight IBM TotalStorage DS4000 EXP100 Expansion Units, this allows for up to 112 disk drives, offering a potential capacity of 28 TB of Serial ATA disk storage. Requires DS4000 Storage Manager 8.41 or later.
- ▶ 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.

3.4.2 Specifications

Table 3-5 IBM DS4300 Storage Server

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 (single controller) or 56 FC drives (using three EXP700/710 Expansion Units) 112 SATA drives (using 8 EXP100 Expansion Units)
Fans and power supplies	Single or Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack

Table 3-6 Physical characteristics

Dimensions	132.3 mm H x 482 mm W x 597 mm D (5.2 in x 19 in x 24 in)
Weight	30 kg (66 lbs)

For the latest specification information, check:

<http://www-1.ibm.com/servers/storage/disk/ds4000/ds4300/index.html>

3.5 DS4300 Turbo Storage Server

To meet growing needs for storage capacity and performance, the DS4300 Storage Server may be upgraded with the 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.

The DS4300 turbo is well suited for cost conscious workgroup SANs as well as small businesses and data centers. Designed specially for open systems environments, this storage system provides heterogeneous platform support by including failover drivers for the Microsoft Windows 2000 operating system and optional support for AIX, Sun Solaris, HP-UX, Novell NetWare and Linux operating systems. The DS4300 turbo can operate with IBM *@server* pSeries® and xSeries servers as well as with Intel processor-based and UNIX based servers.

As part of the IBM DS4000 family, the Model 4300 with turbo uses the same common storage management software and high performance hardware design, providing customers with enterprise like capabilities found in high end models, but at a much lower cost.

3.5.1 Features

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



Figure 3-3 DS4300 turbo storage server

- ▶ Using Storage Manager 9.12, the DS4300 with turbo option performs at more than 75,000 cached I/Os per second a 70 percent improvement and provides up to 768 MB/s throughput.
- ▶ The DS4000 Storage Manager 9.12 enables up to 256 logical volumes (LUNs) per storage partition, definition of array groups greater than 2 TB, and SCSI-3 Persistent Reservation.
- ▶ Higher scalability over base DS4300, scalable to 33TB of Fibre Channel disk by attaching up to seven EXP700 or EXP710 units.
- ▶ Supports eight IBM TotalStorage DS4000 EXP100 Expansion Units, this allows for up to 112 disk drives, offering a potential capacity of 28 TB of Serial ATA disk storage. Requires DS4000 Storage Manager 8.41 or later.
- ▶ Increased cache from 256 MB per controller on base DS4300 to 1 GB per controller on turbo.
- ▶ 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.
- ▶ Host interface on base DS4300 is 2 Gbps, turbo auto senses to connect to 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.
- ▶ Designed to deliver high performance of up to 768 MB/s throughput with turbo option.
- ▶ Up to 64 storage partitions.
- ▶ FlashCopy with VolumeCopy, a function for a complete logical volume copy within the DS4000 Storage Server is available on the DS4300 turbo, DS4400, and DS4500.

Because the DS4300 turbo can scale up to 112 disk drives and 33.6 TB of capacity, 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.

3.5.2 Specifications

Table 3-7 IBM DS4300 turbo Storage Server

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

Table 3-8 Physical characteristics

Dimensions	132.3 mm H x 482 mm W x 597 mm D (5.2 in x 19 in x 24 in)
Weight	30 kg (66 lbs)

For the latest specification information, check:

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

3.6 DS4400 Storage Server

IBM DS4400 Storage Server delivers superior performance with 2 Gbps Fibre Channel technology. The DS4400 is an effective storage server for any enterprise seeking performance without borders. The DS4400 is designed to offer investment protection with advanced functions and flexible features. DS4400 offers advanced replication services to support business continuance. FlashCopy and Remote Mirroring help increase data protection options. Dynamic Volume Expansion is designed to support simplified data management. Centralized storage can ease management workload. Investment protection throughout the DS4000 family of storage systems.

3.6.1 Features

The IBM TotalStorage DS4400 Storage Server is a full-Fibre Channel, Storage Area Network (SAN) solution designed to provide the necessary performance at the right price.

The DS4400 Storage Server is based on the latest 2 Gbps fibre technology. With the DS4000 EXP700, 2 Gbps hard disk drives and DS4000 FC2-133 host bus adapter IBM delivers a complete end-to-end 2 Gb fibre solution. This allows large amounts of data to be consolidated for quicker, more efficient access and management.

Using the Storage Manager Version 9.12, it supports FlashCopy, Dynamic Volume Expansion, Remote Volume Mirroring and VolumeCopy with controller based support for up to 64 storage partitions.

- ▶ Dynamic Volume Expansion resizes logical volumes, typically without interruption of other logical volumes.
- ▶ FlashCopy tool is designed to create a fast, single point-in-time copy of all stored data and to support uninterrupted performance of normal SAN operations.
- ▶ Remote Mirroring creates a mirror of volume across the Fibre Channel fabric between primary and backup DS4400 Storage Servers.
- ▶ VolumeCopy, used with FlashCopy, is a function that allows a complete logical volume copy (clone copy).

RAID levels 0, 1, 3, 5, and 10 are supported, and for performance, it includes a 2 GB battery-backed cache (1 GB per controller).



Figure 3-4 DS4400 storage server (left) and rack-mounted storage servers (right)

The DS4400 scales from 36 GB to 67.2TB of storage using 16 DS4000 EXP700 expansion enclosures. Each expansion enclosure supports up to 14 Fibre Channel disk drives. It can also support the attachment of 16 EXP100 expansion units.

Important: The correct power on sequence is to switch on all expansion enclosures first and when all disks are in the enclosures are ready (steady green LED for each disk) switch on the Storage Server. The correct power off sequence is just the opposite: First, switch off the Storage Server, then the expansion enclosures. Applying the wrong sequence can cause a configuration loss, resulting in a loss of data.

Please refer to: <http://www.ibm.com/support/docview.wss?uid=psg1MIGR-4EZEWN>

3.6.2 Specifications

Table 3-9 DS4400 Storage Server

Model	1742-1RU
RAID controller	Dual active RAID controllers
Controller cache	Total of 2 GB, battery-backed
Drawer option weight	10.0kg (22.0 lb.)
Host interface	Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL)
Connectivity	4 standard, up to 8 via additional mini-hubs
Drive interface	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 and 10
Storage partitioning	Up to 64
Maximum drives supported	224
Fans and power supplies	Dual redundant, hot-swappable
Rack support	Industry standard 19" rack
Supported servers	For the most recent list, view: http://www-1.ibm.com/servers/storage/disk/ds4000/pdf/interop-matrix.pdf
Supported operating system	For the most recent list, view: http://www-1.ibm.com/servers/storage/disk/ds4000/pdf/interop-matrix.pdf

Table 3-10 Physical dimensions

Dimensions	174.5mm H x 444.5mm W x 609.6mm D (6.87 in x 18 in x 24 in)
Weight	38.6 kg (85 lbs)

For the latest specification information, check:

<http://www-1.ibm.com/servers/storage/disk/ds4000/ds4400/index.html>

3.7 DS4500 Storage Server

IBM DS4500 Storage Server delivers breakthrough disk performance and outstanding reliability for demanding applications in compute intensive environments. The DS4500 is designed to offer investment protection with advanced functions and flexible features. Designed for to days on demand business needs, the DS4500 easily scales from 36 GB to over 67 TB of Fibre Channel disk storage to support growing storage requirements. DS4500 offers advanced replication services to support business continuance and disaster recovery.

3.7.1 Features

The DS4500 is an enterprise class storage server that is designed to provide performance and flexibility for today's demanding data-intensive computing environments. It is the same physical size as the DS4400 with new higher performance controllers.

When combined with DS4000 EXP700 disk enclosures and 2 Gb IBM TotalStorage SAN Switches, the DS4500 can be used to help build a complete 2 Gb SAN environment. In addition, the performance capabilities of the DS4500 exceed those of other DS4000 Storage Servers, with up to 795 MB/s disk throughput. It also uses advanced caching algorithms to help optimize disk performance.

- ▶ Uses 2 Gb Fibre Channel connectivity to support high performance (772 MB/s throughput from disk) for faster, more responsive access to data.
- ▶ Facilitates storage consolidation for Storage Area Network (SAN), Network Attached Storage (NAS) and direct-attach storage environments.
- ▶ Supports continuous availability through redundant components and mirrored, battery backed-up cache.
- ▶ Up to 64 storage partitions.
- ▶ Supports 2 to 224 FC hard disk drives and provides more than 67 TB of easily scalable capacity using flexible combinations of 18.2, 36.4, 73.4, 146.8 and 300 GB drives.
- ▶ It attaches to DS4000 EXP700/710 and DS4000 EXP100 disk enclosures, and can support up to 224 Fibre Channel disk drives when using 16 DS4000 EXP700/710 disk enclosures.
- ▶ It can support up to 224 Serial ATA disk drives when using 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 NT®, Novell NetWare, Sun Solaris, HP-UX, Red Hat Linux, VMWare, Linux IA64).
- ▶ Offers advanced replication functions such as FlashCopy and Remote Mirroring for high data availability and protection.

3.7.2 Specifications

Table 3-11 DS4500 Storage Server

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 mini-hubs
Drive interface	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 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

For the latest specification information, check:

<http://www-1.ibm.com/servers/storage/disk/ds4000/ds4500/index.html>

3.8 EXP100 expansion unit

DS4000 Storage Servers with EXP100 SATA disks can be added to Storage Area Networks to satisfy the needs of various fixed content, data reference applications that require large amounts of storage capacity but do not have the high utilization and access characteristics satisfied by Fibre Channel disk drive storage. Users will continue to have the advantages of the DS4000 Storage Manager's common, centralized management interface for handling both classes of storage needs.

The IBM TotalStorage DS4000 EXP100 Storage Expansion Unit is designed to enable the attachment of Serial Advanced Technology Attachment (SATA) disk drives for the DS4500 and DS4300 Storage Servers.

Coupled with the DS4500 Storage Server, it allows you to configure RAID protected storage solutions. The EXP100 is designed with redundant 2 Gbps Fibre Channel connections to the DS4000 Storage Server, supporting increased reliability and performance as compared to older, non-redundant parallel Advanced Technology Attachment (ATA) products.

3.8.1 Features

The EXP100 is designed for mid-range and enterprise 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.

- ▶ 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
- ▶ Scalable to 56 TB on DS4500, 28 TB on DS4300 using 250 GB 7200 rpm (1.5 Gbps) SATA disk drives
- ▶ Support for AIX, Windows, and Linux environments making it the most interoperable solution in its class

- ▶ 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 3-5 EXP100 storage expansion unit

3.8.2 Hardware description

The EXP100 SATA disk enclosure supports up to fourteen 250 GB, 7,200 RPM SATA disk drives, offering up to 3.5 TB per 3U enclosure. It attaches to the DS4500 and DS4300 storage servers 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.

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.

3.8.3 Specifications

Table 3-12 IBM DS4000 EXP100 expansion unit

Model	1710-10U
Number of drives supported	Up to 14 SATA drives (250 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	38.6 kg (85 lbs)

3.8.4 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.

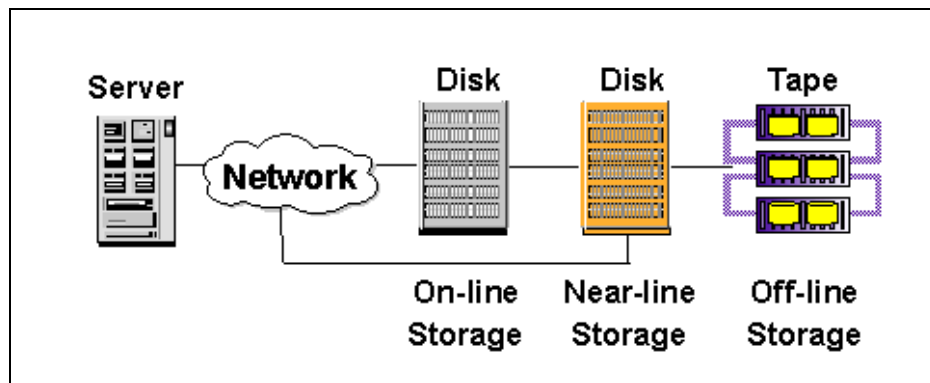


Figure 3-6 Storage data in a networks storage environment

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 networks 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 3-13.

Table 3-13 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	NA
MB/s performance	Highest	High	Lowest
Time to data	Immediate	~ Immediate	Mount time
Media reliability	Highest	Good	Good - Lower
Uptime	24/7	< 24/7	< 24/7
Typical applications	ERP / Oracle	Fixed content	Archive retrieval

3.9 EXP700 expansion unit

The EXP700 (Figure 3-7) is a 3U rack-mountable storage expansion enclosure that is an enhancement on the EXP500.

3.9.1 Features

The EXP700 contains fourteen bays for slim-line hot-swappable Fibre Channel disk drives. It features hot-pluggable and redundant power supplies and fans. Additionally, it contains two Enclosure Service Monitor (ESM) boards that can connect to the DS4000 Storage Server in a fully redundant Fibre Channel loop. This means a maximum of up to 4TB of capacity in a single EXP700.



Figure 3-7 EXP700 expansion unit

The IBM TotalStorage DS4000 EXP700 has been issued a Certificate of Conformance by National Technical Systems, Inc., for conformance to Network Equipment-Building System requirements for Level 3 type 2 and 4 equipment (NEBS 3).

3.9.2 Specifications

Table 3-14 IBM DS4000 EXP700 expansion unit

Model	1740-1RU/1740-1RX
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)

Weight	38.6 kg (85 lbs)
--------	------------------

3.10 EXP710 expansion unit

The EXP710 (Figure 3-7) is a rack-mountable storage expansion enclosure that is an enhancement to the EXP700. The enhancement is that the EXP710 enclosure is a *switched bunch of disks* (SBOD) version of the EXP700. Many SANs use Fibre Channel Arbitrated Loop (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).

3.10.1 Features

The EXP710 supports the 2 Gbps Fibre Channel connection and holds 14 slim-line FC disks. This means a maximum of up to 4TB of capacity in a single EXP710.

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, that the EXP700 is easily upgradeable to the EXP710. The process involves upgrading the storage server to version 9.1 or later, taking the EXP700's offline, replacing the ESMs, re-cabling the expansion unit, then powering it back on.

Important: It is not possible to mix EXP500's and EXP710's behind the DS4000 series disk systems. The EXP710 requires Storage Manager v9.1 (firmware level 06.10 or higher) and this firmware level does not support the EXP500 expansion unit.



Figure 3-8 EXP710 expansion unit

3.10.2 Specifications

Table 3-15 IBM DS4000 EXP710 expansion unit

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)
Weight	38.6 kg (85 lbs)

3.11 DS4000 Storage Manager

The IBM DS4000 Storage Manager software is used to configure, manage and troubleshoot the DS4000 Storage Server. 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 storage server components, updating the firmware of the RAID controllers and managing the storage server. Final, it offers advanced functions such as FlashCopy, Enhanced Remote Mirroring and VolumeCopy (these are premium features that need to be purchased).

Figure 3-9 shows the evolution of the DS4000 Storage Manager.

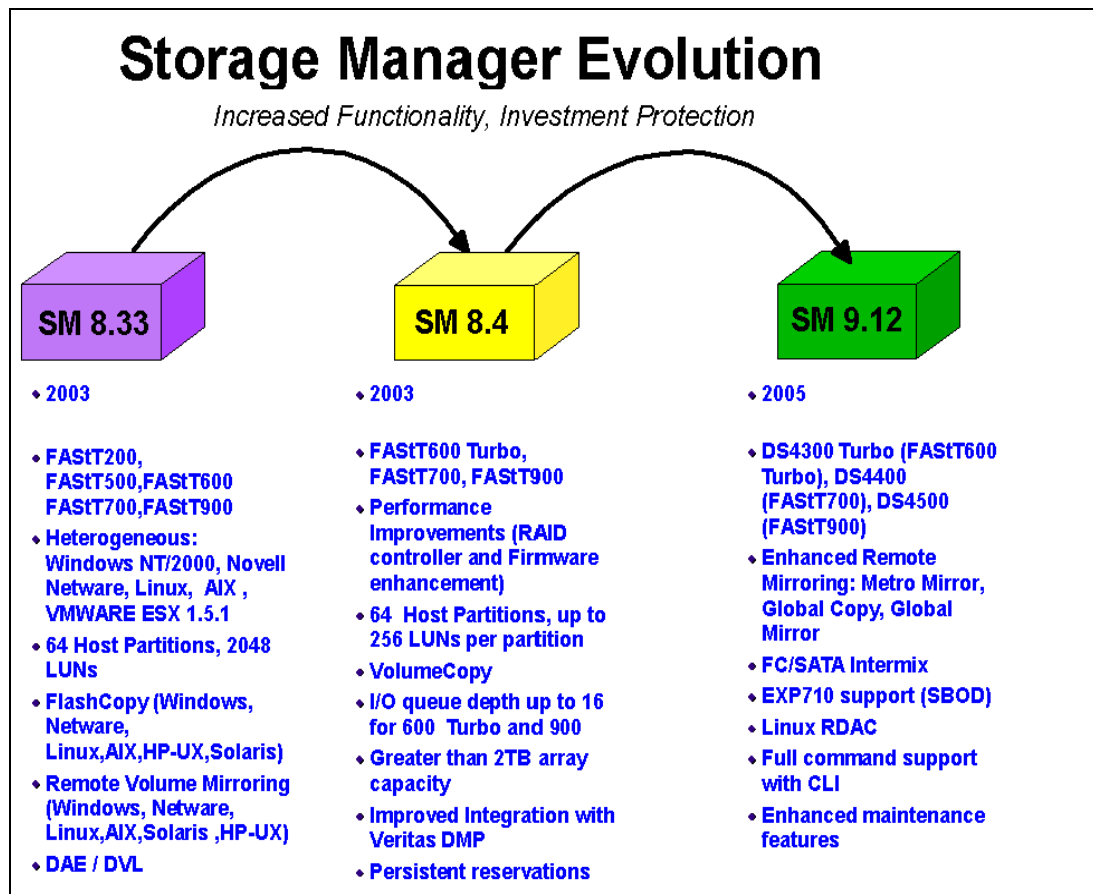


Figure 3-9 IBM DS4000 Storage Manager evolution

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 Storage Server. The flexibility to intermix different drive types (FC and SATA) in one Storage Server 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 Storage Servers. 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 provide Microsoft Exchange or Microsoft SQL Server administrators with the ability to have snapshot and capacity management functions integrated within their applications. Storage Manager 9.12 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

▶ **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 storage server 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 TotalStorage DS4000 Series and Storage Manager 9.10*, SG24-7010.

3.12 Copy services

The following table summarizes the major differences in the copy services between versions.

Table 3-16 Comparison of Storage Manager functionality

8.x Functionality	9.12 Functionality
Synchronous transfers	Synchronous and asynchronous transfers
No resynchronization	Synchronized suspended or broken mirrors
Max of 32 mirrors on DS4400/4500	Max of 64 mirrors on DS4400/4500
No DS4300 turbo support	DS4300 turbo support 32 mirrors
No access to secondary logical drive (mirror)	Read access and snapshot secondary logical drive
N/A	Write consistency option

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.

3.12.1 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 than a physical copy. In addition it 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.

The Storage Management software provides 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.

3.12.2 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.

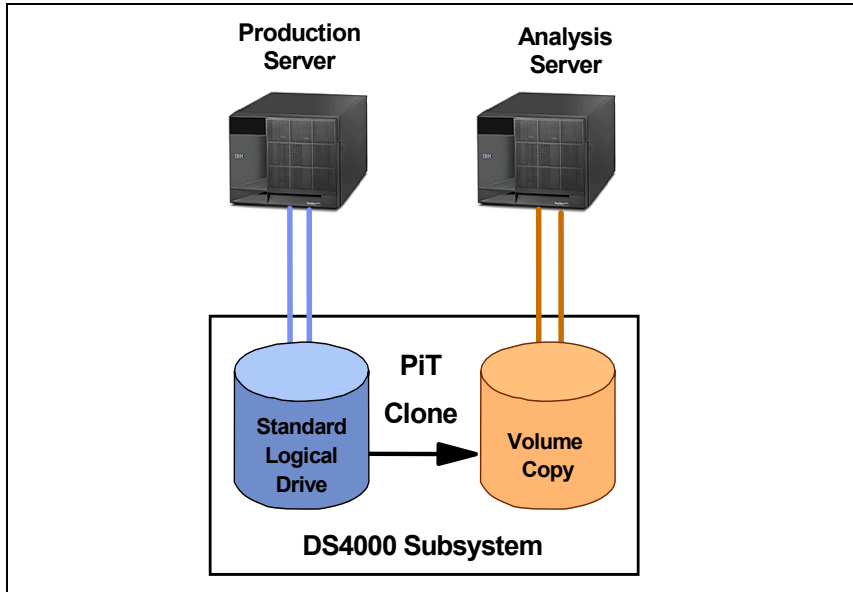


Figure 3-10 VolumeCopy

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 or 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 illustrated in Figure 3-11, 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.

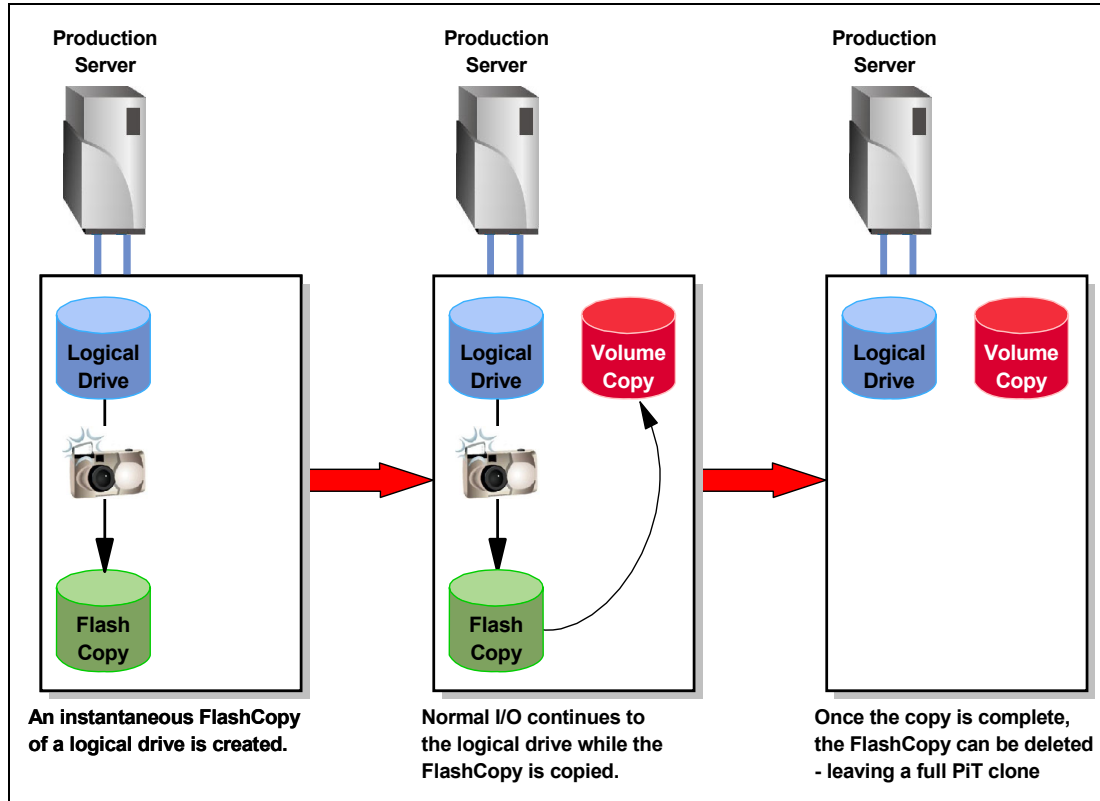


Figure 3-11 VolumeCopy integration with FlashCopy

3.12.3 Enhanced Remote Mirroring

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

An increased number of mirror relationships per Storage Subsystem are now available. This release of Storage Manager offers 64 mirror relationships per Storage Subsystem which is an increase from 32 in previous versions.

The ERM option is used for online, real-time replication of data between storage subsystems over a remote distance (Figure 3-12). 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.

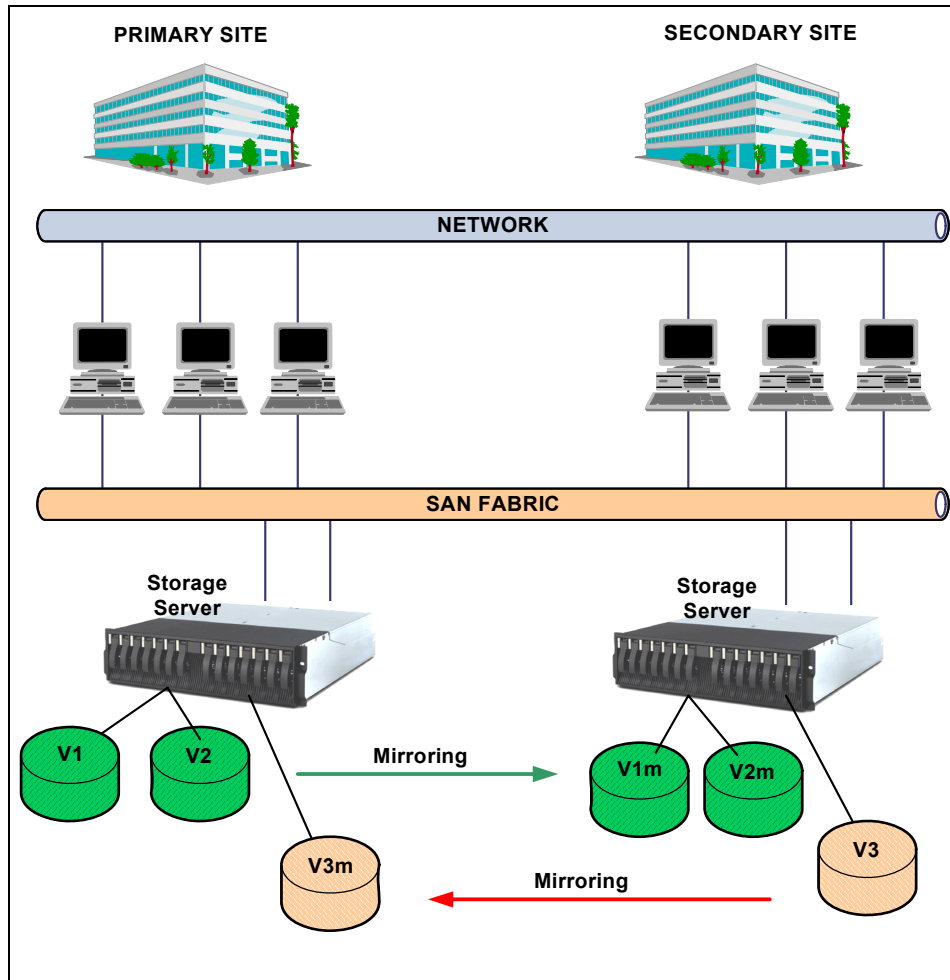


Figure 3-12 Enhanced Remote Mirroring

The Enhanced Remote Mirroring is a redesign of the former Remote Volume Mirroring (RVM) and now 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 resynchronization, 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 resynchronize 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 DS4500, DS4400 and now also the DS4300 turbo. On the DS4300 turbo it supports 32 mirror pairs while on the DS4400 and DS4500 it supports 64 mirror pairs.

3.13 Integrated Backup for Databases

TotalStorage Integrated Backup for Databases offering operates with IBM Tivoli Storage Manager software to provide a policy-based, automated data backup solution for DB2 databases residing on the DS4000 storage system. The solution uses DS4000 Copy Services to create point-in-time images of online databases.

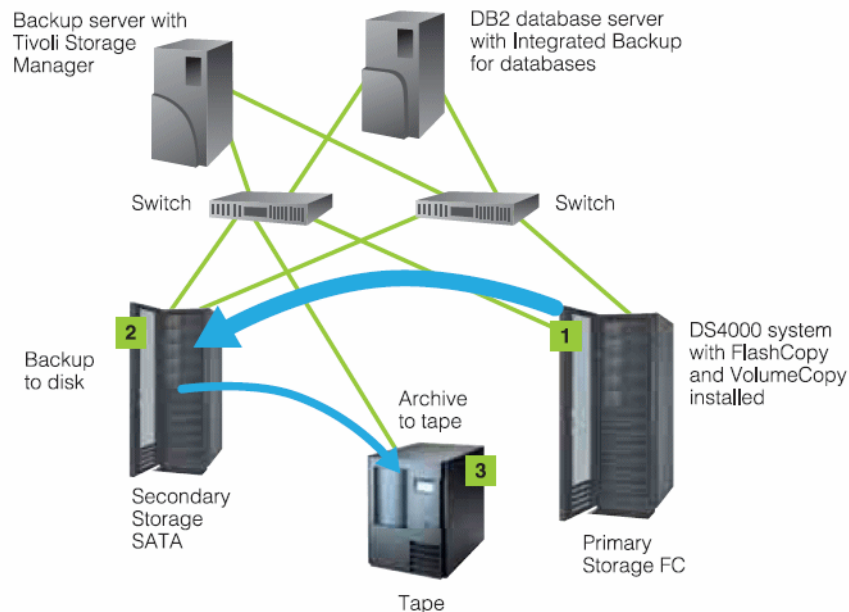


Figure 3-13 Integrated Backup for Databases

The backup process is designed to run automatically with daily applications, and can be scheduled during off-peak hours to help avoid impacting application performance. Because the process is designed to be fully automated, it can help avoid operator errors, enhance backup integrity and reduce system management costs. IBD works with the TSM to help provide a flexible range of database backup strategies. The IBD solution is designed for easy installation and configuration, and administrators can quickly add IBD support for any additional database servers and for any additional databases defined in a database server. IBD is designed to run automatically on a scheduled basis, thereby helping to reduce management costs.

3.14 DS4000 Service Alert

DS4000 Service Alert (hereafter called Service Alert) is a feature of the IBM TotalStorage 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 Storage Server customers.

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

3.15 DS4000 Alert Manager

IBM TotalStorage DS4000 Alert Manager, a services solution designed to support the remote monitoring of installed DS4000s, expands the capability beyond the existing DS4000 Service Alert. In 2005, IBM introduced DS4000 Alert Manager, which is designed to automatically notify IBM Service and Support of problems. The DS4000 Alert Manager enables IBM Service and Support to dial in to the DS4000 information logs to aid problem determination.

The new monitoring appliance, IBM TotalStorage DS4000 Alert Manager will allow 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 IBM TotalStorage 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

3.16 IBM Management Suite Java (FAStT MSJ)

FAStT MSJ is a network-capable application that can connect to and configure remote systems. FAStT MSJ helps you configure IBM Fibre Channel HBAs in a SAN environment. The FAStT Management Suite Java GUI can be used to manage the host bus adapters in servers running the FAStT MSJ agent. It is especially useful if you have hosts that are directly connected to the DS4000 as you can check for errors on the FC link of the HBA as well. Further details can be found in the redbook *IBM TotalStorage DS4000 Series and Storage Manager 9.10*, SG24-7010, at:

<http://www.redbooks.ibm.com/abstracts/sg247010.html?Open>

3.17 More information

More detailed information about the DS4000 range of storage servers can be found in the following publications and Web sites:

<http://www-1.ibm.com/servers/storage/disk/ds4000/index.html>

- ▶ Redbook *Fibre Array Storage Technology A FAStT Introduction*, SG24-6246, which can be found at:

<http://www.redbooks.ibm.com/abstracts/sg246246.html?Open>

- ▶ Redbook *IBM TotalStorage DS4000 Series and Storage Manager 9.10*, SG24-7010, which can be found at:
<http://www.redbooks.ibm.com/abstracts/sg247010.html?0pen>
- ▶ Redpaper *IBM TotalStorage FAStT Best Practices guide*, REDP3690, which can be found at:
<http://www.redbooks.ibm.com/redpapers/pdfs/redp3690.pdf>
- ▶ Redpaper *Introducing the IBM TotalStorage FAStT EXP100 with SATA disks*, REDP3794, which can be found at:
<http://www.redbooks.ibm.com/redpapers/pdf/redp3794.pdf>
- ▶ Publication *Storage Manager version 9 Copy Services User's Guide* GC26-7707-01
ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/gc26-7707-01.pdf
- ▶ Interoperability matrix:
<http://www-1.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's 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:

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

4.1 Introduction

The IBM Enterprise Disk Systems are designed to deliver high-performance, high-availability storage with flexible characteristics that can be configured according to your business requirements. Building on a solid foundation of the IBM TotalStorage Enterprise Storage Server (ESS) environment and reusing IBM's leading technology innovations, the IBM TotalStorage DS8000 series, along with the D6000 series, delivers an enterprise storage continuum of systems with the same functional code, shared replication services and common management interfaces (see Figure 4-1).

One Family: Same Code - Same Functions - Same Interfaces

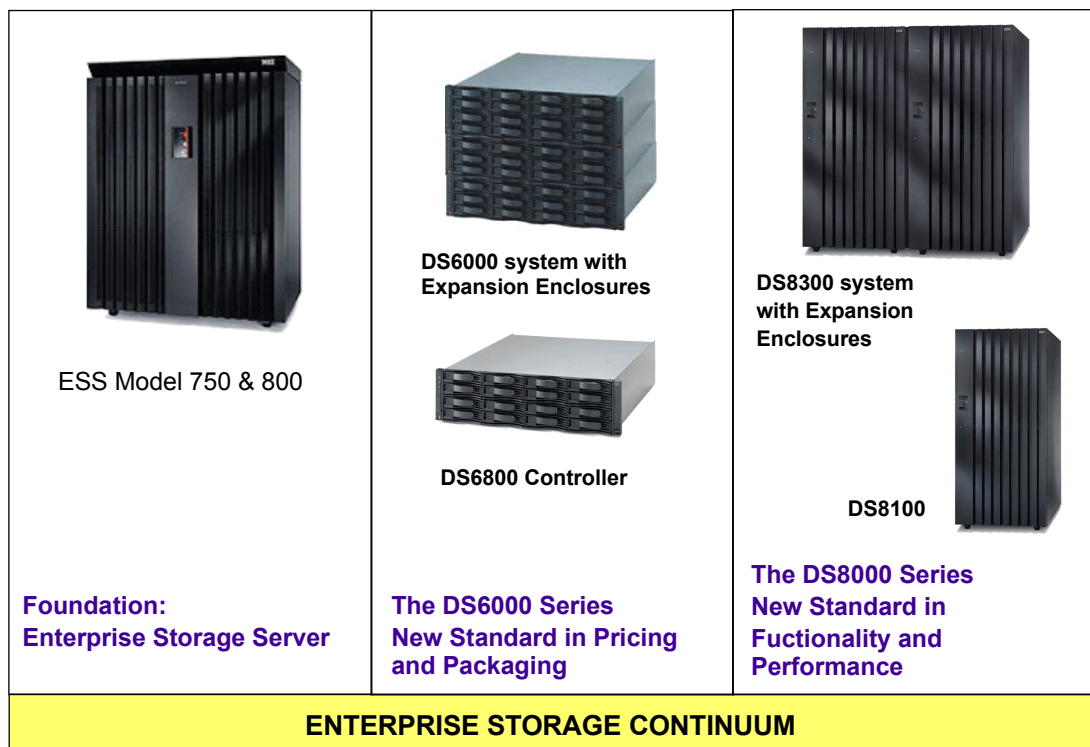


Figure 4-1 Enterprise Disk Systems

Enterprise Storage Server (ESS)

The ESS set a new standard for storage servers back in 1999 when it was first made available. From the initial E models to the succeeding F models, to the current 750 and 800 models, the ESS significantly improved over time levels of performance, throughput and scalability with more powerful hardware and functional enhancements.

The DS6000 Series

The DS6000 series offers true enterprise-class functionality with modular design and reduced price. Clients who currently have IBM TotalStorage ESS models in their enterprise should also consider the IBM TotalStorage DS6000 series when they plan to replace or buy additional storage. Intended for medium and large businesses, the DS6000 series can help simplify data management and enable easy scalability.

The DS8000 Series

The IBM TotalStorage DS8000 series is the next generation of the IBM TotalStorage Enterprise Storage Server (ESS) 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.

This remainder of this section introduces to the IBM Enterprise Disk Systems and describes the key common characteristics of the enterprise-class IBM disk storage products.

4.1.1 Storage consolidation

Consolidation begins with compatibility. The IBM Enterprise Disk Systems can be connected across a broad range of server environments. You can easily split up storage capacity among the attached environments and reduce the number of storage systems you need to use. At the same time, you can construct a disaster recovery solution that makes use of the full range of your Enterprise disk storage. For example, you can mirror a DS8000 series system with a DS6000 series system or an ESS.

4.1.2 Performance

The IBM Enterprise Disk Systems are designed for high performance that takes advantage of IBM's 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 4-2)

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.

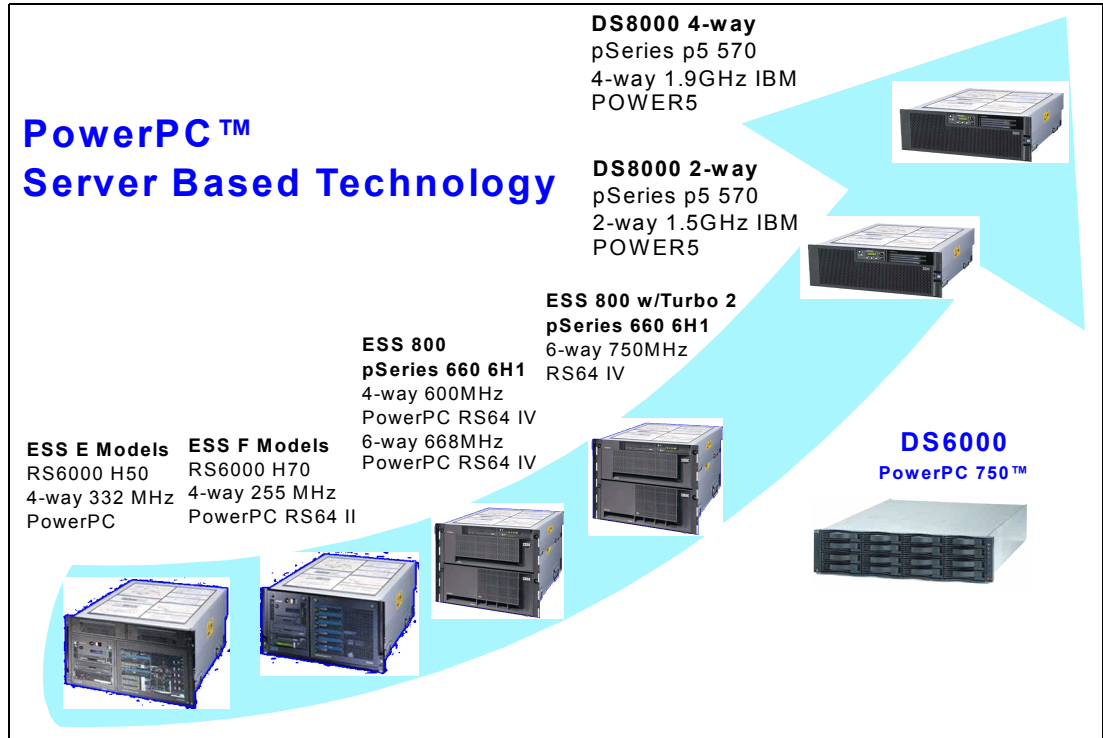


Figure 4-2 The server based architecture brings continuous performance improvements

IBM multipathing software

IBM Multipath Subsystem Device Driver (SSD) 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's 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.

Performance for zSeries®

As is the case for the IBM TotalStorage ESS, the new DS6000 and DS8000 also supports the following IBM performance innovations for IBM @server zSeries environments:

Parallel Access Volumes (PAV)

PAV is an optional feature for zSeries environments which enables a single zSeries server to simultaneously process multiple I/O operations which can help to significantly improve throughput. With Dynamic PAV, storage volumes can be automatically managed to help the workload meet its performance objectives and reduce overall queuing.

Multiple Allegiance

Multiple Allegiance is a standard feature which expands simultaneous I/O access capability across multiple zSeries servers. This function, along with the software function PAV, enables storage systems to process more I/O operations in parallel, helping to dramatically improve performance and enabling greater use of large volumes.

Priority I/O Queuing

Priority I/O Queuing improves performance in z/OS environments with several z/OS images. The z/OS Workload Manager (WLM) controls where work is run and optimizes the throughput

and performance of the total system. The IBM TotalStorage Enterprise Disk Systems provide the WLM with more sophisticated ways to control the processing sequence of I/O operations.

FICON®

2 GB FICON connectivity delivers high bandwidth and provides a high-speed pipe supporting multiplexed operations for zSeries systems. The ESS and the DS8000 series provide ESCON connection for older zSeries hosts that do not support FICON.

4.1.3 Data protection

Many design characteristics and advanced functions of the IBM Enterprise Disk Systems contribute to protect the data in an effective manner.

Fault-tolerant design

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

For the more information regarding RAID concepts, check Appendix B, “Redundant Arrays of Independent Disks (RAID)” on page 463.

4.1.4 Common set of functions

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 TotalStorage DS Command-Line Interface (DS CLI) and the IBM TotalStorage DS open application programming interface (API).

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

4.2 Overview of the DS6000 and DS8000 series

The IBM DS6000 series offers an entirely new era in price, performance and scalability. Now for the first time IBM @server iSeries™ and zSeries customers have the option for a mid-range priced storage subsystem with all the features and functions of an enterprise storage subsystem. Obviously the DS8000 can deliver a higher throughput and scales higher than the DS6000 but not all customers need this high throughput and capacity. You can choose a system that fits your needs.

The maximum storage capability of the DS68000 controller is 4.8 TB. You can use the optional DS6000 expansion enclosures and provide a maximum storage capability of 67.2 TB. The DS8000 series is well prepared to address the exponential growth of data within an enterprise. 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.

Now you have the option of choice, you can build very cost efficient storage systems by adding expansion enclosures to the DS6800 controller or you can 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.

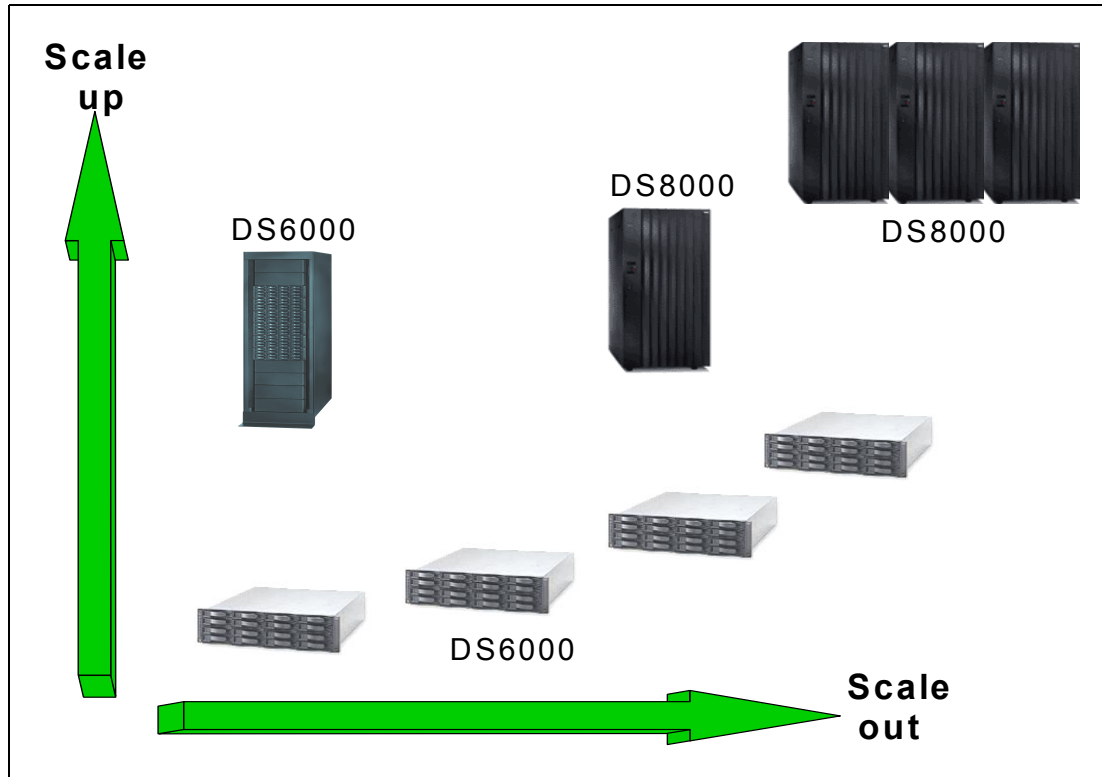


Figure 4-3 Scaling options

Multi-tiered storage

The DS6000/DS8000 series enables you to construct a multi-tiered storage environment to help 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.

4.2.1 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 in comparison to the previous ESS. Instead of SSA loops there is now a switched FC-AL implementation. With this switching technology there is a point-to-point connection to each disk drive and adapter. This allows maximum bandwidth for data movement, eliminates the bottlenecks of loop designs and allows for specific disk drive fault indication. 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.

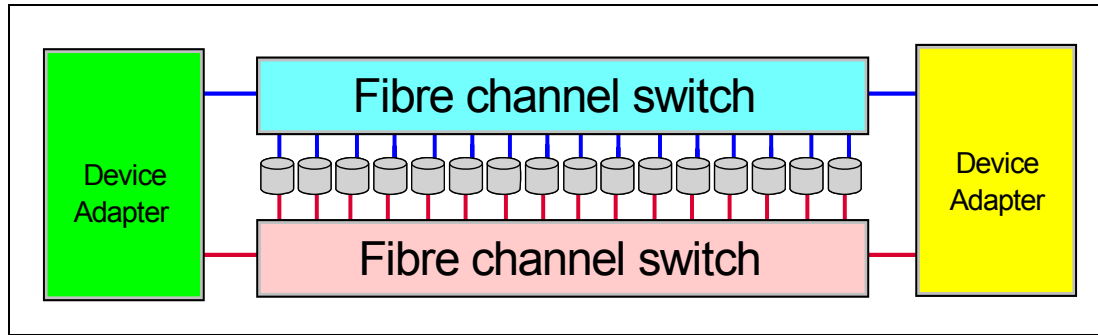


Figure 4-4 Disk enclosure switched connections

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. 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 is connected 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 such that all eight ports on the two device adapters can be concurrently communicating with independent disk drives.

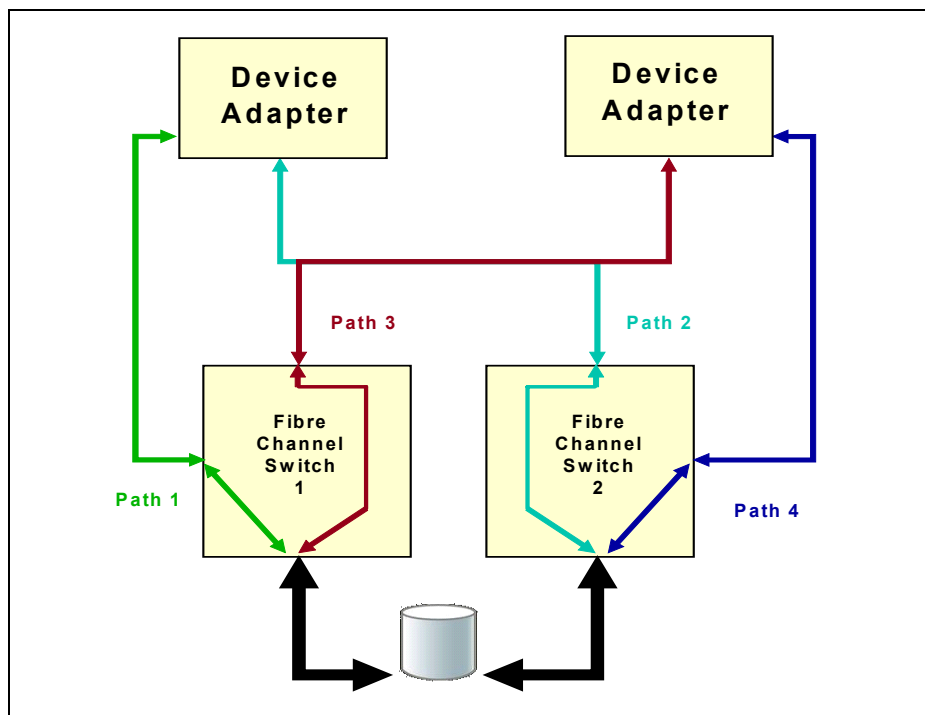


Figure 4-5 Four path to each disk drive

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.

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 4-6.)

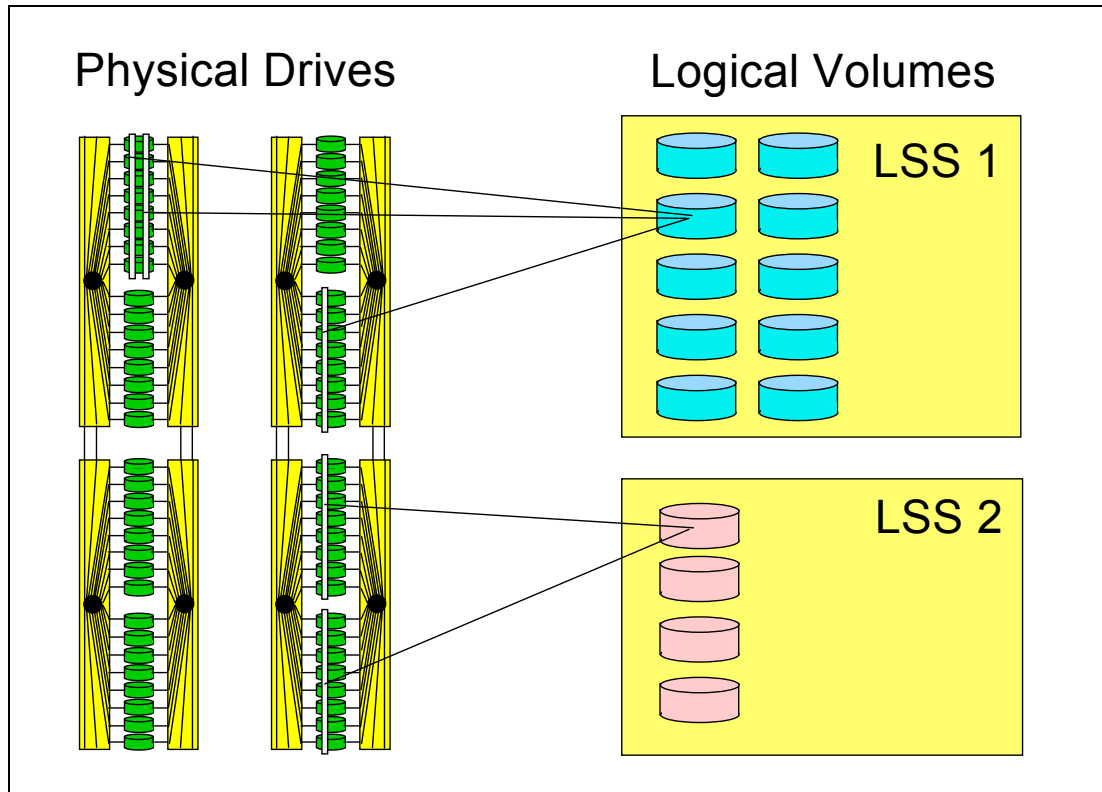


Figure 4-6 Grouping of volumes in LSSs

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.

4.3 The DS6000 series

The DS6000 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 along with your storage needs up to 67.2 TB physical storage.



Figure 4-7 The DS6000 series

The DS6000 consists of the *DS6800 controller enclosure* which has dual Fibre Channel RAID controllers. Capacity can be increased by adding up to 13 *DS6000 expansion enclosures*.

4.3.1 The DS6800 Controller Enclosure

The DS6000 series is based on IBM's 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 4-8).

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 4-8 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 is equipped with 2 GB memory in each controller card, adding up to 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.

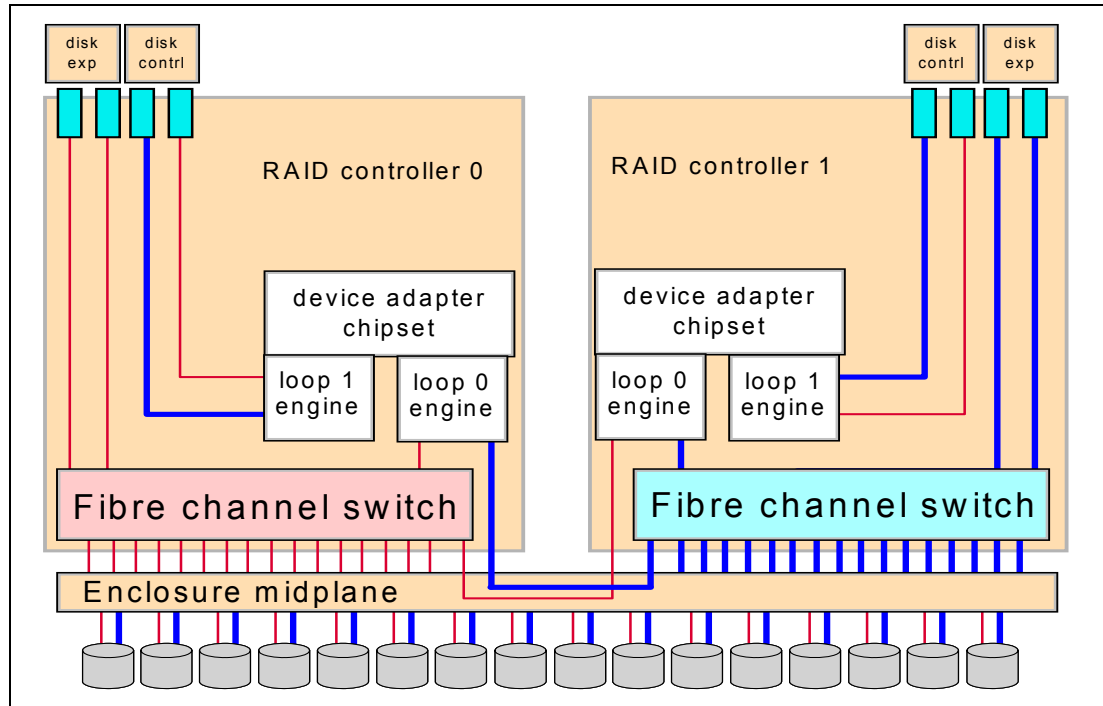


Figure 4-9 DS6000 disk architecture

The configuration process when forming RAID-5 or RAID-10 arrays will require that two global spares be defined in the DS6800 controller enclosure. If you have expansion enclosures, the first enclosure will have another two global spares. More spares could be assigned when drive groups with larger capacity drives are added. If a DDM is not present then its slot must be occupied by a dummy carrier.

Host adapters

The DS6800 has eight 2 Gbps Fibre Channel ports. Each port can be configured individually to operate in Fibre Channel or FICON mode. A 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, whereby the data is written to volatile memory on one controller and 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.

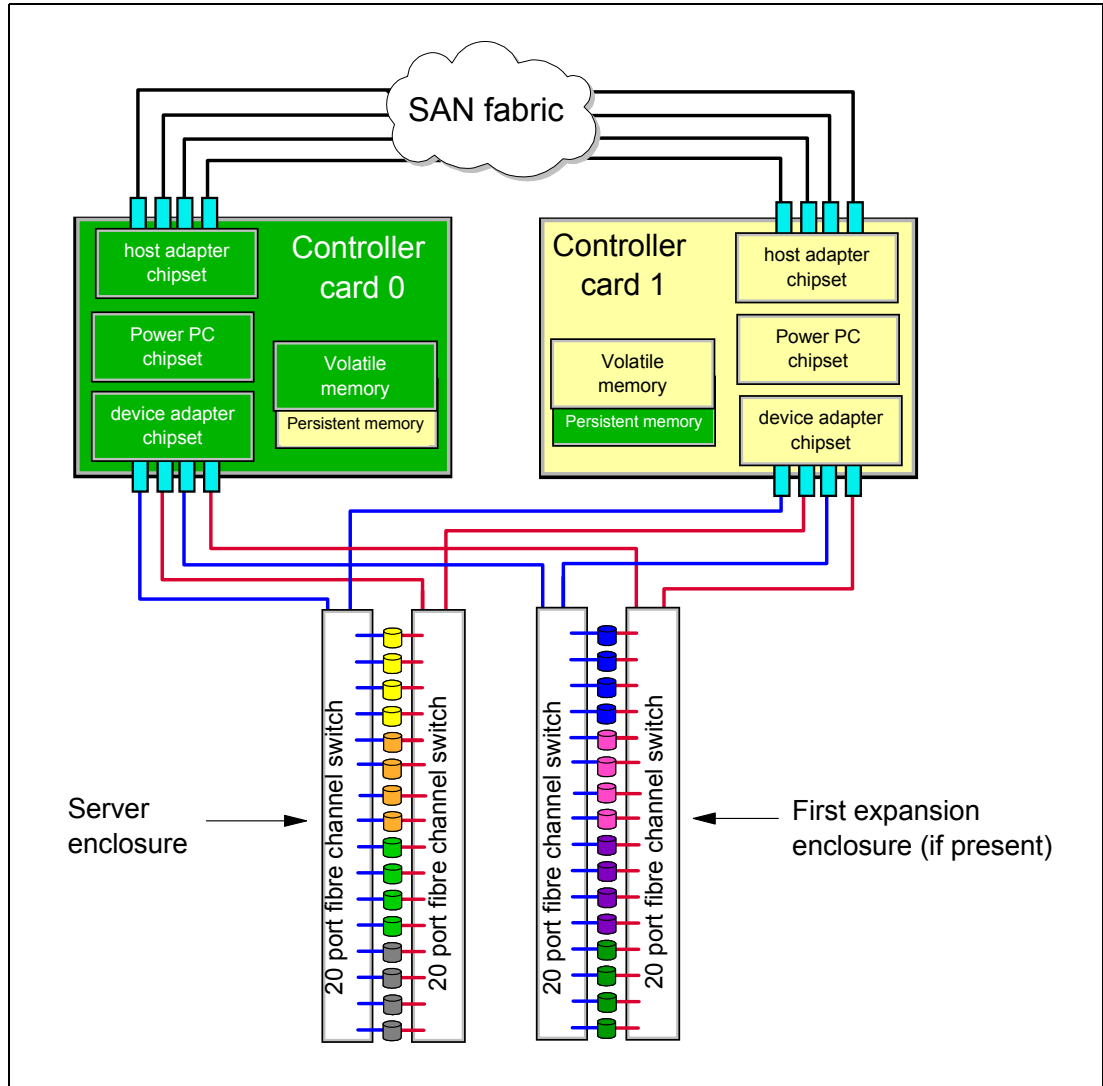


Figure 4-10 DS6800 controller architecture

Power subsystem

The power subsystem of the DS6800 consists of 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 is able to support the power requirements of an entire enclosure.

Dense packaging

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

4.3.2 The DS6000 Expansion Enclosure

The DS6000 expansion enclosure is used to add capacity to an existing DS6800 server enclosure. Up to 13 DS6000 expansion enclosures can be added to a DS6800 controller enclosure.

Expansion enclosure SBOD controller card

The DS6000 SBOD controller card (see Figure 4-11) 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 4-11 SBOD controller card

The DS6800 supports two dual redundant switched loops. The first loop is for the DS6800 and up to six (6) DS6000 expansion enclosures. The second switched loop is for up to seven (7) 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). When you attach the first expansion enclosure you attach it to the DISK CTRL ports to start the second disk loop (loop 1). The DISK EXP ports are used to attach the second expansion enclosure. It joins the same switched loop as the disks in the server enclosure (loop 0).

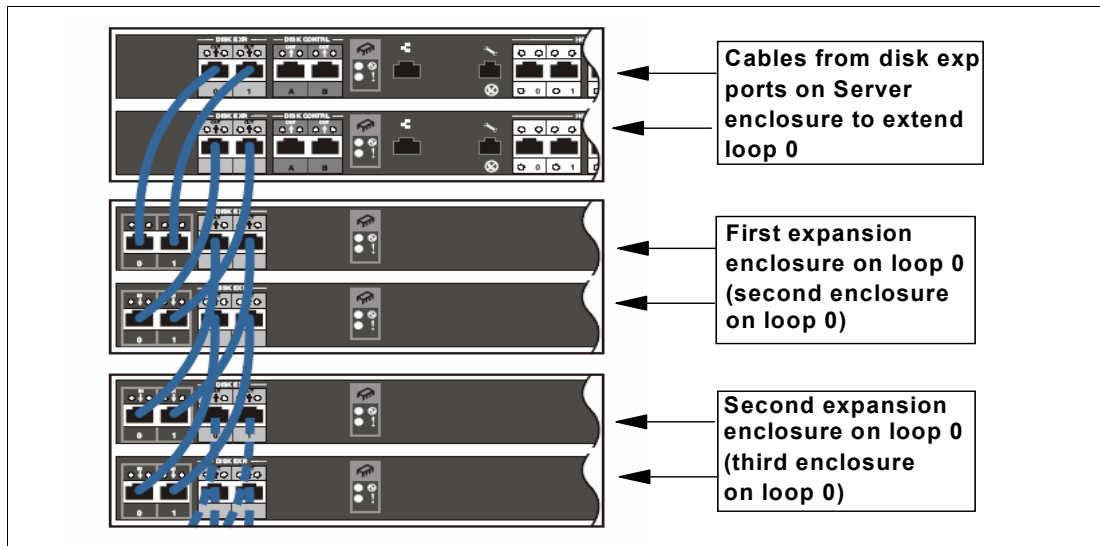


Figure 4-12 Example of expansion enclosure cabling

You add one expansion enclosure to each loop until both loops are populated with seven enclosures each (remembering the server enclosure represents the first enclosure on the first loop). Note that while we use the term disk loops, and the disks themselves are FC-AL disks, each disk is actually attached to two separate Fibre Channel switches.

4.3.3 DS6800 Model 1750-511

The 1750-511 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 4-13 DS6800 Model 1750-511 rear view

The 1750-511 model offers 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 thirteen DS6000 Model 1750-EX1 expansion enclosures
- ▶ Selection of 2 Gbps Fibre Channel disk drives including:
 - 73 GB 15k RPM
 - 146 GB 10k RPM
 - 300 GB 10k RPM
- ▶ 3U in height and mountable in a standard 19-inch rack

4.3.4 DS6000 Model 1750-EX1

DS86000 can expand with the expansion enclosure model 1750-EX1. Up to 13 expansion drawers build up to the maximum configuration of 224 DDMs which is comprised of 14 drawers x 16 DDMs.

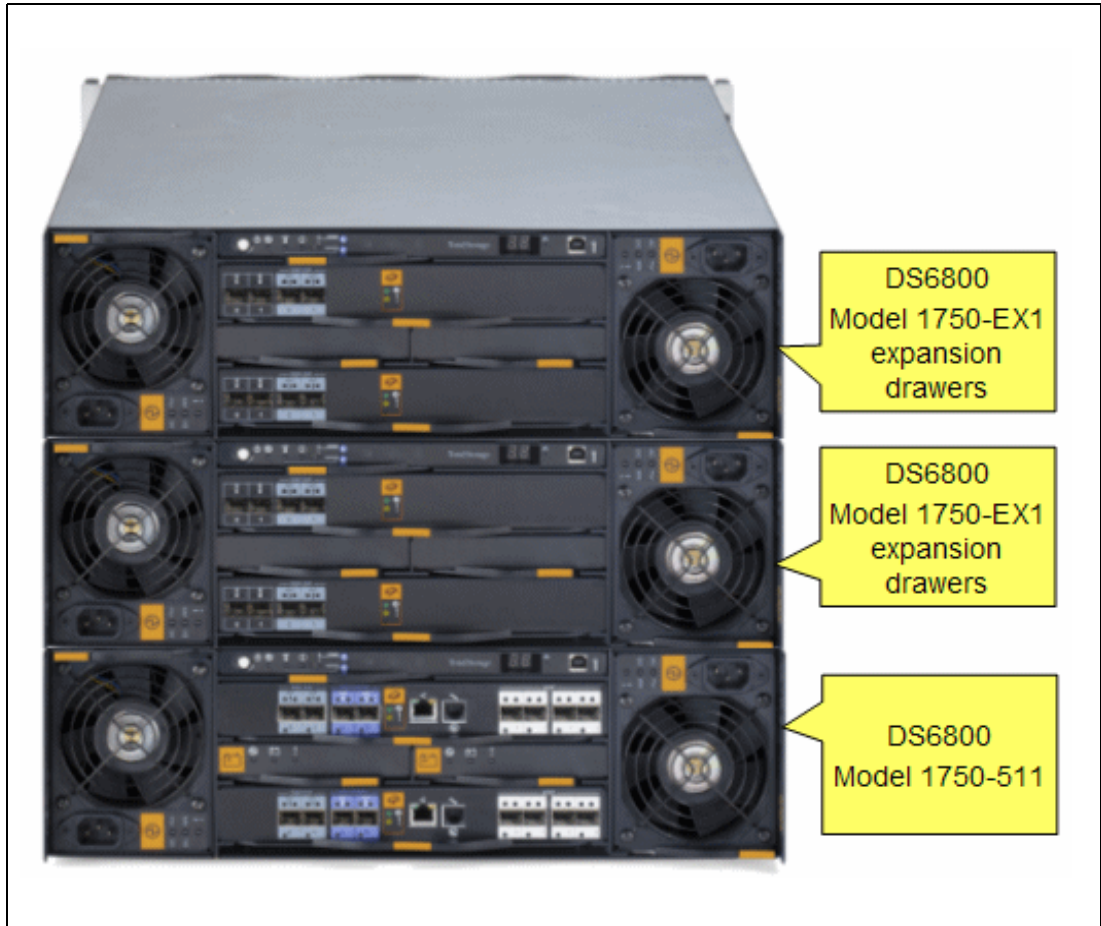


Figure 4-14 DS6000 with 2 expansion enclosures

Each expansion enclosure contains the following features:

- ▶ Two expansion controller cards and each controller card provides the following:
 - 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 imbedded 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

4.3.5 More information

For more information about the DS6000 series architecture, refer to the redbook *IBM TotalStorage DS6000 Series: Concepts and Architecture*, SG24-6471, which can be found at:

<http://www.redbooks.ibm.com/abstracts/sg246471.html>

4.4 The DS8000 series

The DS8000 series 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.

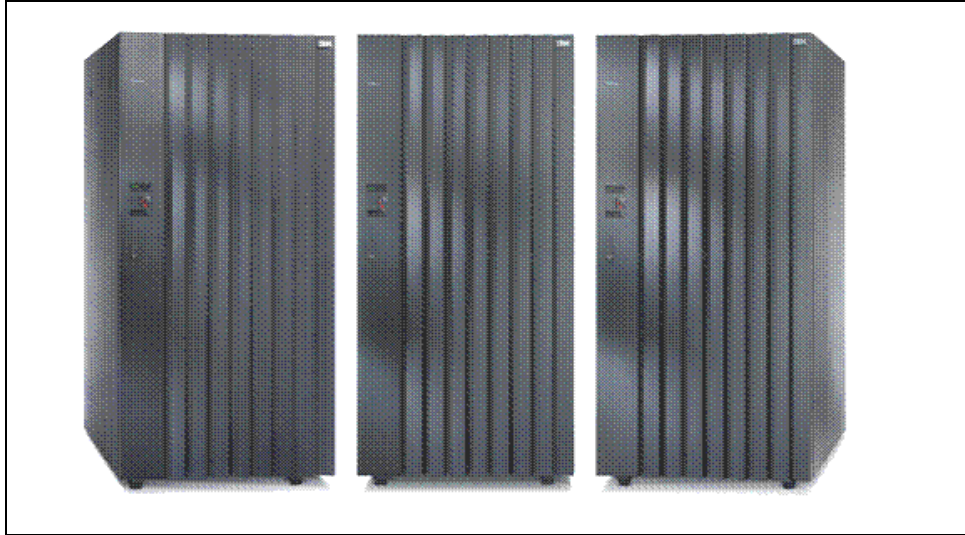


Figure 4-15 The DS8000 series

The hardware has been optimized to provide enhancements in terms of performance, connectivity, and reliability. The DS8000 series features several models in a new, higher-density footprint, than the ESS, however from an architectural point of view the DS8000 series has not changed much in regard to the fundamental architecture of the previous ESS models. Seventy five percent of the operating environment remains the same as for the ESS Model 800. This ensures that the DS8000 can leverage a very stable and well-proven operating environment.

4.4.1 The DS8000 design enhancements

In the following you can find the key enhancements of architecture design:

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. Within the POWER5 servers the DS8000 series offers up to 256 GB of cache, which is up to 4 times as much as the previous ESS models. 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

DS8000 comes with a high bandwidth, fault tolerant internal interconnection, which is also used in the pSeries Server. It is called RIO-2 (Remote I/O) and can operate at speeds up to 1 GHz and offers a 2 Gbps sustained bandwidth per link. The RIO-G design allows the sharing of host adapters between servers and offers exceptional performance and reliability.

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 is designed to enable the creation of two or even more completely separate storage subsystems which run the same or different versions of the 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.

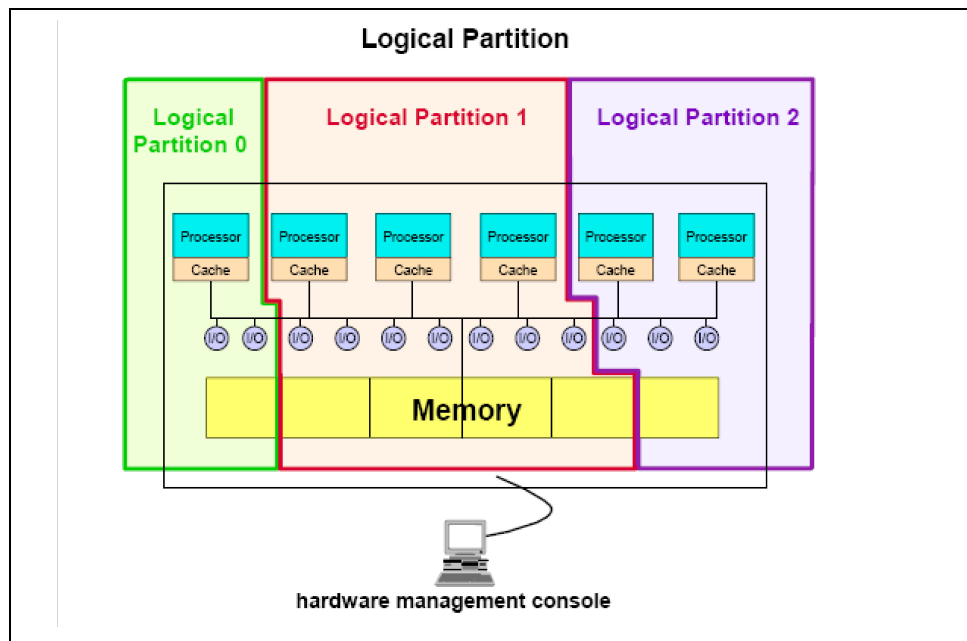


Figure 4-16 Logical partitions (LPARs)

The LPAR functionality is available in the DS8300 Model 9A2. The first application of the pSeries Virtualization Engine™ technology in the DS8000 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)

Note: IBM's plans for the future include offering even more flexibility in the use of storage system LPARs. Not only the processor resources will be more flexible, but plans call for the movement of memory more dynamically between LPARs. Another idea designed to maximize the value of using the storage system LPARs is to have application LPARs. The list of possible applications include Tivoli Storage Manager, Legato, Veritas and other backup/recovery solutions.

4.4.2 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 4-17 shows the components used to create the DS8000 Primary Frame.

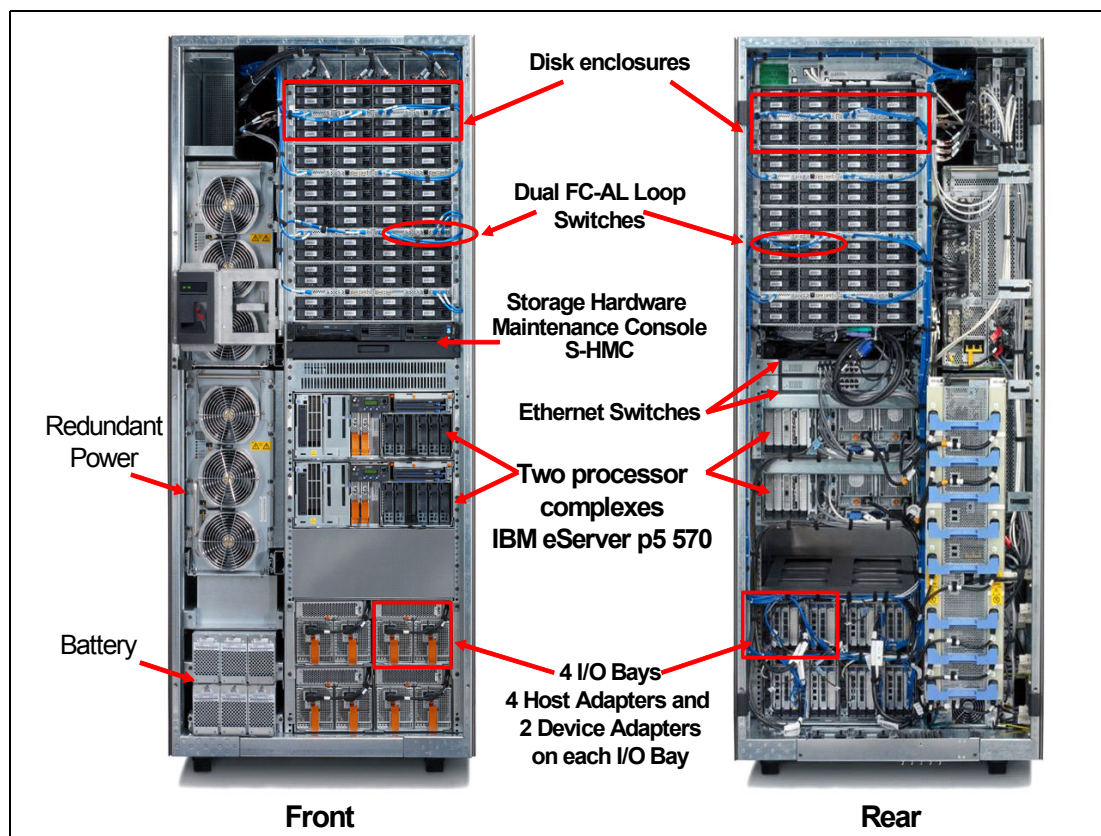


Figure 4-17 DS8000 - Primary Frame

The processor complexes

The DS8000 base frame contains two IBM @server p5 570 processor complexes. These @server p5 570 servers contain the processor and memory that drive all functions within 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.

In the ESS we referred to them as clusters, but this term is no longer relevant. We now have the ability to logically partition each processor complex 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 offer up to 256 GB of processor memory. Half of this will be 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

To access the disk subsystem, each complex uses several four-port fibre-channel arbitrated loop (FC-AL) device adapters. A DS8000 can potentially have up to sixteen of these adapters arranged into 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). With AAL, an array site is actually split 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.

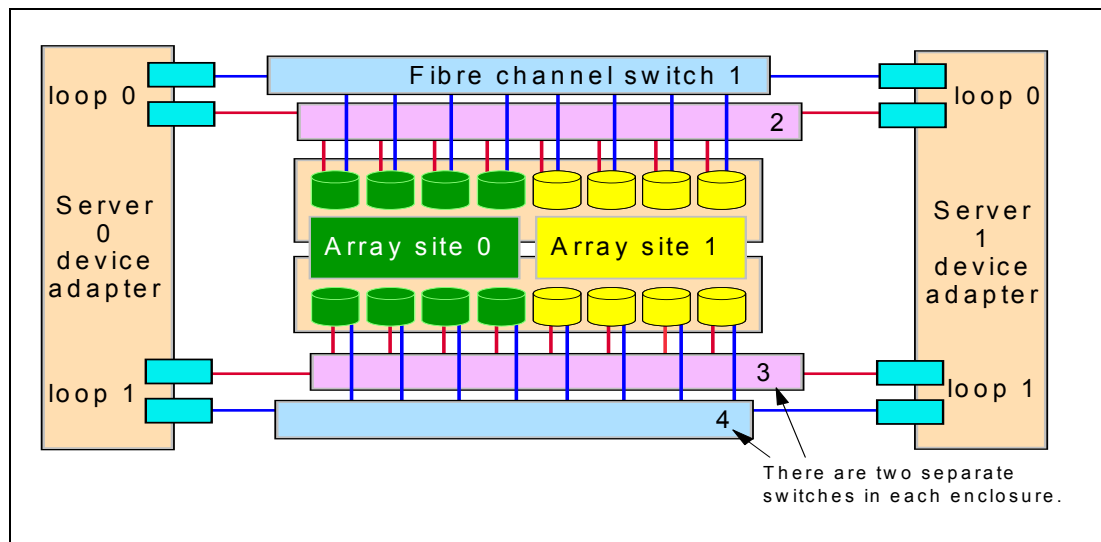


Figure 4-18 Array across loops

The host adapters

The DS8000 offers enhanced connectivity with the availability of four-port Fibre Channel/FICON host adapters. Each port can be independently configured as either fibre-channel protocol (FCP) port for open systems host connection or PPRC FCP links, but also as a FICON port to connect to zSeries hosts. ESCON host connection is also supported, but with ESCON a host adapter contains only two ESCON ports. The mix of ESCON ports and FCP ports on the same HA is not possible. A DS8000 can have both ESCON adapters and FCP/FICON adapters at the same time.

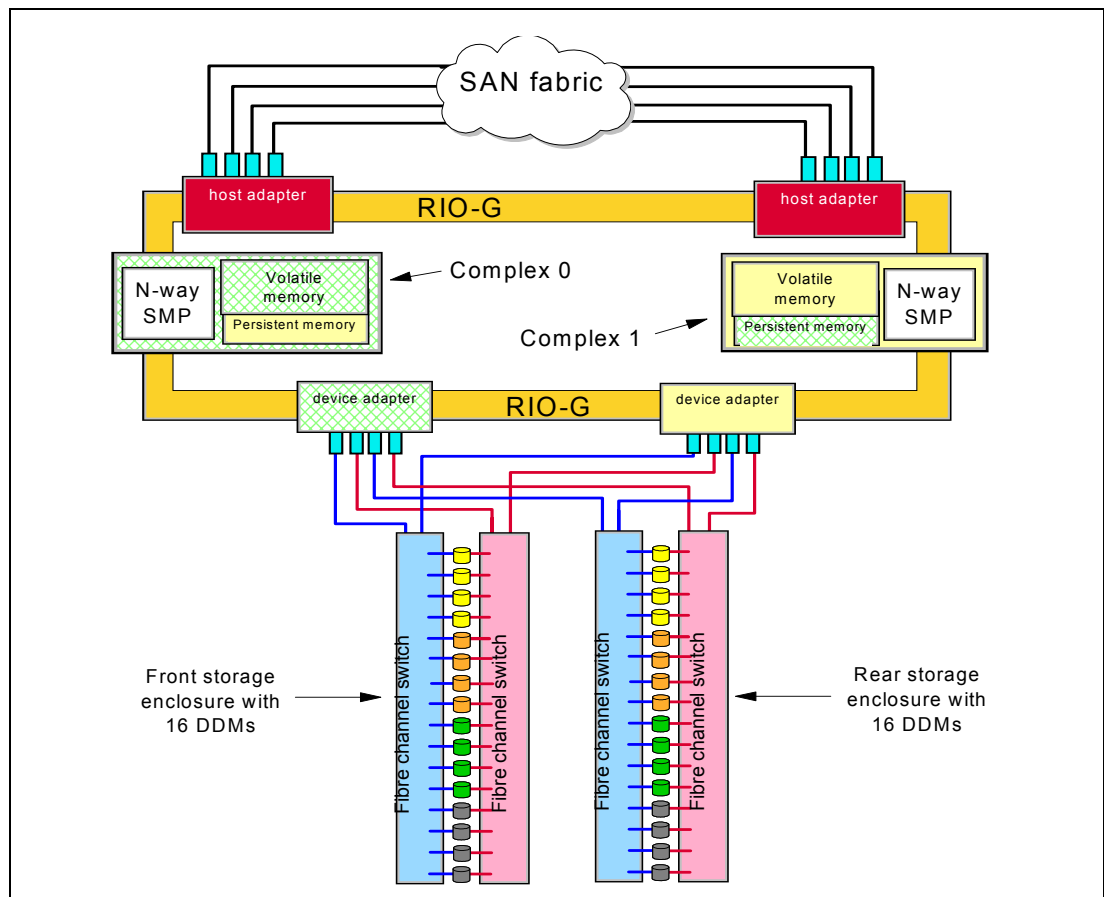


Figure 4-19 The DS8000 architecture

The attached hosts interact with software which is running on the complexes to access data on logical volumes. Each complex will host at least one instance of this software (which is called a *server*), which runs in a logical partition (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. The RIO-G design allows the sharing of host adapters between servers and offers exceptional performance and reliability.

The RIO-G ports are used for I/O expansion to external I/O drawers. RIO stands for remote I/O. 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 is capable of passing 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.

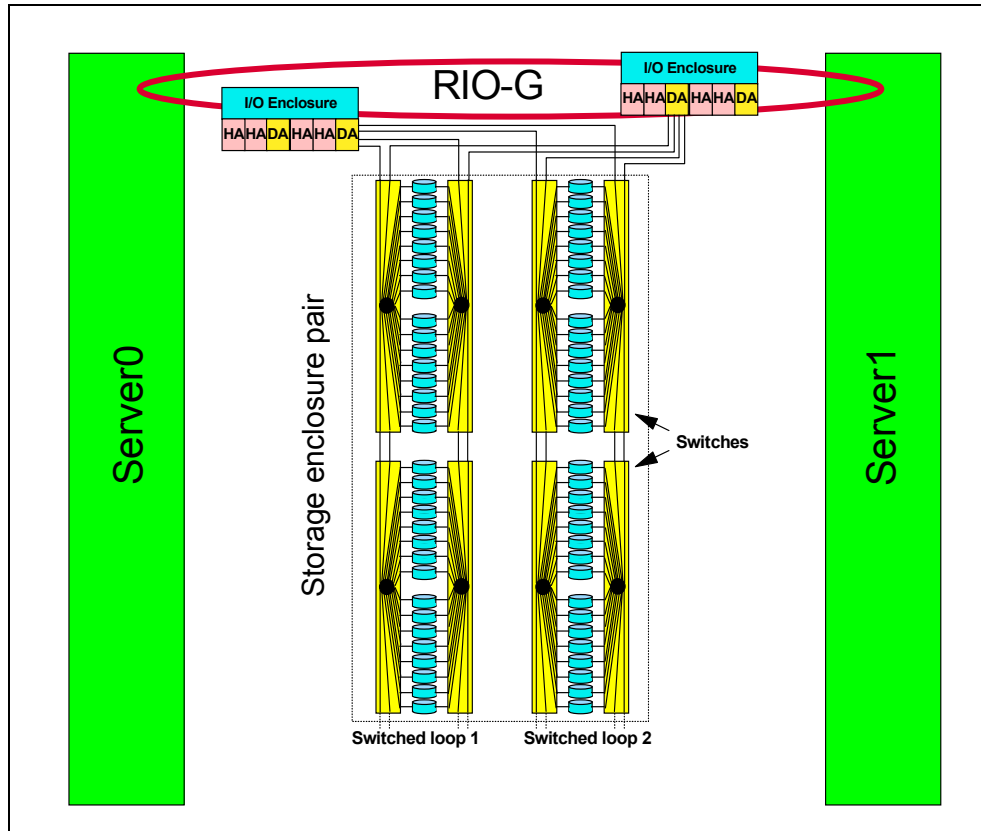


Figure 4-20 RIO-G design

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 will be one management console for one 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

Between the disk enclosures and the processor complexes are two 16-port Ethernet switches. Two switches are supplied to allow the creation of a fully redundant network. Each processor complex has multiple connections to each switch. This allows each server to access each switch. This switch cannot be used for any equipment not associated with 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. Each PPS is capable of powering 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 the event of a 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). In the event of a complete loss of input AC power, 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, prior to 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.

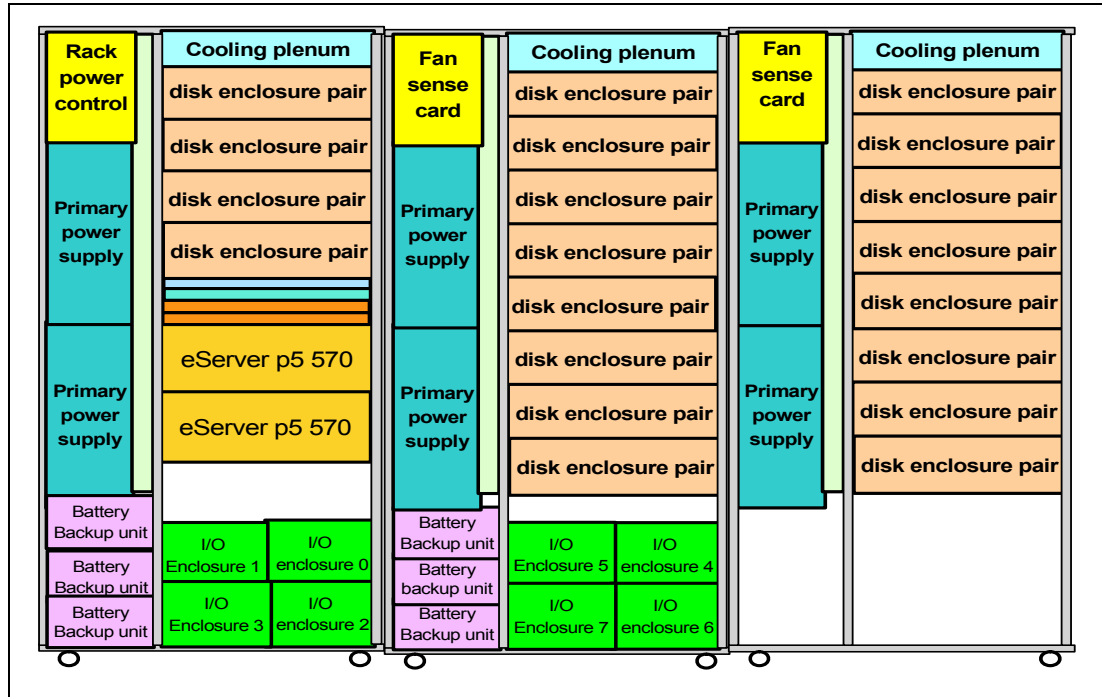


Figure 4-21 DS8000 Frames

Expansion frames can contain I/O enclosures and adapters if they are the first expansion frame that is 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, 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.

4.4.3 DS8000 Models

One of the advantages of the DS8000 series is its linear scalability for capacity and performance. If your customer's 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 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. With expansion frame called Model 92E, you can expand the disk storage capacity as follows:

- ▶ Up to 384 disk drives, for a maximum capacity of 115.2 TB

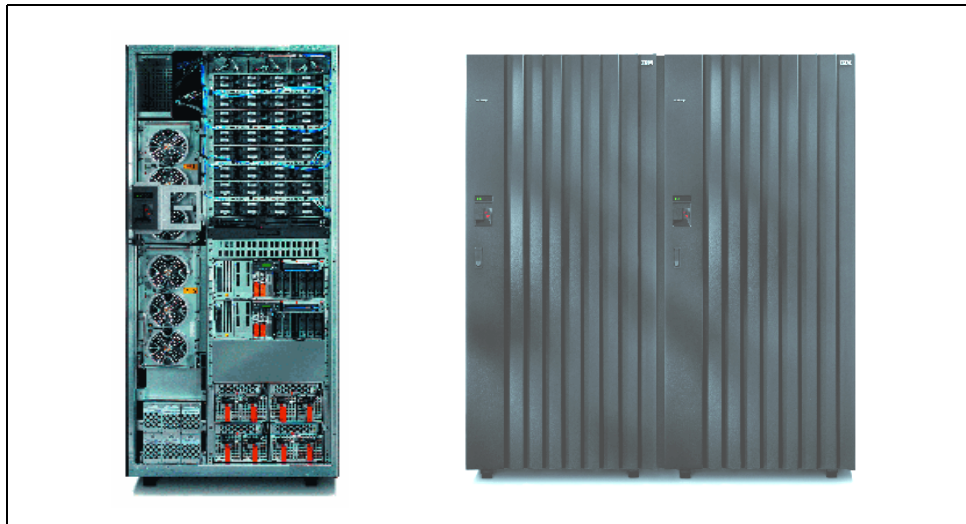


Figure 4-22 DS8100

Note: IBM service representatives can upgrade a Model 921 in the field when you order a model conversion to a Model 922 or Model 9A2.

DS8300 Models 922 and 9A2

The DS8300 Model 922 and 9A2 offer higher capacity and performance than the DS8100.

Model 9A2 is an edition that provides two storage images through two storage system LPARs within the same physical storage unit.

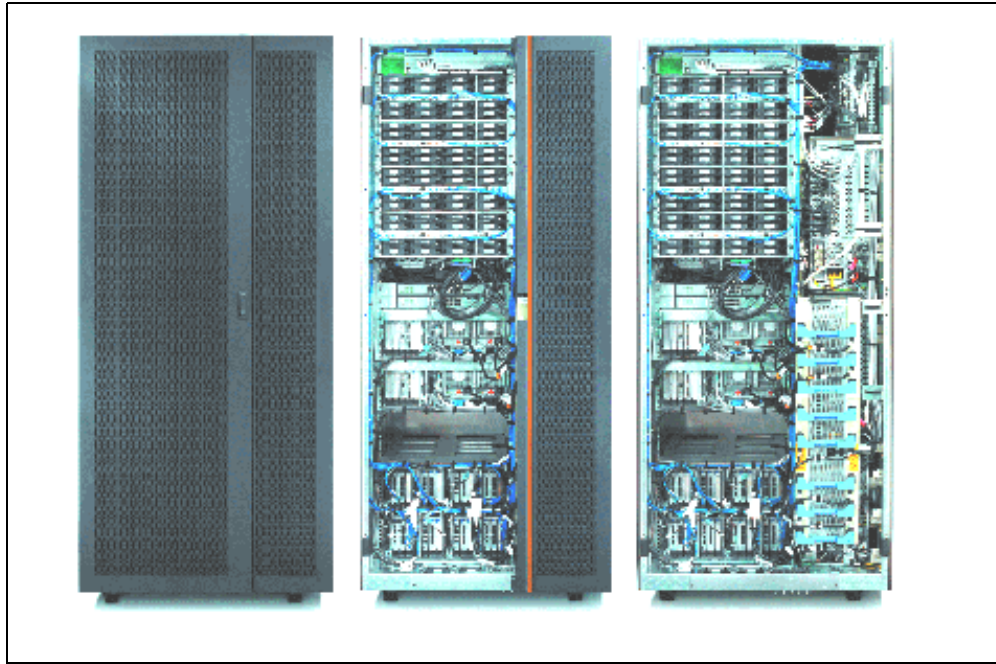


Figure 4-23 Model 922 and 9A2

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

4.4.4 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.

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: ▶ 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)
Maximum Capacity	38.4 TB	115.2 TB	38.4 TB	115.2 TB	192 TB

4.4.5 More information

For more information about the DS8000 series architecture, refer to the redbook *IBM TotalStorage DS8000 Series: Concepts and Architecture*, SG24-6452, which can be found at: <http://www.redbooks.ibm.com/redpieces/pdfs/sg246452.pdf>

4.5 The Enterprise Storage Server Family

Since its initial availability with the ESS Models E10 and E20, and then with the succeeding F10 and F20 models, the ESS has been the storage server solution offering exceptional performance, extraordinary capacity, scalability, 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 back in 1999 when it was first available, and since then it has evolved into the F models and the more recent third-generation ESS Model 800, keeping up with the pace of users’ needs by adding more sophisticated functions to the initial set, enhancing the connectivity options, and powering its performance features.

With the IBM TotalStorage ESS Model 800 important changes have been introduced, with dramatically improved the overall value of the ESS in the marketplace. The ESS Family was expanded in April 2004 with the announcement of the ESS Model 750.

Figure 4-24 provides general descriptions of each model and their design points.

ESS Family of Enterprise Storage Servers

- **ESS Model 750**
 - For moderate capacity needs - supports up to 4.6 TB
 - Designed to meet lower tier customer workload requirements
 - Designed to offer attractive pricing for modest capacity requirements
 - Designed to offer non-disruptive path to ESS Model 800
- **ESS Model 800**
 - Excellent capacity scalability supporting up to 55.9 TB
 - High performance scalability designed to meet/exceed the vast majority of customer workload requirements
 - Wide array of disk capacities and speeds
- **ESS Model 800 with Turbo option**
 - Up to 30% greater throughput than ESS base model 800
 - Designed to meet the needs of very high performance environments
 - Designed to offer non-disruptive Field upgradeability from ESS Model 800 Base

Included in all ESS Models

- | | |
|--|---|
| Designed to avoid single points of Failure/Repair | 20+ autonomic functions |
| Designed to support 24X7 operations | Rich set of optional advanced functions |
| Offers flexibility in efficient management options | Designed to offer low total cost of ownership |
| Wide array of heterogeneous host attachments | |

Figure 4-24 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.

4.6 Storage management

This section provides you information about configuration and management capabilities of the IBM TotalStorage Enterprise Disk Systems.

4.6.1 ESS Specialist

The IBM TotalStorage Enterprise Storage Server Specialist (ESS Specialist) is the Web user interface that is included as standard with the IBM TotalStorage Enterprise Storage Server. You use the ESS Specialist to view machine resources, problem status and machine configuration, and to modify the ESS configuration.



Figure 4-25 ESS Launch Panel

The ESS Specialist also provides the means for invoking ESS Copy Services to establish Peer-to-Peer Remote Copy and FlashCopy without having to involve the operating system running in the host server.

The ESS Specialist can be accessed from the browser provided with the ESS Master Console, which comes with the ESS.

Also using an Internet browser, such as Netscape Navigator or Microsoft Internet Explorer, the storage system administrator can access the ESS Specialist from a desktop or mobile computer as supported by the network.

4.6.2 DS Management Console

The IBM DS6000 and DS8000 series offer new management tools and interfaces which are applicable to both products.

IBM TotalStorage DS Storage Manager GUI

The DS Storage Manager is a Web based graphical user interface (GUI) that is used to perform logical configurations and Copy Services management functions. It can be accessed from any location that has network access using a Web browser.

You have the following options to use the DS Storage Manager:

- ▶ **Simulated (Offline) Configuration**

This application allows the user to create or modify logical configurations when disconnected from the network. After creating the configuration, you can save it and then apply it to a network attached storage unit at a later time

- ▶ **Real-time (Online) Configuration**

This provides real-time management support for logical configuration and Copy Services features for a network attached storage unit.

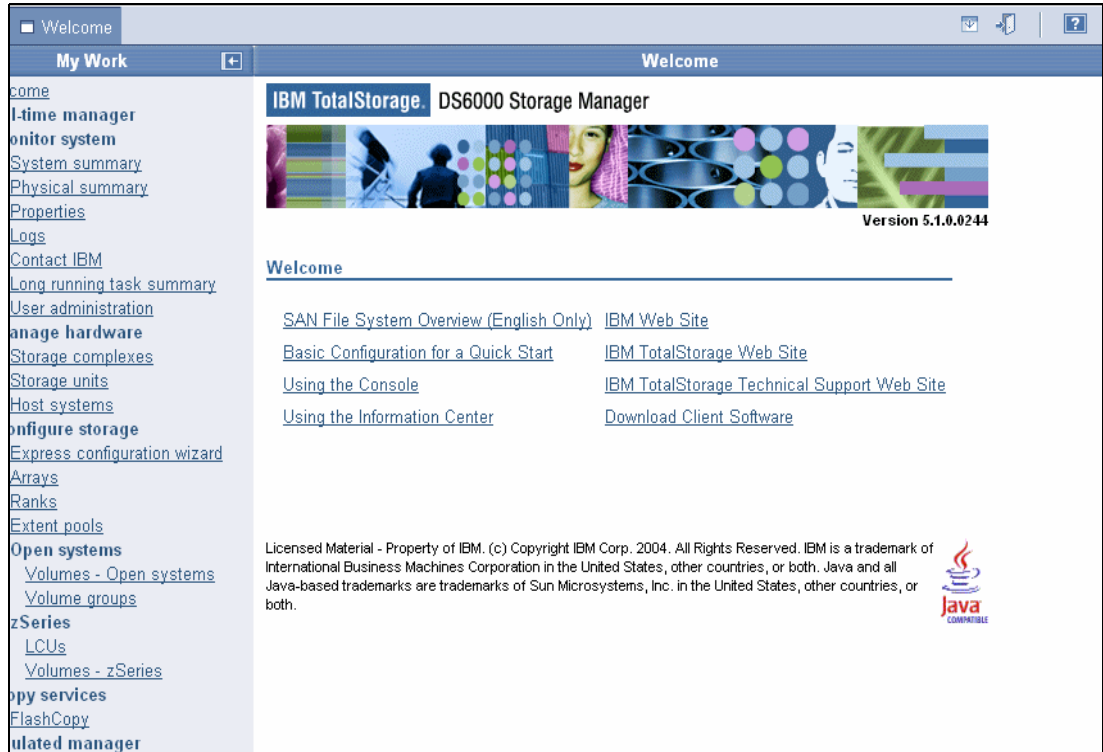


Figure 4-26 DS6000 Storage Manager GUI

The DS8000 Storage Manager contains more diagnostics features than the DS6000 GUI.

IBM TotalStorage DS Command-Line Interface (CLI)

The DS CLI is a single CLI that has the ability to 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, ESS Model 800. It is now possible to combine the DS CLI commands into a script. This can help enhance your productivity since it eliminates the previous (on ESS) requirement for you to create and save a task using the GUI.

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

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.

4.6.3 IBM TotalStorage Expert

The IBM TotalStorage Expert is an optional chargeable product. The IBM TotalStorage ESS Expert gathers performance information and stores it in a relational database. Administrators can then generate reports and make more informed decisions about volume placement and capacity planning. It can significantly help storage administrators manage one or more IBM TotalStorage Enterprise Storage Servers (ESS Expert feature) and Enterprise Tape Libraries (ETL Expert feature). They can also isolate I/O performance to pinpoint bottlenecks, decide where to allocate new space and identify time periods of heavy usage.

For more information about the IBM TotalStorage Expert refer to Chapter 23, “TotalStorage Expert” on page 387.

4.6.4 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 Tivoli Storage Resource Manager
- ▶ IBM Tivoli SAN Manager
- ▶ IBM Tivoli Storage Manager
- ▶ IBM TotalStorage Multiple Device Manager

For more information about the IBM TotalStorage Productivity Center refer to Chapter 15, “IBM TotalStorage Productivity Center” on page 247.

4.7 Copy Services

Advanced Copy Services are enterprise-level, leading-edge functions designed to address an organization’s needs for disaster recovery, data migration, and data duplication.

4.7.1 IBM TotalStorage FlashCopy (Point-in-time Copy feature)

The IBM TotalStorage FlashCopy can help reduce or eliminate planned outages for critical applications. FlashCopy is designed to enable 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 where you need a copy of production data to 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.

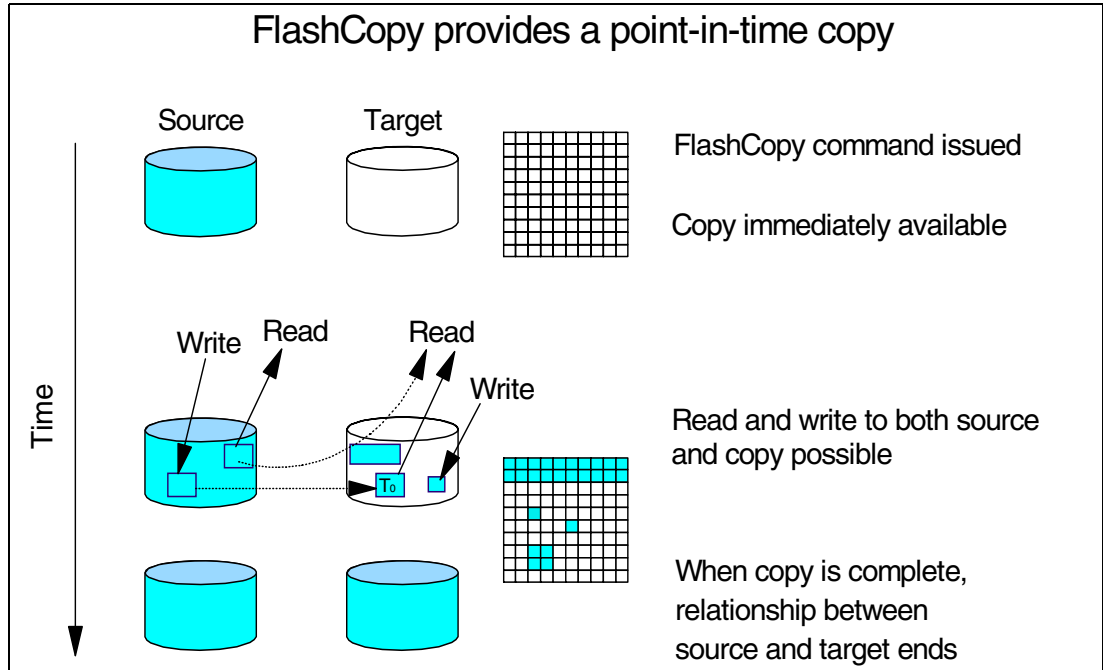


Figure 4-27 FlashCopy

FlashCopy is an additional charged feature. You have to order the Point-in-time Copy feature which includes FlashCopy. The ESS supports FlashCopy V2, which includes all the features of FlashCopy V1 plus enhancements. All functions (with the exception of Metro/Global Copy) are implemented on the DS6000 and the DS8000 series and interoperable with the ESS.

FlashCopy supports many advanced capabilities, including:

► **Data Set FlashCopy**

Data Set FlashCopy allows a FlashCopy of a data set in a zSeries environment.

► **Multiple Relationship FlashCopy**

Multiple Relationship FlashCopy allows a source volume to have multiple targets simultaneously.

► **Incremental FlashCopy**

Incremental FlashCopy provides the capability to update a FlashCopy target without having to recopy the entire volume.

► **Consistency Group function**

Consistency Groups can be 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**

FlashCopy to a Remote Mirror Primary gives you the possibility to use a FlashCopy target volume 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.

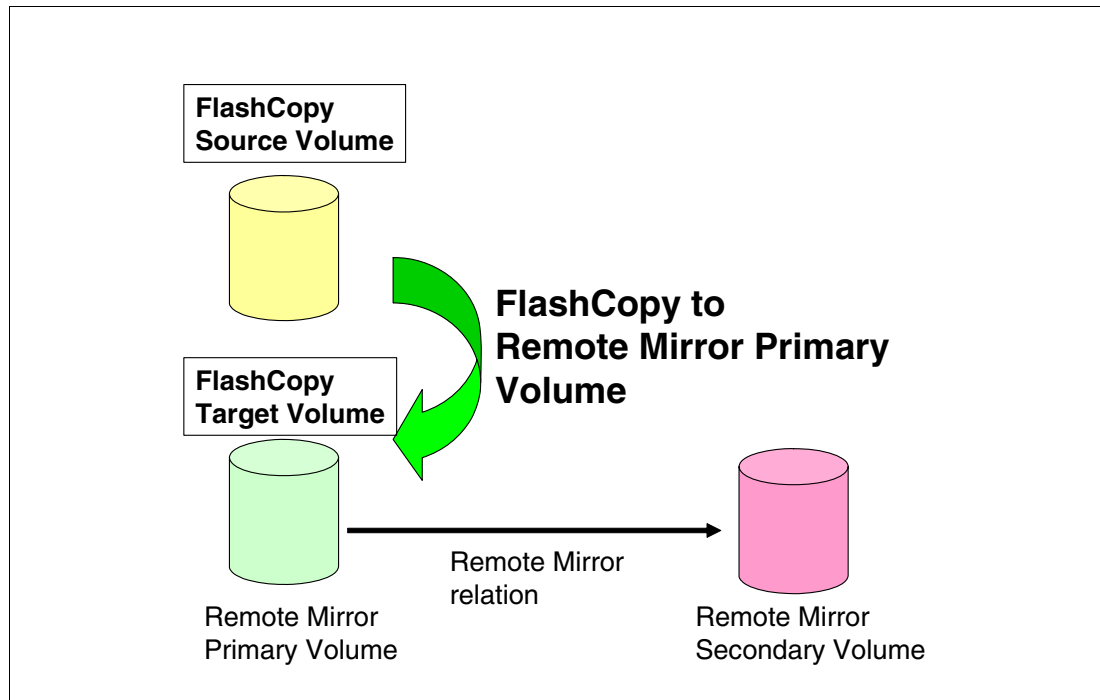


Figure 4-28 Establish FlashCopy on existing Remote Mirror and Copy Primary

4.7.2 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 TotalStorage 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.

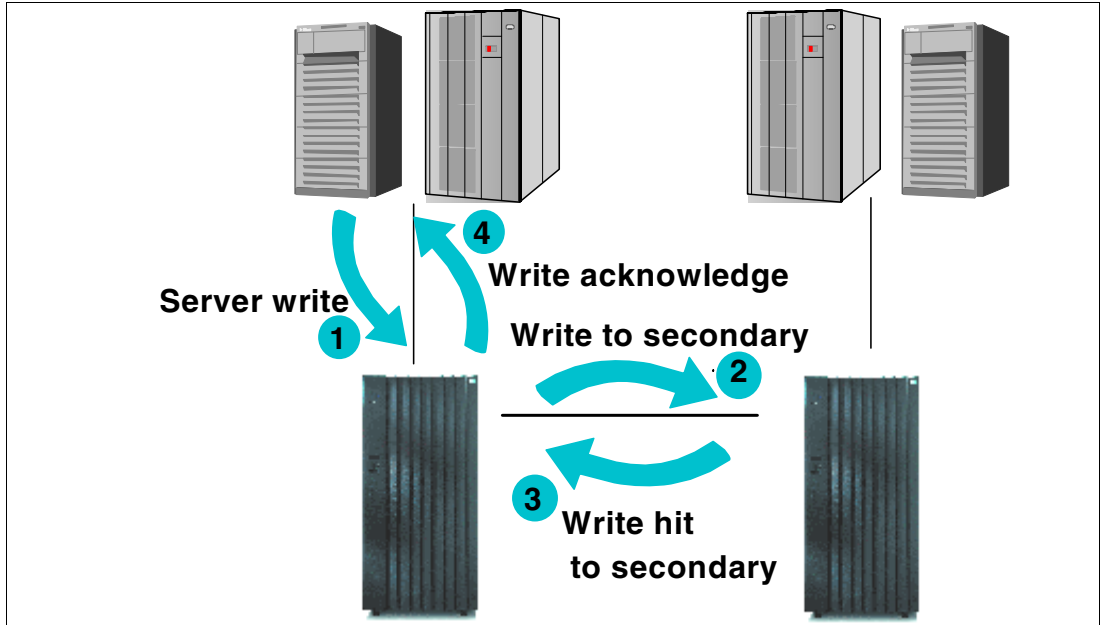


Figure 4-29 Metro Mirror

IBM TotalStorage Global Copy (Asynchronous PPRC-XD)

Global Copy 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.

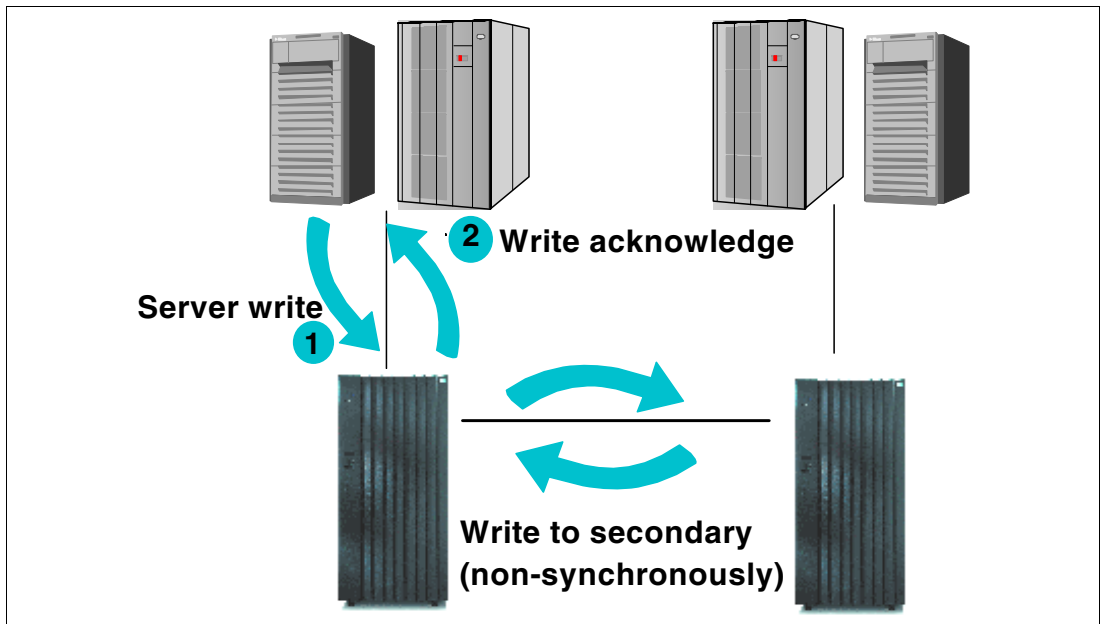


Figure 4-30 Global Copy

IBM TotalStorage Global Mirror (Asynchronous PPRC)

Global Mirror 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.

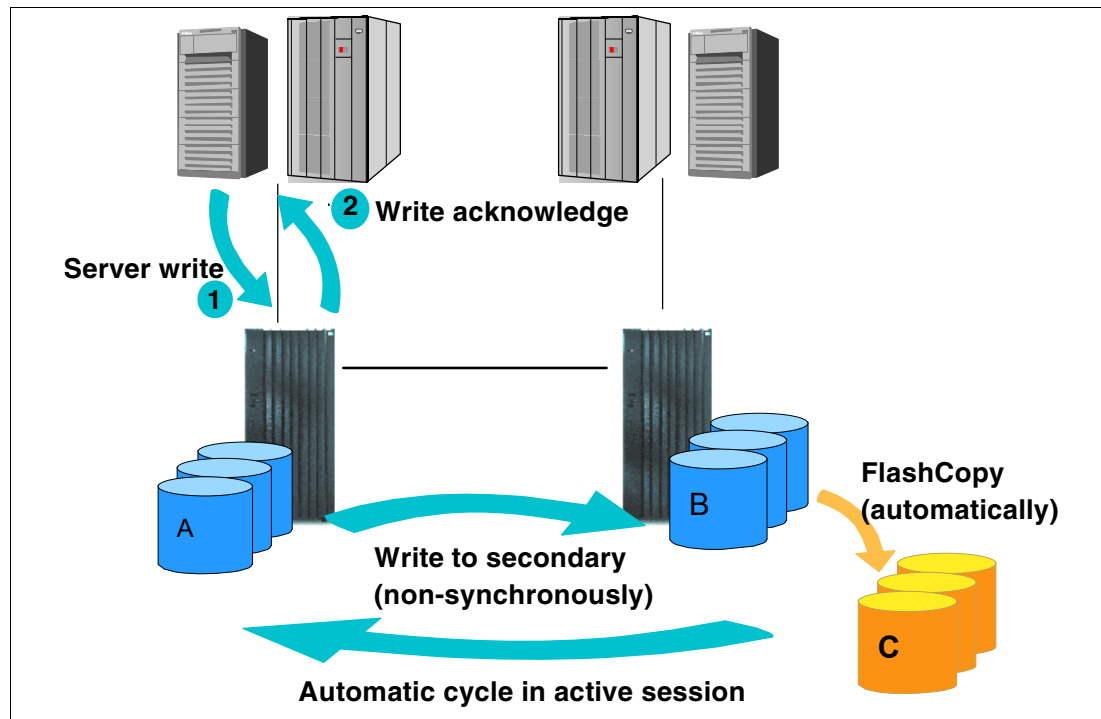


Figure 4-31 Global Mirror

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.

IBM TotalStorage 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.

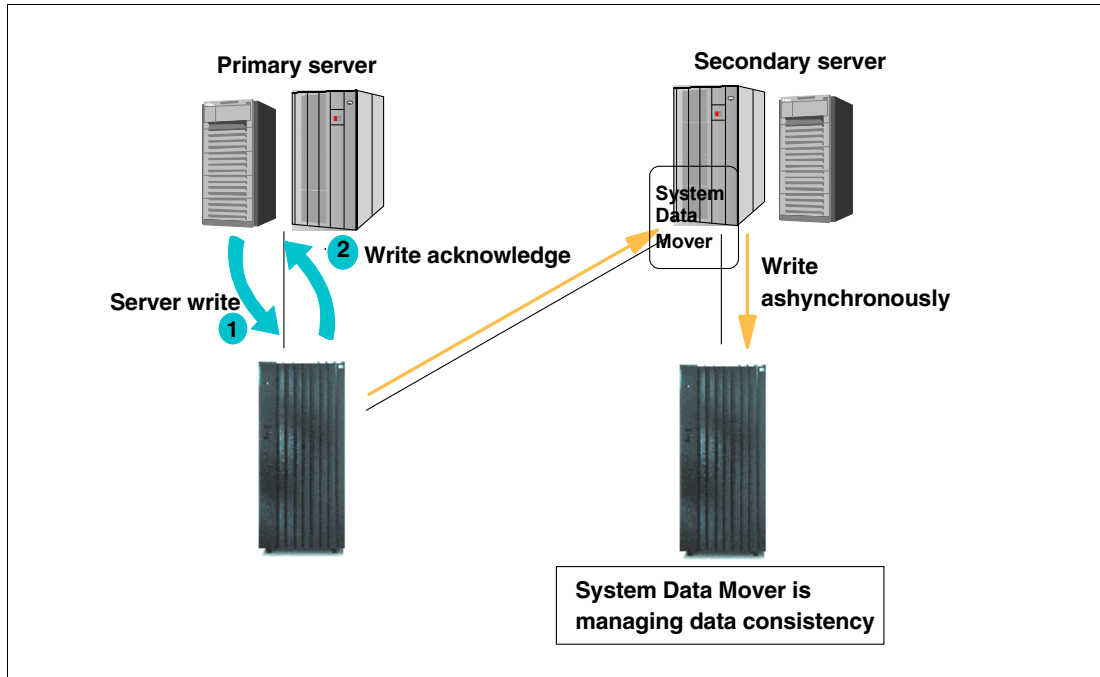


Figure 4-32 z/OS Global Mirror

IBM TotalStorage 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.

Note: z/OS Metro/Global Mirror is not supported on DS6000 series.

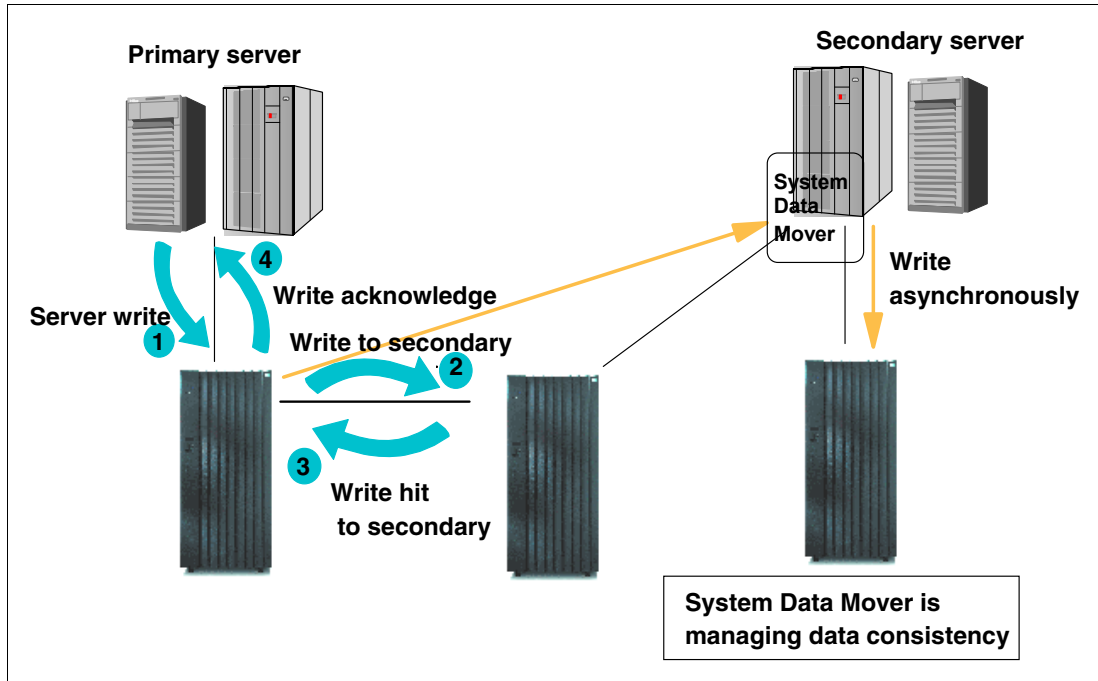


Figure 4-33 z/OS Metro/Global Mirror

IBM TotalStorage Metro/Global Copy (Asynchronous Cascading PPRC)

Metro/Global Copy (previously called Asynchronous Cascading PPRC) enables the creation of three-site configuration by allowing a secondary volume to also serve as a primary volume in another relationship. You first copy your data synchronously to an intermediate site and from there you go asynchronously to a more distant site.

Metro/Global Copy is only available on the ESS Model 800 and Model 750. IBM intends to offer a long distance business continuance solution across three sites allowing recovery from the secondary or tertiary site with full data consistency. The solution may include any mix of DS6000, DS8000 and ESS models.

4.8 Comparison of products

The DS6000 and the DS8000 series are the two new members of the IBM TotalStorage DS Family. In the following we compare the DS6000 and the DS8000 to each other and to the ESS which is the foundation of enterprise storage. We also provide comparison of the DS6000 to Midmarket products.

DS6000 series compared to DS4000 series

Previous DS4000 series (formerly called FASTT) clients will find more of a difference between the DS4000 series and DS6000 series of products. Both product families have about the same size and capacity, but their functions differ. With respect to performance, the DS4000 series range is below the DS6000 series.

The DS6000 series offers enterprise capabilities not found in mid-range offerings. The DS6000 is unique in that it comes from the ESS roots and provides support for zSeries customers as well as iSeries system users.

The implementation of FlashCopy on the DS4000 series is different as compared to the DS6000 series. For example, the target system on a DS4000 series cannot be used for production, it can be used on the DS6000 series.

The DS4000 series also offers remote copy solutions, these functions are not compatible with the DS6000 series.

DS6000 series compared to DS8000 series

You can think of the DS6000 series as the small brother or sister of the DS8000 series. 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 series can deliver a higher throughput and a dramatic leap in performance. It provides high I/O bandwidth and self-healing interconnect with RIO-2 architecture which is only available on the DS8000 series. At the same time, the DS8000 offer extensive virtualization capabilities with the industry's first implementation of storage logical partitioning. Storage system LPARs is not available on the DS6000 series.

DS8000 series compared to ESS

The DS8000 is the next generation of the Enterprise Storage Server (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 point of view, 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.

For the reason that the mirror solutions are compatible between the ESS and the DS8000 series, 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.

DS6000 series compared to ESS

The ESS clients 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.

4.9 Additional information

This section gives complementary information about the support and requirements of the IBM TotalStorage Enterprise Family products.

Licensed Machine Code (LIC)

The IBM TotalStorage 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 TotalStorage DS Manager GUI on DS6000 and DS8000 series.

View machine summary
IBM TotalStorage DS8300

Use this page to verify the functions on your machine. If applicable, take the following actions before viewing your feature activation codes:

- [Assign a function authorization](#) if you received a DS8000 series function authorization reference document, but the function is not shown below.
- [Manage an activation](#) if this is the first time you are activating a particular function or if any unassigned value is larger than 0.0 TB.

If none of the above conditions apply, you can [view your feature activation codes](#).

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 4-34 DS Storage Manager feature activation

Supported environment

The IBM TotalStorage Enterprise Disk Systems can be connected across a broad range of server environment such as:

- ▶ IBM @server zSeries
- ▶ IBM @server iSeries
- ▶ IBM @server pSeries
- ▶ IBM @server xSeries
- ▶ 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.ibm.com/servers/storage/disk/ds8000/pdf/ds8000-matrix.pdf>
- ▶ The DS6000 Interoperability Matrix
<http://www.ibm.com/servers/storage/disk/ds6000/pdf/ds6000-matrix.pdf>
- ▶ The ESS Interoperability Matrix
<http://service.boulder.ibm.com/storage/disk/ess/interop.pdf>



7133 serial disk system

This chapter describes the IBM TotalStorage 7133 serial disk system models D40 and T40.

The IBM TotalStorage 7133 serial disk system is a storage system capable of connecting and sharing storage with multiple UNIX, Microsoft Windows, and Novell NetWare hosts.

5.1 Overview

The IBM TotalStorage 7133 Serial Disk System Advanced Models D40 and T40 provide highly available storage for UNIX, Windows NT, and Novell NetWare servers. Through its implementation of industry-standard Serial Storage Architecture (SSA), the 7133 advanced Models provide outstanding performance, availability, multihost attachment, and economical RAID 5 data-protected storage.

5.2 7133 key features

The 7133 Storage Server can help enterprises to achieve their business objectives in many areas. It provides a disk storage architecture for today and tomorrow, increased throughput for higher performance, high system availability, greater flexibility, and a path for growth, reduced overall cost with flexible characteristics that can be configured according to business requirements.

Serial Storage Architecture

SSA is an open storage interface developed to overcome many of the Small Computer System Interface (SCSI) limitations. SSA technology enables simultaneous communication between multiple devices, subsystems, and local host processors throughout your open systems environment.

While integrating smoothly with existing systems and applications, SSA delivers a significant performance boost over more traditional interfaces. Since its introduction SSA has enabled businesses to expand their environment in a nondisruptive, modular fashion that simplifies storage management and enabled the business to use its computing resources more effectively. In an open environment, SSA is a single I/O interface that can effectively address the storage requirements of a variety of computers - from PCs to supercomputers. Combining the high performance of a direct disk interface with the flexibility of network area attachment, SSA offers a reliable and robust disk technology. Compared to newer interfaces such as FC-AL, SSA offers time-tested reliability and the performance scalability that can only be provided by a non-arbitrated architecture. Unlike Fibre Channel-Arbitrated Loop, multiple I/Os can occur at one time on an SSA loop.

Increased throughput for higher performance

One of the most impressive benefits SSA offers over SCSI is the high 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 MB/s conversations at once (one inbound and one outbound), to enable a total of 160 MB/s 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.

Greater flexibility and a path for growth

Because SSA is an architecture in which performance scales up to match the business requirement, it is ideal for all system sizes and organizations. SSA helps organizations to dynamically add subsystems and grow incrementally, easily, and inexpensively. SSA supports flexible configurations, enabling an assortment of system connection options that can help organizations to improve their system throughput by isolating sub-tasks from production tasks. SSA is also a low-cost storage interface that delivers the high level of performance previously available only in a mainframe environment. The price/performance of SSA gives organizations the freedom to run their applications on a variety of platforms, depending on specific application requirements. The vast improvement in throughput over older SCSI technology is ideally suited for the new generation of large database applications and massively parallel processors that can provide organization with business competitive advantage.

Reduced overall cost

By helping organizations to expand the power of their existing computing resources without requiring additional staffing or tools, SSA provides an affordable alternative to using traditional computing implementations. SSA's overall costs can be much lower than these previous implementations because of reduced maintenance, simplified management, and more efficient use of existing resources.

And even though SSA is an entirely different architecture than SCSI, it still maps the SCSI command set, so existing applications can migrate transparently to SSA-based subsystems. This saves time and cost of rewriting applications while giving those applications an impressive performance boost.



Figure 5-1 7133 Serial Advanced Models D40 and T40

The 7133 Advanced Models 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 MB/s per adapter (therefore, up to 160 MB/s per subsystem).

The 7133 Advanced Models subsystem 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 Advanced Model D40 is designed for integration into an industry standard 19 inch rack, while the 7133 Advanced Model T40 is a free-standing deskside tower unit.

Both 7133 Advanced 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, providing the flexibility to build storage capacity from gigabytes to terabytes. 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 Advanced 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.

5.3 More information

7133 products are supported in the environments shown in the Web links that follow. The sales representative and/or customer is responsible for ensuring that the specific host system configuration used (such as, server model, operating system level, and host adapter combination) is a valid and supported configuration by the server manufacturer:

<http://www.ibm.com/servers/storage/support/disk/7133/index.html>

- ▶ Redbook *Understanding SSA Subsystems in Your Environment*, SG24-5750, which can be found at:

<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg245750.html?Open>



Expandable Storage Plus

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

The IBM 2104 Expandable Storage Plus provides flexible, scalable, and low-cost disk storage in a compact package for RS/6000® and pSeries servers. This disk enclosure is 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 320 models

The IBM 2104 Expandable Storage Plus offers the following:

- ▶ 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 MB/s 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.

The IBM 2104 Storage Plus family consists of the 320 models DS4 or TS4 (Figure 6-1).



Figure 6-1 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 MB/s throughput. The 320 models are designed for high availability.

6.1.1 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.

6.1.2 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 Models DS4/DU3 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 56TB.

The standalone tower Models TS4/TU3 are ideal for one- or two-server environments, providing up to 4TB of capacity in a small desk-side tower unit.

The rack-mounted DS4/DU3 models and the desk-side tower TS4/TU3 models all 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.

6.2 More information

- ▶ For information on the 320 models, refer to:
<http://www-1.ibm.com/servers/storage/disk/expplus>
- ▶ For information on the U3 models and their supported servers, refer to:
<http://www-1.ibm.com/servers/storage/support/disk/2104/installing.html>
- ▶ For the supported adapters, refer to:
<http://www-1.ibm.com/servers/storage/disk/expplus/pdf/interop-matrix.pdf>



Part 2

Tape products

In this part we review IBM TotalStorage offerings for tape drives, tape libraries, and virtual tape solutions in midmarket and enterprise product groups.



IBM TotalStorage LTO products

This chapter provides information about IBM Midmarket tape products. It provides an overview of the Linear Tape-Open (LTO) technology and related Ultrium 1, Ultrium 2 and Ultrium 3 specifications.

The chapter successively covers the following products:

- ▶ **IBM TotalStorage External Tape Drive 3580:** A stand-alone desktop single-drive unit without autoloader. Tapes are mounted manually one at a time.
- ▶ **IBM TotalStorage Tape Autoloader 3581:** 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 Tape Library 3582:** 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 Tape Library 3583:** 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.
- ▶ **IBM TotalStorage Tape Library 3584:** A larger modular library with the potential to house a maximum of 192 tape drives in as many as 16 frames.

We provide information about each model covering their characteristics in terms of capacity, performance, scalability, and reliability.

7.1 Overview

All IBM Midmarket tape products support various platforms including iSeries, pSeries, other UNIX platforms (HPUX, Sun Solaris), Windows NT, Windows 2000, and Windows 2003.

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:** Accelis is a fast-access specification that offers data retrieval in less than ten seconds, with smaller native capacity than Ultrium.
- ▶ **Ultrium:** Ultrium is a high-capacity specification that offers much more capacity than Accelis but a longer access time.

Some technical standards defined for LTO technology are as follows:

- ▶ 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 Ultrium formats:
 - Ultrium format generation 1 allows for 200 GB capacity (2:1 compression) in a single compact cartridge.
 - Ultrium format generation 2 allows for 400 GB capacity (2:1 compression) in a single compact cartridge.
 - Ultrium format generation 3 allows for 800 GB capacity (2:1 compression) in a single compact 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>

7.2 The IBM LTO Ultrium 3 Technology



IBM LTO Ultrium 3 tape drive is the latest in IBM LTO Ultrium Tape Drive Technology available.

7.2.1 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 MB/s 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 new IBM TotalStorage 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(70m). 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.
- ▶ **2Gbps Fibre Channel attachment** — The 3588 Model F3A, a new drive for the 3584 library supporting the LTO3 technology, comes with a 2 Gbps Fibre Channel interface for connection to a wide spectrum of open system servers. They are supported on AIX, OS/400®, i5/OS™, Sun Solaris, HP-UX, Microsoft Windows 2000, Windows 2003, Linux, and other open 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.

7.2.2 Enhanced features

Enhanced features in the IBM LTO Ultrium 3 Tape Drive include:

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's patented Surface Control Guiding Mechanism is designed to guide the tape along the tape path in the IBM 3588 Tape 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 is designed to perform dynamic speed matching (at one of five speeds, 40, 50, 60, 70, 80 MB/s) 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 MB/s versus 17.5 to 35 MB/s 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's 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 MB/s). 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 7-1 shows the compatibility.

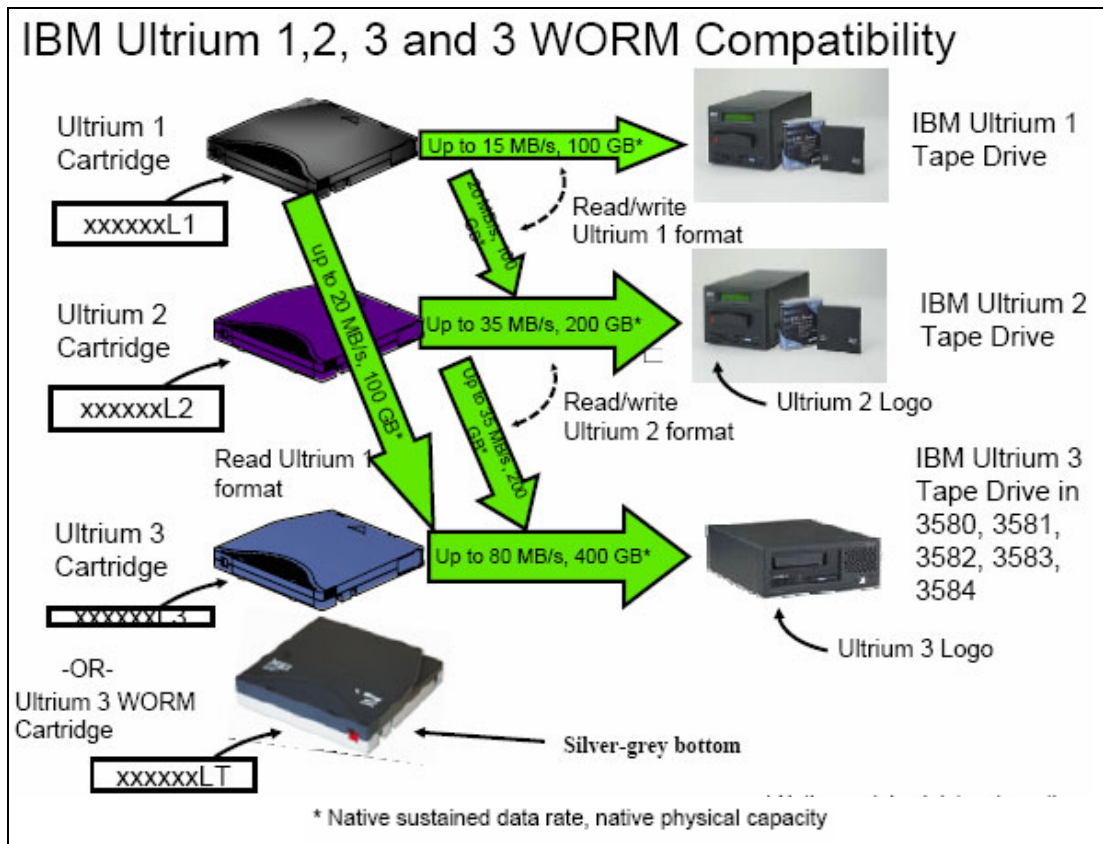


Figure 7-1 Ultrium 1, 2 and 3 compatibility

Note: At the time of writing, IBM was preparing the integration of Ultrium generation 3 WORM tapes into the LTO tape products line.

For more information, refer to this Web site:

<http://www.storage.ibm.com/tape/lto/index.html>

7.3 IBM LTO tape family product overview

The IBM Ultrium product family (Figure 7-2) provides a new level of capacity, performance, scalability, and reliability in the network server and mid-range marketplace. It includes:

- ▶ IBM TotalStorage Ultrium External Tape Drive 3580
- ▶ IBM TotalStorage LTO Ultrium Tape Autoloader 3581
- ▶ IBM TotalStorage LTO Ultrium Tape Library 3582 Model L23
- ▶ IBM TotalStorage LTO Ultrium Tape Library 3583

► IBM TotalStorage LTO Ultrium Tape Library 3584



Figure 7-2 IBM LTO Ultrium family of tapes and libraries

7.3.1 IBM TotalStorage Ultrium External Tape Drive 3580

The IBM TotalStorage 3580 model L33 Tape Drive Figure 7-3 is an external drive incorporating the 3rd and latest generation of IBM LTO technology. 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 Tape Drive provides an excellent migration path from digital linear tape (DLT or SDLT), 1/4-in., 4mm, or 8mm tape drives. The 3580 model L33 can read and write LTO Ultrium 2 Data Cartridges and read LTO Ultrium 1 Data Cartridges.

The IBM TotalStorage 3580 Tape Drive Model L33 features a SCSI Ultra160 Low Voltage



Figure 7-3 The IBM TotalStorage Ultrium External Tape Drive 3580 Model L33

Product description

There are three model types available:

- ▶ L23 model has 200 GB native capacity with a Low-Voltage Differential (LVD) Ultra2 SCSI attachment that connects to LVD fast/wide adapters.
- ▶ H23 model has 200 GB native capacity with a High-Voltage Differential (HVD) Ultra SCSI attachment that connects to HVD fast/wide adapters.
- ▶ L33 model has 400 GB native capacity with a Differential (LVD) Ultra160 SCSI interface that connects to a wide spectrum of open system servers.

Note: There is another model available, 3580 model L3H. It has a new 19-inch rack mount shelf option which accommodates a single 3580 Tape Drive.

The L23 and H23 models use the 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 MB/s (uncompressed).

The L33 model use the 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 MB/s (uncompressed), the L33 model can read and write to Ultrium 2 LTO Cartridge and only read on the Ultrium Cartridge.

The IBM TotalStorage Ultrium External Tape Drive 3580 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, and provides an excellent migration path from DLT, 1/4-inch, 4 mm, or 8 mm tape drives.

The IBM TotalStorage Ultrium External Tape Drive 3580 integrates with tape management solutions such as IBM Tivoli Storage Manager or other third party storage software. With such software, the tape drive can support highly automated, centrally scheduled, policy-managed backup and archive facilities to protect data in a distributed environment.

Note: Although the IBM TotalStorage Ultrium External Tape Drive 3580 provides the capability for high tape performance, other components of the system may limit the actual performance achieved. Also, the actual degree of compression achieved is highly sensitive to the characteristics of the data being compressed.

Ultrium 3 - Model L33

Native device driver support is available for:

- ▶ IBM AIX
- ▶ OS/400 and i5/OS
- ▶ Microsoft Windows 2000 (build 2195 or greater)
- ▶ Windows Server 2003 (build 3790 or greater)
- ▶ Sun Solaris 8 and 9
- ▶ HP-UX 11.0 and 11i
- ▶ Red Hat Enterprise Linux 3
- ▶ SUSE LINUX Enterprise Server 8.

Ultrium 2 - Models L23 and H23

The IBM TotalStorage Ultrium External Tape Drive 3580 is shipped with device drivers to support Ultrium 2 drives in the following operating environments at the minimal levels shown:

- ▶ AIX Versions 4.3.3, 5.1

- ▶ Sun Solaris 7, 8, and 9
- ▶ Windows NT 4.0 with Service Pack 6
- ▶ Windows 2000 (build 2195)
- ▶ Windows 2003
- ▶ HP-UX 11.0, 11i
- ▶ OS/400 V5R1, V5R2
- ▶ Linux Red Hat 7.2 (32 and 64 bit kernels), Red Hat 7.3, Red Hat Advanced Server 2.1, SuSE Linux Enterprise Server Update

Tip: The device driver CD or diskette that is shipped with the IBM TotalStorage Ultrium External Tape Drive 3580 may not contain the device drivers with the most recent level or the device drivers for all supported systems. For the latest device drivers, go to:

<http://ftp.software.ibm.com/storage/devdrv/>

For more information, refer to this Web site:

<http://www.storage.ibm.com/tape/lto/index.html>

For specific information on the IBM TotalStorage Ultrium External Tape Drive 3580, refer to:

<http://www.storage.ibm.com/tape/lto/3580/index.html>

7.3.2 IBM TotalStorage LTO Ultrium Tape Autoloader 3581

The IBM TotalStorage Ultrium Tape Autoloader 3581 is the second product in the IBM LTO Ultrium family (Figure 7-4).

The IBM TotalStorage 3581 Tape Autoloader can provide an excellent solution for businesses looking for high density and performance tape autoloader storage in constrained rack or desktop environments. The 3581 offers a single Ultrium 3 tape drive and storage for up to eight tape cartridges in the new Models L38 and F38, and is available in both Low Voltage Differential (LVD) 160Ultra SCSI and 2 Gbps Fibre channel interface.



Figure 7-4 IBM TotalStorage LTO Ultrium Tape Autoloader 3581

The IBM LTO Ultrium 3 Tape Drives are designed to support up to 80 MBps native data transfer rates. In addition, with the use of the new IBM TotalStorage LTO Ultrium 400 GB Data Cartridge, the IBM Ultrium 3 Tape drive doubles tape cartridge capacity up to 400 GB native physical capacity (800 GB with 2:1 compression), providing an IBM 3581 Autoloader capacity of up to 6.4TB with 2:1 compression or 3.2TB native.

The IBM Ultrium 3 Tape Drives are designed to read and write second-generation LTO Ultrium Data Cartridges and read LTO Ultrium 1 Data Cartridges at original capacities.

The IBM 3581 Tape Autoloader, 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. The Model L28 can also attach to High Volume Differential (HVD) interfaces with an optional HVD converter kit (Feature Code 3104).

Product description

The IBM TotalStorage LTO Ultrium Tape Autoloader 3581 is a high-performance Ultrium tape autoloader. It is available as four separate model types, depending on the capacity and required SCSI2 or fibre interface. The four models are:

- ▶ L28 models have eight cartridge slots, a 1.6 TB native (3.2TB compressed) data capacity with a LVD Ultra160 SCSI (also High-Voltage Differential (HVD) with converter) which can connect to a wide variety of open system servers.
- ▶ F28 models have eight cartridge slots, a 1.6 TB native (3.2TB compressed) data capacity with a native switched 2 GB Fibre Channel which can connect to a wide variety of open system servers.
- ▶ L38 models have eight cartridge slots, a 3.2 TB native (6.4 TB compressed) data capacity with a LVD Ultra160 SCSI which can connect to a wide variety of open system servers.
- ▶ F38 models have eight cartridge slots, a 3.2TB native (6.4TB compressed) data capacity with a native switched 2 GB Fibre Channel which can connect to a wide variety of open system servers.

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 MB/s native (or 70 MB/s 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 MB/s native(or 160 MB/s compressed).

Note: Both Model L38 and F38 are designed to read and write second-generation LTO Ultrium Data cartridges and read LTO Ultrium 1 Data cartridges at original capacities.

The IBM TotalStorage LTO Ultrium Tape Autoloader 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 TotalStorage LTO Ultrium Tape Autoloader 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 on the IBM TotalStorage LTO Ultrium Tape Autoloader 3581, refer to:

<http://www.storage.ibm.com/tape/lto/3581/index.html>

7.3.3 IBM TotalStorage LTO Ultrium Tape Library 3582

The IBM TotalStorage LTO Tape Library 3582 is shown in Figure 7-5. is a high-performance, reliable, scalable tape subsystem. Designed for tape automation, it can be attached to iSeries, pSeries, xSeries, Intel (running Windows or Linux), Sun SPARC, Hewlett-Packard, and other open systems using SCSI or Fibre Channel attachment. Each aspect of the library subsystem has been optimized for repeated, reliable unattended tape handling.



Figure 7-5 IBM TotalStorage LTO Ultrium Tape Library 3582

The IBM TotalStorage 3582 Model L23 provides support for Ultrium 3. It supports 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 TotalStorage LTO Ultrium 400 GB Data Cartridge.

The Ultrium 3 tape drive can read and write second-generation LTO Ultrium data cartridges and read original LTO Ultrium data cartridges at original Ultrium 1 and 2 capacities with improved Ultrium 1 performance.

Product description

The compact 4U IBM 3582 Tape Library is designed with a variety of advanced features. The IBM 3582 Tape Library 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 9.6TB native (19.2TB 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 on the IBM TotalStorage LTO Ultrium 3 Tape Library 3582, refer to:

<http://www.storage.ibm.com/tape/lto/3582/index.html>

7.3.4 IBM TotalStorage LTO Ultrium Tape Library 3583

The fourth member of the IBM Ultrium family is the IBM TotalStorage LTO Ultrium Scalable Tape Library 3583, shown in Figure 7-6. It is a high-performance, reliable and scalable tape subsystem. Designed for tape automation, it can be attached to iSeries, pSeries, xSeries, 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 tape

drives for fast data transfer and reliability in automated library service. Each aspect of the library subsystem has been optimized for repeated, reliable unattended tape handling.



Figure 7-6 IBM 3583 Ultrium Scalable Library

Product description

The IBM TotalStorage 3583 Tape Library now supports the IBM LTO Ultrium 3 Tape Drive. The IBM Ultrium 3 drive has more than double the drive performance of the second-generation LTO Ultrium Tape Drives. The IBM Ultrium 3 Tape Drives are designed to support up to 80 MBps native data transfer rates. In addition, with the use of the 400 GB Data Cartridge, the IBM Ultrium 3 Tape drive doubles tape cartridge physical capacity up to 400 GB native capacity (800 GB with 2:1 compression) over the Ultrium 2 data cartridges.

The IBM Ultrium 3 Tape Drives come in Low Voltage Differential (LVD) Ultra160 SCSI and 2 Gbps Fibre Channel to attach to a variety of open system servers. The IBM TotalStorage Ultrium 3 Tape Drives can read and write second-generation LTO Ultrium data cartridges and read original LTO Ultrium data cartridges at original Ultrium 1 and 2 capacities with improved Ultrium 1 performance.

The Ultrium 3 tape drives and cartridges can be resident in the same IBM 3583 library with first and second-generation Ultrium tape drives and cartridges. The IBM 3583 Tape Library is designed with a variety of advanced features.

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 Tape Library 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 7.2TB,

The IBM TotalStorage Tape Library 3583 comes in three different models as seen in Table 7-1 on page 120. The major difference between the models is the number of storage cells shipped with the initial order:

- ▶ Model L18 is supplied with space for 18 cartridges
- ▶ Model L36 is supplied with space for 36 cartridges
- ▶ Model L72 is supplied with space for 72 cartridges

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 7-1 IBM TotalStorage LTO Ultrium Scalable Tape Library 3583 model summary

Models	Cartridge Slots	Data Capacity (Native)	Data Capacity (Comp.)	IBM Ultrium Tape Drives
L18	18	3.6TB	7.2TB	1-6
L36	36	7.2TB	14.4TB	1-6
L72	72	14.4TB	28.8TB	1-6

Interface

Each IBM Ultrium tape drive contains the electronics and logic for reading and writing of data, control of the tape drive, management of the data buffer, and error recovery procedures. All tape drives are packaged as a common assembly that is a Field Replaceable Unit (FRU), designed for quick removal and replacement. The cartridge capacities are unaffected by the number of drives installed.

Cartridge storage

As well as the installed tape drives, the library enclosure contains cartridge storage slots that are arranged in columns (from one to three depending on the library model). Each of the three storage columns has provision for an additional fixed slot located at the top of each column. This slot is reserved for future use. The tape cartridges are stored in removable magazines, which are installed into the columns in the library (three magazines to each column). 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 reader

A barcode reader is provided as standard with all models, and it does not affect the slot capacity of the libraries. The barcode reader is used during the inventory process to locate all cartridges inserted in the library. This action is repeated every time the front door is opened to ensure that the inventory is updated, should a cartridge have been manually added, moved, or removed while the door was open.

I/O station

This facility allows the insertion and ejection of cartridges without interrupting the normal operation of the library. There are two types of I/O station:

- ▶ The single-cartridge station, where the cartridge can be inserted or ejected by opening the I/O station door
- ▶ The 12-cartridge station, where 12 cartridges are contained in two removable magazines that can be inserted or ejected by opening the I/O station door

Models L18 and L36 have the single-cartridge I/O station for tape cartridges and can be upgraded to a 12-cartridge station. Model L72 has the 12-cartridge station installed as standard equipment.

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 barcode reader 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
- ▶ A picker assembly for mounting the cartridge picker and the bar code scanner
- ▶ Y-axis and Z-axis drive motors for rotating the picker assembly, and moving it vertically, inside the library enclosure

Multi-Path support

The patented Multipath feature of the 3583 Tape Library supports sharing of the library robotics. This is accomplished by partitioning the library into up to three multiple logical libraries, and providing each logical library its own separate and distinct drives, storage slots, and control paths.

Note: An 18-slot storage column needs to be dedicated to each logical partition. I/O slots are shared on a first-come-first-served basis. This type of partitioning allows heterogeneous applications to share the library robotics independent of each other. Cartridges under library control are not shared between logical libraries, nor allowed to be moved between logical libraries. An example of heterogeneous sharing is a Microsoft Windows 2000 application using the drive and storage slots of one logical library while a UNIX application uses the drive and slots of another logical library.

Logical libraries can also be used for separating Ultrium 3 Tape Drives and cartridges from Ultrium 2 Tape Drives and cartridges, for applications that do not support mixing the drives in the same logical library.

Control Path Failover

For enhanced availability you may use Control Path Failover. This optional feature provides automatic control path failover to a preconfigured redundant control path in the event of a loss of a host adapter or control path drive, without aborting the current job in progress. Support is provided under AIX, LINUX, and Solaris for both SCSI and Fibre Channel attachments when the IBM tape device driver is used.

Data Path Failover and Load Balancing

Data Path Failover and Load Balancing exclusively supports native Fibre Channel Ultrium 3 or Ultrium 2 Tape Drives in the 3583 Tape Library using the IBM tape device driver for AIX, LINUX, and Solaris (Load Balancing is not supported for Solaris). Data path failover is designed to provide a failover mechanism in the IBM device driver, which enables you to configure multiple redundant paths in a SAN environment. In the event of a path or component failure, the failover mechanism is designed to automatically provide error recovery to retry the current operation using an alternate, preconfigured path without aborting the current job in progress.

When accessing a tape drive device that has been configured with alternate pathing across multiple host ports, the IBM device driver automatically selects a path through the host bus adapter (HBA) that has the fewest open tape devices, and assigns that path to the application. This autonomic self-optimizing capability is called Load Balancing. The dynamic load balancing support is designed to optimize resources for devices that have physical connections to multiple HBAs in the same machine. The device driver is designed to dynamically track the usage on each HBA as applications open and close devices, and balance the number of applications using each HBA in the machine. This may help optimize HBA resources and improve overall performance. Further, Data Path Failover provides

autonomic self-healing capabilities similar to Control Path Failover, with transparent failover to an alternate data path in the event of a failure in the primary host-side path. Data Path Failover and Load Balancing for Ultrium 3 Tape Drives requires the Control Path Failover feature.

Upgrades and optional features

You can add features to each model type in the IBM TotalStorage LTO Ultrium Scalable Tape Library 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.

Table 7-2 IBM TotalStorage LTO Ultrium Scalable Tape Library 3583 capacity upgrade

Model	Extra tape storage columns	Single cartridge I/O station	12 cartridge I/O station defined as storage	12 cartridge I/O station defined as I/O
L18	0	18 storage 1 I/O	36 storage 0 I/O	24 storage 12 I/O
L18	1	36 storage 1 I/O	54 storage 0 I/O	42 storage 12 I/O
L18	2	54 storage 1 I/O	72 storage 0 I/O	60 storage 12 I/O
L36	0	36 storage 1 I/O	54 storage 0 I/O	42 storage 12 I/O
L36	1	54 storage 1 I/O	72 storage 0 I/O	60 storage 12 I/O
L72	Not available	Not available	72 storage 0 I/O	60 storage 12 I/O

Ultrium Device Drivers

The IBM TotalStorage LTO Ultrium Tape Library 3583 is shipped with the device drivers to support Ultrium 1; 2 and 3 drives in most of the operating environments listed below:

- ▶ AIX
- ▶ OS/400
- ▶ i5/OS
- ▶ Windows 2000
- ▶ Windows Server 2003
- ▶ Sun Solaris
- ▶ HP-UX
- ▶ Red Hat
- ▶ SUSE LINUX

Note: For more information on the operating systems supported, refer to this Web site:

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

RMU with TotalStorage Specialist

The Remote Management Unit (RMU) provides remote access to the IBM TotalStorage LTO Ultrium Scalable Tape Library 3583 over a network. It can be attached to the network through a 10/100 Ethernet port on the RMU. All available functions are accessible without the need of

a dedicated server or separate software. Any server on the network can access the library through its IP address or IP name in a Web browser.

For more information, refer to this Web site:

<http://www.storage.ibm.com/tape/lto/index.html>

For specific information on the IBM TotalStorage LTO Ultrium Scalable Tape Library 3583, refer to:

<http://www.storage.ibm.com/tape/lto/3583/index.html>

7.3.5 IBM TotalStorage Ultrium Tape Library 3584

IBM TotalStorage LTO Ultrium Tape Library 3584 shown in Figure 7-7 on page 123, is the largest member of the IBM Ultrium family of tape library storage.



Figure 7-7 IBM TotalStorage LTO Ultrium Tape Library 3584

It is an excellent choice for:

- ▶ Rapid growth in online storage requirements
- ▶ Automated backup using tape autoloaders and tape libraries with software for automatic backup, archive, or fast-access tape operation to accommodate growth and reduce manual operations
- ▶ Using standardized IBM LTO Ultrium format tape
- ▶ Large cartridge capacity and fast data streaming transfer capability requirements
- ▶ Cross-platform connectivity for networked storage environments
- ▶ High data availability is achieved with hot swap drives, redundant drive power supplies and *call home* feature for improved service and support
- ▶ Mixed media support enables ease of migration from DLT to LTO technology
- ▶ High data integrity and reliability

- ▶ Consolidating tape workloads from multiple heterogeneous open systems servers.
- ▶ Flexibility as it supports the 3592 tape drives which have WORM support.

The IBM TotalStorage 3584 Tape Library is a stand-alone device that provides reliable, automated tape handling and storage for unattended mid-range systems and network servers in an Open Systems environment. Scalability for the UltraScalable Tape Library is a maximum of 16 frames and 192 drives.

You can order the 3584 Tape Library with the new **IBM TotalStorage 3588 Tape Drive Model F3A**. The drive, known as the Ultrium 3 Tape Drive, features a native capacity of 400 GB when used with the IBM TotalStorage LTO Ultrium 400 GB Data Cartridge (800 GB at 2:1 compression). It has a native data rate of 80 MB/s (160 MB/s at 2:1 compression). A library of Model L52 and D52 frames can contain up to 6887 cartridges, with a library capacity of 2754 TB. A library of Model L32 and D32 frames can contain up to 6881 cartridges, with a library capacity of 2752 TB (5504 TB at 2:1 compression).

In addition to being available with the Ultrium 3 Tape Drive, the 3584 Tape Library comes with several tape drive and frame options to meet your needs. Drives and frames include the:

- ▶ **Ultrium 2 Tape Drive**, which features a native capacity of 200 GB with the IBM TotalStorage LTO Ultrium 200 GB Data Cartridge (400 GB at 2:1 compression). This drive also has a native data rate of 35 MB/s (70 MB/s at 2:1 compression). A library of Model L52 and D52 frames can contain up to 6887 cartridges, with a library capacity of 1377 TB. A library of Model L32 and D32 frames can contain up to 6881 cartridges, with a library capacity of 1376 TB (2752 TB at 2:1 compression).
- ▶ **Ultrium 1 Tape Drive**, which features a native capacity of 100 GB with the IBM LTO Ultrium Data Cartridge (200 GB at 2:1 compression). This drive has a native data rate of 15 MB/s (30 MB/s at 2:1 compression). A library of Model L52 and D52 frames can contain up to 6887 cartridges; a library of Model L32 and D32 frames can contain up to 6881 cartridges. The capacity for each type of library is 688 TB (1376 TB at 2:1 compression).
- ▶ **IBM TotalStorage 3592 Tape Drive Model J1A**, which features a native capacity of 300 GB with the IBM TotalStorage 3592 Enterprise Tape Cartridge (900 GB at 3:1 compression). This drive also has a native data rate of 40 MB/s (80 MB/s at 2:1 compression). A library of Model L22 and D22 frames can contain up to 6260 cartridges, with a library capacity of 1878 TB (5634 TB at 3:1 compression).

Structure of the 3584 Tape Library

This section defines the base and expansion frames, and describes the models of the 3584 Tape Library.

The basic 3584 Tape Library is a single storage unit known as the base frame (Model L22 for 3592 Tape Drives, or Models L32 or L52 for LTO Ultrium Tape Drives). The library's scalability allows you to increase capacity 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 Ultrium Tape Drives or 3592 Tape Drives, but may not contain a mix of both.

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 TotalStorage 3584 Tape Library Introduction and Planning Guide.

The 3584 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 D22 or Model D52 frame.

The models of the 3584 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:

- ▶ **Model 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** 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.) shorter in depth than Model D32.
- ▶ **Model 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.) longer in depth than Model L52
- ▶ **Model 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.) longer in depth than Model D52.
- ▶ **Model 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.) shorter in depth than Model L32.
- ▶ **Model 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.) shorter in depth than Model D32.
- ▶ **Model HA1** Required when you order the optional second accessor, a service bay frame that contains no tape drives or cartridge capacity. Model HA1 contains only slots for diagnostic cartridges. This frame is always configured as service bay A. The Model HA1 is approximately 307 mm (12 in.) shorter in depth than Models L32 and 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 means 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, currently available only for AIX 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. For example, consider a simple multi-path architecture connection consisting of two HBAs in an AIX 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. This simple configuration provides two physical control paths to the library for redundancy if one path from an HBA to the library fails.

Reliability

The IBM TotalStorage LTO Ultrium Ultrascapable Tape Library 3584 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 for AIX.
- ▶ 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.

Dual accessors and service bays in the 3584 Tape Library

This section explains that when an optional second accessor is installed, the 3584 Tape Library features enhanced availability through its ability to use that accessor and to operate without disruption when any component of the working accessor fails.

Note: To use dual accessors and service bays, the 3584 Tape Library must be installed with the Advanced Library Management System (ALMS).

If the library is installed with the optional second accessor, cartridge mount performance is also optimized. (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 new D22 or D52 frame as service bay B. If you order a second accessor you must also order a 3584 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.

Advanced Library Management System (ALMS)

This section gives an overview of the Advanced Library Management System (ALMS), which virtualizes the locations of cartridges in the 3584 Tape Library. Logical libraries can then consist of unique drives and ranges of volume serial numbers instead of fixed locations.

ALMS is an extension of IBM's patented Multi-Path Architecture. With ALMS, the 3584 Tape Library 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 3584 Tape Library 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, L32, or L52 storage capacity without taking the library offline
- ▶ Cartridge assignment policy

Tape drives supported by the 3584 Tape Library

This section gives information about the types of drives that can be installed in the 3584 Tape Library.

The LTO Ultrium Tape Drives and the 3592 Tape Drive are high-performance, high-capacity data-storage units that can be installed in the 3584 Tape Library. Up to 12 drives may be installed in each frame of the library, but the two types of drives may not be mixed in the same frame.

LTO Ultrium tape drives

You can identify the Ultrium 2 or Ultrium 3 Tape Drive by the logo at the front of the drive and the label at the rear of the drive's canister. You can identify the Ultrium 1 Tape Drive by the label at the rear of its canister.

Ultrium Tape Drives do not read or write to 3592 Tape Cartridges, and 3592 Tape Drives do not read or write to Ultrium Tape Cartridges. The 3584 Tape Library supports Ultrium 3 Tape Drives and Ultrium 3 tape cartridges in newly purchased L52 and D52 frames, as well as in installed L32, D32, L52, and D52 frames.

The Ultrium Tape Drives can read tapes that have been written by non-IBM Ultrium drives. They also write to tapes that can be read by non-IBM Ultrium drives.

Ultrium 1, Ultrium 2, and Ultrium 3 Tape Drives and cartridges can reside in the same frame. When a cartridge is labeled according to proper IBM bar code label specifications, the last character of its volume serial (VOLSER) number indicates the generation of the media. For example, a cartridge with a VOLSER of 000764L3 is an Ultrium 3 cartridge; a cartridge with a VOLSER of 003995L2 is an Ultrium 2 cartridge.

To enhance library performance, Ultrium 3 and Ultrium 2 Tape Drives include speed matching, channel calibration, and power management. Speed matching dynamically adjusts the drive's native (uncompressed) data rate to the slower data rate of a server. Channel calibration customizes each read/write data channel for optimum performance. The customization enables compensation for variations in the recording channel transfer function, media characteristics, and read/write head characteristics. Power management reduces the drive's power consumption during idle power periods.

3592 tape drives

The 3592 tape drive has a native data rate of up to 40 MB/s. With the use of the IBM TotalStorage 3592 Enterprise Tape Cartridge, it provides a native cartridge capacity of up to 300 GB (900 GB at 3:1 compression). It uses an optimal dynamic compression method called byte level compression scheme swapping, which is designed to achieve maximum data compression. Unlike other tape drive compression methods, it is also designed to prevent data expansion. The 3592 Tape Drive is built for automation and uses a tape cartridge that is similar in size to the 3590 or 3480 tape cartridges.

The 3592 Tape Drive has dual-ported, 2 Gbps native switch fabric Fibre Channel interfaces. This offers attachment flexibility in an Open Systems environment. The drives can be directly attached to Open Systems servers with Fibre Channel attachments. The 3592 Tape Drive uses the Statistical Analysis and Recording System to assist in isolating failures between media and hardware. It uses the cartridge performance history (saved in the cartridge) and the drive performance history (kept in the drive) to determine the cause of the failure. This arrangement causes the drive to mark the media as degraded, and to indicate that the hardware has degraded.

Tape cartridges supported by the 3584 tape library

This section gives information about the tape cartridges that you can use in the 3584 Tape Library.

In the 3584 Tape Library, frames that are installed with Ultrium Tape Drives use Ultrium Tape Cartridges; frames that are installed with 3592 Tape Drives use 3592 Tape Cartridges. A frame cannot house both Ultrium Tape Drives and 3592 Tape Drives. However, in a library that includes both types of frames, you may insert 3592 Tape Cartridges into the lower I/O station of a Model L52 or L32 frame for transport (by the cartridge accessor) to a Model D22 frame (your library must contain an I/O station that will accept 3592 Tape Cartridges). Similarly, you may insert Ultrium Tape Cartridges into the lower I/O station of a Model L22 frame for transport (by the cartridge accessor) to a Model D52 or D32 frame (again, your library must contain an I/O station that will accept Ultrium Tape Cartridges).

The capacity for supported cartridges is:

- ▶ IBM TotalStorage LTO Ultrium 400 GB Data Cartridge (native capacity of 400 GB; compressed capacity of 800 GB at 2:1 compression).
- ▶ IBM TotalStorage LTO Ultrium 200 GB Data Cartridge (native capacity of 200 GB; compressed capacity of 400 GB at 2:1 compression).
- ▶ IBM LTO Ultrium Data Cartridge (native capacity of 100 GB; compressed capacity of 200 GB at 2:1 compression)
- ▶ IBM TotalStorage 3592 Enterprise Tape Cartridge (native capacity of 60 or 300 GB; compressed capacity of up to 900 GB at 3:1 compression)

For more information, refer to this Web site:

<http://www.storage.ibm.com/tape/lto/index.html>

For specific information on the IBM TotalStorage LTO Ultrium Scalable Tape Library 3583, refer to:

<http://www.storage.ibm.com/tape/lto/3584/index.html>

7.4 More information

For more information on LTO technology and IBM mid-range tape products family, refer to the following Web sites:

- ▶ <http://www.storage.ibm.com/tape/lto/index.html>
- ▶ <http://www.storage.ibm.com/tape/lto/3580/index.html>
- ▶ <http://www.storage.ibm.com/tape/lto/3581/index.html>
- ▶ <http://www.storage.ibm.com/tape/lto/3582/index.html>
- ▶ <http://www.storage.ibm.com/tape/lto/3583/index.html>
- ▶ <http://www.storage.ibm.com/tape/lto/3584/index.html>
- ▶ <http://www.lto-technology.com/>

For in-depth information about IBM mid-range tape products family, refer to these Redbooks and documentation:

- ▶ *Using IBM LTO Ultrium with Open Systems*, SG24-6502:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246502.html?Open>
- ▶ *Implementing IBM LTO in Linux and Windows*, SG24-6268:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246268.html?Open>

- ▶ *The IBM LTO Ultrium Tape Libraries Guide*, SG24-5946:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg245946.html?Open>



IBM Enterprise Tape

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 3590 Enterprise Tape Drive and controller
- ▶ IBM 3592 Enterprise Tape Drive and controller
- ▶ IBM 3494 Enterprise Tape Library
- ▶ IBM 3494 Virtual Tape Server

8.1 Overview

All IBM Enterprise Tape products are characterized by their ESCON and FICON interface connection to zSeries and S/390®.

Note: At the time of writing this book, IBM just announced the IBM TotalStorage 3953 Tape Frame and Library manager, which enables zSeries and VTS connectivity for the 3584 tape library running 3592 drives.

Table 8-1 depicts the comparative features of IBM Enterprise Tape products.

Table 8-1 Comparative features of IBM Enterprise Tape products

	3592 Enterprise Tape System	3590 Enterprise Tape System	3494 Library	3494-VTS
Description	1/2" Enterprise Tape Systems	1/2" Enterprise Tape Systems	1/2" Enterprise Tape Library	Virtual Tape Server
Machine Model	3592 J1A - Library	3590 B1A/E1A/H1A - Library B11/E11/H11- Rack	3494 (frames) L12, L14, L22 Library D12, D14, D22 Library (Frame) S10-Storage HA1-High Availability	3494 B10, B20
Product strengths	Multipurpose drive, Capacity, Performance, Write Once Read Many (WORM) support	Capacity, performance, reliability	Modular, reliable automation for high volume storage	Virtual tape, fast access, peer-to-peer copy
Technology	Longitudinal Serpentine	Longitudinal Serpentine		Disk: Emulated 3490E Tape: 3590 1/2"
# Heads/tracks	8/512	B: 16/128 E: 16/256 H: 16/384	Drive Dependent	N/A
Number of drives	1	1	1-32 SCSI/FC 1-76 ESCON/ FICON or SCSI/Fiber mix	64-256 (virtual)
Max number of cartridges	N/A	10 per autoloader	6,240 (max #of cartridges decrease as tape/control units are added)	500,000 (virtual volumes)
Cartridge capacity native/compressed	300/900 GB	B: 10/30 GB BX: 20/60 GB E: 20/60 GB EX: 40/120 GB H: 30/90 GB HX: 60/180 GB	Drive Dependent	400/1200 MB 800/2400 MB (depends on virtual cartridge configuration)

	3592 Enterprise Tape System	3590 Enterprise Tape System	3494 Library	3494-VTS
Max system capacity native/compressed	900 GB	B11: 300/600 GB E11: 600/1200 GB H11: 900/1800 GB	H: 561 TB HX: 1,122 TB J1A: 5,616 TB	Disk: up to 5.2 TB Tape: 748 TB
Max drive data rate/second native/compressed	40/120 MB	B: 9/27 MB E: 14/42 MB H: 14/42 MB	Drive Dependent	VTS: up to 280 MB/s
Time to data	Cartridge Dependent	Model Dependent	Drive Dependent	VTS: 1-3 second if data in VTS Cache
Interface	FC, ESCON, FICON, 2 Gb FICON	SCSI-2 F/W Diff PCI, Ultra SCSI FC, ESCON, FICON, 2 Gb FICON	SCSI-2 F/W Diff PCI, Ultra SCSI FC, ESCON, FICON, 2 Gb FICON	Ultra SCSI, ESCON, FICON
SAN-ready	pSeries, iSeries, RS/6000, Sun, HP, Windows NT, Linux, Windows 2000	pSeries, iSeries, RS/6000, Sun, HP, Windows NT, Linux, Windows 2000, Tru64 UNIX	pSeries, iSeries, RS/6000, Sun, HP, Windows NT, Linux, Windows 2000, Tru64 UNIX	pSeries, RS/6000, Sun, Windows NT, Windows 2000
Supported platform	iSeries, pSeries, zSeries, S/390, Linux, Windows NT, Windows 2000, Windows 2003 server, HP, Sun	iSeries, AS/400, pSeries, RS/6000, zSeries, S/390, Tru64 UNIX, Linux, Windows NT, Windows 2000, Windows 2003 server, HP, Sun	iSeries, AS/400, pSeries, RS/6000, zSeries, S/390, Linux, Windows NT, Windows 2000, Windows 2003 server, HP, Sun	pSeries, RS/6000, Windows NT, Windows 2000, Windows 2003 server, Sun, HP

8.2 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 and tape controller 3592 model J70.

8.2.1 IBM 3590 Enterprise Tape Drive

The IBM 3590 Enterprise Tape Drive (Figure 8-1 on page 134) is designed for high data integrity by using sophisticated error correction codes and factory-written servo tracks for

precise head positioning. The IBM 3590 Enterprise Tape Drive uses a bidirectional longitudinal serpentine recording technique and a fourth-generation magneto-resistive head that reads and writes 16 data tracks at a time. The drives use IBMLZ1 hardware compression to give a greater compression ratio than the IDRC compression technique used in 3490 drives.



Figure 8-1 IBM TotalStorage Enterprise Tape Drive 3590 in a rack

The 3590 Tape Drives are attachable to a small computer system interface (SCSI) host interface or a Fibre Channel interface on selected IBM and non-IBM mid-range and open systems. The tape drives can also attach to Enterprise Systems CONnection (ESCON) and Fibre CONnection (FICON) channels on S/390 or zSeries systems through an IBM TotalStorage Enterprise Tape Controller 3590 model A60.

With Fibre Channel interfaces, the IBM Enterprise Tape Drive 3590 can be directly attached at greater distances and provide additional performance. Uniquely, the 3590 can be attached both to open systems host servers using a native Fibre Channel interface, or it can be attached to a S/390 host using ESCON links or the newer FICON links. FICON links are based on Fibre Channel transport layers with a different I/O protocol instead of SCSI in the top layers.

The IBM Enterprise Tape Drive 3590 has a choice of native Fibre Channel attachment or UltraSCSI attachment. Each drive has two interfaces, both of the same type although only one interface is ever active at any moment. This provides additional availability if paths to the drive are configured correctly. If the drives are attached to a 3590 model A60 for S/390 attachment or are attached to AIX, failover to the second path is automatic.

Predictive failure analysis 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.

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.

The IBM Enterprise Tape Drive 3590 H11 is frame-mounted or rack-mounted and incorporate a 10-cartridge Automated Cartridge Facility (ACF) autoloader for high-capacity (up to 1.2 TB for Model H11s) unattended operation.

The IBM Enterprise Tape Drive 3590 H1A is designed to be used in an IBM Enterprise Tape 3494 Library, IBM Enterprise Tape 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. An IBM Enterprise Tape Drive 3590 model H tape drive provides an uncompressed data rate of up to 14 MB/s, which is the same rate as a 3590 model E drive. With 3590 H models, a total of 384 tracks are written across the tape, increasing the cartridge capacity compared to E models by 50%. 3590 H models may read 128-track, 256-track or 384-track cartridges and write only 384-track cartridges. There is an increase in mount times due to the requirement for the head to lock into the servo tracks. The 3590 H model is only supported by the 3590 A60 controller.

These are some product highlights of the IBM TotalStorage Enterprise Tape System 3590:

- ▶ Attaches to IBM @server iSeries (AS/400), pSeries (RS/6000), xSeries, zSeries (S/390), Hewlett-Packard, Sun Microsystems, Intel-based systems running Linux, Windows NT, Windows 2000 and more.
- ▶ 3590 is supported for native attachment to Ultra SCSI, Fibre Channel, ESCON and FICON servers.
- ▶ Features an uncompressed drive data rate of up to 14 MB/second.
- ▶ Uses 3590 extended length tape cartridge for up to 180 GB of 3:1 compressed data per cartridge.
- ▶ 3590 tape drive technology includes advanced Error Correction Code (ECC), servo track mechanisms, resident diagnostics, redundant parity checking, and read after write verification.

Note: The 3590 has one additional feature which differs from other tape drives. The Alternate Pathing Support allows Automatic Failover and Dynamic Load Balancing.

8.2.2 IBM 3592 Enterprise Tape Drive

The IBM TotalStorage Enterprise Tape Drive, 3592 model J1A, is the first offering in the new 3592 tape drive family designed for enterprise applications. IBM 3592 is designed to provide higher levels of performance, reliability, and cartridge capacity than the IBM TotalStorage Enterprise Tape Drive 3590, which it will ultimately replace. The IBM 3592 represents the next leap in technology in the same way that 3590 replaced the 3490 drive family.

The IBM 3592 family maintains the investment in IBM 3590 automation products because it uses the same shaped cartridge as the 3590 and 3490 product lines and therefore can be used in libraries designed for those products. The new cartridges use a new media label to describe the cartridge type, (the “JA” media type).

Additionally, the 3592 J1A tape drive offers Write Once, Read Many (WORM) functionality that is designed to help support data retention needs and applications requiring an audit trail.

Important: The data format on the 3592 cartridge is unlike any of the previous IBM tape products and so the new drive cannot read or write the previous generation 3590 tapes. Because the IBM 3592 cannot process 3590 cartridges, migration is supported within a shared automation environment.

The IBM 3592 Model J1A is designed to meet the needs of tape customers across a broad range of computing environments, including selected IBM @server zSeries, S/390, iSeries, AS/400, pSeries, RS/6000, xSeries, Sun, and Hewlett Packard servers, as well as Intel-compatible servers running Microsoft Windows and Linux.

3592 Model J1A tape drive

Like the IBM 3590 tape drive, the IBM 3592 model J1A is a modular unit, designed to fit in a library frame or a stand-alone rack. The drive module is 198mm (7.8 in) wide, 467mm (18.4 in) deep, and 95mm (3.8 in) in height. It weighs 5.7 kg (12 lb 9 oz). Though the drive technology and mechanics are new, the basic mechanisms and logic in the new drive are built on the proven high reliability IBM 3590 and IBM 3490 designs. Figure 8-2 shows the IBM 3592 in comparison with the IBM 3590 enclosure, demonstrating the potential benefits of the smaller scale drive.



Figure 8-2 3592 model compared to 3590 model

Enhancements incorporated into the IBM 3592 model J1A tape drive include:

- ▶ Up to 40 MB/s native data rate
- ▶ Up to 300 GB native cartridge capacity (900 GB at 3:1 compression)

Note: It uses an optimal dynamic compression method called *byte level compression scheme swapping*, which is designed to achieve maximum data compression, and unlike other tape drive compression methods, it is designed to prevent data expansion.

- ▶ Dual ported switched fabric 2 Gbps Fibre Channel attachments
- ▶ High reliability and availability design
- ▶ Additional performance and access improvements
- ▶ Smaller form factor that allows more tape drives in a frames and racks
- ▶ Support for capacity scaling, enabling very fast locate and read times
- ▶ 3592 J1A tape drive offers Write Once, Read Many (WORM) functionality that is designed to help support data retention needs and applications requiring an audit trail.

Additional enhancements over the 3590 drives to improve availability, performance, and capacity include:

- ▶ Large 128 MB internal data buffer.
- ▶ Redundant power supplies
- ▶ Digital speed matching to adjust the drives native data rate to the net host data rate (after data compression been factored out) throughput performance
- ▶ Channel calibration which is designed to allow for customization of each read/write data channel for optimum performance
- ▶ High resolution tape directory plus enhanced search speed to improve nominal and average access times
- ▶ Streaming Lossless Data Compression (SLDC) algorithm which is designed to offer an improvement over previous IBM lossless compression algorithms

The IBM 3592 drive provides as standard two switched fabric 2 Gbps Fibre Channel attachments, for attachment to multiple servers or a single server with redundancy.

Note: The IBM 3592 drive will auto-negotiate to match the port to which it is attached. It will attempt to connect at 2 Gbps (FC-2) and will then negotiate down to 1 Gbps (FC-1) if the system/switch it is connected to cannot support 2 Gb. Note that there is no SCSI attachment option.

The 3592 J1A tape drive can also be used in the 3584 Library; see 7.3.5, “IBM TotalStorage Ultrium Tape Library 3584” on page 123.

8.2.3 IBM TotalStorage Tape Controller 3590 Model A60

Both ESCON and FICON attachments are supported on the 3590 Model A60 Controller, enabling 3590 Tape Drives to be shared between ESCON and FICON servers.

The Model A60 Controller includes up to four dual-ported ESCON channel attachment features for a maximum capacity of eight ESCON channels. Up to two FICON long-wavelength or shortwave length attachment features attach to one or two FICON channels. The FICON maximum channel link speed of 2 Gbps full duplex (versus 17 MB/s simplex for ESCON) helps enhance tape drive performance.

The 3590 Model A60 Controller supports high-speed location to relative file position. This capability can significantly improve read performance for filestacking applications that access by file number rather than block ID. The Model A60 also offers significantly improved write performance for applications that stack multiple small files on a tape. The performance gain is achieved through the elimination of costly redundant, rewritten tape marks and the associated back hitch. These performance capabilities of the hardware can be exploited by z/OS and are designed to be transparent to the host application.

This controller is designed to provide:

- ▶ Native FICON Attachment for more distance, performance and reduced tape storage infrastructure
- ▶ ESCON Attachments for increased flexibility
- ▶ Drive sharing and tape storage consolidation for reduced solution costs
- ▶ Full media capacity utilization when 3490E emulation mode is used
- ▶ More overall performance for reduced infrastructure costs and enhanced data protection

- ▶ Support for the IBM TotalStorage SAN Switch Model F16

These are some of tape controller 3590 model A60 features:

- ▶ Provides 1 or 2 Native 1 Gbps or 2 Gbps FICON server connections
- ▶ Provides up to 8 ESCON attachments
- ▶ Up to 12 Fibre Channel attached 3590 tape drives or 8 Ultra SCSI attached drives to a single A60 Tape Controller
- ▶ Multiple S/390 and zSeries connectivity
- ▶ Drive sharing between ESCON and FICON-attached servers

8.2.4 IBM TotalStorage Tape Controller 3592 Model J70

The Enterprise Tape Controller 3592 model J70 as seen in Figure 8-3 is the follow-on to the highly successful Enterprise Tape Controller 3590 model A60 and is designed for environments with ESCON or FICON attached servers where customers depend on the mainframe-class reliability, performance, and capacity that the IBM 3592 model J70 provides.

The IBM 3592-J70 controller is designed to be installed in the same frames as the current IBM 3590-A60 controller. These include IBM 3494 frames, the IBM 3590 model C10, the 3590 model A14, and stand-alone racks.

As well as supporting the new IBM 3592-J1A drives the IBM 3592-J70 will attach IBM 3590 tape drives. It is a suitable replacement for users of IBM 3590-A00 and A50 controllers that need the additional performance and attachment capabilities that the IBM 3592-J70 controller provides.

Enhancements incorporated into the IBM 3592 model J70 tape controller include:

- ▶ Support for up to four FICON attachments, (twice the number supported by the IBM 3590 model A60 controller)
- ▶ Support for up to eight ESCON attachments
- ▶ Support for an intermix of ESCON and FICON attachments
- ▶ Support for up to 16 3592 Model J1A or 12 3590 Tape Drives with a new 2 Gb Fibre Channel Switch in either single or redundant configurations.
- ▶ Direct connection of up to two 3592 Model J1A Tape Drives to the 3592 Model J70 Controller.
- ▶ Up to 1.5 times the throughput of the IBM 3590-A60, with FICON attachment using the IBM 3592-J1A tape drive
- ▶ Support for capacity scaling with the IBM 3592 model J1A tape drive
- ▶ Support for IBM 3592 drive hot swap capabilities

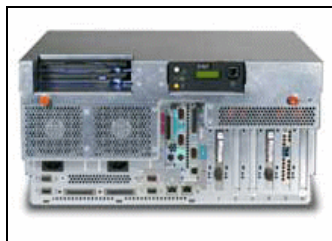


Figure 8-3 IBM 3592 model J70 tape controller

8.2.5 IBM 3590 cartridges

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.

8.2.6 IBM 3592 cartridges

The IBM TotalStorage Enterprise Tape Cartridge 3592 is an advanced metal particle tape specifically optimized for the enterprise tape environment, providing a native cartridge capacity of 300 GB (or up to 900 GB with 3:1 compression), which is five times greater than the maximum capacity of the 3590 H Model. *This can be very beneficial in space savings and economy of data storage since it can help lower the cost of storage per megabyte.* Cartridge memory is built into every data cartridge to enhance the functionality and reliability of the media. The 3592 tape cartridge is designed for automation, with a form factor similar to the 3590 tape cartridges that allows it to be used in the IBM TotalStorage Enterprise Tape Library 3494 and StorageTek Automated Cartridge System (ACS) solutions.

Note:3592 tape cartridges are not compatible with 3590 tape drives and, likewise, 3590 tapes cannot be used in the 3592 drives.

Economy cartridges with a native cartridge capacity of up to 60 GB(or up to 180 GB with 3:1 compression) are available. In addition, WORM cartridges in both sizes are available. WORM data cartridges provide non-alterable, non-rewritable tape media for long-term records retention.

Capacity Scaling

The 3592 Model J1A is designed to support capacity scaling of individual tape cartridges to 60 GB. The effect of capacity scaling to 60 GB capacity is to reduce the average locate time of a random record (from load point) to 30% of the normal locate time. The 3592 Model J1A tape drive allows an application to issue a command to scale an individual cartridge to 60 GB. When data is written to a scaled tape, on subsequent readings, the tape and drive will yield very fast locate and read times.

8.3 IBM 3494 Enterprise Tape Library

The IBM TotalStorage 3494 tape library Figure 8-4 on page 140 is an excellent solution for large storage requirements. The 3494 tape library consists of individual frame units for modular expansion that provides a wide range of configurations. This flexibility enables organizations to start small and grow in an affordable and incremental manner.



Figure 8-4 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 3494 Enterprise Tape Library attaches to multiple systems, including zSeries and S/390 enterprise servers, iSeries, selected systems from Sun Microsystems, Hewlett-Packard, and Microsoft Windows 2000 and many other non-IBM platforms.

The main components of the 3494 Enterprise Tape Library are:

- ▶ Tape Library Base Frame
- ▶ Tape Drive Expansion Frame
- ▶ Tape Storage Frame
- ▶ High Availability Frames
- ▶ Virtual Tape Server

8.3.1 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 Enterprise Virtual Tape Server, and a high-availability model to create a maximum configuration of 16 frames and two service bays.

Tape library base frame is a central and important component of the Enterprise Tape Library subsystem. Each Tape Library Base Frame contains:

- ▶ Tape subsystem (3592 Tape Drives or 3590 Tape Drives or IBM 3490E 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 available:

- ▶ L10 - contains zero to two IBM 3490E tape drives (SCSI, ESCON, or Parallel attached)
- ▶ L12 - contains zero to two 3590 Tape Drives (SCSI or FC attached)
- ▶ L14 - contains zero to two 3590 Tape Drives (ESCON attached)
- ▶ L22 - contains zero to six 3592 Tape Drives (FC attached)

One control unit frame is required in each Enterprise Tape Library.

8.3.2 Tape Drive Expansion Frame

Tape Drive Expansion Frame is designed to provide additional tape drives and tape cartridge storage. There are three models of the drive frame to give you flexibility in choosing the correct tape drive technology to fit your server platform.

Each Tape Drive Expansion Frame contains:

- ▶ Zero to twelve drives per frame depending on the Drive Frame model
- ▶ Up to 400 cartridge storage cells
- ▶ Up to fifteen drive units can be attached to a single Lxx control unit

The five Tape Drive Expansion Frames D10, D12, D14, D22 and D24 differ in cartridge storage cell capacity, tape subsystem supported, and host platform attachment capability:

- ▶ D10 is designed to attach from zero up to two IBM 3490E tape drives and will store from 270 to 400 tape cartridges
- ▶ D12 is designed to attach up to six native Fibre Channel SCSI attached 3590 tape drives and will store from 230 up to 400 tape cartridges
- ▶ D14 is designed to attach up to four ESCON/FICON attached 3590 Tape Drives and will store from 275 to 400 tape cartridges
- ▶ D22 is designed to attach up to twelve native Fibre Channel SCSI attached 3592 tape drives and will store 250 to 400 tape cartridges
- ▶ D24 is designed to attach up to eight ESCON/FICON attached 3592 Tape Drives and will 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 Enterprise Tape Drive 3590-E1A
- ▶ IBM Enterprise Tape Drive 3590-H1A

8.3.3 Tape Storage Frame

A 3494-S10 Tape Storage Frame is designed to provide additional tape cartridge storage capacity. Each storage unit allows for the attachment of up to 400 additional tape cartridges. The cartridges can be a mix of 3490E, 3590 or 3592 tape cartridges.

Up to fifteen storage frames can attach to any Lxx control unit frame.

8.3.4 High Availability Frame

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.

It 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.

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.

IBM 3494 Library Manager

The library manager, the operational focal point of the IBM 3494, provides the support to set up, maintain, configure, and operate the 3494 Tape Library. It consists of a controller, a personal computer (PC), a display, a keyboard, and the library manager application. This section describes the major functions and features of the IBM 3494 Library Manager.

The library manager graphical user interface (GUI) is provided by the library manager application, which runs on the library manager PC and is delivered as Licensed Internal Code

(LIC). As such, backup, restoration, and maintenance of the application are the sole responsibility of an IBM Service Representative.

The library manager provides three user interfaces:

- ▶ **Library Manager Console:** Is installed on the rear side of the IBM 3494 Model L12 frame and allows to monitor the 3494 Tape Library and to take actions
- ▶ **IBM TotalStorage Enterprise Tape Library (ETL) Specialist:** A Web browser interface that allows you to monitor the 3494 Tape Library and to take specific actions for an installed VTS or Peer-to- Peer (PtP) VTS.
- ▶ **IBM TotalStorage Master Console:** Integrates service monitoring of the 3494 library.
- ▶ **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.

8.4 3494 Virtual Tape Server

The IBM TotalStorage Virtual Tape Server (VTS) is an enterprise tape solution designed to enhance performance and provide the capacity required for today's backup requirements. The adoption of this solution can help reduce batch processing time, or the total cost of ownership and management cost. The VTS initially creates a virtual volume in a buffer known as the Tape Volume Cache (TVC), a RAID-5 disk array. If the host needs to reference these virtual volumes again, they are accessed in most instances from the TVC, helping to eliminate many of the physical delays associated with tape I/O, and improving the performance of the tape process. The virtual volume is also written to an attached IBM TotalStorage Enterprise tape drive as a logical volume. A copy of the virtual volume remains in the cache, which may provide fast access to critical data.

A VTS Peer-to-Peer (PtP) configuration can provide redundancy for greater disaster tolerance with features like remote dual copy for use with Geographically Dispersed Parallel Sysplex™ (GDPS®), disaster backup and recovery, and remote tape vault. Advanced policy management allows selective dual copy for noncritical data. The Virtual Tape Server and the new configurations can be installed at extended distances. These enhancements have also addressed planned outages such as maintenance and upgrades, as well as unplanned outages by eliminating single points of failure. The data availability improvement this brings further enhances the Virtual Tape Server as a Business Continuation solution.

All IBM Virtual Tape Servers (VTS) are connected to a 3494 tape library. They appear to the host as a 3494 with their own library name, their own set of tape volumes (each VTS can have up to 500,000 standard 400 MB volumes or enhanced capacity 800 MB volumes per 3494) and their own set of tape drives.

The host sees 32, 64, 128 or 256 virtual 3490E drives, attached by up to:

- ▶ Sixteen Enterprise System Connection (ESCON) paths
- ▶ Eight SCSI paths that transfer data concurrently

Up to eight FICON channels are also supported up to a maximum of eight on a model B20 and four on a VTS model B10. However these channel types cannot be intermixed with SCSI.

The VTS can be shared between S/390 and Open Systems by specifying up to four additional SCSI Host Attachment features. Each feature provides two SCSI buses and includes data

compression compatible with the EHPO. Up to 64 virtual drives can be assigned to Open Systems.

Currently two VTS subsystems can be integrated into a 3494 sharing the maximum number of 500,000 logical volumes equally and providing up to 256 virtual drives. The VTSs can coexist effectively in the same IBM Enterprise 3494 Tape Library with native 3490E, native 3590 and native 3592 IBM Enterprise tape cartridges and drives.

With access to a virtual library with 500,000 emulated cartridges and up to 256 emulated drives, the host can process much more tape workload in parallel without the expense of actually installing 256 real tape drives. With so many virtual tape drives, there is the opportunity to dedicate drives to different hosts and platforms rather than managing an often complex shared environment.

The VTS is managed by the Library Manager in the IBM 3494, so that the host sees a logical library, separate from any other drives within the IBM 3494, with its own library name.

A host system processes the data and sends it to the Virtual Tape Server. Once the data has been transferred to the Virtual Tape Server, the host processor is finished, until there is a further request for information.



Figure 8-5 3494 Virtual Tape Server

The Virtual Tape Server stores the data received from the host, as a virtual volume, in cache on hard disk drives (called virtual drives), not on tape. As more data is received from any attached system, the VTS continues to store this data as virtual volumes on the disk cache, known as the tape volume cache. When a cache occupancy threshold is reached, the Virtual Tape Server transfers multiple virtual volumes to an IBM Enterprise 3590 tape drive as well as 3592 tape drives depends which are currently available.

Note: The 3494 VTS Model B10 and B20 can now utilize the increased capacity of the recently announced IBM TotalStorage Enterprise Tape Drive 3592 Model J1A.

When a virtual volume is written to the physical tape, it is then called a logical volume. Figure 8-6 show all these processes.

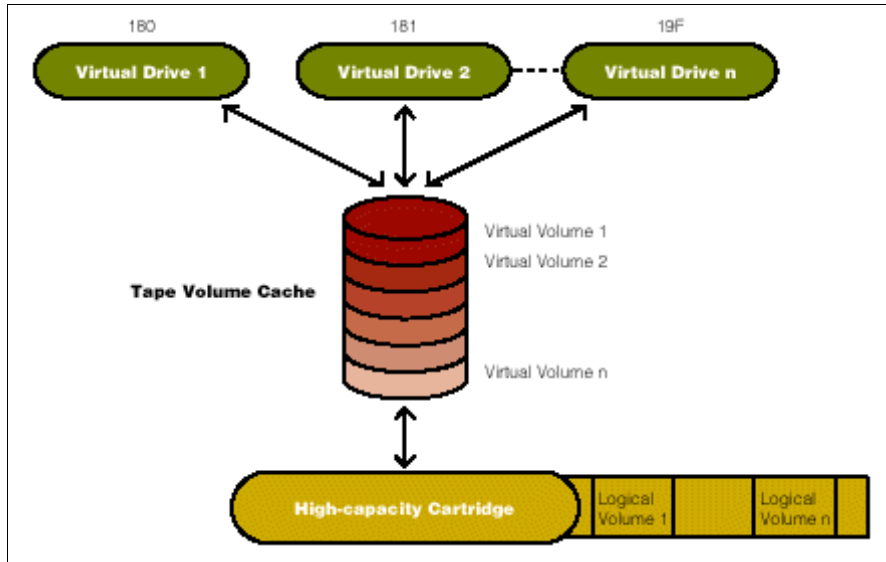


Figure 8-6 VTS Virtual Drives and Hard disk caching

Each tape cartridge in the library is filled, greatly reducing the number of cartridges needed. Today, many cartridges stored in traditional tape automation solutions contain one file per tape and are, on average, only 15% utilized. By fully utilizing cartridge capacity, the Virtual Tape Server can greatly reduce tape hardware, floor space and personnel expenses.

Most data retrievals in a tape library occur within 24 hours of when that data was created. When a retrieval request is made by a host server, the virtual volume is usually still resident in the tape volume cache.

Up to 500,000 logical volumes per VTS and an uncompressed tape storage capacity from 20 TB to 240 TB per library provides the potential to replace multiple traditional libraries and automate the data on over six thousand tape cartridges. With a VTS, reduced floor space, power requirements and maintenance costs compared to non-virtualized solutions may be realized.

The VTS Tape Volume Cache (TVC) is designed to eliminate many of the physical delays associated with tape read/write activities. With up to 256 virtual devices, most of the current bottlenecks associated with the physical movement of data to tape can be eliminated. A VTS runs its own, fully automated storage management subsystem to make sure that the tape data and VTS cache is optimally managed at a minimum of cost.

The VTS is also designed to enhance your current level of data protection and disaster tolerance through either the introduction of an automation device or the installation of a remote VTS through the ability to eliminate tape cartridge transportation and management.

A configuration consisting of two VTSs and coupled by peer-to-peer controllers automatically duplicates your chosen tape data to better support disaster tolerance and Business Continuity. The coupled VTSs can be installed in two sites up to thousands of kilometers apart using channel extender technology. This capability can help minimize your critical system outages in case of a disaster.

Some benefits of implementing 3494-VTS are:

- ▶ Reduce operating costs, such as power, maintenance, and operations and support staff
- ▶ Reduce floor space consumed by tape operations
- ▶ Improve operational quality

- ▶ Increase the level of tape automation
- ▶ Improve tape data availability
- ▶ Reduce the backup window
- ▶ Provide fast access to data
- ▶ Share data between diverse servers
- ▶ Provide remote dual copy for use with Geographically Dispersed Parallel Sysplex (GDPS), disaster backup and recovery and remote tape vault
- ▶ Provide Storage Area Network (SAN) attachment

8.4.1 VTS models

In this section we describe the VTS and introduce the Model B10 VTS and Model B20 VTS. We explain how they are designed and how they work.

This new generation of VTS products provides an entry model; the Model B10 VTS and an enhanced model; the Model B20 VTS. These models incorporate data compression and Performance Accelerator features as a standard, as well as an improved controller. The VTS drives reside in a D12 Frame, containing 3590-B1A, E1A or H1A drives as well as a D22 frame containing 3592 drives.

The Model B10 VTS and Model B20 VTS are based on the IBM server pSeries 660 architecture which features powerful processors and expanded I/O capability. The B10 VTS is configured as a cost effective solution for modest throughput requirements while the B20 VTS establishes higher standards for throughput performance as well as for the number of virtual devices managed.

Note: Both Models B16 and B18 have been withdrawn from marketing.

Model B10 VTS

The Model B10 VTS supports up to 432 GB of TVC (1.3 TB assuming 3:1 compression), four to six 3590 drives, up to 64 virtual devices and up to four ESCON channels or eight SCSI ports or four FICON channels.

Model B20 VTS

The Model B20 VTS supports up to 1.7 TB of TVC (5.2 TB assuming 3:1 compression), six to twelve 3590 drives and up to 16 ESCON channels or eight SCSI ports or eight FICON channels, as well as up to 256 virtual devices.

Peer-to-Peer (PtP) VTS

The IBM TotalStorage Peer-to-Peer VTS is designed to eliminate all single points of failure and provide higher performance than the stand-alone VTS. IBM TotalStorage Peer-to-Peer VTS couples two separate VTSs into a single image with additional components and features. The two VTSs can be located at the same site or can be geographically separated for disaster tolerance. An IBM Model B10 VTS or Model B20 VTS can all be used in IBM TotalStorage Peer-to-Peer VTS configurations.

Additionally with feature code 4001-4004 (Advanced Policy Management), the PtP VTS configuration can be further tailored to suit individual site requirements. Control is gained over logical volume copy modes (either immediate or deferred). Media errors can also be eliminated as a single point of failure within each VTS with the ability to create dual copies of logical volumes on separate physical media.

The combined libraries require the addition of one to four 3494 CX1 frames which contain a Virtual Tape Controller (VTC) each. In turn, each VTC provides a total of four FICON channel

attachments; two attached to the host systems and one attached to each of the Model Bxx VTSs in the Peer-to-Peer configuration. The host systems no longer attach directly to the Model Bxx VTS. A VTC presents itself to the hosts as one 3490E tape control unit with sixteen 3490E tape drives.

Model B20 PtP has support for four VTC's

The enhancement provides for a lower entry cost PtP configuration when the VTS Model B20 is used. Previously, only an eight VTC configuration was supported whenever PtP configuration used two Model B20's. With this support, a four VTC PtP Model B20 configuration and update paths are now available.

This configuration is independent whether the interface used for the PtP are ESCON or FICON.

Note: Conversion from an existing eight VTC configuration to a four VTC configuration is not supported.

In summary the enhancements that have been made to the VTS is as follows:

- ▶ Attachment support for 3592 Tape Drives - which can provide native cartridge capacity up to 300 GB (900 GB with 3: 1 compression) with the use of the IBM TotalStorage Enterprise Tape Cartridge 3592.
- ▶ Enhanced reclamation policies - to help migrate data between pools and provide efficient storage of long-term data.
- ▶ Functional enhancements for improved monitoring and control.
- ▶ PtP VTS Model B20 support for four VTC's - to provide a low entry cost PtP configuration when the VTS Model B20 is used.

Note: It is important to note that in a mixed Peer-to-Peer (PtP) configuration the Composite Library will assume the attributes of the smaller VTS in relation to virtual drives.

8.5 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 these Redbooks for in-depth information about IBM Enterprise tape products:

- ▶ *IBM TotalStorage Peer-to-Peer Virtual Tape Server Planning and Implementation Guide*, SG24-6115, which can be found at:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246115.html?Open>
- ▶ *Magstar and IBM 3590 High Performance Tape Subsystem Technical Guide*, SG24-2506, which can be found at:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg242506.html?Open>
- ▶ *IBM TotalStorage Enterprise Tape A Practical Guide*
<http://www.redbooks.ibm.com/redbooks/pdfs/sg244632.pdf>



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 4mm, 8mm, VXA-2, SDLT, and SLR/MLR QIC to suit all business needs. Two additional products (DVD-RAM and disk) are also briefly mentioned.

9.1 Overview

Table 9-1 presents a summary of the entry-level storage products that are available as part of the TotalStorage portfolio. Details on each of the drives are presented below.

Table 9-1 Tape storage products

Product	External Disk Drive	4mm DDS-4 4mm Gen 5 (DAT72)	8mm VXA-2 Tape Drive	4mm DDS-3,4 Tape Autoloader
Machine	7204	7206	7206	7332
Model	518-18 GB 536-36 GB 573-73 GB 646-146 GB	220-DDS4 336-DDS5	VX2	220-DDS4
Product Strengths	External deskside disk storage	Cost-effective streaming tape drive	Low-cost, high-capacity VXA-2 technology	Autoload tape cartridges into DDS-4 drive
Technology	N/A	Helical Scan	Helical Scan	Helical Scan
Number of Heads/Tracks	8 heads 4 disks	Rotating Drum	Rotating Drum	Rotating Drum
Number of Drives	1	1	1	1
Max Number of Cartridges	N/A	1	1	6
Cartridge Capacity Native/Compressed	N/A	220: 20/40 GB 336: 36/72 GB	20/40 GB 59/118 GB 80/160 GB	20/40 GB
Max System Capacity Compressed	518: 18 GB 536: 36 GB 573: 73 GB 646-146 GB	220: 40 GB 336: 72 GB	160 GB	240 GB
Max Drive Data Rate Native/Compressed	80 MB/s	220: 3/6 MB/s 336: 3/6 MB/s	6/12 MB/s	3/6 MB/s
Time to Data	70 ms	50 seconds	50 seconds	40 seconds
Interface	SCSI-3 ULTRA, LVD/SE, 160/320	SCSI-2 F/W SE, LVD/SE	SCSI-3 ULTRA, LVD/SE, 160/320	SCSI-2 F/W SE, LVD/SE
SAN-Ready	pSeries, RS/6000	pSeries, RS/6000	pSeries, RS/6000, iSeries, AS/400, Windows NT	pSeries, RS/6000
Supported Platforms	pSeries, RS/6000	pSeries, RS/6000	pSeries, RS/6000, iSeries, AS/400	pSeries, RS/6000
Warranty	IBM Onsite Repair 1 year, 24x7	IBM Onsite Repair 1 year, 24x7	IBM Onsite Repair 1 year, 24x7	IBM Onsite Repair 1 year, 24x7

Product	8mm Mammoth Tape Drive	SDLT Drive Tape Drive	SLR (QIC) Compatible External Tape Drive	Storage Device Enclosure	VXA-2 Autoloader
Machine	7208	7205	7207	7212	7212
Model	345-Mammoth-2	550	122 330	102	312
Product Strengths	High-performance 8mm technology	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	Helical Scan	Longitudinal Serpentine	SLR (QIC format)	DDS, DVD, VXA	Helical Scan
Number of Heads/Tracks	Rotating Drum	8/536	1/1	DDS/VXA: Rotating Drum	Rotating Drum
Number of Drives	1	1	1	1-2	1
Max Number of Cartridges	1	1	1	2	10
Cartridge Capacity Native/Compressed	60/150 GB	160/320 GB	122: 4/8 GB 330: 30/60 GB	DDS-4: 20/40 GB DAT72: 36/72 GB VXA-2: 80/160 GB	80 GB/160 GB
Max System Capacity Compressed	150 GB	320 GB	122: 4/8 GB 330: 30/60 GB	160 GB with two VXA-2 drives 18.8 GB with two DVD-RAM drives	1600 GB
Max Drive Data Rate Native/Compressed	12/30 MB/s	16/32 MB/s	122: 38/.76 MB/s 330: 4/8 MB/s	DDS/DAT72: 3/6 MB/s VXA: 6/12 MB/s	6/12 MB/s
Time to Data	93 seconds	70 seconds	122: 85 seconds 330: 50 second	VXA-2: 40 seconds VXA: 40 seconds DAT72: 50 seconds	50 seconds
Interface	SCSI-2 Ultra2 LVD	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
SAN-Ready	pSeries, RS/6000, iSeries, AS/400, Windows NT	pSeries, RS/6000	iSeries, AS/400, pSeries, RS/6000		pSeries/RS6000
Supported Platforms	pSeries, RS/6000, iSeries, AS/400	pSeries, RS/60003	iSeries, AS/400, pSeries, RS/6000	pSeries, RS/6000, iSeries, AS/400	pSeries/RS6000

Product	8mm Mammoth Tape Drive	SDLT Drive Tape Drive	SLR (QIC) Compatible External Tape Drive	Storage Device Enclosure	VXA-2 Autoloader
Warranty	IBM Onsite Repair 1 year, 24x7	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

9.2 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 MB/second. The IBM 7204 External Disk Drives use a 68-pin LVD interface and attach to the host IBM p5, pSeries and RS/6000 via SCSI cable to a variety of SCSI adapters. Obviously, multiple 7204 devices can be installed on a single system.



Figure 9-1 IBM 7204 external disk drive

There are a number of different models which differ solely on the size of the disk drive that is installed in the enclosure (see Table 9-1 on page 150).

9.3 IBM 7205 SDLT tape drive

The IBM 7205 Model 550 is an external, stand-alone, SCSI-attached, DLT drive that attaches to both IBM RS/6000 and IBM - pSeries families of workstations and servers via a SCSI Ultra2 LVD interface. The IBM 7205 Model 550 features a native transfer rate of up to 16 MB/s, 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 MB/second.

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 9-2 IBM 7205 SDLT tape drive

9.4 IBM 7206 4mm tape drive

Several models are available.

9.4.1 Model 110

The IBM 7206 Model 110 is a stand-alone 3.5-inch half-high external 4 mm DDS-3 cartridge tape drive that provides a native media capacity of 12 GB per cartridge and a data transfer rate of 1.1 MB/se

9.4.2 Model 220

The IBM 7206 Model 220 external 4mm DDS-4 Tape Drive that connects to the IBM RS/6000 family of workstations and servers. The IBM 7206 Model 220 is a stand-alone external tape drive that provides a native media capacity of 20 GB per cartridge and a sustained data transfer rate of 21.6 GB per hour. Using Digital Data Storage (DDS), you can achieve a media capacity of 40 GB with 2:1 compression and a data transfer rate of up to 3.0 MB/s for unattended save and restore applications.

In addition to reading and writing on 20 GB DDS-4 tape cartridges, the 7206 Model 220 also reads and writes on other DDS-2 and DDS-3 tape cartridges.



Figure 9-3 IBM 7206 4mm tape drive

9.4.3 Model 336

The IBM 7206 Model 336 External DDS Gen 5 (DAT72) Tape Drive is designed to be attached to RS/6000 and pSeries 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).

9.4.4 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 models 220 and 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 pSeries or iSeries servers is via a standard Ultra2 SCSI LVD connector.

9.5 IBM 7207 SLR tape drive

The following models are available:

9.5.1 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 iSeries/AS400 internal QIC drives and can also be attached to RS/6000 and pSeries equipment 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 MB/second. 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 MB/second. 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 9-4 IBM 7207 SLR tape drive

9.5.2 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.

9.6 IBM 7208 8mm tape drive

The IBM 7208 Model 345 8mm tape drive provides an excellent solution to users of 8mm tape who require a larger capacity or higher performance tape backup. The IBM 7208 tape drive is based on enhanced 8mm Mammoth-2 (M2) tape drive technology and delivers fast and dependable save and restore function on both pSeries and iSeries servers using an industry standard SCSI Ultra2 LVD interface.



Figure 9-5 IBM 7208 8mm tape drive

The IBM 7208 is a standalone, SCSI-based streaming device that attaches to the IBM RS/6000 family of workstations and servers. Native cartridge capacity comes in three sizes of 20, 40, or 60 GB depending on the tape length and provides a sustained data transfer rate of 12 MB/s (up to 30 MB/s with 2.5:1 compression). The 7208 Tape Drive reads and writes to Advanced Metal Evaporated (AME) data cartridges with SmartClean technology. SmartClean cartridges are designed to combine the extremely reliable AME recording media with a short segment of tape drive head cleaning material. The new SmartClean cartridges are easily identified by their cobalt-blue color and are available in 75m, 150m, and 225m lengths. Although extensive use is not recommended, M2 can write to AME cartridges designed for the first-generation Mammoth Tape technology tape drives. When these cartridges are used, M2 requires regularly scheduled manual cleaning using a Mammoth Tape cleaning cartridge.

9.7 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 9-6 IBM 7210 DVD-RAM drive

9.7.1 Model 025

The 7210 Model 025 DVD-RAM Drive is an external storage device. This self-powered stand-alone drive is designed for the open systems environment, which include the IBM @server i5, iSeries, and AS/400, and p5, pSeries, and RS/6000 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 MB/second
- ▶ DVD-ROM media maximum speed is 2.7 MB/second
- ▶ DVD-RAM media maximum speed is 1.3 MB/second Write, 2.7 MB/second 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 3600KB/second
- ▶ DVD-ROM speeds of up to 8115KB/second
- ▶ DVD-RAM speeds of 1352 to 2705KB/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

9.7.2 Model 030

Model 030 is similar in most respects to the model 025, apart from only being able to read and write to 4.7/9.4 GB DVD-RAM, and supporting Ultra-3 SCSI (LVD) attachment.

9.8 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 iSeries and pSeries servers and RS/6000 workstations and servers.

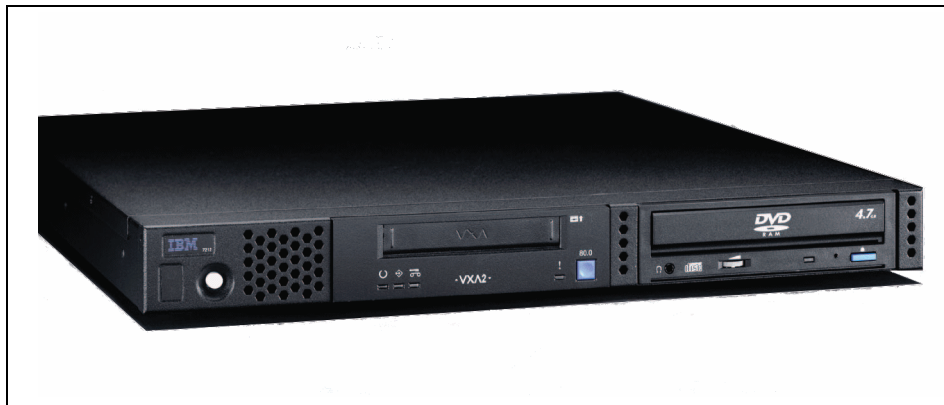


Figure 9-7 IBM 7212-102 storage device enclosure

It offers several storage device options including the following:

- ▶ Low cost VXA-2 technology

- ▶ DDS Gen 5 (DAT72) and DDS4 4mm,
- ▶ SLR60 and SLR100 (QIC format) tape drives
- ▶ DVD-RAM and DVDROM optical drives

It provides a solution for environments in which cabling space and server storage bays are limited.

9.9 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 MB/s, is designed to offer excellent capacity and performance at a low price.



Figure 9-8 IBM 7212-312 VXA-2 autoloader

9.10 IBM 7332 4mm autoloader

The following models are available.

9.10.1 Model 110

The IBM 7332 Model 110 4 mm DDS-3 Tape Cartridge Autoloader provides an integrated, cost-effective, entry-level tape storage solution with a four 12 GB cartridge capacity. The IBM 7332 Model 110 uses the same 12 GB tape drive implemented in the IBM 7206 Model 110 4 mm DDS-3 Tape Drive, providing all of its functional and performance characteristics. The 7332 Model 110 is designed to operate in streaming mode by using the autoloader to sequentially feed cartridges. Random autoloader operation is supported through industry-standard software applications.

9.10.2 Model 220



Figure 9-9 IBM 7332 4mm autoloader

The 7332 Model 220 is an external 4mm DDS-4 tape autoloader that operates in streaming mode and attaches to the host RS/6000 via a LVD/SE Wide SCSI cable. This tabletop storage solution includes a cartridge magazine with slots for six 4mm cartridges. The 7332 Model 220 has a maximum capacity of 120 GB with the standard six-cartridge magazine. With data compression, the maximum capacity for the standard configuration is 240 GB. The IBM 7332 Model 220 uses the 20 GB tape drive implemented in the IBM 7206 Model 220 4mm DDS-4 Tape Drive, providing all of its functional and performance characteristics. The 7332 Model 220 is designed to operate in streaming mode by using the autoloader to sequentially feed cartridges. Random autoloader operation is supported through industry-standard software applications.

9.11 More information

For more detailed information on these products refer to:

Media:

http://www-1.ibm.com/servers/storage/media/tech_tapecartridges/index.html

Tape drives:

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

Entry-level tape storage:

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

IBM 7204 disk:

<http://www-1.ibm.com/servers/storage/disk/7204/index.html>



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.

10.1 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.

There are basically two modes for attaching storage for the network: Storage Area Network (SAN) and Network Attached Storage (NAS). Both the modes 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 Local Area Networks: 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.

10.2 Storage Area Network (SAN)

A SAN is a dedicated network for storage devices and the processors that access those devices. Figure 10-1 shows a picture of 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.

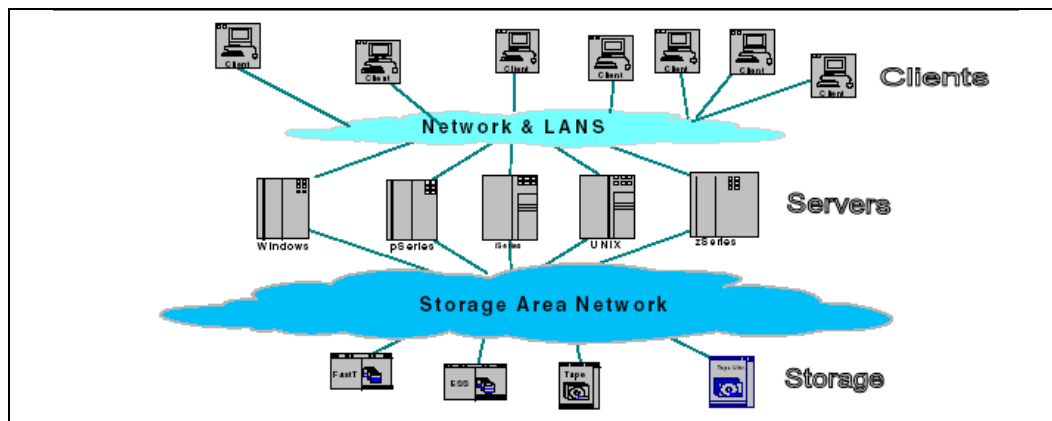


Figure 10-1 A picture of SAN

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.

10.2.1 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 is able to provide speeds of up to 2 Gbps and products for 4 Gbps are now becoming available.

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 you can see in Figure 10-2, 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.

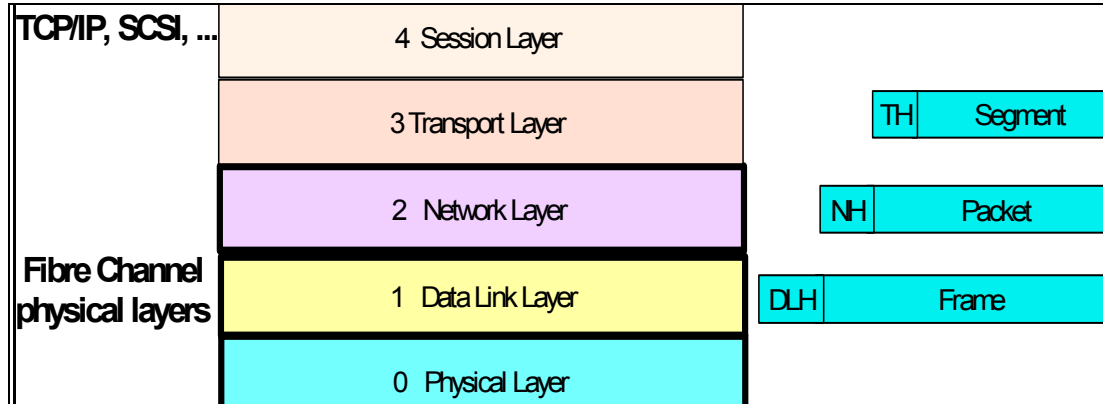


Figure 10-2 Fibre Channel protocol layers

10.2.2 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 10.3, “Extended distance fabrics” on page 170).

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 10-3. A node is any device with one or more Fibre Channel ports.

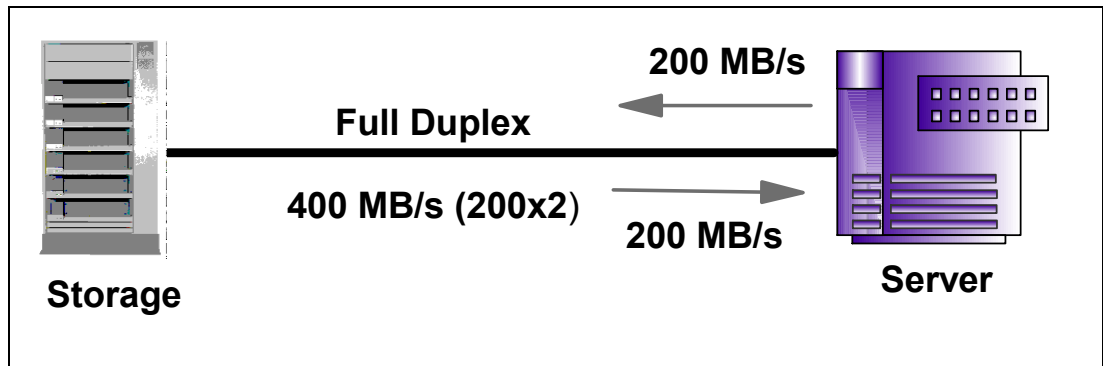


Figure 10-3 Fibre Channel point-to-point topology

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.

Other characteristics which should be considered are:

- ▶ Arbitration: Only one device can be *talking* (writing or reading data) at any one time:
This is particularly important if multiple servers are to be attached to the same loop for tape library sharing. Once a device has arbitrated and has a virtual connection, other servers are locked out and must wait.
Because storage devices usually cannot write and read at the same time, only one half of the full duplex bandwidth is being used effectively.
- ▶ The hub is a single point-of-failure because the loop has to be complete for it to work.

More advanced FC hub devices support FC loop connections while offering some of the benefits of switches. Figure 10-4 shows an FC loop using a hub.

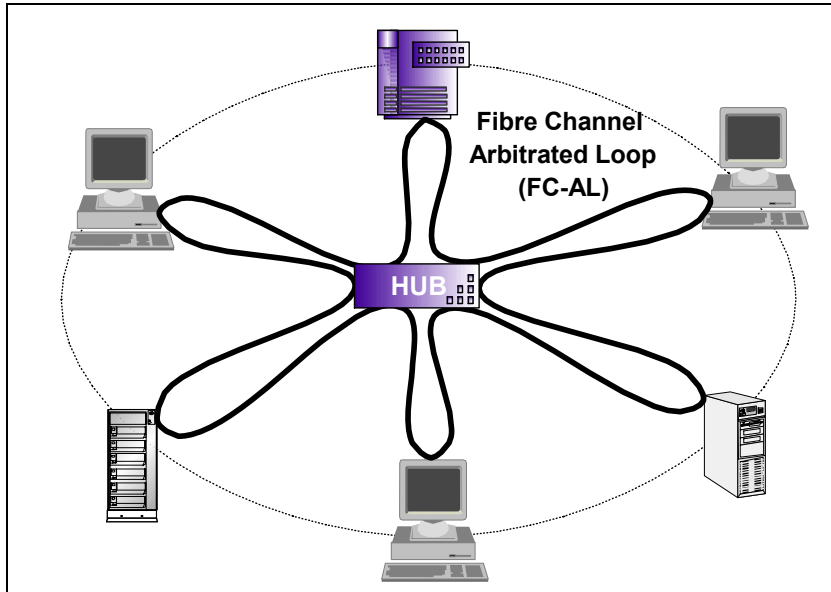


Figure 10-4 Fibre Channel loop topology

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 10-5. When switches are connected, each switch's configuration information has to be copied (*cascaded*) into all the other participating switches.

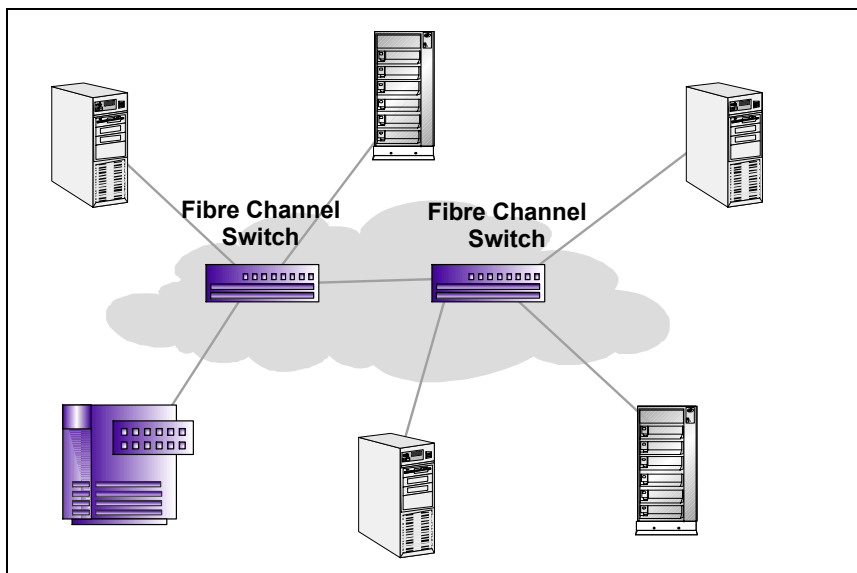


Figure 10-5 Fibre Channel switched topology

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 or cascaded (as shown in Figure 10-5) to provide extra paths and larger networks.

Switched fabric without cascading

A switched point-to-point fabric without cascading can support the configurations required by very large systems and 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 fabric with non-cascaded switches is shown in Figure 10-6.

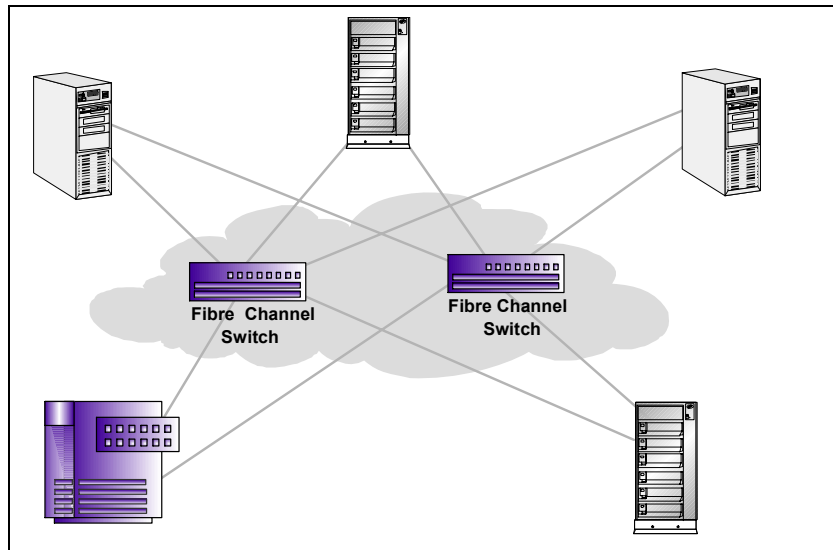


Figure 10-6 Fibre Channel switched topology (non-cascaded switches)

Switched fabric with cascading

A switched fabric with cascading 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.

The primary advantage of a switched cascaded 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 10-7.

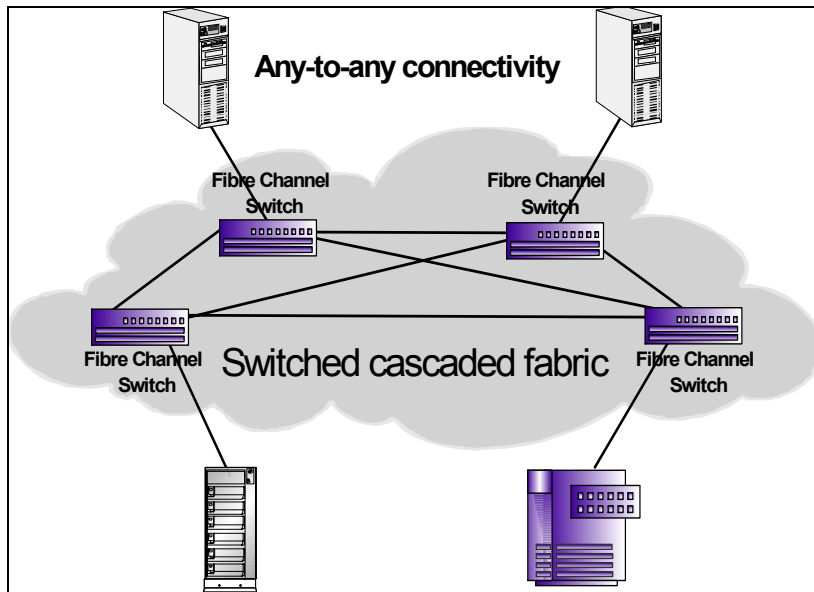


Figure 10-7 Fibre Channel switched topology (cascaded switches)

10.2.3 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 10-8 shows FC connectors and cables.

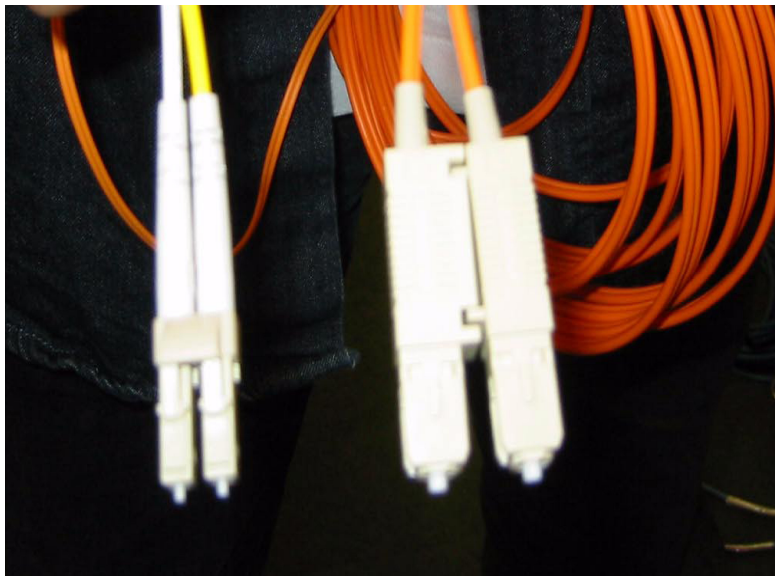


Figure 10-8 LC and SC connectors

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, the transceivers are called SFPs (Small Form Factor Pluggable Media). Figure 10-9 shows SFPs and a GBIC.

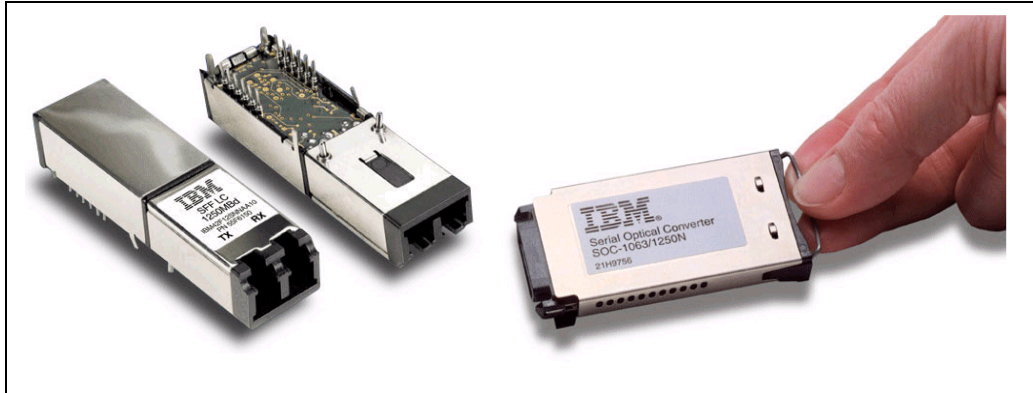


Figure 10-9 SFP (Small Form Factor Pluggable Media) and GBIC (Gigabit Interface Converter)

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 10-10 shows HBA.

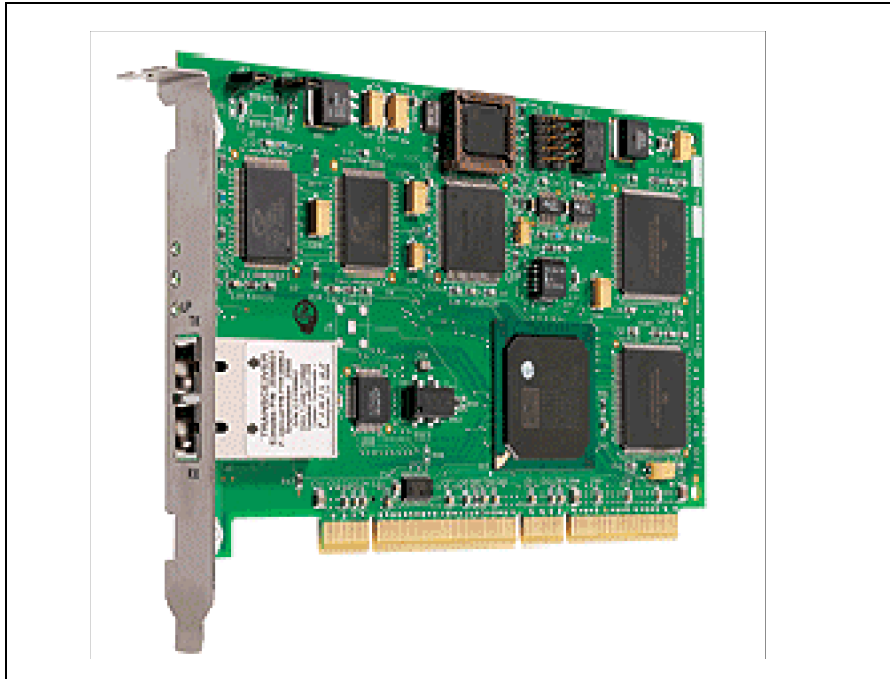


Figure 10-10 Host Bus Adapter

10.2.4 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 on Naming and addressing, refer to the redbook *Introduction to Storage Area Network*, SG24-5470.

10.2.5 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's previous fiber optic channel standard.

Note: For more information on FICON, refer to the redbook *FICON Native Implementation and Reference Guide*, SG24-6266

10.3 Extended distance fabrics

Extending distance across a wide area network (WAN) is a huge topic, which is covered only briefly here in order to show what can be achieved, not how. As stated earlier, the distance of a SAN can be extended beyond the normal 10 km achieved with Fibre Channel long-wave GBICs to about 100 km for synchronous transfers and to virtually unlimited distances for asynchronous transfers over a WAN. In both cases, additional components are required to extend the fabric.

Synchronous transfer is limited by the speed of light in glass. A Fibre Channel frame in which data is packaged spans just over 2 km of fibre, therefore, at 10 km, there could be almost 5 frames in the pipeline and, at 100 km, almost 50 frames. The Fibre Channel architecture includes a feature known as *buffer credits*, where the sender is only allowed to send as many frames as he has received buffer credits from the receiver.

To maintain performance at distances above 10 km, and fully utilize the potential bandwidth, a switch can use multiple frame buffers, in which it remembers the data until the data has been successfully acknowledged, together with a greatly increased number of buffer credits for the port that is attached to the remote switch. The number of buffer credits is normally controlled by a software feature in the switch's operating system. For example, in the IBM 2109-F16 switch, the Extended Fabrics feature must be licensed.

In addition to extra frame buffers, there must also be a mechanism to physically send the light over longer distances. There are many products available to do this, but two examples are the Finisar Optilinx 2000 Optical Link Extender and the IBM 2029 Fibre Saver (Figure 10-11).

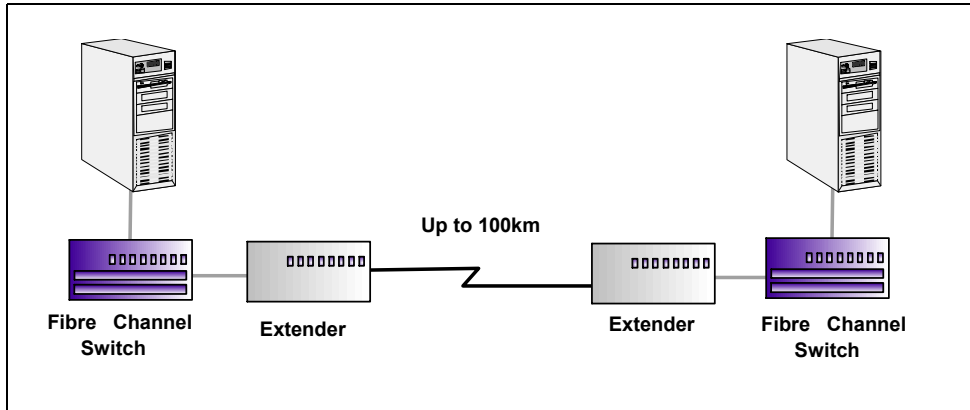


Figure 10-11 Extending the distance of fabric using extender

Even greater distances can be achieved by connecting to an Asynchronous Transfer Mode (ATM) WAN. This requires the use of additional components such as the CNT Open Systems Gateway to convert the SAN to WAN and back again. A simple example is shown in Figure 10-12.

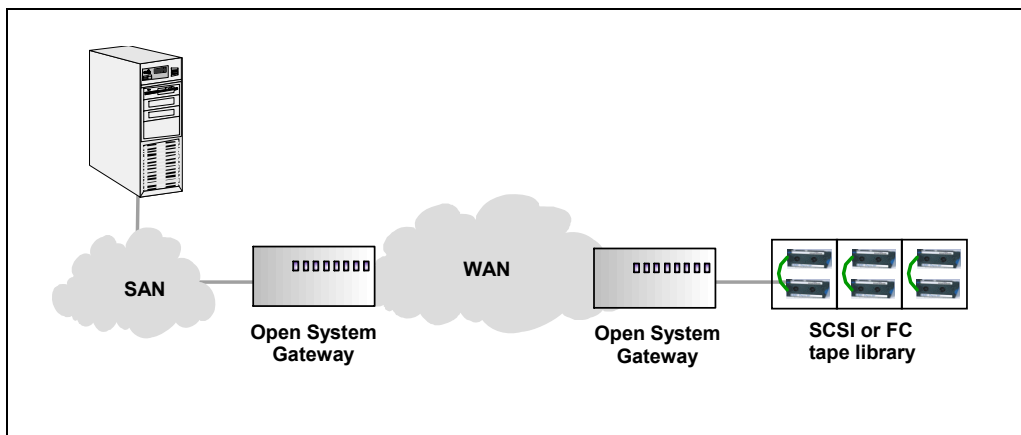


Figure 10-12 Extended distance with a WAN

10.4 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 10-13.

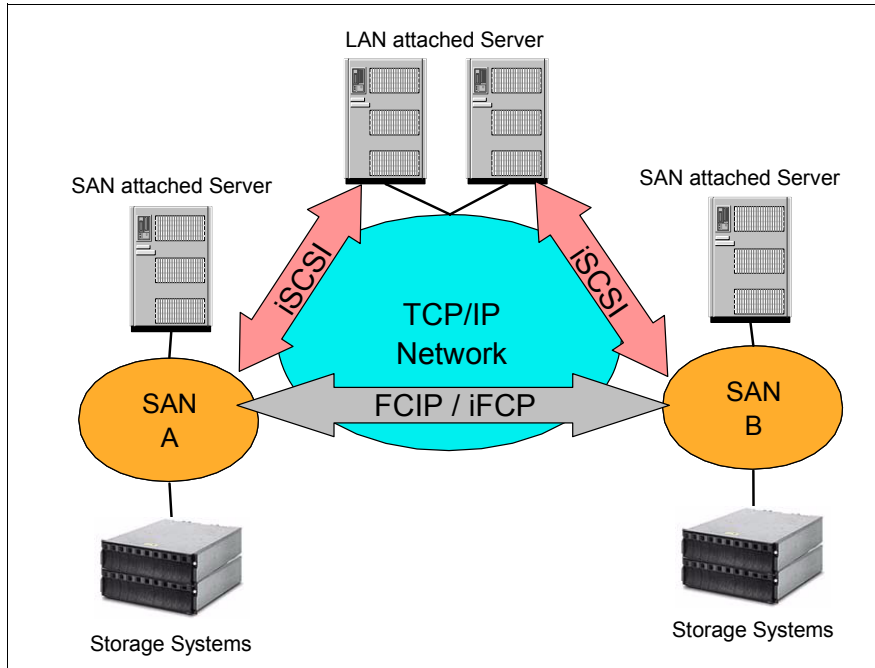


Figure 10-13 IP Storage networking technology

10.4.1 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.

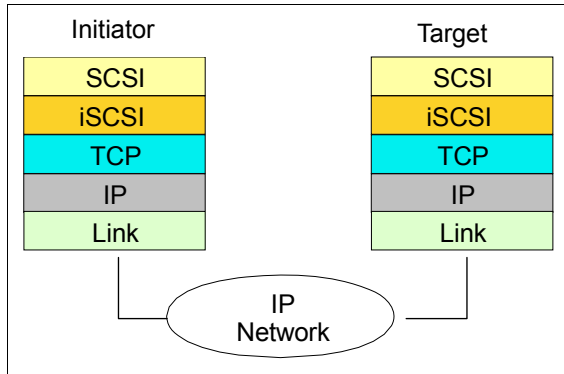


Figure 10-14 iSCSI protocol layering model

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.

10.4.2 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 10-15. Using FCIP, customers can utilize their current wide-area networking infrastructure for connecting remote SAN islands over long distances.

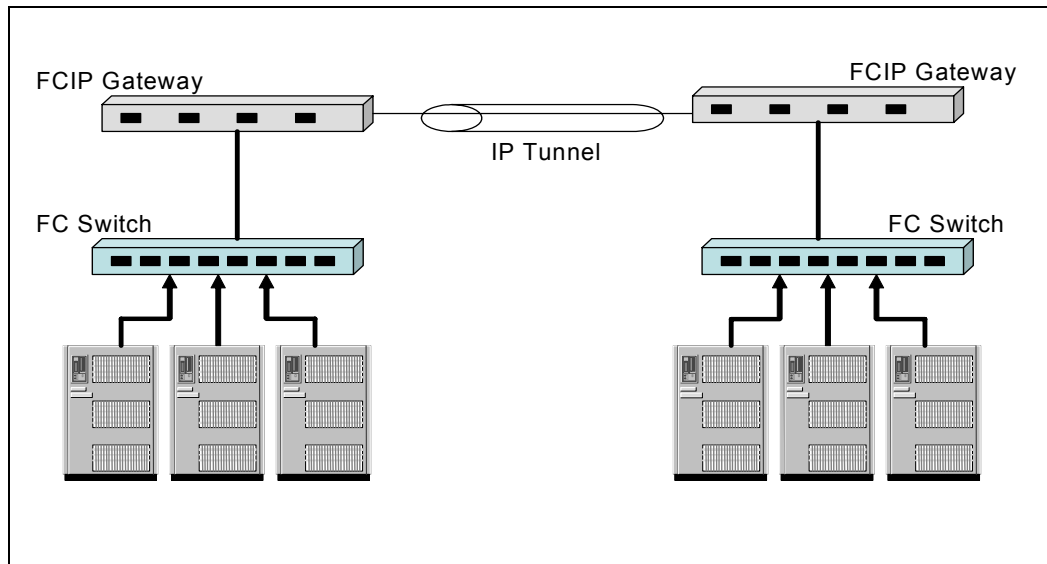


Figure 10-15 FCIP with IP tunneling

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.

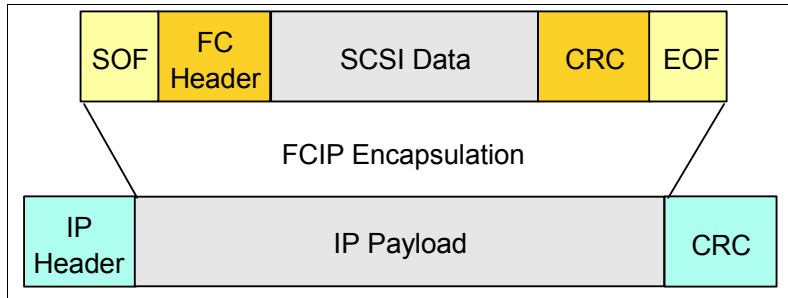


Figure 10-16 FCIP Encapsulation model

Today's Fibre Channel-over-IP solutions encapsulate Fibre Channel and transport it over a TCP socket (Figure 10-16). 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.

10.4.3 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 10-17).

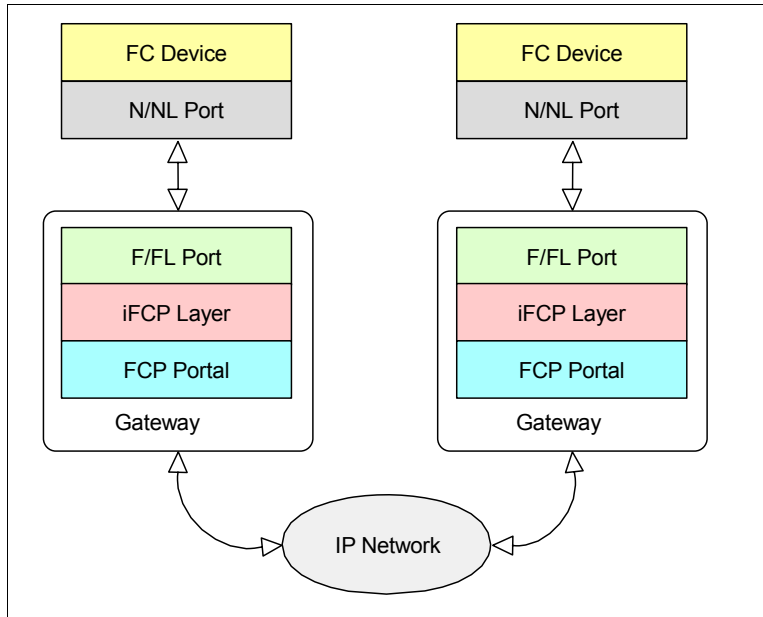


Figure 10-17 iFCP Gateway architecture

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.

10.4.4 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 10-18, the different structure of the protocol stacks as well as the different key attributes in Table 10-1 on page 176.

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

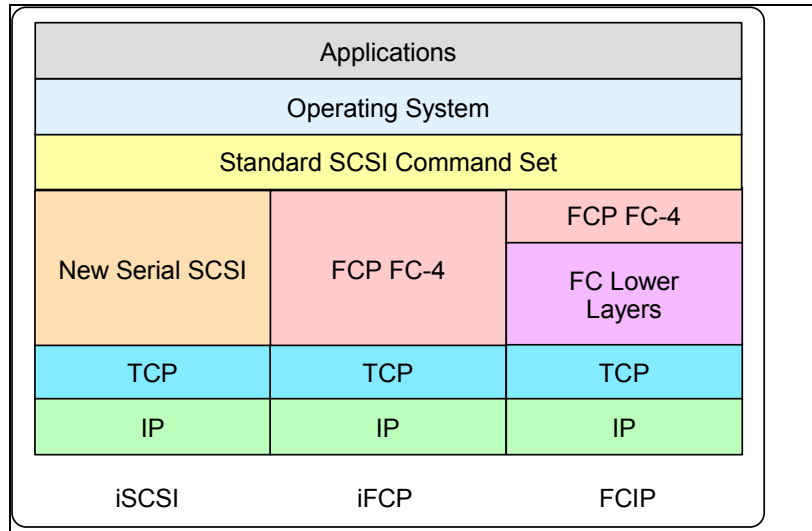


Figure 10-18 iSCSI, FCIP and iFCP protocol stacks

Table 10-1 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

10.5 Network Attached Storage (NAS)

An 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 10-19 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.

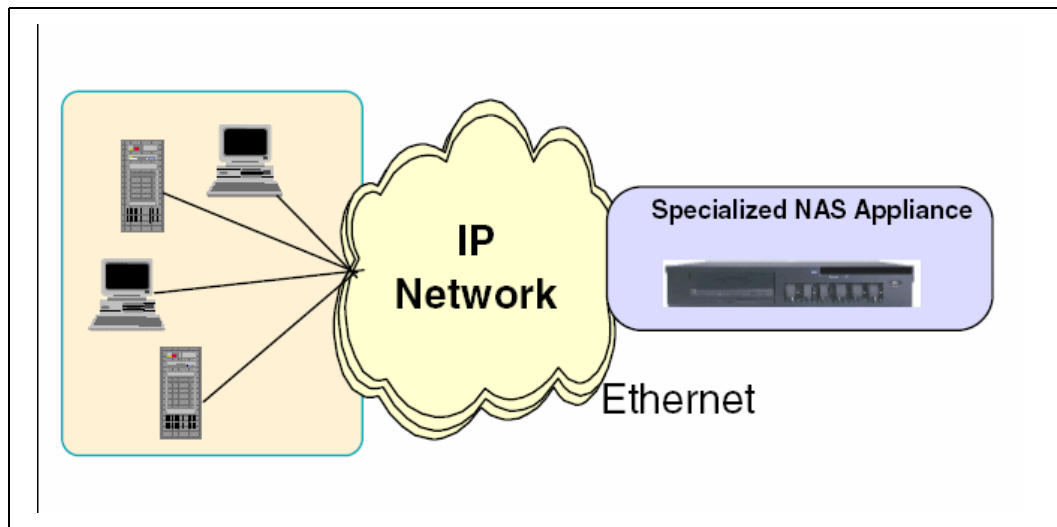


Figure 10-19 A picture of NAS appliance

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.

10.5.1 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.

10.6 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 10-20). In summary WDM enables many signals to be concentrated into a single fibre, all being sent at different wavelengths.

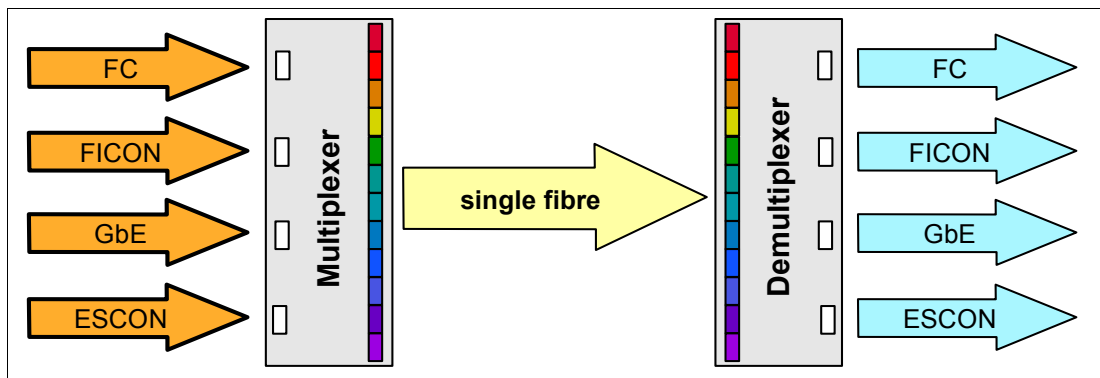


Figure 10-20 Wave@ Division Multiplexer concept

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.

IBM offers a solution named “CISCO CWDM Extended Distance Solution” in this area. The Cisco CWDM solution increases the data capacity of existing fiber by up to eight times that of standard optical modules while extending the reach of Inter-Switch Links (ISLs) to up to 100 kilometers, significantly lowering cost of SAN connectivity. For more information, refer to:

<http://www.ibm.com/storage/cisco>

10.7 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.

10.8 More information

For a better understanding of Fibre Channel, SAN technology, FICON and NAS technology refer to these Redbooks:

- ▶ *Introduction to Storage Area Network*, 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 System/390® FICON*, SG24-5176
- ▶ *IBM S/390 FICON Implementation Guide*, SG24-5169
- ▶ *Planning for S/390 Fibre Optic Links*, GA23-0367
- ▶ *The IBM TotalStorage NAS Integration Guide*, SG24-6505
- ▶ *IP Storage Networking: IBM NAS and iSCSI Solutions*, SG24-6240

You can download these books as well as others from the IBM Redbooks Web site at:

<http://publib-b.boulder.ibm.com/redbooks.nsf/portals/Storage>



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 TotalStorage 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 TotalStorage SAN b-type switches and directors
- ▶ IBM TotalStorage SAN m-type family
- ▶ IBM TotalStorage 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 storage solutions” on page 9. Additionally, these products can be used with tape drives and libraries, as discussed in Part 2, “Tape products” on page 105.

For the latest SAN products and information, go to:

<http://www.ibm.com/servers/storage/san/>

11.1 SAN switches and connectivity products

IBM TotalStorage 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), CNT (n-type) and Cisco switches and directors.

IBM TotalStorage 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 TotalStorage SAN switch offerings can be positioned into three broad product groups: entry, mid-range, and enterprise. Suitable products for every Storage Area Network project can be found among these.

11.1.1 Common characteristics

With all IBM TotalStorage 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
- ▶ 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

11.1.2 Other switch features

Here are other features available in selected IBM TotalStorage 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. Advance zoning of the 2 Gbps switches compared with the zoning of 1 Gbps switches enlarges the range of hardware enforcement and so provides the switch with more security access control functions as before, preventing an unauthorized devices from accessing the fabric.

Common IBM SAN Switch capabilities

IBM TotalStorage 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 TotalStorage 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.

11.1.3 More information

- ▶ For the interoperability matrix, which shows adapter lists, go to:
<http://www-1.ibm.com/servers/storage/support/san/index.html>
Click *Plan and Upgrade* under the desired switch product
- ▶ For technical support information, go to:
<http://www-1.ibm.com/servers/storage/support/san/index.html>
- ▶ For additional information, refer to the Draft redbook *IBM TotalStorage: SAN Product, Design and Optimization Guide* SG246834.pdf
<http://www.redbooks.ibm.com/redpieces/pdfs/sg246384.pdf>
- ▶ For additional information, refer to the redbook *IBM SAN Survival Guide Featuring the IBM 3534 and 2109*, SG24-6127, which can be found at:
<http://www.redbooks.ibm.com/redpieces/pdfs/sg246116.pdf>
- ▶ See also the redbook *Implementing an Open IBM SAN*, SG24-6116, which can be found at:
<http://www.redbooks.ibm.com/redbooks/pdfs/sg246116.pdf>

11.2 Entry-level SAN switches

The IBM TotalStorage SAN switch entry-level products offer a new level of price performance for first-time storage area network users. Used in small and medium business, infrastructure simplification and Business Continuity solutions, IBM entry fabric switches provide 2 Gbps fully non-blocking performance, and advanced intelligence features. Scalability ranges from 4 to 20 ports and can be attached in a two to four IBM TotalStorage SAN Switch fabric for Microsoft Windows NT, 2000 and UNIX server clustering.

Entry-level products include:

- ▶ IBM TotalStorage SAN Switch L10
- ▶ IBM TotalStorage SAN Switch H08
- ▶ IBM TotalStorage SAN Switch SAN12M-1

In this section, we introduce IBM TotalStorage Switch solutions ideal for entry-level applications.

11.2.1 IBM TotalStorage Switch L10

IBM TotalStorage Storage Switch L10 - provides simple-to-use infrastructure simplification and Business Continuity solutions for IBM @server xSeries and IBM @server BladeCenter. Entry ten-port storage switch includes zoning and integrated Web server. Two storage

switches may be cascaded for expanded Microsoft Windows NT, 2000 Server and Cluster Service (MSCS); Red Hat and SUSE LINUX and Novell NetWare solutions.

Highlights

- ▶ Simple-to-use infrastructure simplification and Business Continuity solutions for xSeries and IBM BladeCenter
- ▶ Designed to support high availability with a minimal number of components and intelligent management of network changes
- ▶ An intuitive integrated management Web server with smart settings helps simplify administration
- ▶ Designed for high performance with 1 or 2 gigabit per second (Gbps) throughput on all ports
- ▶ Offers high-density packaging and small form-factor pluggable (SFP) transceivers that provide ten ports in a one-half rack width with a 1U rack height



Figure 11-1 IBM TotalStorage Switch L10

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/e_type/110/

11.2.2 IBM TotalStorage SAN Switch H08

Next generation entry fabric switch provides 4-8 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers zoning, WEBTOOLS and attachment in a two switch fabric. Full fabric feature supports maximum core/edge fabrics.

Highlights

- ▶ Entry-fabric switch with advanced zoning and ability to attach to one IBM TotalStorage SAN Switch for Microsoft Windows NT, Windows 2000 and UNIX server clustering
- ▶ Upgrade capability to full-enterprise capabilities for use in large core-to-edge fabrics
- ▶ Offers 8-port Fibre Channel switch with wide range of Small Form-Factor Pluggable (SFP) transceiver features
- ▶ Improved availability capabilities include hot code activation and single field replaceable unit, FRU design
- ▶ Provides up to 2 Gbps non-blocking performance
- ▶ Uses WEBTOOLS, Fabric Manager and open fabric interfaces which can help simplify management of small SAN fabrics and large core-to-edge SAN fabrics

- ▶ Offers advanced fabric service features such as end-to-end performance monitoring and Inter-Switch Link (ISL) Trunking with aggregate speed up to 8 Gbps
- ▶ Offers Advanced Security feature with extensive, policy-based security capabilities



Figure 11-2 IBM TotalStorage SAN Switch H08

The IBM H08 SAN Switch is well suited to address small and medium business (SMB) customer requirements for infrastructure simplification and improved Business Continuity.

The H08 SAN Switch is designed to be interoperable with other members of the IBM TotalStorage SAN Switch family. You can configure a wide range of scalable solutions that help address SMB demands for affordable, simple to operate and integrated heterogeneous open server SANs.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/b_type/h08/

11.2.3 IBM TotalStorage SAN Switch SAN12M-1

IBM TotalStorage SAN12M-1 entry fabric switch (2026-E12, replaces McDATA 4300) provides 4, 8, 12-ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers zoning and SANpilot. Full fabric feature supports maximum IBM SAN m-type core/edge fabrics.

Highlights

- ▶ Simple-to-use entry-level infrastructure simplification and Business Continuity solutions for xSeries, iSeries and pSeries servers
- ▶ Designed to support high availability with minimum number of components and HotCAT online code activation
- ▶ Designed for high-performance with 1 Gbps and 2 Gbps ports
- ▶ Integrated browser-based SANpilot management for first-time SAN users
- ▶ Pay-as-you-grow with flexible scalability with FlexPort and full fabric capabilities
- ▶ Offers excellent price/performance for first-time SAN users



Figure 11-3 IBM TotalStorage SAN Switch SAN12M-1

A wide range of IBM TotalStorage storage area network (SAN) infrastructure simplification and Business Continuity solutions can be created with the IBM TotalStorage SAN12M-1 entry fabric switch. Infrastructure simplification solutions for xSeries, iSeries and pSeries servers include high-availability server clustering and storage consolidation with IBM TotalStorage disk storage arrays. Business Continuity solutions include data protection with IBM TotalStorage Linear Tape-Open (LTO) or IBM TotalStorage 3590 and 3592 Tape Drives and IBM Tivoli Storage Manager data protection software.

More information

- ▶ For additional product details, go to:

http://www-1.ibm.com/servers/storage/san/m_type/san12m/

11.3 Mid-range SAN switches

IBM TotalStorage mid-range SAN solutions provide additional capability, features and benefits beyond the simple entry solutions. The IBM TotalStorage 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 8, 16, 24, 32, and 40 port models.

IBM TotalStorage 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 TotalStorage 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 SAN Switch H16
- ▶ IBM TotalStorage SAN Switch F32
- ▶ IBM TotalStorage SAN Switch SAN32B-2 fabric switch
- ▶ IBM TotalStorage SAN24M-1
- ▶ IBM TotalStorage SAN32M-1
- ▶ Cisco MDS 9120 and 9140 Multilayer Fabric Switch
- ▶ Cisco MDS 9216 Multilayer Fabric Switch

In this section, we introduce IBM TotalStorage Switch solutions ideal for mid-range applications.

11.3.1 IBM TotalStorage SAN Switch H16

IBM TotalStorage SAN Switch H16, with next generation switch technology, provides improved availability capabilities, 16-port, 2gigabit per second fully non-blocking performance, and advanced intelligence features.

Highlights

- ▶ Entry-fabric switch with advanced zoning and ability to attach to four IBM TotalStorage SAN Switch fabric for Microsoft Windows NT, Windows 2000 and UNIX server clustering
- ▶ Upgrade to full-enterprise capabilities for use in large core-to-edge fabrics
- ▶ Offers 16-port Fibre Channel switch with wide range of Small Form-Factor Pluggable (SFP) transceiver features
- ▶ Improved availability capabilities include hot code activation and single field replaceable unit, FRU design
- ▶ Provides up to 2 Gbps non-blocking performance
- ▶ Uses WEBTOOLS, Fabric Manager and open fabric interfaces which can help simplify management of small SAN fabrics and large core-to-edge SAN fabrics
- ▶ Offers advanced fabric service features such as end-to-end performance monitoring and Inter-Switch Link (ISL) Trunking with aggregate speed up to 8 Gbps
- ▶ Offers Advanced Security feature with extensive, policy based security capabilities



Figure 11-4 IBM TotalStorage SAN Switch H16

Entry-fabric switch with advanced zoning can be attached in a four IBM TotalStorage SAN Switch fabric for Microsoft Windows NT, 2000 and UNIX server clustering. Full fabric upgrade feature enables use in large core-to-edge fabrics.

The IBM H16 SAN Switch is well suited to address small and medium business customer requirements for infrastructure simplification and improved Business Continuity.

The H16 SAN Switch is designed to be fully interoperable with other members of the IBM TotalStorage SAN Switch family. You can configure a wide range of scalable solutions that help address SMB demands for afford ability, simple to manage and integrated heterogeneous open servers SANs.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/b_type/h16/

11.3.2 IBM TotalStorage SAN Switch F32

IBM TotalStorage SAN Switch F32 provides 32-port, 2gigabit per second open server and mainframe mid-range and enterprise solutions designed with a common architecture and an

integrated enterprise SAN management capability. IBM SAN Switches are supported by the broadest range of IBM open server and storage devices. Compatibility with the Brocade SilkWorm family of switches can enable interoperability with a wide range of non-IBM server and storage devices.

IBM SAN Switches provide the performance, scalability, high availability and manageability required to scale SAN fabrics up to thousands of devices. Customers have the opportunity to initially deploy separate SAN fabrics at the departmental and data center levels and then to interconnect them into an integrated IBM TotalStorage Enterprise SAN as their experience and requirements grow and change.

Highlights

- ▶ Can be used as a departmental switch, a core fabric switch or an edge switch with large enterprise SANs
- ▶ Designed to provide high performance with up to 2 Gbps throughput per port and Inter-Switch Link (ISL) Trunking with aggregate speed up to 8 Gbps
- ▶ Offers Advanced Security with comprehensive, policy-based security capabilities
- ▶ Provides broad range of IBM open server and storage support including fabric, loop and private loop attachments
- ▶ Offers 32-port Fibre Channel switch with rack space saving Small Form-Factor Pluggable (SFP) transceivers
- ▶ Offers advanced fabric services such as end-to-end performance monitoring and advanced zoning



Figure 11-5 IBM TotalStorage SAN Switch F32

The SAN Switch F32 is built upon a next-generation switch technology that supports link speeds of 1 and 2 Gbps.

Each port supports either 100 MB/s or 200 MB/s, full-duplex data transfers. Auto-sensing ports automatically negotiate to the highest speed supported by the attached server, storage or switch.

The F32 switch is suited for disaster tolerance solutions such as remote tape vaulting and remote disk mirroring. F32 Switches can provide up to twice the throughput of prior generation SAN Switches. This performance capability can be used either to reduce the number of ISL connections or to improve the performance with the same number of connections.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/b_type/f32/

11.3.3 IBM TotalStorage SAN Switch SAN32B-2 Fabric Switch

High performance mid-range fabric switch provides 16, 24 and 32-port, 4 Gbps fabrics switching for Windows NT/2000 and UNIX server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers Advanced Zoning, Fabric Watch, WEBTOOLS, dual replaceable power supplies and 16-ports activated. Ports on Demand features support “pay-as-you-grow” scalability.

Highlights

- ▶ Simple-to-use mid-range and enterprise infrastructure simplification and Business Continuity SAN solutions
- ▶ Designed for high-performance with 4 Gigabit per second (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
- ▶ Multiple-management options for first-time storage area network (SAN) users and complex enterprise SAN consolidation solutions
- ▶ Interoperability with IBM TotalStorage SAN b-type switch family helps protect switch investment

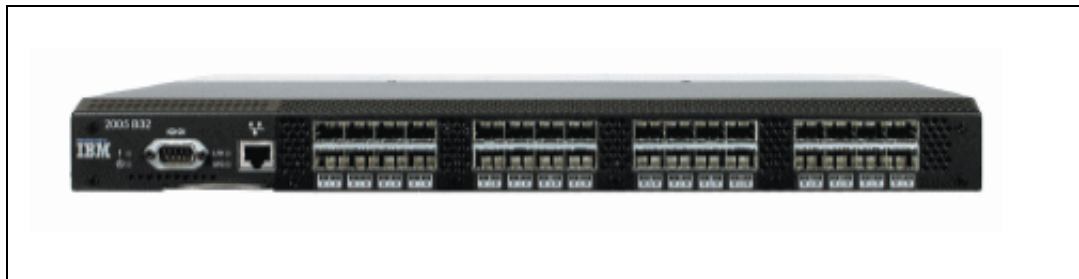


Figure 11-6 IBM TotalStorage SAN32B-2 fabric switch

More information

- ▶ For additional information, go to:
http://www-1.ibm.com/servers/storage/san/b_type/san32b/

11.3.4 IBM TotalStorage SAN Switch SAN24M-1

IBM TotalStorage SAN24M-1 (2026-224, replaces McDATA 4500) fabric switch provides 8, 16, 24-ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers zoning, SANpilot and attachment in IBM SAN m-type core-to-edge fabrics. With the Element Manager feature, it provides FC-AL tape attachment to IBM SAN m-type directors under McDATA Enterprise Fabric Connectivity Manager

Highlights

- ▶ Simple-to-use mid-range infrastructure simplification and Business Continuity solutions for IBM @server xSeries, iSeries and pSeries servers
- ▶ Enterprise-to-edge storage area network (SAN) management for advanced enterprise infrastructure simplification and Business Continuity solutions, including disk and tape SAN island consolidation

- ▶ Designed to support high availability with redundant, hot swappable fans and power supplies and HotCAT online code activation
- ▶ Designed for high-performance with 1 Gbps and 2 Gbps ports
- ▶ Integrated browser-based SANpilot management for first-time SAN users
- ▶ Pay-as-you-grow with flexible scalability with FlexPort feature IBM TotalStorage mid-range solutions



Figure 11-7 IBM TotalStorage SAN24M-1 SAN Switch

Enterprise Fabric Connectivity Manager

Enterprise Fabric Connectivity Manager (EFCM) software is designed to support interconnection of multiple IBM TotalStorage SAN m-type switches and directors for the creation of enterprise-to-edge SANs.

The **SAN24M-1 Element Manager** feature is designed to enable enterprise- to-edge management of SAN24M-1 switches when interconnected with multiple IBM TotalStorage SAN m-type switches and directors in an enterprise SAN solution. EFCM software, with the 1U Rack Mount Management Server, is designed to centralize the management of multiple, distributed switches and directors in an enterprise-wide Fibre Channel fabric. EFCM software is designed to provide advanced functions such as nondisruptive code activation, call-home and e-mail alerts as events occur. The fabric switch is connected through an Ethernet hub to the 1U Rack Mount Management Server with EFCM software.

Open Trunking is designed to optimize the total throughput between two switches automatically by redirecting traffic from high-utilization links to links with available bandwidth. Open Trunking is designed to continuously monitor loads on all links, detect congestion and automatically balance traffic across available ISLs without operator interaction. Based on constant traffic monitoring, Open Trunking acts to help improve throughput and reroute traffic efficiently.

Open System Management Server (OSMS) is an ANSI (American National Standards Institute)-based feature that supports SAN management software from vendors such as VERITAS, Tivoli and BMC. OSMS is designed to extend the switch's capability to include in-band management with an open systems host-based application. OSMS allows the fabric switch and devices attached to it to be "discovered," or seen in a fabric through a framework software application.

SANtegrity Security Suite is designed to help enhance Business Continuity by reducing the impact of human influences on networked data. SANtegrity Zoning, a standard capability of Enterprise Operating System (E/OS) 5.0, provides hardware-enforced world wide name (WWN) and port zoning. SANtegrity Secure Management Zone (SMZ) is designed to secure management access to local and remote SAN devices over a secure connection.

SANtegrity Binding is designed to create multiple layers of access control, including port, switch and fabric binding.

Standards-based **SANtegrity Authentication** is designed to force each device in the SAN to prove who they are in order to avoid unauthorized access and unauthenticated devices. SANtegrity Binding is required in order to use the Authentication feature.

Preferred Path Option is designed to allow a customer to define routes across a fabric. The Preferred Path feature modifies the behavior of Open Trunking by providing guidance for the balancing function. The Preferred Path configuration helps address user preference regarding exit port assignment but is subject to the standard rules regarding the Fabric Shortest Path First (FSPF) protocol.

Full Volatility is designed to support high-security environments, which require no customer data be retained after power-off. The Full Volatility feature is designed to configure a switch or director so that no frame data is stored after a power-off.

High-performance

The IBM TotalStorage SAN24M-1 switch, featuring a non-blocking switch core, is designed to provide 1 Gbps and 2 Gbps ports. Each switch port auto-negotiates to either 1 Gbps or 2 Gbps depending upon the device attachment. With 24-port configurations, up to 96 Gbps throughput is possible.

FlexPort scalability

FlexPort technology is designed to support scalable switch upgrades in eight-port increments without fabric disruption. Each FlexPort upgrade consists of eight shortwave SFP transceivers and an activation key that can add eight ports to the fabric switch. Connectivity on demand, or the ability to increase switch capacity while maintaining service, helps allow a customer to implement a pay-as-you-grow strategy.

Configuration flexibility

The IBM TotalStorage SAN24M-1 fabric switch is offered with 8, 16 or 24 ports. Each port is self-configuring as a fabric, fabric loop or expansion port. Longwave SFP transceivers can be added to any port for 2 Gbps connections up to 10 km.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/m_type/san24m/

11.3.5 IBM TotalStorage SAN32M-1

IBM TotalStorage SAN32M-1 (2027-232, replaces McDATA 3232) fabric switch provides 8-32 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX; and FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers zoning, SANpilot and attachment in IBM SAN m-type core-to-edge fabrics. With Element Manager feature, provides attachment to IBM SAN m-type switches and directors under McDATA Enterprise Fabric Connectivity Manager.

Highlights

- ▶ Simple-to-use, mid-range infrastructure simplification and Business Continuity solutions for IBM @server xSeries, iSeries, pSeries and zSeries servers
- ▶ Enterprise-to-edge storage area network (SAN) management for advanced enterprise infrastructure simplification and Business Continuity solutions, including mainframe FICON disk and tape storage

- ▶ Designed to support high availability with redundant, hot swappable fans and power supplies and HotCAT online code activation
- ▶ Designed for high-performance with 1 Gbps and 2 Gbps ports
- ▶ Integrated browser-based SANpilot management for first-time SAN users
- ▶ Pay-as-you-grow with flexible scalability with FlexPort feature IBM TotalStorage mid-range solutions

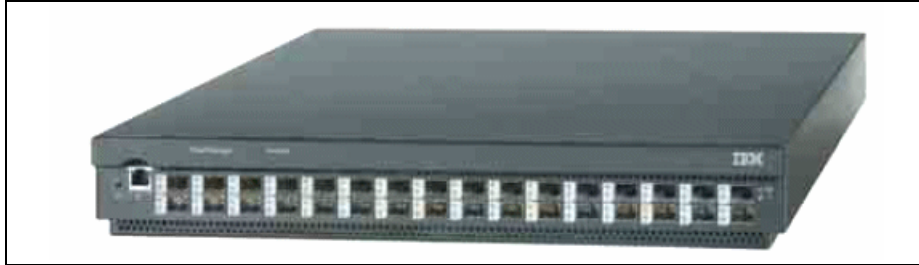


Figure 11-8 IBM TotalStorage SAN32M-1 SAN Switch

FICON In-Band Management Server supports management of the SAN32M-1 switch using System Automation for z/OS and System Automation for OS/390 and zSeries 900 servers. IBM 9032 ESCON director is a type of dynamic in-band management. System Automation—for z/OS or OS/390—is designed to concurrently manage IBM 9032 ESCON directors as well as McDATA FICON directors.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/m_type/san32m/

11.3.6 Cisco MDS 9120 and 9140 Multilayer Fabric Switch

Fabric switch provides 4-20 ports and 4-40 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers four “target-optimized” ports and sixteen “host-optimized” ports, Virtual SAN (VSAN) and Cisco Fabric Manager.

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 TCO
- ▶ 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
- ▶ Offers interoperability with a broad range of IBM servers as well as disk and tape storage devices

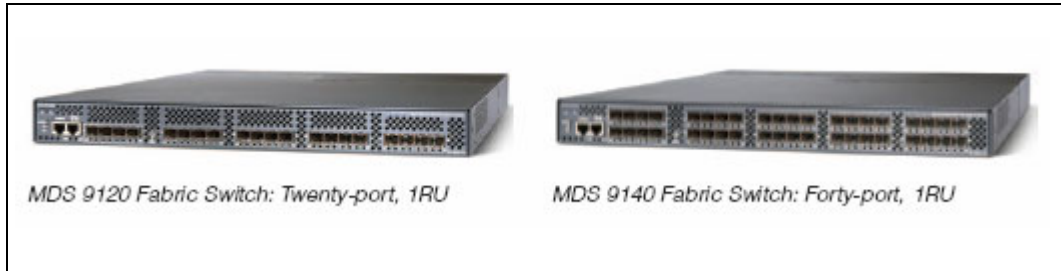


Figure 11-9 Cisco MDS 9120 and 9140 Multilayer Fabric Switch

High performance and MDS 9000 family compatibility

The Cisco MDS 9120 and 9140 Multilayer Fabric Switches are designed with 1 Gbps and 2 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 9120 and 9140 switches to 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

options Fibre Channel ports to support an auto-sensing 1 Gbps and 2 Gbps interface for high-performance connectivity and compatibility with legacy 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-1.ibm.com/servers/storage/san/c_type/9100/

11.3.7 Cisco MDS 9216 Multilayer Fabric Switch

Modular fabric switch provides 16 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX server clustering, infrastructure simplification and Business Continuity solutions. The base switch offers sixteen Fibre Channel ports, Virtual SAN (VSAN) and Cisco Fabric Manager. Features include 4 and 8-port IPS Modules with iSCSI and FCIP capabilities, 16-port FC Switch Module, 32-port FC Switch Module with “host-optimized” ports, Caching Services Module for IBM SAN Volume Controller Software, and Mainframe Package for 16 or 32 port FICON switching.

Highlights

- ▶ Supports throughput of up to 2 Gbps per port and up to 32 Gbps with each Port Channel Inter-Switch Link (ISL) connection
- ▶ Offers scalability from 14 to 48 Fibre Channel ports
- ▶ Offers two to ten Gigabit Ethernet ports for iSCSI or FCIP connectivity
- ▶ Features modular design with excellent availability capabilities
- ▶ Uses intelligent network services to help simplify Storage Area Network (SAN) management and reduce total cost
- ▶ Helps provide security for large enterprise SANs
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Offers compatibility with a broad range of IBM servers as well as disk and tape storage devices



Figure 11-10 Cisco MDS 9216 Multilayer Fabric Switch

A modular switch designed for high availability

The Cisco MDS 9216 Multilayer 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.

Multiprotocol support and traffic management features

The unique architecture of the Cisco MDS 9216 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 Storage Services Module features offer four or eight 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. The Tri-Rate Longwave SPF Transceiver feature is required for each port to be

used. The FCIP Activation for 4 or 8 Port IP Storage Services Module features adds Fibre Channel over IP (FCIP) support. **Multiprotocol Services Module** offers two Gigabit Ethernet iSCSI ports and fourteen Fibre Channel ports. FCIP Activation for the Multiprotocol Fabric Switch and for Multiprotocol Service Module features add FCIP support.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/c_type/9216/

11.4 Enterprise SAN switches

The IBM TotalStorage 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 TotalStorage 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 SAN Director M14
- ▶ IBM TotalStorage SAN16B-R multiprotocol router
- ▶ IBM TotalStorage SAN SAN140M
- ▶ IBM TotalStorage SAN16M-R multiprotocol SAN router
- ▶ IBM TotalStorage SANC40M
- ▶ IBM TotalStorage SAN256M
- ▶ Cisco MDS 9506 Multilayer Director
- ▶ Cisco MDS 9509 Multilayer Director
- ▶ IBM TotalStorage SAN256N
- ▶ CNT UltraNet Multi-service Director

In this section, we introduce IBM TotalStorage Switch solutions ideal for enterprise applications.

11.4.1 IBM TotalStorage SAN Director M14

IBM TotalStorage SAN Director M14 (2109-M14) high availability enterprise director provides 32-128 ports in a single fabric, 2 Gbps fabric switching for Windows NT/2000 and UNIX; and FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. The base director includes zoning, WEBTOOLS, Fabric Watch, ISL-Trunking and Performance Monitoring. Fabric Manager feature can simplify complex fabric management.

Highlights

- ▶ High availability director, scalable to 128-ports, is designed to support large enterprise SAN island consolidation
- ▶ Mainframe FICON, FICON Cascading, FICON and Fibre Channel intermix capabilities help support enterprise infrastructure simplification
- ▶ Intelligent fabric management simplifies deployment, management and network growth.

Advanced security with comprehensive, policy-based security capabilities can improve availability and simplify operation



Figure 11-11 IBM TotalStorage SAN Director M14

Enhanced scalability

The IBM M14 SAN Director is scalable from 32 to 128 ports in a single domain. This provides twice the scalability of the prior generation IBM M12 SAN Director. The M14 Switch Blade feature expands connectivity in sixteen port increments. A mixture of shortwave and longwave ports can be configured by adding SFP optical transceiver features.

FICON Director Operation

FICON Director switching includes FICON servers, intermixed FICON and Open servers and FICON cascading between two Directors. FICON CUP Activation provides Control Unit Port (CUP) in-band management function designed to allow mainframe applications to perform configuration, monitoring, management and statistics collection.

More information

- ▶ For additional product details, go to:

http://www-1.ibm.com/servers/storage/san/b_type/m14/

11.4.2 IBM TotalStorage SAN16B-R multiprotocol router

IBM TotalStorage SAN16B-R (2109-A16) multiprotocol router provides 8 or 16 Fibre Channel and IP ports with Fibre Channel to Fibre Channel, FC-FC Routing Service for SAN island consolidation, Fibre Channel over IP, FCIP Tunneling Service for metro and global Business Continuity solutions and iSCSI Gateway Service for low cost infrastructure simplification solutions. The 16-port router includes 8-ports enabled, WEBTOOLS, Advanced Zoning, Extended Fabric capability, iSCSI Gateway Service and router-to-router exchange based ISL-Trunking. Optional features include Ports on Demand, FC-FC Routing Service and FCIP Tunneling Service. Integrated router and switch management helps simplify installation and operation.

Highlights

- ▶ Enables SAN island consolidation for infrastructure simplification without compromising security
- ▶ Utilizes existing IP WAN infrastructures for metro and global Business Continuity solutions
- ▶ Offers iSCSI server SAN connectivity for low cost infrastructure simplification solutions

- ▶ Designed for high performance with 2 Gigabit per second (Gbps) Fibre Channel and Gigabit Ethernet (GbE) on all ports
- ▶ Integrated router and IBM TotalStorage SAN b-type (Brocade) switch management helps simplify installation and operation and helps provide switch investment protection
- ▶ Ports on Demand for pay-as-you-grow scalability from 8 to 16 ports



Figure 11-12 IBM TotalStorage SAN16B-R multiprotocol router

The IBM TotalStorage SAN16B-R multiprotocol SAN router provides iSCSI Gateway Service to integrate low-cost Ethernet-connected servers to existing SAN infrastructures. It also provides Fibre Channel, FC-FC Routing Service to interconnect multiple SAN islands without requiring the fabrics to merge into a single large SAN. This can help create a tiered enterprise SAN infrastructure without having to redesign or reconfigure the entire environment.

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.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/b_type/san16b-r/

11.4.3 IBM TotalStorage SAN140M

IBM TotalStorage SAN140M (2027-140), replaces McDATA 6064 and 6140) high availability enterprise director provides 16-140 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX; and FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. IBM TotalStorage SAN24M-1 fabric which provides FC-AL tape attachment. McDATA Enterprise Fabric Connectivity Manager provides integrated management of complex IBM SAN m-type core-to-edge fabrics.

Highlights

- ▶ Easy-to-manage enterprise infrastructure simplification and Business Continuity solutions for IBM @server xSeries, iSeries, pSeries and zSeries 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 11-13 IBM TotalStorage SAN 140M

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/m_type/san140m/

11.4.4 IBM TotalStorage SAN16M-R multiprotocol SAN router

IBM TotalStorage SAN16M-R multiprotocol SAN router (2027-R16) 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
- ▶ Offers iSCSI server SAN connectivity for low cost infrastructure simplification solutions
- ▶ Provides SAN routing over distance for metro and global Business Continuity solutions
- ▶ Designed for high throughput with 1 and 2 Gigabit per second (Gbps) Fibre Channel and Gigabit Ethernet (GbE) with Fast Write and compression

- ▶ Interoperability with IBM TotalStorage SAN m-type (McDATA) family provides switch investment protection
- ▶ Includes SANvergence Manager for router and network management



Figure 11-14 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-1.ibm.com/servers/storage/san/m_type/san16m_r/

11.4.5 IBM TotalStorage SANC40M Cabinet

IBM TotalStorage SANC40M (2027-C40, replaces McDATA Fabriccenter FC-512) provides space saving cabinet for IBM TotalStorage SAN m-type (McDATA) directors and switches. Redundant power distribution for high availability directors, space for 1U rack mount server and 39U for directors and switches and cabling flexibility to support up to 512 Fibre Channel ports.

Highlights

- ▶ Space-saving cabinet for IBM TotalStorage SAN m-type (McDATA) directors and switches
- ▶ Dual power distribution system designed for high availability
- ▶ Space for 1U rack mount management server and 39U for directors and switches
- ▶ Future ready design
- ▶ Flexible configuration options

Efficient data center design and rapidly growing Fibre Channel SAN infrastructures place a premium on data center floor space and overhead cable raceway capacity. The IBM TotalStorage SANC40 cabinet provides vertical space for 1U rack mount management server and 39U for IBM TotalStorage m-type (McDATA) directors and switches, while occupying only seven square feet (0.7 square meter) of valuable floor space. The cabinet supports internal cable management for up to 512 fiber cables and includes a 24-port Ethernet.

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/m_type/sanc40m/

11.4.6 IBM TotalStorage SAN256M

IBM TotalStorage SAN256M high availability enterprise director (2027-256) provides 10 Gbps backbone connections, 64-256 ports 2 Gbps fabric switching for Windows NT/2000 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 IBM @server xSeries, iSeries, pSeries and zSeries
- ▶ 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 Gigabit per second (Gbps) links up to 190 km
- ▶ EFCM and FICON Management Server (CUP) software can help simplify management of complex SAN infrastructures



Figure 11-15 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 Small Form-factor Pluggable, 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-1.ibm.com/servers/storage/san/m_type/san256m/

11.4.7 Cisco MDS 9506 Multilayer Director

Cisco MDS 9506 Multilayer Director (2062-D04) high availability enterprise director provides 16-128 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX and 16-64 ports FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. The base director offers Virtual SAN (VSAN), Cisco Fabric Manager and four feature slots. Features include 4 and 8-port IPS Modules with iSCSI and FCIP capabilities, 16-port FC Switch Module, 32-port FC Switch Module with “host-optimized” ports, Caching Services Module for IBM SAN Volume Controller Software, and Mainframe Package for FICON switching.

Highlights

- ▶ Provides throughput of up to 2 Gbps per port and up to 32 Gbps with each PortChannel Inter-Switch Link (ISL) connection
- ▶ Offers scalability from 14 to 128 Fibre Channel ports
- ▶ Offers 2 to 24 Gigabit Ethernet ports for iSCSI or FCIP connectivity
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Uses intelligent network services to help simplify Storage Area Network (SAN) management and reduce total cost of ownership (TCO)
- ▶ Helps provide security for large enterprise SANs
- ▶ Includes high-availability design with nondisruptive firmware upgrades
- ▶ Offers compatibility with a broad range of IBM servers as well as disk and tape storage devices



Figure 11-16 Cisco MDS 9506 Multilayer Director

Routing and load balancing capabilities

The Cisco MDS 9506 Multilayer Director includes two Supervisor Modules designed for high availability and performance. The Supervisor 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.

A single Supervisor Module has 720 Gbps of internal non-blocking switching capacity; a dual Supervisor Module configuration provides a total bandwidth of 1.44 Tbps. Because this bandwidth capability is more than twice that required for current 2 Gbps modules, the Cisco MDS 9506 has built-in capacity to support future 10 Gbps modules.

More information

- ▶ For additional product details, go to:

http://www-1.ibm.com/servers/storage/san/c_type/9506/

11.4.8 Cisco MDS 9509 Multilayer Director

Cisco MDS 9509 Multilayer Director (2062-D07) high availability enterprise director provides 32-224 port, 2 Gbps fabric switching for Windows NT/2000 and UNIX and 32-112 ports FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. The base director offers Virtual SAN (VSAN), Cisco Fabric Manager and seven feature slots. Features include 4 and 8-port IPS Modules with iSCSI and FCIP capabilities, 16-port FC Switch Module, 32-port FC Switch Module with “host-optimized” ports, Caching Services Module for IBM SAN Volume Controller Software, and Mainframe Package for FICON switching.

Highlights

- ▶ Supports throughput of up to 2 Gbps per port and up to 32 Gbps with each PortChannel Inter-Switch Link (ISL) connection
- ▶ Offers scalability from 28 to 224 Fibre Channel ports
- ▶ Offers 2 to 40 Gigabit Ethernet ports for iSCSI or FCIP connectivity
- ▶ 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



Figure 11-17 Cisco MDS 9509 Multilayer Director

More information

- ▶ For additional product details, go to:
http://www-1.ibm.com/servers/storage/san/c_type/9509/

11.4.9 IBM TotalStorage SAN256N

IBM TotalStorage SAN256N (2045-N16) high availability enterprise director provides 32-256 ports, 2 Gbps fabric switching for Windows NT/2000 and UNIX; and FICON switching for mainframe server clustering, infrastructure simplification and Business Continuity solutions. IBM TotalStorage SAN Switches provide FC-AL tape attachment. Enterprise Manager provides integrated management of IBM SAN n-type and CNT director fabrics.

Highlights

- ▶ Designed to avoid single points of failure and to provide concurrent maintenance and upgrades
- ▶ Provides ultra-scalability from 32 to 256 ports
- ▶ Designed to simplify Fibre Channel and mainframe FICON SAN consolidation using logical domain capability
- ▶ Integrated management of IBM TotalStorage SAN256N, CNT UMD and CNT FC/9000 directors in a common fabric backbone helps provide CNT director investment protection
- ▶ Provides IBM TotalStorage infrastructure simplification and Business Continuity solutions for IBM open and mainframe servers as well as Sun and HP servers



Figure 11-18 IBM TotalStorage SAN256N

More information

- ▶ For additional product details, go to:

http://www-1.ibm.com/servers/storage/san//n_type/san256.html

11.4.10 CNT UltraNet Multi-service Director

CNT UltraNet Multi-service Director (UMD) family provides ultra-scalable director-class switching for enterprise-wide storage and SAN consolidation, virtualization and automation. UMD Model N16 provides scalability from 32 to 256 ports, logical domain capability and integrated enterprise management of UMD and FC/9000 networks.

Highlights

- ▶ Designed to avoid single points of failure and to provide concurrent maintenance and upgrades
- ▶ Provides ultra-scalability from 32 to 256 ports
- ▶ Designed to simplify Fibre Channel and mainframe FICON SAN consolidation using logical domain capability
- ▶ Helps provide CNT FC/9000 director investment protection through CNT Enterprise Management System integration of UMD and FC/9000 Directors networks
- ▶ Provides improved disaster recovery solutions with extended distance SFP transceivers.
- ▶ Provides IBM TotalStorage infrastructure simplification and Business Continuity solutions for IBM open and mainframe servers as well as Sun and HP servers



Figure 11-19 CNT UltraNet Multi-service Director

UMD logical domains for SAN consolidation

The UMD physical director, with up to 256-ports in a single domain, may be partitioned, on a blade basis, into up to four logical domains, or switches. A logical domain consists of two or more I/O modules. For example, the UMD could be partitioned into three logical domains consisting of 32-port, 64-port and 128-port logical switches with the remaining 32 ports available for future expansion.

Logical domain partitioning provides the capability to scale down by segmenting the director into multiple subswitches with independent firmware loads and redundant fabric services per domain. It also provides the capability to scale up with virtual interswitch links (ISLs) between logical domains. Virtual ISLs require no physical cabling and utilize no director ports. Virtual ISLs can free up additional ports for servers and storage and do not incur any ISL latency. SAN services bundle feature is required for logical domain and virtual ISL capabilities.

Partitioning the director into logical domains can help simplify SAN consolidation by isolating different protocols, like mainframe FICON and open Fibre Channel, and different applications such as SAP and storage management systems. This isolation can help enable consolidation of multiple SANs without sacrificing availability and performance. It can also help simplify testing of new protocols by not requiring additional separate directors.

More information

- For additional product details, go to:

http://www.ibm.com/servers/storage/san/n_type/model_n16.html

11.5 SAN solutions

IBM TotalStorage 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 TotalStorage 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; and data sharing with IBM SAN File System.
- ▶ 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 TotalStorage SAN switches.

More information

- ▶ For additional IBM SAN solution information, go to:

<http://www.ibm.com/servers/storage/solutions/san/index.html>

11.5.1 IBM TotalStorage entry-level SAN solutions

An *entry-level storage consolidation solution* consists of up to ten servers attached to one 12-port SAN12M-1 fabric switch with two disk storage arrays.

A *high-availability server clustering solution* can be created with redundant switches. An entry-level server clustering solution consists of up to 10 servers, each with dual Fibre Channel adapters, cross-connected to the separate switches with two disk storage arrays.

This server clustering solution can be scaled up with the full-fabric feature to 18 servers by cascading second switches, each with two Inter-Switch Links (ISLs) for resiliency (see Figure 11-20). Additional IBM TotalStorage SAN m-type switches can be added for additional scalability.

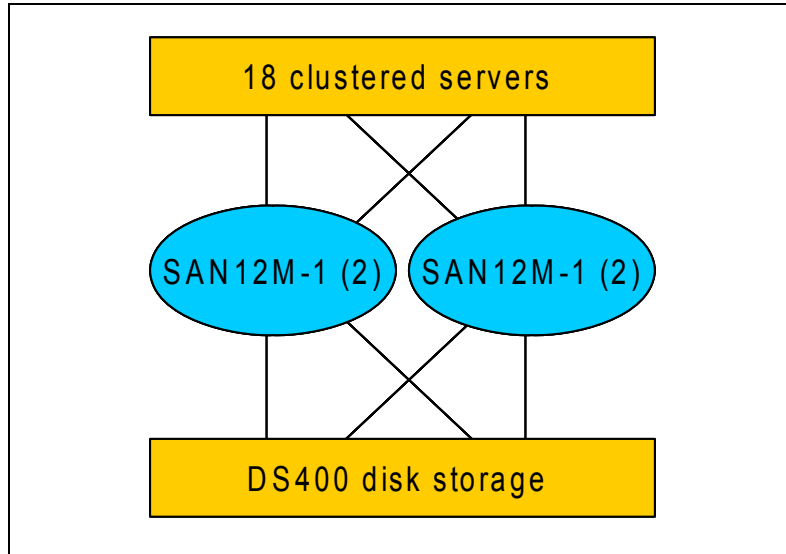


Figure 11-20 Storage consolidation and high availability server clustering solution

An entry-level data protection solution consists of up to 10 servers attached to one SAN12M-1 switch with one disk storage array and a library with one tape drive.

An expanded solution, designed to provide data protection and disaster tolerance, can be created by cascading two local-site 12-port switches, each with two ISLs for resiliency, with one remote site 4-port switch. The local and remote sites are connected with longwave transceivers that support up to 10 km distances. The local site can support up to 17 servers with two disk storage arrays and the remote site can support a library with up to three tape drives. Each switch requires the full-fabric feature for cascading and longwave transceivers.

11.5.2 IBM TotalStorage mid-range SAN solutions

A mid-range storage consolidation solution consists of up to 14 servers attached to one 16-port SAN32B-2 switch with two disk storage arrays. Ports on Demand feature is designed to enable this solution to scale-up to a mid-range 32-port switch without taking the switch offline.

A high-availability server clustering solution can be created with redundant switches. A mid-range server clustering solution consists of up to 30 servers—each with dual Fibre Channel adapters—cross-connected to redundant SAN32B-2 switches with two disk storage arrays, each with dual adapters. This server clustering solution can be scaled-up to 60 servers by cascading two additional SAN32B-2 switches, each with two 4 Gbps Inter-Switch Links (ISLs) for optimum performance and resiliency. Additional IBM TotalStorage SAN b-type switches (IBM 2005 and IBM 2109) can be added for almost unlimited scalability.

A mid-range SMB data protection solution can be created using an entry 16-port SAN32B-2 switch installed in the production site (see Figure 11-21). The switch may be scaled up with the Ports on Demand feature to 32-ports as requirements grow. The SAN32B-2 provides longwave an extended longwave SFP transceiver options. The backup site uses the SAN Switch H08 to concentrate six LTO drives onto two ISLs.

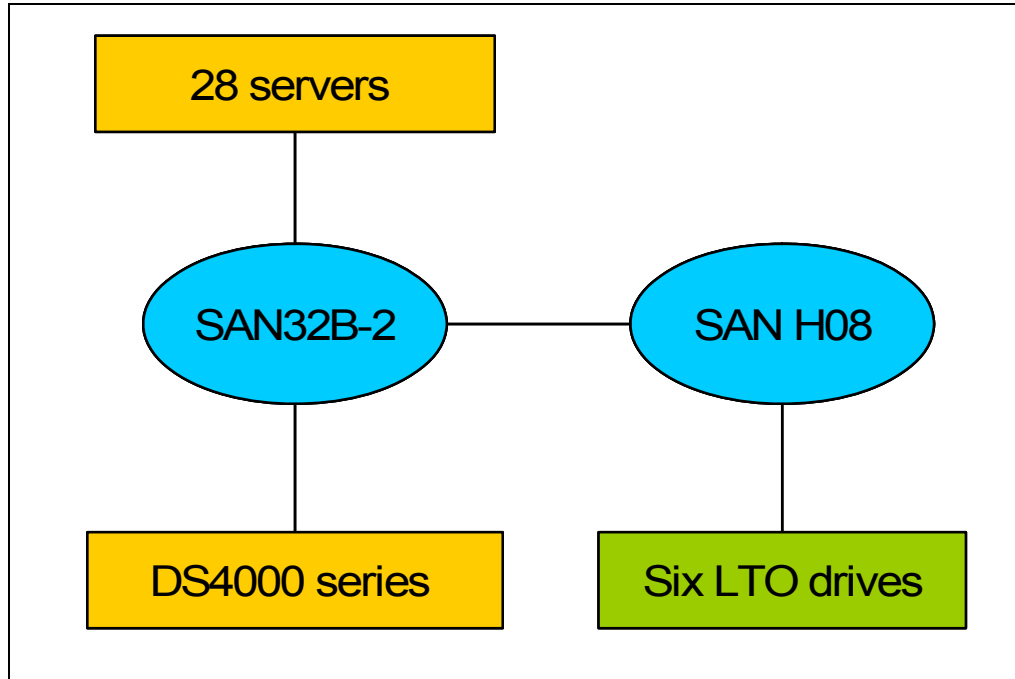


Figure 11-21 Disaster tolerant metro mirror and data protection solutions

An expanded Metro Mirror solution designed to help protect data and provide disaster tolerance can be created by cascading two 32-port SAN32B-2 switches, each with two 4 Gbps shortwave ISLs to support resiliency, in both sites. The local and recovery sites are connected with two 2 Gbps longwave ISLs that can be supported at 10 km, 35 km or 80 km distances with standard transceiver features. The Extended Fabric Activation feature helps maintain full 4 Gbps operation at distances up to 100 kilometers. The Remote Switch Activation feature enables Business Continuity solutions over wide area networks (WANs).

The local site and the remote site can support a combination of up to 60 server, disk storage array and tape library drive connections. Additional IBM TotalStorage SAN b-type switches can be added for almost unlimited scalability.

11.5.3 IBM TotalStorage enterprise SAN solutions

IBM TotalStorage *infrastructure simplification and Business Continuity solutions* with IBM SAN256N enterprise director integrate IBM @servers, IBM TotalStorage storage and software components, and IBM TotalStorage SAN switches into solutions that have been extensively tested and are designed to provide high availability, scalability, security and simplicity of management. IBM TotalStorage solutions are offered with IBM service and end-to-end solution support. These solutions may be customized by IBM Business Partners or IBM Global Services to address individual customer requirements.

The *enterprise infrastructure simplification solution* in Figure 11-22 shows two IBM SAN 256N directors cross connected to FICON and FC server and FICON and FC disk storage using DS8000. The solution supports multiprotocol SAN consolidation including IBM @server mainframe FICON and open servers Fibre Channel servers and IBM TotalStorage disk and tape devices. The SAN256N scales from 32 to 256 ports and it may be combined with IBM n-type and CNT directors, for almost unlimited scalability.

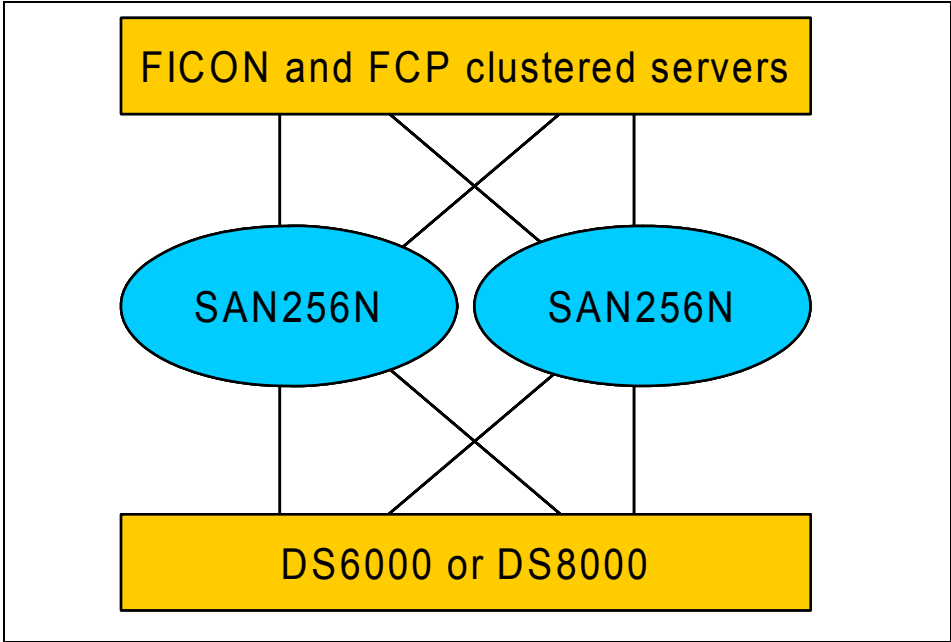


Figure 11-22 Storage and SAN island consolidation and high availability server clustering

IBM TotalStorage open software

IBM TotalStorage 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 can be divided into three categories:

- ▶ IBM TotalStorage Virtualization:
 - IBM TotalStorage SVC
 - IBM TotalStorage SFS
- ▶ IBM TotalStorage Productivity Center and Tivoli Provisioning Manager:
 - TPC data
 - TPC fabric
 - TPC disk
 - TPC replication
- ▶ IBM Tivoli Storage Manager, including:
 - TSM for Data Retention
 - TSM for Space management
 - other TSM products



IBM TotalStorage 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 TotalStorage virtualization solution.

12.1 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.

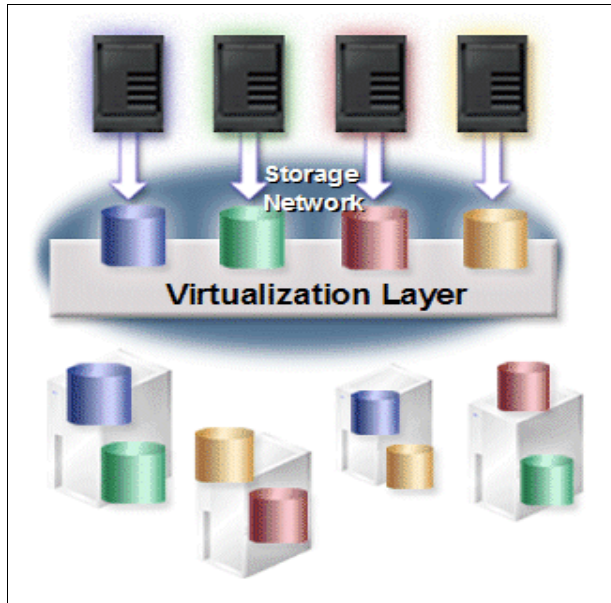


Figure 12-1 SAN with virtualization

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.

12.1.1 Levels of storage virtualization

Virtualization can be achieved at different levels in a storage network, as illustrated in Figure 12-2.

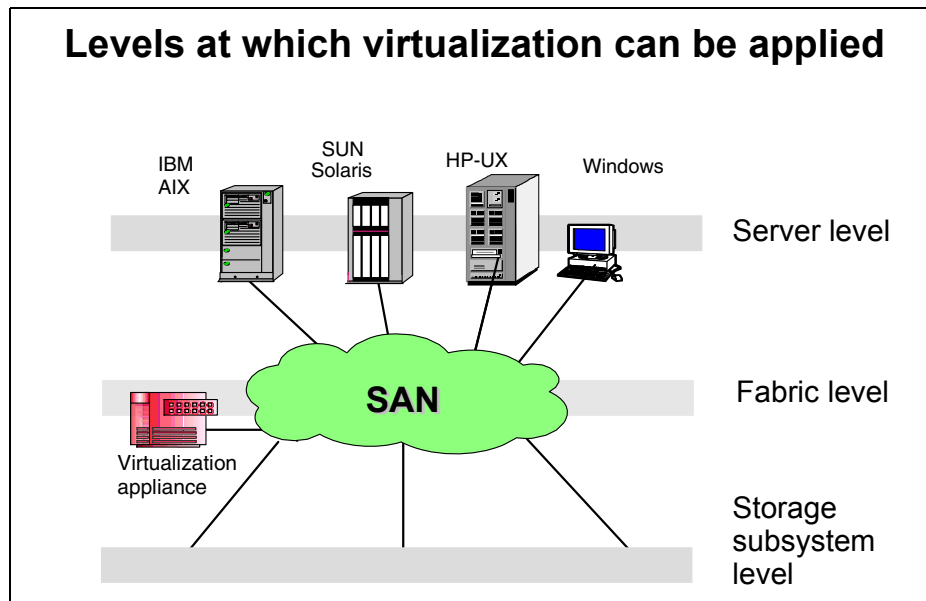


Figure 12-2 Levels of application of storage virtualization

Server level

Abstraction at the server level is by means of the logical volume management of the operating systems on the servers. 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 12-3. These models are not mutually exclusive. In many environments a combination of both may be desired.

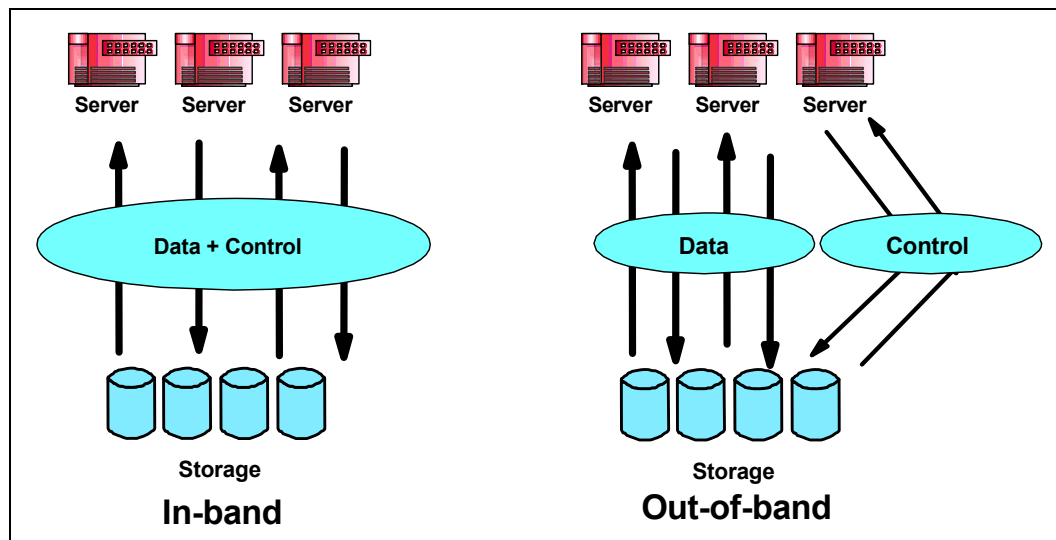


Figure 12-3 In-band and out-of-band virtualization models

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 TotalStorage **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

The IBM TotalStorage **SAN File System** is a SAN-wide file system for accessing data on storage networks across multiple application operating system platforms and heterogeneous storage devices. The SAN File System also provides centralized, policy-based management of the data in the SAN. It is an example of an out-of-band virtualization implementation.

12.2 IBM TotalStorage 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 TotalStorage Enterprise Storage Server (ESS). 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*.

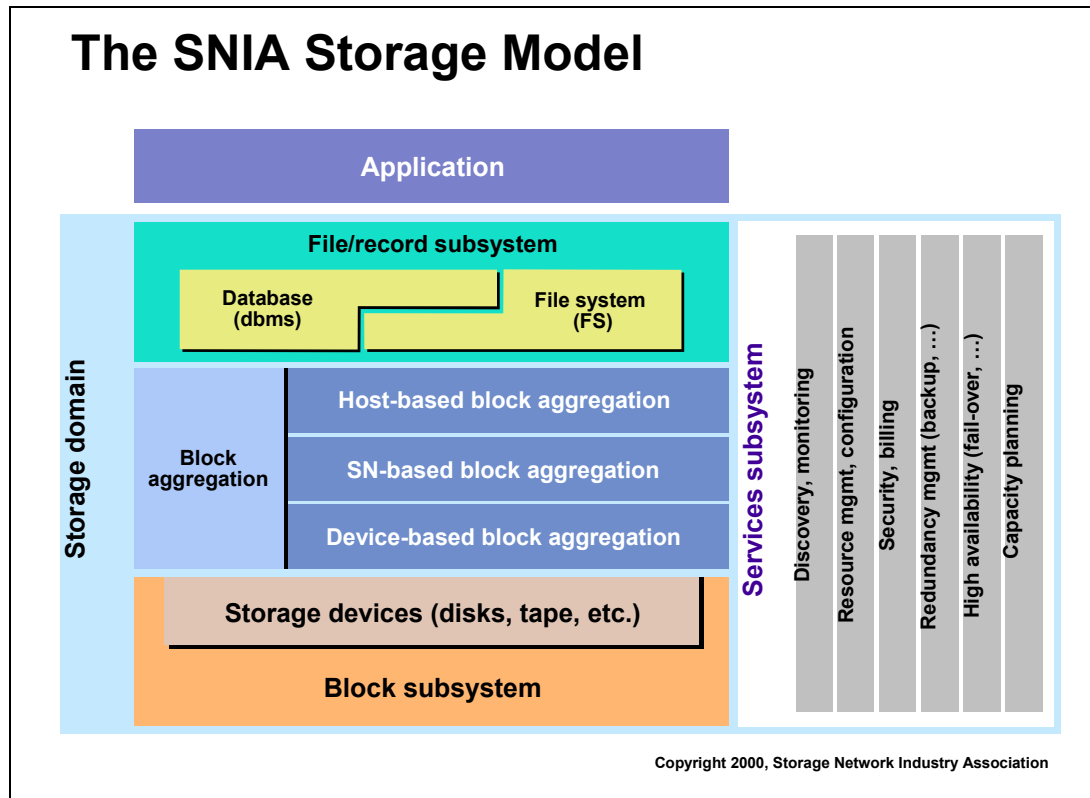


Figure 12-4 SNIA storage model

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 12-4, 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 TotalStorage 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. IBM's 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.

12.3 IBM TotalStorage virtualization family

The IBM TotalStorage 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 TotalStorage Virtualization family of products can help you to:

- ▶ Simplify the deployment and administration, while increasing utilization, of physical storage asset:

IBM TotalStorage SAN Volume Controller

- ▶ Bring local data (files and databases) management capabilities to the entire SAN using:

IBM TotalStorage SAN File System

- ▶ 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 IBM's 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 TotalStorage 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.

12.4 More information

For more information about the IBM TotalStorage Virtualization family, refer to these links:

<http://www.ibm.com/totalstorage/virtualization>
<http://www.storage.ibm.com/software/roadmap/index.html>

For more information on SAN virtualization, refer to the Redpaper *Virtualization in a SAN*, which can be found at:

<http://www.redbooks.ibm.com/redpapers/pdfs/redp3633.pdf>

For more information on *SNIA Shared Storage Model*, refer to:

http://www.snia.org/tech_activities/shared_storage_model/



SAN Volume Controller

IBM TotalStorage SAN Volume Controller (SVC) is designed to reduce the complexity and costs of managing storage networks. It allows users to virtualize their storage and helps increase the utilization of existing capacity and centralize the management of multiple controllers in an open-system SAN environment. The IBM TotalStorage 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 TotalStorage SAN Volume Controller (SVC).

13.1 IBM TotalStorage SAN Volume Controller (SVC)

The IBM TotalStorage SAN Volume Controller is designed to allow the use of storage resources more efficiently, helping to maximize return on investment (ROI), improve personnel productivity and increase application availability.

13.1.1 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 TotalStorage 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 TotalStorage SAN Volume Controller is designed to:

- ▶ Combine storage capacity from multiple vendors into a single reservoir of capacity that can be managed from a central 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

SAN Volume Controller combines hardware and software into a comprehensive, modular appliance. Using xSeries server technology in highly reliable clustered pairs, SVC is designed to avoid single points of failure. SAN Volume Controller 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.

High scalability

An I/O group is formed by combining a pair of high-performance, redundant Intel processor-based servers. Each I/O group contains 4 GB of mirrored cache memory. Highly available I/O groups are the basic configuration of a cluster. Adding another I/O group can help increase cluster performance and bandwidth.

At its base level, SAN Volume Controller contains a single I/O group. It can scale-up to support four I/O groups. For every cluster, SAN Volume Controller supports up to 4096 virtual disks.

Personnel productivity

SAN Volume Controller is designed to help improve administrator productivity by enabling management at the cluster level, and it is designed to provide a single point of control over all the storage it manages.

SAN Volume Controller provides a comprehensive, easy-to-use graphical interface for central management. This simple interface incorporates the Storage Management Initiative Specification (SMIS) application programming interface (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 disparate storage controllers. SAN Volume Controller is designed to allow administrators to map disk storage volumes to virtual pooled volumes to help better use existing storage.

Application availability

By pooling storage into a single reservoir, SAN Volume Controller is designed to help insulate host applications from physical changes to the storage pool. This helps allow applications to continue to run without disruption.

SAN Volume Controller includes a dynamic data-migration function that is designed to help administrators migrate storage from one device to another, without taking it offline. This helps administrators to reallocate and scale storage capacity without disrupting applications.

The solution supports both local area network (LAN)-free and server-free backups. Through the IBM FlashCopy function, administrators can copy point-in-time mission critical data to lower cost storage devices, such as Serial Advanced Technology Attach (SATA) devices. SAN Volume Controller also supports IBM TotalStorage Multipath Subsystem Device Driver (SDD). This mature multipathing software is designed to provide failover and load-balancing capabilities.

Tiered storage

In most IT environments, inactive data makes up the bulk of stored data. SAN Volume Controller is designed to help administrators control storage growth more effectively by moving low-activity or inactive data into a hierarchy of lower-cost storage. And administrators can free disk space on higher-value storage for more-important, active data. SAN Volume Controller is designed to enable you to match the cost of storage to the value of data.

13.2 Glossary of commonly used terms

Here is a short glossary of the most commonly used terms (in alphabetic order) to help better understand the TotalStorage SAN Volume Controller.

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

SAN Volume Controller 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 SAN Volume Controller.

Grain

A grain is the unit of data represented by a single bit in a FlashCopy bitmap, 256 K in SAN Volume Controller.

I/O group

An input/output (I/O) group contains two SAN Volume Controller nodes defined by the configuration process. Each SAN Volume Controller 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 an 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 TotalStorage Enterprise Storage Server (ESS), IBM TotalStorage DS4000, DS6000 and DS8000 series 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 on which the software used to manage the SAN Volume Controller runs.

Node

Node is the name given to the individual servers in an SAN Volume Controller cluster on which the SAN Volume Controller software runs.

SAN Volume Controller

The SAN Volume Controller 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 a SAN Volume Controller device that appears to host systems attached to the SAN as a SCSI disk. Each vDisk is associated with exactly one I/O group.

13.2.1 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 that require minimal porting effort to leverage a new hardware platform. 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. Performance growth over time is assured by the ability to port Compass to just about any platform and remain current with the latest processor and chipset technologies on each.

The IBM TotalStorage SAN Volume Controller is a modular solution that consists of a Master Console for management, up to eight cluster nodes (added in pairs), and dual UPS for write cache data protection (Figure 13-1). 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 GB of high-speed memory, which serves as the cluster cache. A management card is installed in each server to monitor various parameters which the cluster uses to determine the optimum and continuous data path. The cluster is protected against data loss by uninterruptible power supplies.

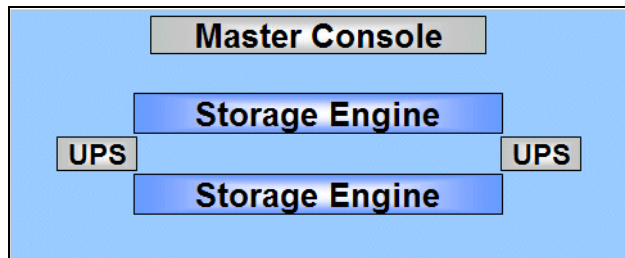


Figure 13-1 SVC components

The storage engines (or storage nodes) are always installed in pairs and combined into a high availability cluster.

During normal operation of the cluster 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. Should a node fail 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).

Note: Although the SAN Volume Controller 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 SAN Volume Controller application software.

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 will represent the preferred path for I/O to a given vDisk - the other node representing 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.

Beyond automatic configuration and cluster administration, the data transmitted from attached application servers is also treated in the most reliable manner. When data is written by the host, the preferred node within the I/O Group stores a write in 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 in the event of a node failure, the

surviving node empties its write cache and proceeds in write-through mode until the cluster is returned to a fully operational state.

13.2.2 SVC virtualization

The SVC provides block aggregation and volume management for disk storage within the SAN. In simpler terms, 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 conflict between SVC and the application servers both trying to manage the back-end storage.

The SVC I/O Groups are connected to the SAN in such a way 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 mDisks are collected into groups, known as 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 13-2.

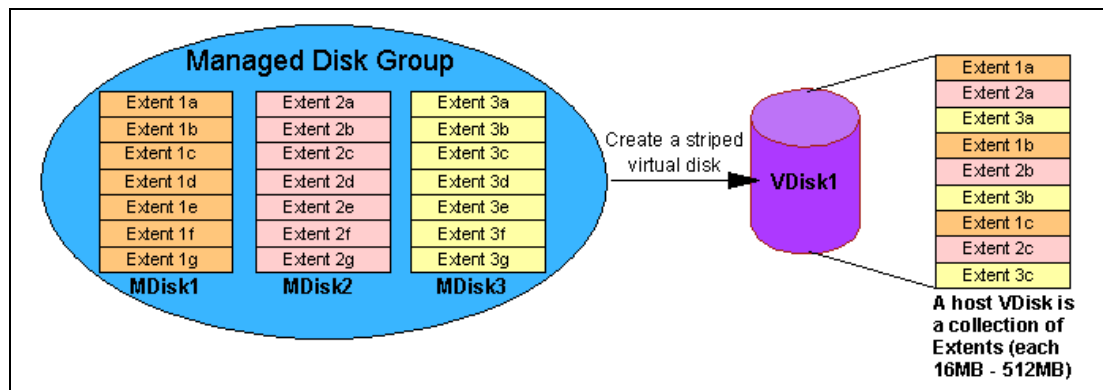


Figure 13-2 Extents being used to create a vDisk

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. Figure 13-3 on page 227 summarizes the various relationships that bridge the physical disks through to the virtual disks within the SVC architecture.

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's Subsystem Device Driver (SDD). Note that the SDD code has been updated to support both the SVC and the ESS, and provided the latest version is used, IBM supports the concurrent connections of a host to both an SVC and "native" ESS environment.

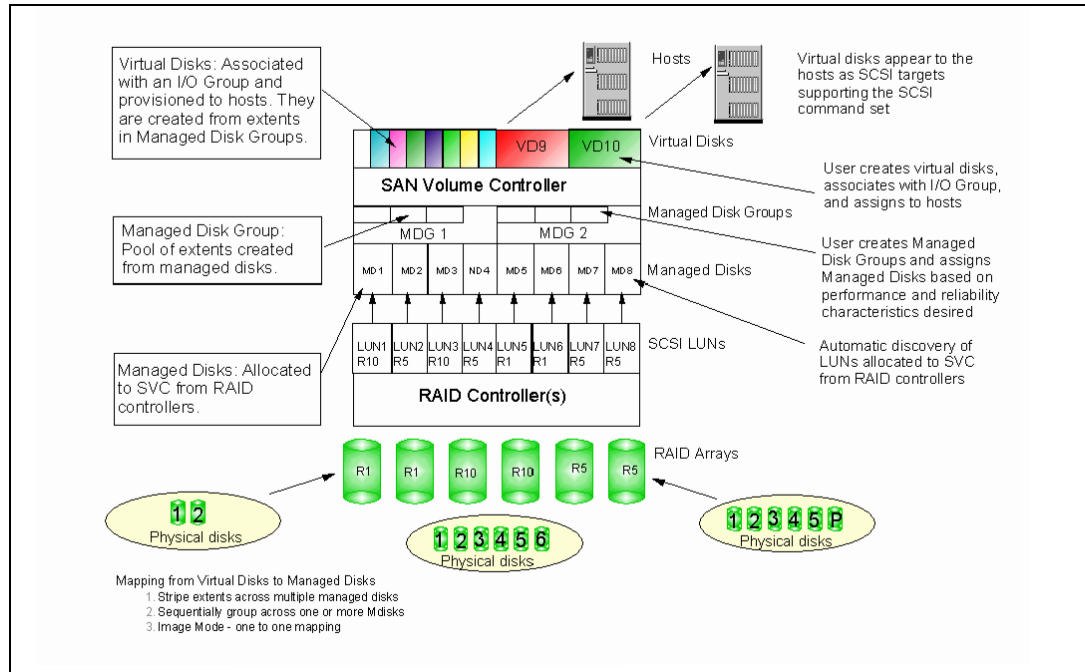


Figure 13-3 Relationship between physical and virtual disks

.An example of a SVC configuration is shown in Figure 13-4.

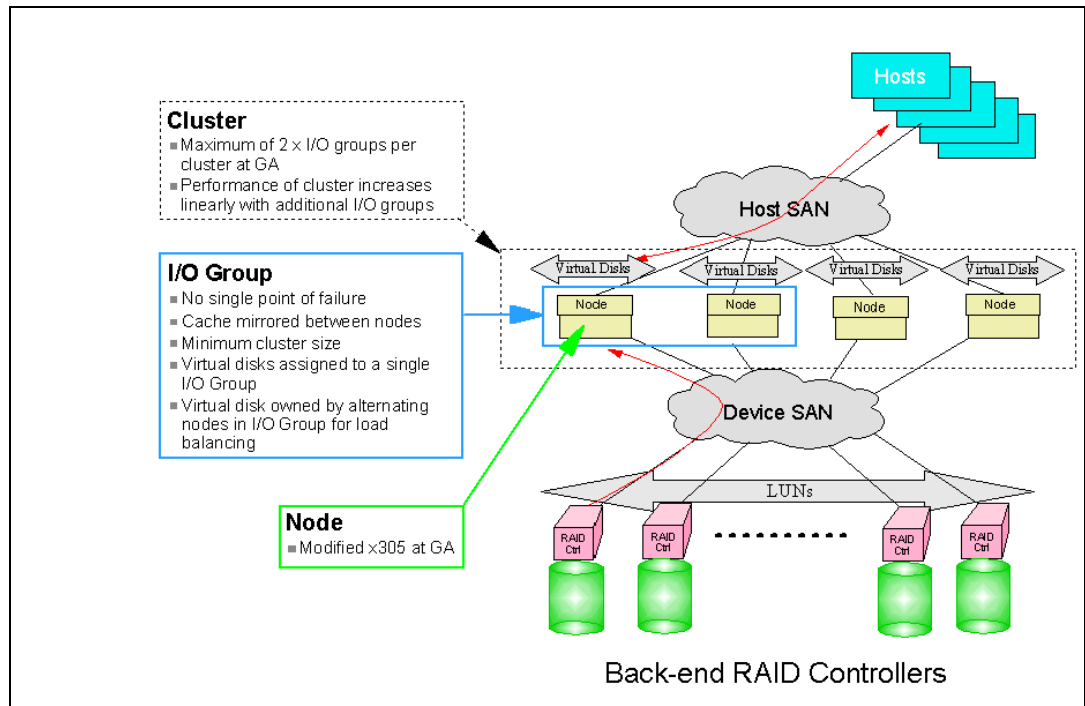


Figure 13-4 SVC logical view

Notes:

- ▶ The Fibre Channel SAN connections between the SAN Volume Controller and the switches are optical fibre preferably running at 2 Gbps. However, the SAN Volume Controller is also supported in 1 Gbps Fibre Channel fabrics.
- ▶ To provide high availability, the SAN Volume Controller nodes should be configured in redundant SAN fabrics
- ▶ The Fibre Channel switches need to be zoned to permit the hosts to see the SAN Volume Controller nodes and the SAN Volume Controller nodes to see the RAID Controllers. The SAN Volume Controller nodes within a cluster must be able to see each other and the master console. In addition, if there are two SAN Volume Controller clusters with Metro Mirror (formerly referred to as Peer-to-Peer Remote Copy (PPRC)) services between them, zoning must be set so that all the nodes in both clusters see all the other nodes in both clusters.
- ▶ In addition to 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.

13.2.3 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.

- ▶ 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.

- ▶ The amount of virtualized storage to Metro Mirror to another system.

The amount of virtualized storage that you want to Metro Mirror to another system: You may be managing 8 TB of storage and FlashCopy copying 1 TB (2 TB Fibre Channel 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.

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.

13.2.4 SVC usage

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.

Replication

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 synchronous PPRC eliminates this problem making it significantly simpler to deploy for disaster recovery.

Data migration

It takes long time to complete migration from one storage device to another using the current SAN available technologies. It involves lots of down time as workloads are shifted. Some storage systems do not support remote copy function forcing host-based volume drain, back up to tapes or restore from tapes during the migration. In addition data is unavailable during the migration process. See Figure 13-5 which explains what is involved in data migration using the SAN technology.

- ▶ 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

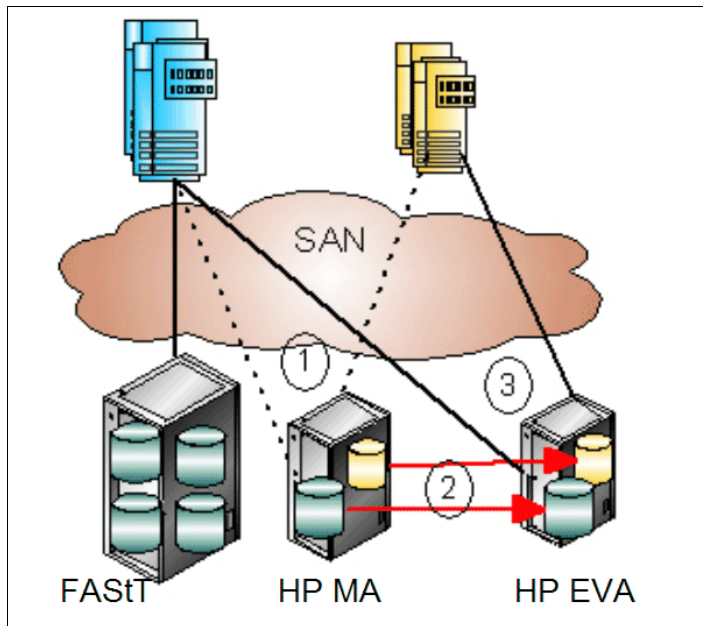


Figure 13-5 Data migration using SAN

Using SVC's block virtualization, data is migrated without an application outage due to the virtual mapping capability of the SVC (Figure 13-6 on page 230). This architecture could also be used for better lifecycle management by not only migrating off data but also to move data to different device classes as the characteristics of its management policies change.

As shown in Figure 13-6, servers have access to virtual disks allowing migration of managed disks and updates to mapping tables to occur in the background. After migration, old arrays are reused appropriately (that is, non mission-critical data). The SVC facilitates nondisruptive maintenance, upgrades, and data backups.

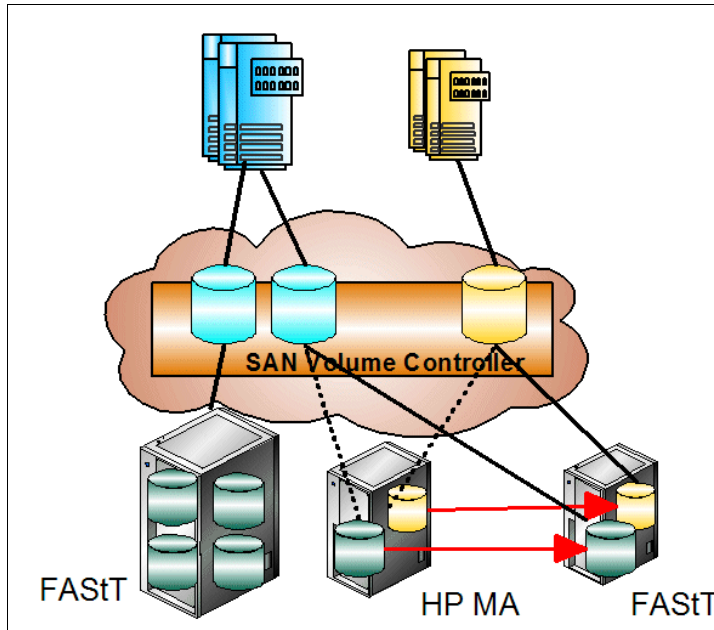


Figure 13-6 Data migration using SANVC

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 SANVC) 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.

13.2.5 SVC compatibility

The SAN Volume Controller is capable of supporting Windows NT, 2000 and 2003, AIX, SUSE SLES 8 and Red Hat EL AS Linux, Novell NetWare 6.5 with clustering, VMWare ESX, Sun Solaris, and HP-UX hosts. The SAN switch support is also broad and includes the IBM TotalStorage SAN Switches and members of the McDATA, CNT, and Cisco families of SAN Switches and Directors.

Support of disk subsystems includes the IBM TotalStorage DS4000, DS6000 and DS8000 series servers and the IBM TotalStorage Enterprise Storage Server. It also supports several models of Hitachi Thunder and Lightning, EMC CLARiiON and Symmetrix, and Hewlett-Packard StorageWorks Modular Arrays. Future releases of the SAN Volume Controller will include support for other third-party disk subsystems from manufacturers such as HDS, HPQ and EMC.

Refer to:

<http://www-1.ibm.com/servers/storage/support/software/sanvc/installing.html>

13.3 More information

For more information, refer to these Web sites and Redbooks.

- ▶ For the latest SVC product information, go to:

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

- ▶ Redbook *IBM TotalStorage SAN Volume Controller*, SG24-6423

<http://www.redbooks.ibm.com/redpieces/abstracts/sg246423.html?open>

- ▶ Redbook *IBM TotalStorage, Introducing the SAN Volume Controller and SAN Integration Server*, SG24-6423

<http://www.redbooks.ibm.com/abstracts/sg246423.html?open>

- ▶ Redbook *IBM TotalStorage: Integration of the SAN Volume Controller, SAN Integration Server and the SAN File System*, SG24-6097

<http://www.redbooks.ibm.com/abstracts/sg246097.html?open>



SAN File System

The IBM TotalStorage SAN File System (SFS) is designed to simplify file and data management in your storage area network by consolidating file systems across UNIX, Windows, and Linux servers. With the SAN File System, files are viewed and managed as a centralized IT resource with a single point of administrative control. By providing a common SAN-wide file system, SFS can support secure, heterogeneous data sharing and centralized policy based storage management in an open environment.

This chapter introduces IBM TotalStorage SAN File System and discusses its architecture and key features.

14.1 Overview

The IBM TotalStorage SAN File System architecture makes it possible to bring the benefits of the existing mainframe system-managed storage (DFSMS) to the SAN environment. Features such as policy-based allocation, volume management, and file management have long been available on IBM mainframe systems. However, the infrastructure for such centralized, automated management has been lacking in the open systems world of Linux, Windows, and UNIX. On conventional systems, storage management is platform dependent. IBM TotalStorage SAN File System provides a single, centralized point of control to better manage files and data, and is platform independent. Centralized file and data management dramatically simplifies storage administration and lowers the Total Cost of Ownership (TCO).

SAN File System is a common file system specifically designed for storage networks. By managing file details (via a metadata controller) on the storage network instead of in individual servers, the SAN File System design moves the file system intelligence into the storage network where it can be available to all application servers. Figure 14-1 shows the file level virtualization aggregation, which provides immediate benefits: a single global namespace and a single point of management. This eliminates the need to manage files on a server-by-server basis. A global namespace is the ability to access any file from any client system using the same name.

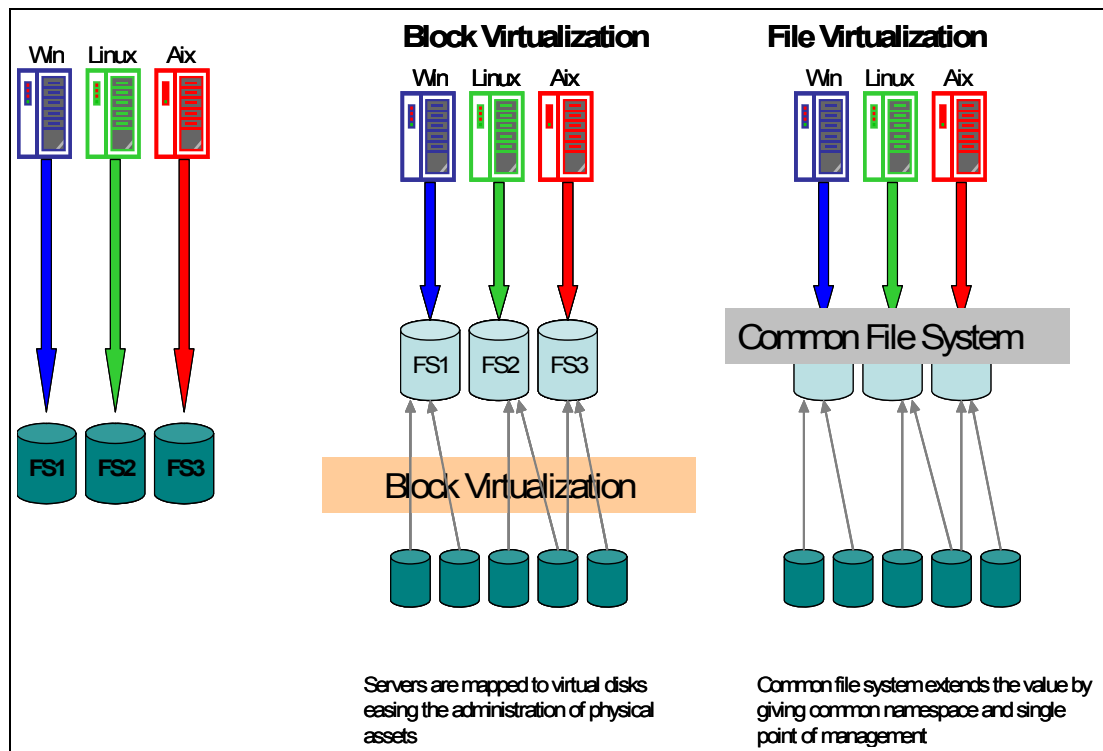


Figure 14-1 File level virtualization

IBM TotalStorage SAN File System automates routine and error-prone tasks, such as file placement, and monitors out of space conditions. IBM TotalStorage SAN File System allows true heterogeneous file sharing — where read and write on the exact same data can be done by different operating systems.

IBM TotalStorage SAN File System is configured in a high availability configuration with clustering for the metadata servers, providing redundancy and fault tolerance.

IBM TotalStorage SAN File System is designed to provide policy-based storage automation capabilities for provisioning and data placement, nondisruptive data migration, and a single point of management for files on a storage network.

14.2 What a SAN File System is

File systems can be broadly classified into:

- ▶ Local file system
- ▶ LAN file system
- ▶ SAN file system

Local file system

A local file system is tightly integrated with the operating system, and is therefore usually specific to that operating system. All data and metadata are served over the system's internal I/O path. Some examples of local file systems are NTFS, Ext2-fs, AIX JFS.

LAN file system

LAN file system allows computers attached via a LAN to share data. They use the LAN for both data and metadata flow. Some LAN file systems also implement a global namespace, like Andrew file system (AFS®). Examples of LAN file systems are Network file system (NFS), AFS, Distributed file system (DFS™), and Common Internet file system (CIFS).

SAN file system

SAN file system allows 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. The IBM TotalStorage SAN File System, as the name suggests, is a SAN File System.

Figure 14-2 shows the different file systems types.

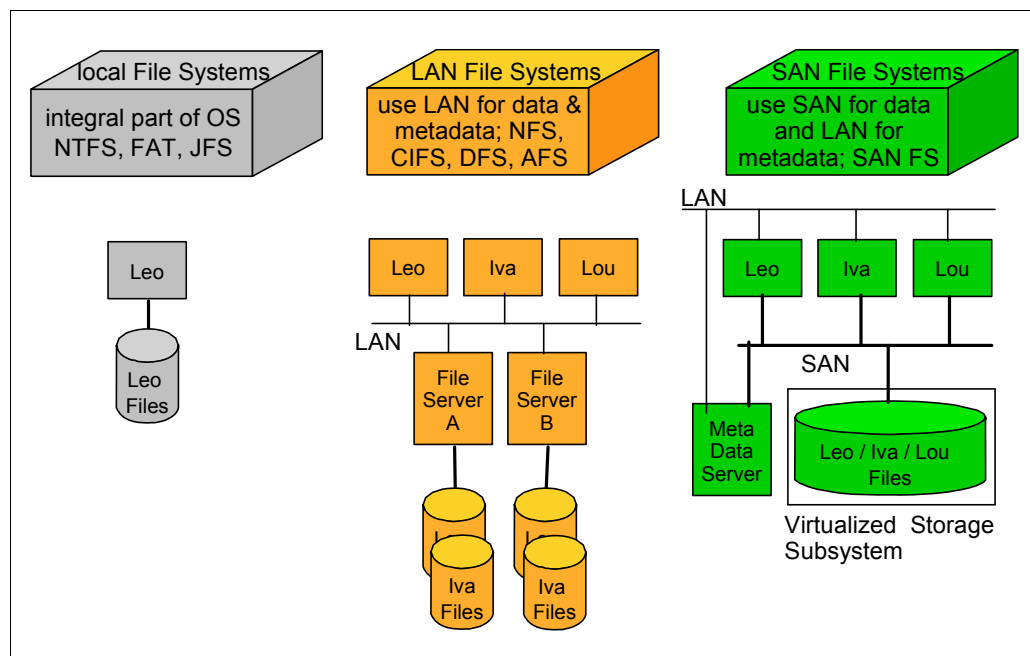


Figure 14-2 File system types

14.3 SAN File System architecture

The SAN File System differs from conventional distributed file systems in that it uses a data-access model that requires clients to contact servers to obtain only information they need to locate and access the storage devices (metadata), not the data itself. The SAN File System clients access data directly from storage devices via their own direct connection to the SAN. Direct data access eliminates server bottlenecks and provides the performance necessary for data-intensive applications. Thus, SAN File System does separation of data and metadata. Metadata operations are done over IP network and data operations happen over SAN. This is called “out-of-band virtualization” in which the data flow is separated from the control flow.

SAN File System architecture and components are illustrated in Figure 14-3.

In Figure 14-3, we show five clients, each running a SAN File System currently supported client operating system.

A SAN File System client is a computer that accesses and creates data that is stored in the SAN File System global namespace. The SAN File System client software enables them to access the global namespace through a virtual file system (VFS) on UNIX/Linux systems and an installable file system (IFS) on Windows systems. This layer (VFS/IFS) is built by the OS vendors for use specifically for special-purpose or newer file systems.

The machines which are called clients in SAN file system terms, will actually act as database servers, application servers or file servers in the enterprise.

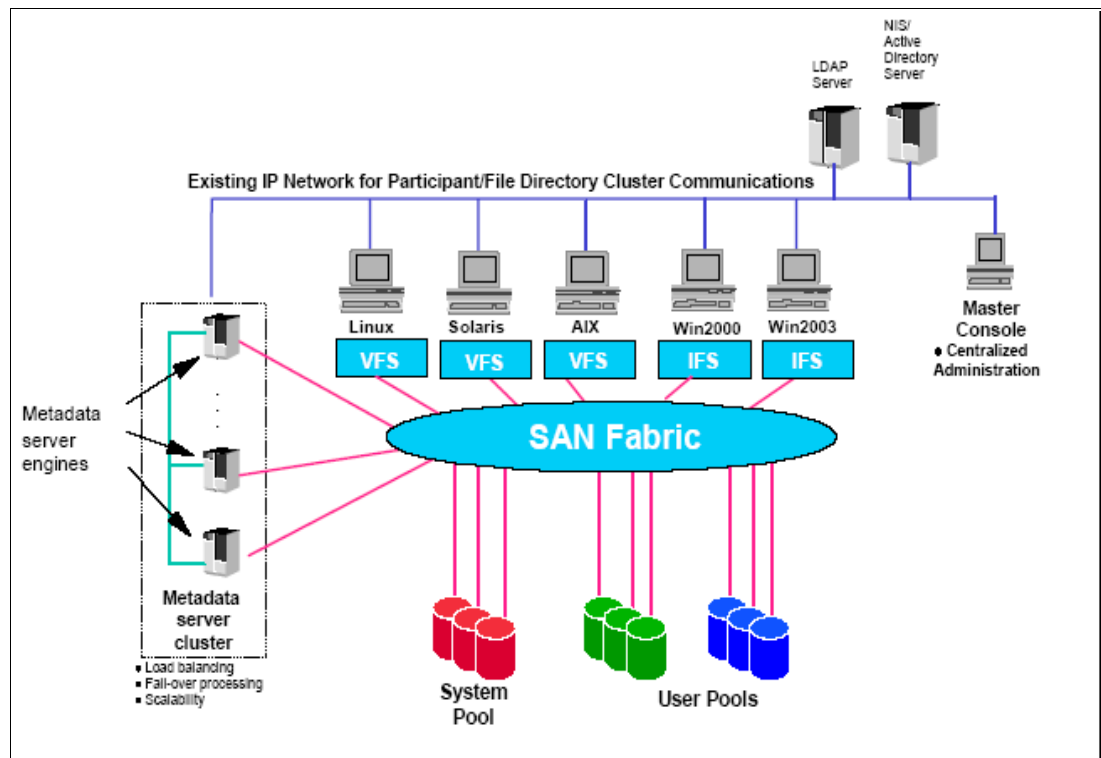


Figure 14-3 SAN File System architecture

Metadata server (MDS) engines, that run the Metadata server software are shown on the left side of the figure. The MDS engine is the physical hardware (which is based on the xSeries platform) on which the MDS and administrative servers run. It runs SUSE Linux operating system.

A metadata server (MDS) is a software server that runs on an engine and performs metadata, administrative, and storage management services.

MDSs are clustered for scalability and availability of metadata operations and are often referred to as the Metadata server cluster. In a SAN File System server cluster, there is one master Metadata server and one or more subordinate Metadata servers. Each MDS runs on a separate physical engine in the cluster. Additional Metadata servers can be added as required if the workload grows — providing solution scalability.

Figure 14-3 on page 236 also shows user pools that store the SAN File System clients' user/file data and System pool that stores metadata.

The Administrative server allows SAN File System to be remotely monitored and controlled through a Web-based user interface, called the SAN File System console. The Administrator server also processes requests issued from an administrative command line interface (CLI - which can also be accessed remotely). This means the SAN File System can be administered from almost any system with suitable TCP/IP connectivity. The Administrative server uses an LDAP server (customer supplied) to look up authentication and authorization information about the administrative users. The primary Administrative server runs on the same engine as the master Metadata server. It receives all requests issued by administrators and also communicates with Administrative servers that run on each additional server in the cluster to perform routine requests.

Figure 14-4 shows the overall administrative interface structure of SAN File System.

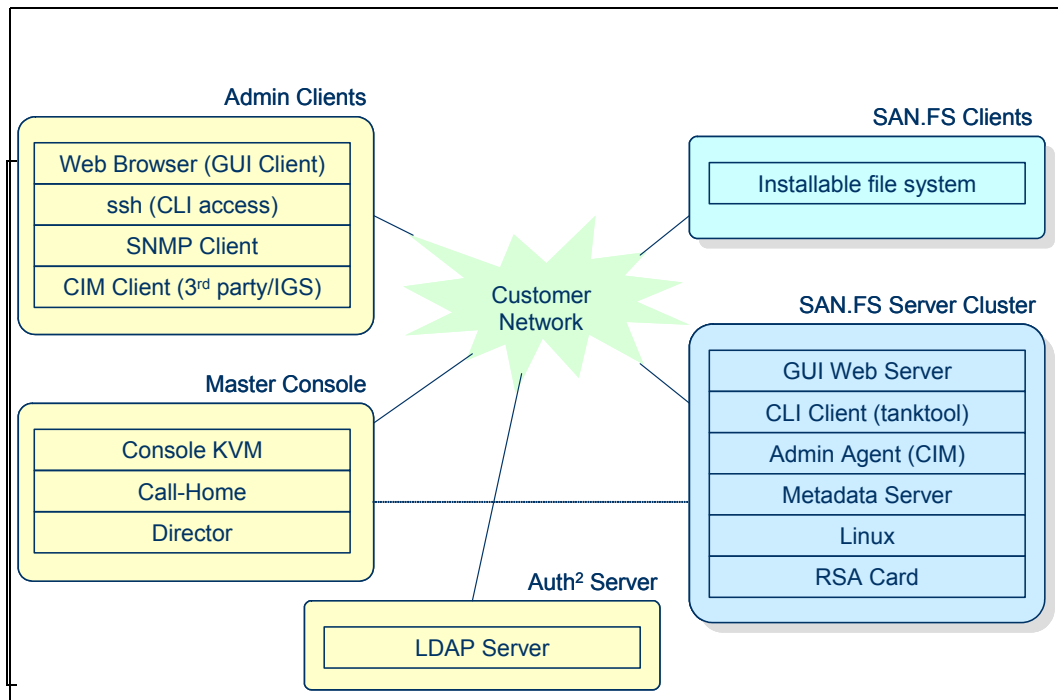


Figure 14-4 SAN File System administrative structure

The SAN File System software includes a task oriented GUI (SAN File System Console) that can be accessed from a Web browser, a command line interface that can be used to create scripts for frequently used commands, and an application programming interface that can be used by other storage and data management applications to manage or monitor the SAN File System. SAN File System administrative structure. The Application Programming Interface (API) is based on the Common Information Model (CIM) and is encoded in XML with common

methods and syntax to the Storage Networking Industry Association's (SNIA) Storage Management Interface Specification (SMI-S).

Note: For more information on SMI-S and CIM, refer to the Appendix A, "Introduction to SMIS and CIM/WBEM" on page 459.

Figure 14-5 shows the SAN File System Console GUI.

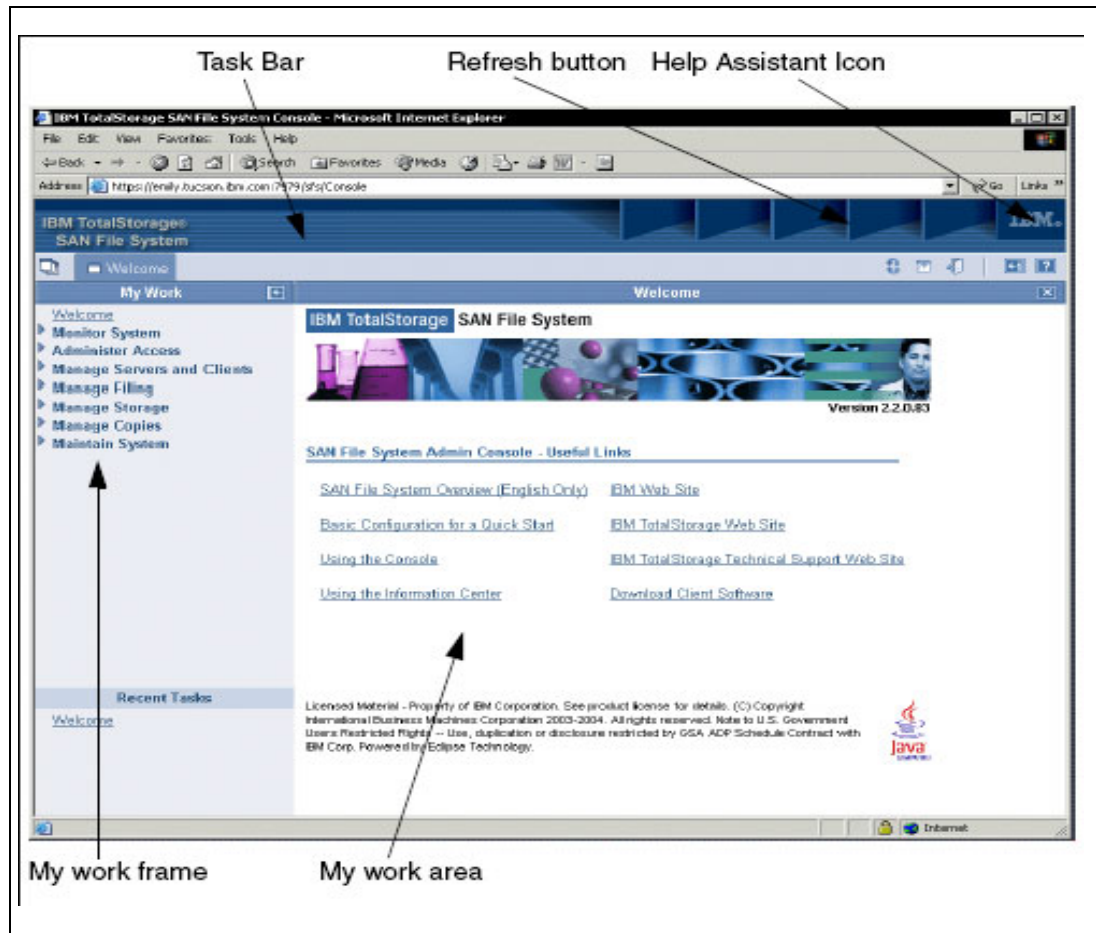


Figure 14-5 GUI welcome screen

14.4 SAN File system terminology

In this section we explain some of the commonly encountered terms used in the SAN File System environment.

Global namespace

A key concept for the SAN File System is the *global namespace*. Traditional file systems and file sharing systems operate separate namespaces — that is, each file is tied or mapped to the server which hosts it, and the clients must know which server this is. For example, in Figure 14-2, in a LAN File System, user Iva has files stored both on File Server A and File Server B. She would need to specify the particular File Server in the access path for each file.

The SAN File System, by contrast, presents a global namespace - there is one file structure (subdivided into parts called filesets) which is available simultaneously to all the clients. A global namespace is the ability to access any file from any client system using the same name.

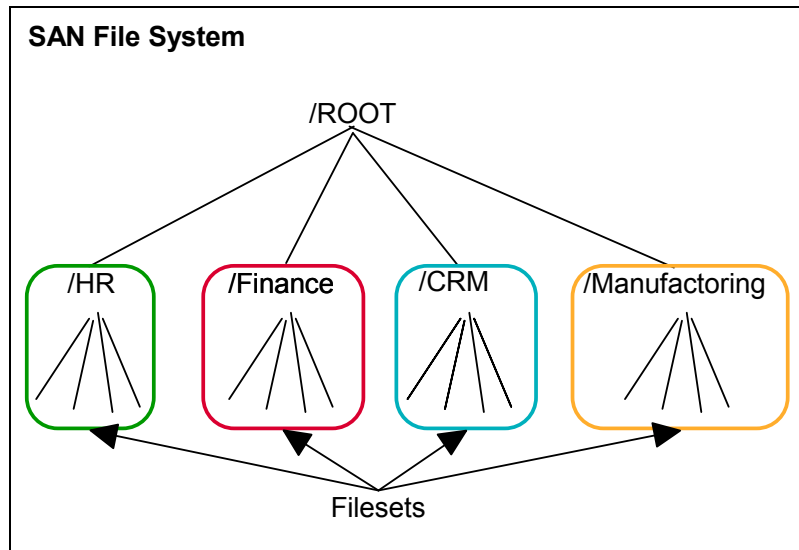


Figure 14-6 Global namespace

Fileset

Filesets are subsets of the entire global namespace, and serve to organize the namespace for all the clients. A fileset serves as the unit of workload for the metadata server (MDS). Each MDS has a number of filesets assigned to it.

To the clients, the filesets appear as normal directories where they can create their own subdirectories, place files, and so on. But from the SAN File System server perspective, the fileset is the building-block of the global namespace structure, which can only be created and deleted by SAN File System administrators.

Storage pool

A storage pool is a collection of SAN File System volumes that can be used to store either metadata or file data. A storage pool consists of one or more volumes (LUNs from the back-end storage system perspective) that provide for example, the desired quality of service. An administrator must assign one or more volumes to a storage pool before it can be used. There are two types of storage pools:

- ▶ System pool
- ▶ User pool

System pool

The system pool contains the system metadata (system attributes, configuration information, and metadata server state) and file metadata (file attributes and locations) that are accessible to all MDSs in the server cluster. There is only one system pool that is created automatically when the SAN File System is installed with volumes that are specified as input parameters during the install process.

User pool

User pools contain the blocks of data that make up user files. Administrators can create one or more user storage pools, and then create policies that contain rules that cause the MDS

servers to store data for specific files in the appropriate storage pools. A special user pool is the default user pool. This is used to store the data for a file if the file is not assigned to a specific storage pool by a rule in the active policy. One user pool is automatically designated as the default user pool, and it is created when the SAN File System is installed. This can be changed by creating another user pool, and setting it to the default user pool.

Volume

A volume is the principal component of a pool. The nature of the volume is dependent on what kind of storage subsystem is connected to the SAN File System through a SAN. In the most common sense, a volume is represented by a LUN (vDisk in a SVC) accessed by the MDS or the client.

Master console

The Master Console provides a remote support interface (remote access) and service alert (call home) capability. In the SAN File System V2.2 the Master Console is an optional component.

Policy

A policy is a list of rules that determines where the data for specific files is to be stored. A rule is an SQL-like statement that tells a SAN File System MDS the specific storage pool to place the data for a file. A rule can apply to any file being created, or only to files being created within a specific fileset, depending on how it is defined.

The rules in a file placement policy are processed in order until the condition in one of the rules is met. The data for the files is then stored in the specified storage pool. If none of the conditions specified in the rules of the policy is met, the data for the file is stored in the default storage pool.

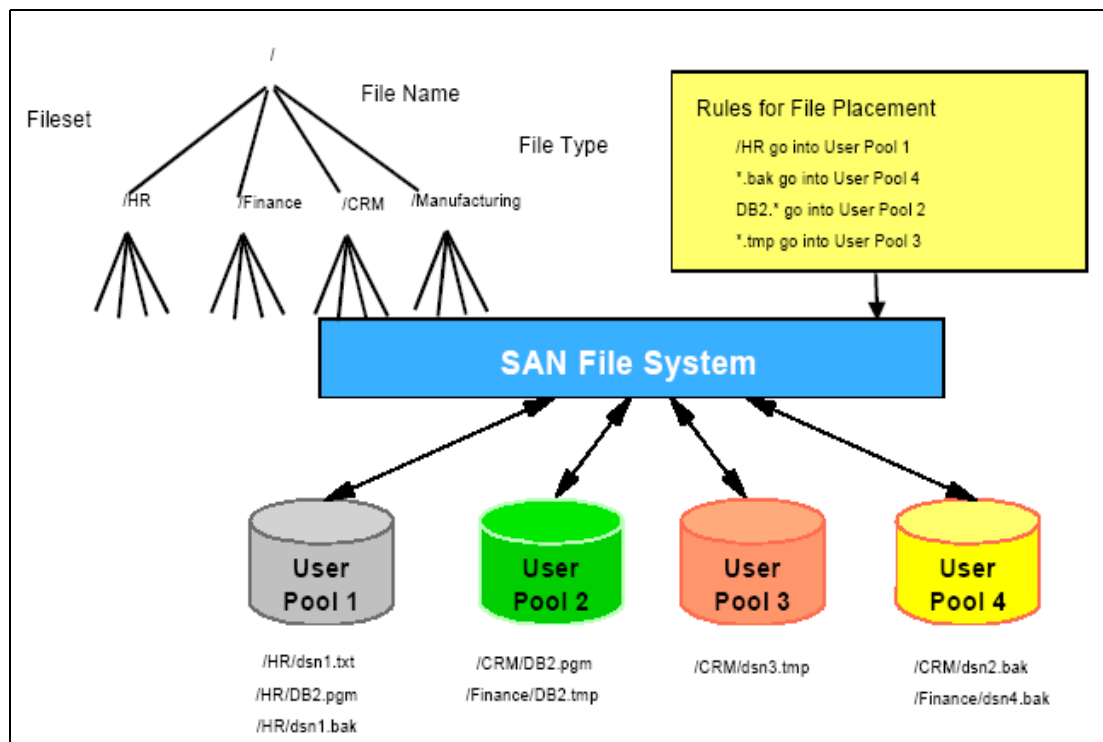


Figure 14-7 Policy execution for file placement

FlashCopy

A FlashCopy image is a space-efficient, read-only copy of the contents of a fileset in a SAN File System global namespace at a particular point in time.

Note that while a FlashCopy image is being created, all data remains online and available to users and applications. The FlashCopy image does not include any nested filesets within it. You can create and maintain a maximum of 32 FlashCopy images of any fileset.

Figure 14-8 shows how a FlashCopy image is viewable at a Windows client. In this case, a FlashCopy image was made of the fileset container_A.

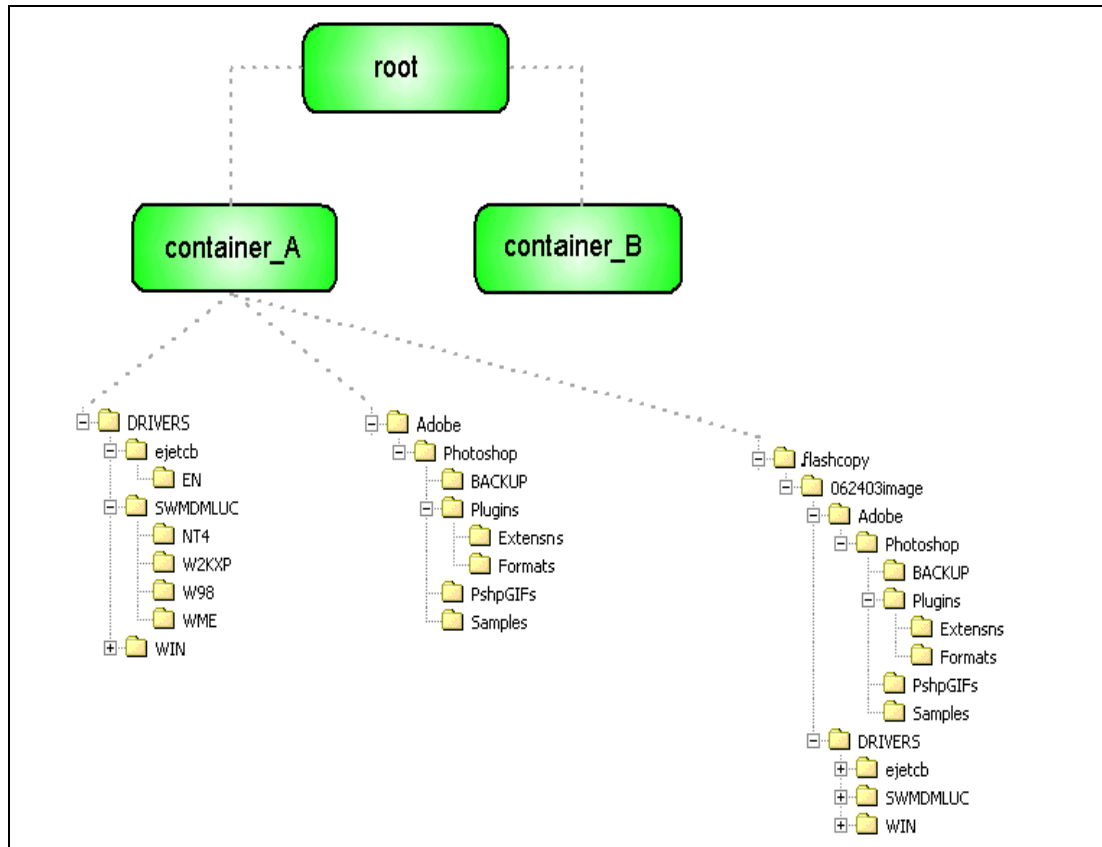


Figure 14-8 FlashCopy Image viewable on Windows client

14.5 Supported systems and planning considerations

The **SAN File System** is delivered as a software only package for **V2.2**. The SAN File System software bundle consists of three components:

- ▶ Software that runs on the SAN File System Metadata servers
- ▶ Software that runs on your application servers, called the SAN File System Client software
- ▶ Optional software that is installed on the Master Console

Metadata server

MDS engine runs SUSE Linux operating system. SAN File System V2.2 supports from two to eight MDSs running on hardware that must be **supplied by the customer**. SAN File System V2.2 will support a cluster of metadata servers consisting of both 4146-1RX engines, xSeries

servers, and equivalents. If not using the IBM 345, 346, or 4146-Model 1RX, the following hardware components are required for each metadata server:

- ▶ Two processors, minimum 3 GHz each
- ▶ Minimum of 4 GB of system memory
- ▶ Two internal hard disk drives with mirroring capability, minimum 36 GB each
- ▶ Two power supplies (optional, but highly recommended for redundancy)
- ▶ One 10/100/1000 Gb port for Ethernet connection (Fibre or Copper)
- ▶ Two 2 Gb Fibre Channel host bus adapter (HBA) ports: These must be compatible with the SUSE operating system and the storage subsystems in your SAN environment. They must also be capable of running the QLogic 2342 device driver. Suggested adapters: QLogic 2342 or IBM part number 24P0960.
- ▶ CD-ROM and diskette drives
- ▶ Remote Supervisory Adapter II card (RSA II): This must be compatible with the SUSE operating system. Suggested card: IBM part number 59P2984.

Master Console

The SAN File System V2.2 Master Console is an *optional* component of a SAN File System configuration for use as a control point. If deployed, it requires a customer-supplied, high performance, and highly reliable rack-mounted Intel Pentium® 4 processor server. This can be an xSeries 305 server, File System V1.1 or V2.1 Master Console, 4146-T30 feature #4001, a SAN Volume Controller Master Console, a SAN Integration Server Master Console, or equivalent Intel server with the following capabilities:

- ▶ At least 2.6 GHz processor speed
- ▶ At least 1 GB of system memory
- ▶ Two 40 GB IDE hard disk drives
- ▶ CD-ROM drive
- ▶ Diskette drive
- ▶ Two 10/100/1000 Mb ports for Ethernet connectivity (Copper or Fiber)
- ▶ Two Fibre Channel Host Bus Adapter (HBA) ports
- ▶ Monitor and keyboard: IBM Netbay 1U Flat Panel Monitor Console Kit with keyboard or equivalent

If a SAN Volume Controller or SAN Integration Server Master Console is already available, it can be shared with SAN File System, since it meets the hardware requirements.

SAN File System Client

SAN File System currently supports clients that run these operating systems:

- ▶ AIX 5L™ Version 5.1, 5.2, 5.3
- ▶ Windows 2000 Server and Windows 2000 Advanced Server
- ▶ Windows 2003 Server Standard and Enterprise Editions
- ▶ Red Hat Enterprise Linux 3.0 AS, ES, and WS
- ▶ SUSE Linux Enterprise Server 8.0
- ▶ Sun Solaris 9

SAN File System supports clients running in the clustered environment, specifically AIX running with HACMP, Solaris running with Sun clustering and Windows 2000 and 2003 with MSCS.

Supported storage

SAN-attached storage is required for both metadata volumes as well as user volumes.

Metadata volumes

These storage systems can be used for metadata storage with SAN File System V2.2:

- ▶ IBM TotalStorage IBM DS 4300 Turbo, DS4400, and DS4500 (non SATA drives)
- ▶ IBM TotalStorage Enterprise Storage Server (ESS) Models F20, 750, and 800
- ▶ IBM TotalStorage SAN Volume Controller V1.1 or V1.2 with the storage devices supported by the SAN Volume Controller

User volumes

The SAN File System can be configured with any SAN storage device for the user data storage, providing it is supported by the operating systems running the SAN File System client (including having a compatible HBA) and that it conforms to the SCSI standard for unique device identification.

The SAN File System also supports storage devices for user data storage attached through iSCSI. The iSCSI attached storage devices must conform to the SCSI standard for unique device identification and must be supported by the SAN File System client operating systems.

14.6 Key features

Here is a summary of the key features of the IBM TotalStorage SAN File System.

▶ **Uniform global namespace**

All of the supported application servers connected to the SAN File System are able to access the same data using SAN File System's uniform global namespace. A uniform global namespace provides the capabilities for all clients to have a consistent view of the SAN File System directory and files which can enable enhanced workflow and collaboration between applications running on heterogeneous servers.

▶ **Storage pools**

The SAN File System provides the capability to assign storage volumes to storage pools where the storage pools can represent different classes of service for storage. Volumes can be added to a storage pool without impacting application processing. Data can be moved between volumes without impacting availability or performance.

▶ **Policy based file space provisioning**

Storage administrators will be able to create rules for the SAN File System to determine what storage pool is used to allocate space for a file when the file is created. The storage administrator can use any of the file attributes (such as file name, file type, date created, user ID or group ID, and so on) to create these rules. Through the use of the policy-based rules, the SAN File System automates the task of allocating space on the desired storage volumes.

▶ **Policy based movement of files between storage pools and policy based file deletion**

The SAN File System provides information lifecycle management capabilities allowing administrators to specify how files should be automatically moved among storage pools during their lifetime, and optionally, specify when files should be deleted. This feature can potentially lower the overall costs of storage and improve storage space utilization enabling a balanced use of premium and inexpensive storage matching the objectives of the enterprise. This feature also has the potential to reduce manual storage administration for managing space utilization and can reduce the cost of storage management.

► **Storage quota management**

The SAN File System provides storage capacity to applications and servers creating new files or extending existing files. The SAN File System allows the storage administrator to define the amount of storage to be used by a specific server or by a specific application (defined as a fileset) and can help the storage administrator monitor the storage utilization by alerting them when the storage reaches a predefined capacity. The use of storage quotas can prevent an application from using more storage than permitted. The SAN File System allows for storage utilization quotas to be set higher than the physical storage available to the system or applications. This allows for simplified management of the system by allowing administrators to dynamically add storage without having to reset application level policies. Both hard and soft quotas are supported along with threshold alerts for integration with systems management tools.

► **Locking**

The SAN File System is designed to provide file level locking to the SAN File System clients running on your application servers. This function provides data sharing with integrity across heterogeneous application servers. The file level locking uses the same locking permissions (read, write, execute, and so on) that are provided by the local operating systems for locking within one server. In addition to locking the entire file, the SAN File System also supports the ability to do a lock on a byte range within a file.

► **Heterogeneous file sharing between Windows and UNIX servers**

The heterogeneous file sharing feature enables flexible, yet secure, sharing of files between UNIX and Windows servers. The SAN File System provides cross-environment access checking such that files created on Windows can be accessed by authorized users on UNIX, and vice versa. To control cross-environment authorization, the administrator manages a set of user map entries using the administrative CLI or GUI. Each user map entry specifies a UNIX domain-qualified user and a Windows domain-qualified user that are to be treated as equivalent for the purpose of checking file access permission in the cross-platform case. The SAN File System metadata server cluster accesses the customer's UNIX and Windows directory services, as needed, to obtain user ID and group membership information.

► **Caching**

The SAN File System clients cache the metadata that they get from the SAN File System metadata servers. This allows the applications running on a server to continue to access the file data without additional requests to the metadata servers and also allows other applications on the same server to share files without additional requests to the metadata servers. The caching of the metadata provides for high performance data access over the SAN without requests flowing to the metadata server. It can also provide continuous access to open files even if a temporary condition prohibits communication with the metadata servers.

► **Volume drain**

The SAN File System provides a volume drain function that allows you to remove or upgrade storage volumes without disrupting application processing. The volume drain function moves the data off the specified volume onto other volumes in the same storage pool. The SAN File System manages the access to the data that is being moved so there is no need to stop application processing during the volume drain. Data can be moved between volumes without impacting availability or performance.

► **FlashCopy**

The SAN File System provides a file-based point-in-time copy function called FlashCopy. The FlashCopy image is a logical point-in-time image of a Fileset, a set of files. A copy of the data block is made upon the next update after the FlashCopy image. The FlashCopy image only requires storage for changed blocks, therefore, reducing the amount of

storage needed. The SAN File System supports 32 FlashCopy images for each Fileset. The files can be restored to any one of the FlashCopy images.

File-based backup and restore products can back up files using the FlashCopy function while the applications are online and updating a later version of the files. The file-based FlashCopy function can help significantly reduce maintenance windows for data backups and can also help reduce the amount of storage space needed for online data backups.

▶ **File copy migration tool**

The SAN File System provides a file copy migration tool that you can use to copy files into the SAN File System. The file copy migration tool runs on your application servers that are the SAN File System clients. The file copy migration tool includes a planning option that creates an estimate of the time needed to copy the files into the SAN File System.

▶ **iSCSI attached clients and storage**

The SAN File System supports host application servers and storage devices attached through iSCSI. The iSCSI attached storage devices must conform to the SCSI driver interface standards, including the unique device identification calls.

14.7 More information

- ▶ You can find more information about IBM SAN File System from the following Web site:

<http://www.storage.ibm.com/software/virtualization/sfs/index.html>

- ▶ You can refer to these redbook for more in-depth information about IBM SAN File System:

– *IBM TotalStorage: Introducing the SAN File System*, SG24-7057

<http://www.redbooks.ibm.com/redbooks/pdfs/sg247057.pdf>

– *IBM TotalStorage: Integration of the SAN Volume Controller, SAN Integration Server and the SAN File System*, SG24-6097

<http://www.redbooks.ibm.com/redbooks/pdfs/sg246097.pdf>



IBM TotalStorage Productivity Center

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 improve administrative efficiency. It is designed to enable a flexible storage infrastructure that can respond to on demand storage needs.

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

IBM TotalStorage Productivity Center with Advanced Provisioning comes with an additional product for provisioning. The solution is designed to allow you to automate capacity provisioning capabilities through automated workflows.

This chapter introduces IBM TotalStorage Productivity Center and its constituents.

15.1 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.

IBM TotalStorage Productivity Center is comprised of the following four products.

- ▶ **IBM TotalStorage Productivity Center for Data** (formerly IBM Tivoli Storage Resource Manager) can provide over 300 enterprise-wide reports, monitoring and alerts, policy-based action, and file system capacity automation in an heterogeneous environment.

For details, see Chapter 16, “IBM TotalStorage Productivity Center for data” on page 255.

- ▶ **IBM TotalStorage Productivity Center for Fabric** (formerly IBM Tivoli Storage Area Network Manager) can provide automated device discovery, topology rendering, error detection fault isolation, SAN error predictor, zone control, real-time monitoring and alerts, and event management for heterogeneous enterprise SAN environments.

For details, see Chapter 17, “IBM TotalStorage Productivity Center for Fabric” on page 271.

- ▶ **IBM TotalStorage Productivity Center for Disk** (formerly IBM TotalStorage Multiple Device Manager — Performance Manager feature) can enable device configuration and management of supported SAN-attached devices from a single console. It can discover storage and provides configuration capabilities for the IBM TotalStorage Enterprise Storage Server (ESS), IBM TotalStorage SAN Volume Controller (2145), and the IBM TotalStorage DS4000 (FAStT) storage. IBM TotalStorage Productivity Center for Disk manages the performance for the IBM TotalStorage Enterprise Storage Server (ESS) and SAN Volume Controller (SVC).

For details, see Chapter 18, “TotalStorage Productivity Center for Disk” on page 297.

- ▶ **IBM TotalStorage Productivity Center for Replication** (formerly IBM TotalStorage Multiple Device Manager — Replication Manager feature) can provide copy services management for the IBM TotalStorage Enterprise Storage Server (ESS). IBM TotalStorage Productivity Center for Replication provides configuration and management of the Point-in-Time copy (FlashCopy) and Metro Mirror (synchronous point-to-point remote copy) capabilities of the ESS in supported configurations.

For details, see Chapter 19, “TotalStorage Productivity Center for Replication” on page 307.

15.2 Launchpad

The IBM TotalStorage Productivity Center provides an integration point for storage and fabric management and replication. The IBM TotalStorage Productivity Center user interface allows storage administrators to launch IBM TotalStorage Open Software Family products, which help to support managing the storage volume lifecycle, device configuration, performance, replication, storage network fabric, data backup, data availability, and data recovery, as well as enterprise policies for managing host, application, database, and file system data.

The integration point which IBM TotalStorage Productivity Center provides is a launchpad which is designed to simplify storage administration by providing a task-based organization to the broad range of capability available with the TotalStorage Open Software Family products.

Figure 15-1 on page 249 shows the TotalStorage Productivity Center launchpad after installation of the TotalStorage Productivity Center.

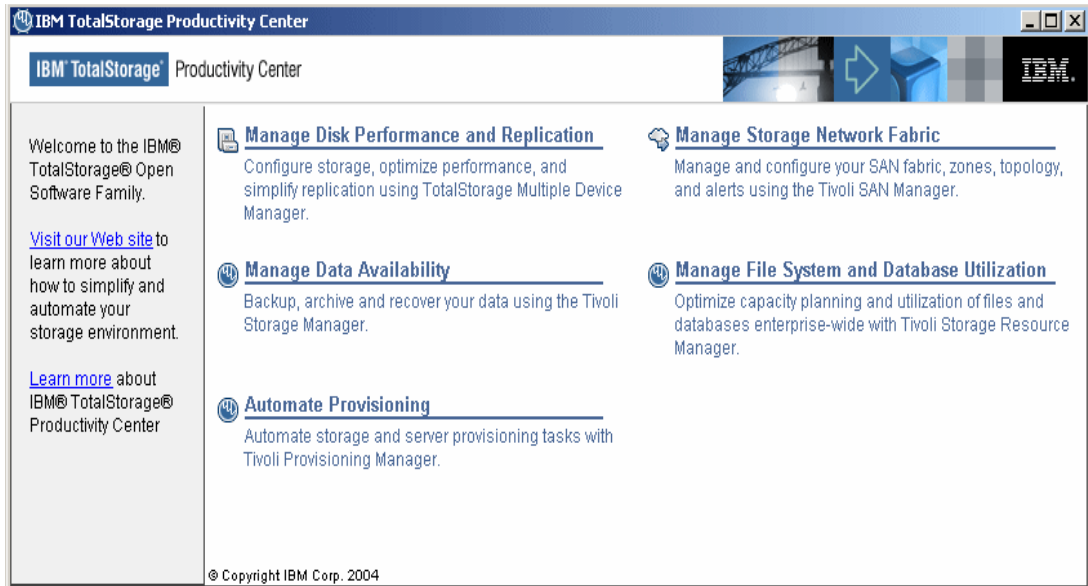


Figure 15-1 IBM TotalStorage Productivity Center Launchpad

The launchpad allows you to centrally manage your TotalStorage Productivity Center functions. It can be customized to point to nonstandard implementations, or you can even add your own applications to the selection options to help you have a central location for key management applications.

15.3 Rebranding to IBM TotalStorage Productivity Center

The IBM TotalStorage Multiple Device Manager and its feature names - Device Manager, Performance Manager, and Replication Manager are still used within the IBM TotalStorage Productivity Center documentation, the product user interface and some command names. The Productivity Center for Data and Productivity Center for Fabric products are also referred to by their previous names. In this book we use the new names. Some screen shots may however still show the product names in use prior to IBM TotalStorage Productivity Center V2.1.

Table 15-1 Rebranded names in the IBM TotalStorage Productivity Center suite

Former name	Rebranded name
Multiple Device Manager components: Device Manager (DM) and Performance Manager (PM)	IBM TotalStorage Productivity Center for Disk (DM and PM)
Multiple Device Manager components: Device Manager (DM) and Replication Manager (RM)	IBM TotalStorage Productivity Center for Replication (DM and RM)
IBM Tivoli Storage Resource Manager	IBM TotalStorage Productivity Center for Data
IBM Tivoli Storage Area Network Manager	IBM TotalStorage Productivity Center for Fabric

15.4 Productivity Center with Advanced Provisioning

IBM's vision for on demand has at its heart, the automation and provisioning strategies provided by products like IBM Tivoli Provisioning Manager. The move from "just-in-case" manual provisioning to "just-in-time" automated provisioning is brought one major step closer. The IBM software product suite offers a foundation for responding flexibly to internal and external changes and streamlining business operations while dynamically aligning your IT resources with your business priorities. IBM TotalStorage Productivity Center with Advanced Provisioning does just that for your storage environment.

The IBM TotalStorage Productivity Center with Advanced Provisioning offering is designed to allow assisted capacity provisioning capabilities through automated workflows.

Storage provision using IBM TotalStorage Productivity Center with Advanced Provisioning allows you to save time and human errors by automating the many steps involved in allocating storage to an application or user.

IBM TotalStorage Productivity Center with Advanced Provisioning provides you with one additional product; that is, the IBM Tivoli Provisioning Manager, which provisions and configures servers, operating systems, middleware, applications, storage, and network devices, including your Fibre Channel switches. Specific to IBM TotalStorage Productivity Center with Advanced Provisioning, the solution provisions storage and grants access to that storage to the application server or user.

Managing complex IT processes, such as allocating storage, and configuring the network or SAN, often require significant IT staff time, can be slow to execute, and can result in human error. Imagine executing these complex tasks in hours versus days. Provisioning automates the way IT resources are made available. IBM Tivoli Provisioning Manager helps clients significantly reduce the time it takes to accomplish complex IT tasks while eliminating the risk of human error by automating complex IT processes using best practices.

Figure 15-2 on page 251 depicts the IBM TotalStorage Productivity Center with Advanced Provisioning. It shows the components and basic interactions with agents, hardware, and other products that comprise the IBM TotalStorage Productivity Center with Advanced Provisioning offering. IBM TotalStorage Productivity Center manages the agents and communication to the storage devices, while IBM Tivoli Provisioning Manager communicates with the various components of IBM TotalStorage Productivity Center to provide the automated provisioning.

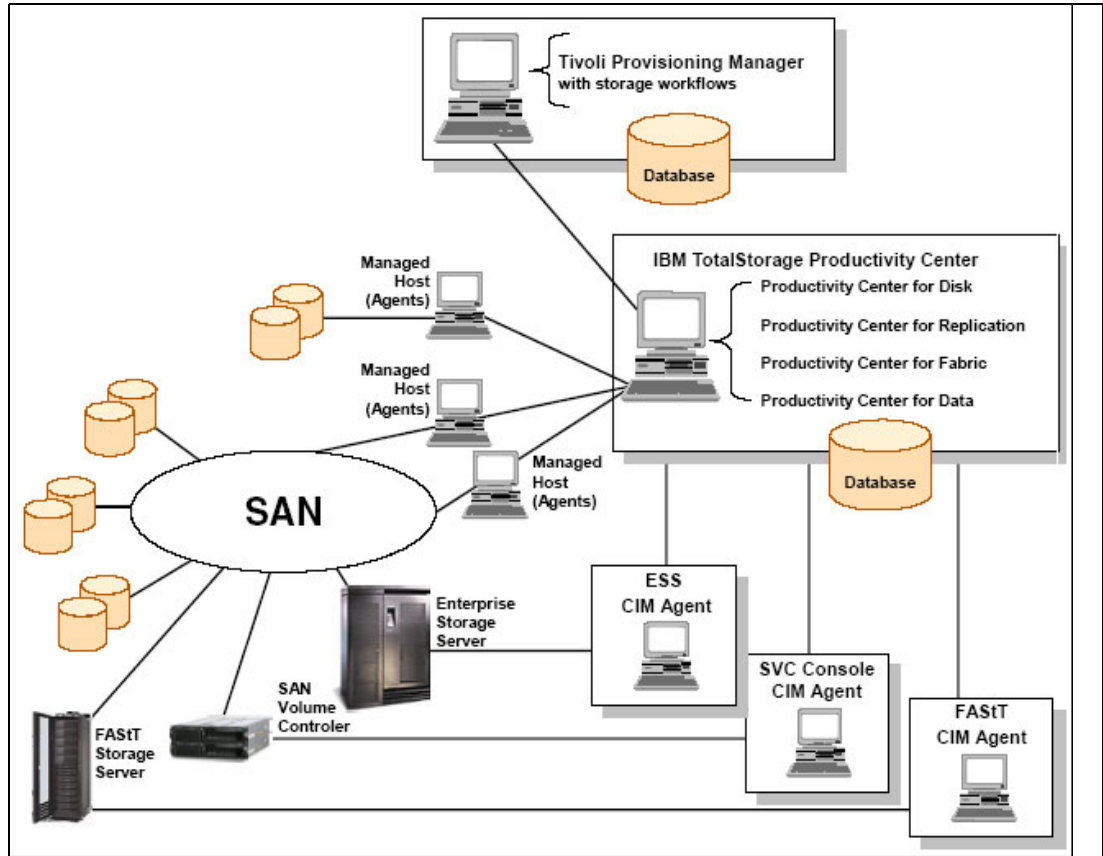


Figure 15-2 A typical IBM TotalStorage Productivity Center configuration

15.5 Architecture

The architectural overview diagram in Figure 15-3 on page 252 is used to illustrate the governing ideas and building blocks of the product suite which makes up the IBM TotalStorage Productivity Center. It provides a logical overview of the main conceptual elements and relationships in the architecture, components, connections, users and external systems.

IBM TotalStorage Productivity Center and Tivoli Provisioning Manager are presented as building blocks in the diagram. Actually these two products are single applications but rather complex environments by themselves.

The diagram also shows the different methods used to collect information from multiple systems to give an administrator the necessary views on the environment, i.e.

- ▶ software clients (agents)
- ▶ standard interfaces and protocols (for example, SNMP, CIM Agent)
- ▶ proprietary interfaces (for only a few devices)

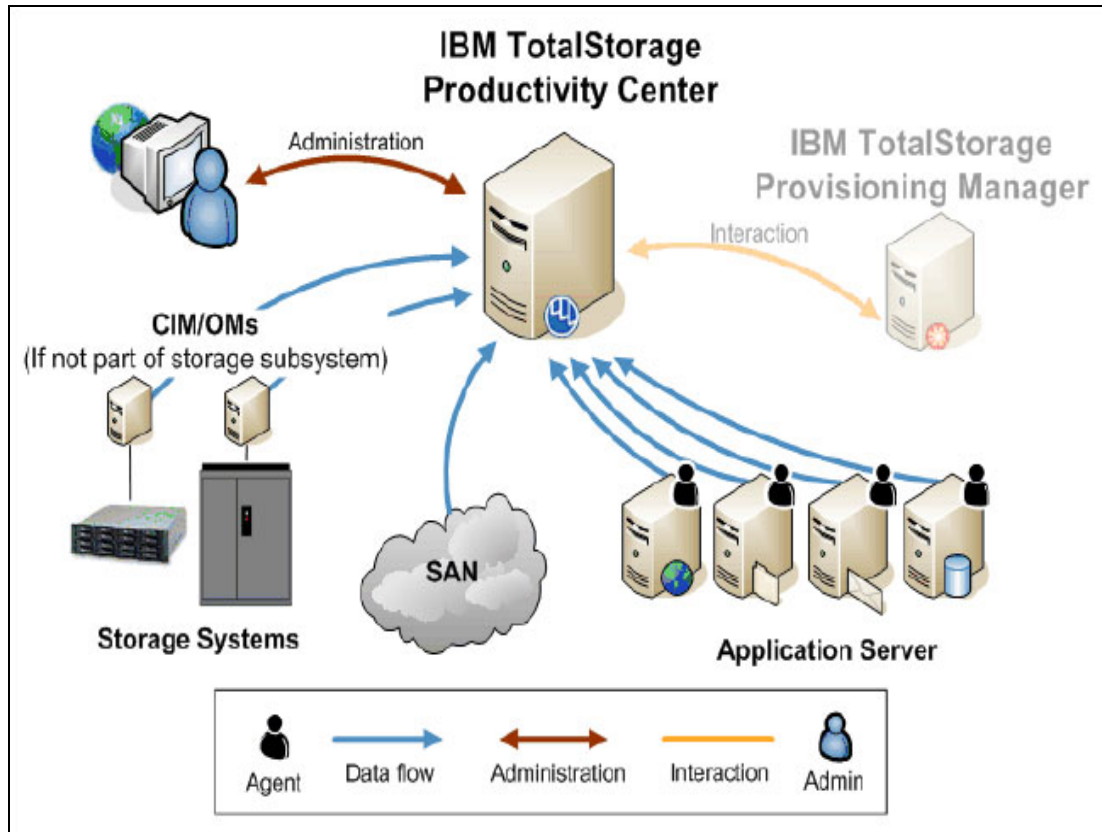


Figure 15-3 IBM TotalStorage Productivity Center architecture overview diagram

In addition to the central data collection Productivity Center provided a single point of control for a storage administrator, even though each of the manager still come with its own interface. A program called the launchpad is provided to start the individual applications form a central point.

Tivoli Provisioning Manager relies on Productivity Center in order to make provisioning possible.

Architectural layers

The IBM TotalStorage Productivity Center architecture can be broken up in three layers as shown in Figure 15-4 on page 253.

Layer one represents a high level overview, there is only IBM TotalStorage Productivity Center instance in the environment. In layers two and three we drill down into the TotalStorage Productivity Center environment to see the managers and the pre-require components.

Layer two consist of the individual products that are part of the product suite: IBM TotalStorage Productivity Center for Disk, IBM TotalStorage Productivity Center for Replication, IBM TotalStorage Productivity Center for Fabric and IBM TotalStorage Productivity Center for Data. These products are referred to as managers.

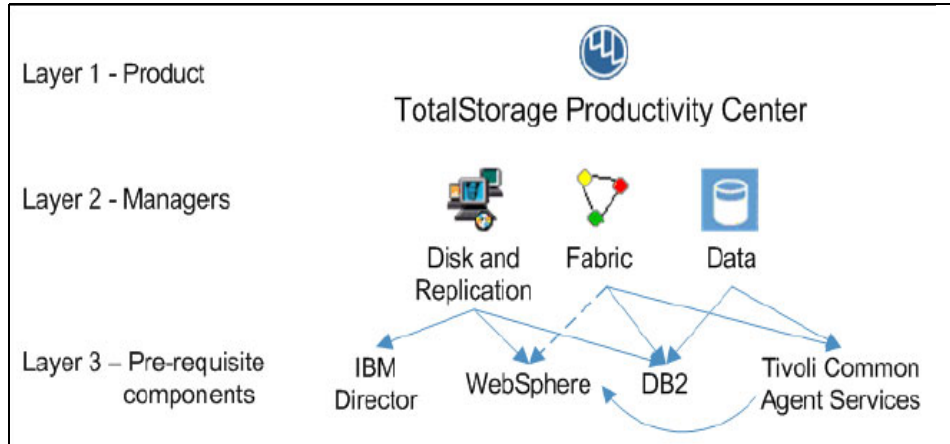


Figure 15-4 Architectural layers

The agents are not shown in Figure 15-4, but they play an important role in the TotalStorage Productivity Center environment. There are two types of agents: CIM Agents and agents to belong to one of the managers:

- ▶ CIM Agents - agents that offer a CIM interface for management applications for example, for Enterprise Storage Server, SAN Volume Controller, and FASTT)
- ▶ Data Agents - agents to collect data for the Data Manager
- ▶ Fabric Agents - agents that are used by the Fabric Manager for in-band SAN operations

15.6 Key benefits and highlights

To summarize, the IBM TotalStorage Productivity Center helps you:

- ▶ Monitor and track the performance of SAN attached SMI-S compliant storage devices.
- ▶ Manage the capacity utilization and availability of file systems and databases.
- ▶ Monitor, manage and control (zone) SAN fabric components.
- ▶ Manage advanced storage replication services (Peer-to-Peer Remote Copy and FlashCopy).
- ▶ Automate capacity provisioning to help improve application availability.

15.7 More Information

Refer to these Redbooks for more information on IBM TotalStorage Productivity Center:

- ▶ *Exploring Storage Management Efficiencies and Provisioning*, SG24-6373:
<http://www.redbooks.ibm.com/redbooks/pdfs/sg246373.pdf>
- ▶ IBM TotalStorage Productivity Center: Getting Started
<http://www.redbooks.ibm.com/redpieces/pdfs/sg246490.pdf>

For latest information about IBM TotalStorage Productivity Center, refer to following link:

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



IBM TotalStorage Productivity Center for data

As a component of the IBM TotalStorage Productivity Center, IBM TotalStorage Productivity Center for Data (formerly IBM Tivoli Storage Resource Manager) is designed to 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

This chapter introduces and positions the IBM TotalStorage Productivity Center for Data and discusses its architecture and key features.

16.1 Overview

IBM TotalStorage Productivity Center for Data helps you discover, monitor, and create enterprise policies for your disks, storage volumes, file systems, files, and databases. Knowing where all your storage is located, places you in a better position to act intelligently on your data. Application availability is improved because you get early warnings when file systems are running out of space and you have the ability to provision more disk automatically. If you are thinking about server consolidation, you can use IBM TotalStorage Productivity Center for Data to help efficiently utilize your accumulated storage resources.

Architected for efficiency and ease-of-use, IBM TotalStorage Productivity Center for Data uses a single agent per server to provide detailed information without high consumption of network bandwidth or CPU cycles.

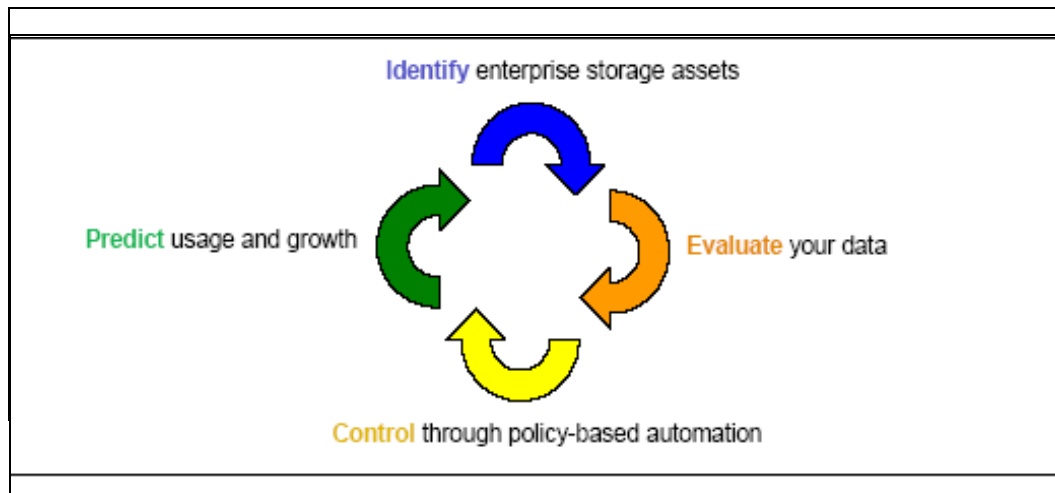


Figure 16-1 Storage life cycle resource management

Figure 16-1 shows the concept of storage resource management from a life cycle perspective. The idea is to establish a base understanding of your storage environment, with an emphasis on discovering areas where simple actions can deliver rapid return on investment. Ideally, you want to identify potential areas of exposure, evaluate the data residing on your servers, set up control mechanisms for autonomic management, and start the capacity planning process by predicting growth.

TPC for Data monitors storage assets, capacity, and usage across an enterprise. It can look at:

- ▶ Storage from a host perspective: Manage all the host-attached storage, capacity and resources attributed to file systems, users, directories, and files, as well as the view of the host-attached storage from the Storage Subsystem perspective.
- ▶ Storage from an application perspective: Monitor and manage the storage activity inside different database entities including instance, tablespace, and table.
- ▶ Storage utilization and provide chargeback information allowing to justify, or account for, storage usage.

Productivity Center for Data provides over 300 standardized reports (and the ability to customize your own reports) about file systems, databases, and storage infrastructure. These reports provide the storage administrator information about:

- ▶ Assets
- ▶ Availability
- ▶ Capacity

- ▶ Usage
- ▶ Usage violation
- ▶ Backup

With this information, the storage administrator can:

- ▶ Discover and monitor storage assets enterprise-wide
- ▶ Report on enterprise-wide assets, files and file systems, databases, users, and applications
- ▶ Provide alerts (set by the user) on issues such as capacity problems and policy violations.
- ▶ Support chargebacks by usage or capacity

16.2 Product highlights

This section describes the significant features of IBM TotalStorage Productivity Center for Data.

▶ **Policy-based management**

IBM TotalStorage Productivity Center 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. IBM TotalStorage Productivity Center for Data can provide automated solutions through event management.

For example, if IBM TotalStorage Productivity Center 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**

IBM TotalStorage Productivity Center 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, IBM TotalStorage Productivity Center 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 IBM TotalStorage Productivity Center for Data. The policy is set through a File System Extension Policy Panel. A probe will run on agents and send file system statistics to the server. The server will compare the current file system stats against the policy, and invoke provisioning and extension as necessary.

▶ **Direct Tivoli Storage Manager integration**

With this feature, IBM TotalStorage Productivity Center 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 IBM Tivoli Storage Manager, and then delete the files to free up the disk space.

▶ **Subsystem reporting**

IBM TotalStorage Productivity Center for Data gathers and reports on disk subsystems. Information includes physical characteristics such as the drive's manufacturer, model, serial number, capacity, and rotational speed. Also included is how that drive's storage is

allocated to logical volumes, snapshot copy volumes, and free space. This feature allows users to perform the following (subject to vendor's implementation of Storage Management Initiative Specification (SMIS)):

- Display the physical disks behind what the host sees as a disk drive
- Show the allocated and free capacity of subsystems on the network
- List subsystem volumes that have been allocated but aren't in use
- Show which hosts have access to a given subsystem volume
- Show which hosts have access to a given disk drive (within the subsystem)
- Show which subsystem volumes (and disks) a host has access to
- IBM SAN Volume Controller reporting

► **Reporting for IBM SAN File System (SFS) implementation**

IBM TotalStorage Productivity Center for Data can help identify environments where IBM TotalStorage SAN File System implementation may be appropriate. It does this by utilizing network-wide statistics to show metrics including, but not limited to, temporary space, duplicate file data, and storage utilization by disk and file systems, to help plan for IBM SFS implementation. IBM TotalStorage Productivity Center for Data can help create an inventory of assets, offers suggestions for policies, and helps you plan for implementation.

► **NAS Support**

IBM TotalStorage Productivity Center for Data can enable storage administrators to monitor, report on, and manage NAS resources. IBM TotalStorage Productivity Center for Data is designed to provide a universal view of direct-attached and network-attached storage, from a file system or application perspective. For Network Appliance files, IBM TotalStorage Productivity Center for Data will monitor, report on, and manage physical disk information such as total disk capacity and disk usage information. By providing this information as part of a universal view of enterprise storage, storage administrators can manage storage from a logical (filesystem) perspective as well as a physical (disk) perspective. Administrators can also view information for a single filer, a group of filers, a user or group of users, or all filers enterprise-wide, enabling them to manage all enterprise storage resources from a single Web-based interface, to help them reduce costs and increase availability by pinpointing and solving problems more effectively.

► **Advanced provisioning**

IBM TotalStorage Productivity Center for Data is an integral piece of the IBM TotalStorage Productivity Center with Advanced Provisioning solution and is designed to allow you to automate capacity provisioning capabilities through automated workflows.

16.3 Architecture

IBM TotalStorage Productivity Center for Data has four major components: the server, scheduler, repository, and agent. There is also a Common Information Model/Object Messaging (CIM/OM) component used for collecting data from storage subsystems, such as the IBM Enterprise Storage Server and DS4000.

Tip: For more information on CIM infrastructure, refer to Appendix A., “Introduction to SMIS and CIM/WBEM” on page 459.

The high-level architecture and agent interaction is shown in Figure 16-2. The three layers in the graphic represent the various component areas, such as the server, the agent, and the proxy agents.

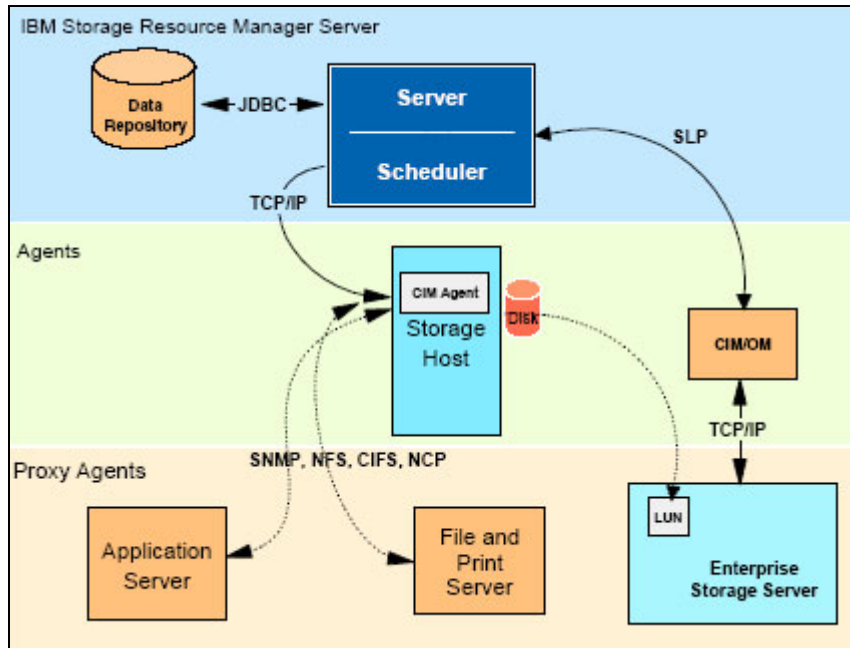


Figure 16-2 IBM TotalStorage Productivity Center for Data high level architecture

Server

The Server component is the center of IBM TotalStorage Productivity Center for Data's architecture and directs all of the activities related to:

- ▶ The monitoring performed on the Monitored Agent computers
- ▶ The reporting performed on the Client GUI computers

The Server interacts with the components described in the following sections.

Enterprise Repository

IBM TotalStorage Productivity Center for Data monitors your environment to accumulate information about its storage assets and their usage over time. It records this information in the centralized Enterprise Repository. The repository is organized into a series of relational database tables and is accessed using Java Database Connectivity (JDBC).

Job Scheduler

IBM TotalStorage Productivity Center for Data deploys monitoring activities via an embedded Job Scheduler. This Job Scheduler controls when to run Discovery, Ping, Probe, Scan, and Quota checking jobs. It provides robust scheduling functionality and the flexibility to specify exactly when to run monitoring jobs, as well as other functions such as running scripts and creating reports.

Agents

The Agent component is installed on each machine containing storage that you want to monitor. The Agent component runs in the background and deploys monitoring jobs such as Pings, Probes, and Scans. It interacts with the Job Scheduler and deploys monitoring jobs according to the schedules that you have defined.

This allows an Agent to:

- ▶ Detect low-level problems, such as grown defects in the hardware containing the storage information.

- ▶ Keep track of information such as the uptime/downtime of the machine on which it is running.
- ▶ Return information to the server from the scans and probes.

Agents receive jobs to be run from a server's scheduling service and then contact the server for the job definition. Once this has been successfully completed, no further communication to the server is required until the job is completed.

Probes

A probe gathers local hardware and general filesystem information on a machine where an Agent is installed. Use probes to discover information about new or removed disks and filesystems and gather CPU, memory, and operating system level information. The probe relies on local system calls and APIs.

A proxy probe is required for systems on which IBM TotalStorage Productivity Center for Data does not support an Agent (for example, on most NAS devices and NetWare servers). A proxy probe on these devices will use SNMP queries to gather hardware information and CIFS, NFS, NCP commands to gather volume and file system information. Again, see Figure 16-2 on page 259 for a depiction of this.

Scans

Where a probe will gather higher level information about a file system, a scan will gather file and directory level information. The scan traverses the file systems on a local host reading attributes of each file and directory. For the NTFS file system, all file system stream space is read and recorded.

A proxy scan requires the selection of an Agent or Agents to scan each file system on an NAS device or NetWare volume. A single Agent can scan many NAS or NetWare servers or you can select one Agent per volume. This allows for a large NetWare or NAS device to be scanned with several Agents simultaneously for increased performance. The proxy scan uses NFS, CIFS or NCP commands to access the remote servers.

Storage Subsystems

IBM TotalStorage Productivity Center for Data uses a Common Information Model (CIM) agent to gather information on storage subsystem controllers to display Storage Subsystem reports. In order to use this capability, you will need to connect to a CIM/OM that exists, and is successfully configured to communicate with the desired storage subsystem that you want to collect information from.

16.4 Planning considerations and prerequisite

We look successively at planning considerations for the server, storage subsystems, and security.

16.4.1 Server recommendations

The IBM TotalStorage Productivity Center for Data server component acts as a traffic officer for directing information and handling requests from the agent and UI components installed within an environment. You need to install at least one server within your environment. We recommend you do not manage more than 1000 agents with a single server. If you need to install more than 1000 agents, we suggest you install an additional server for those agents to maintain optimal performance.

16.4.2 Supported systems and databases

We list here some of the currently supported storage subsystems and databases or filesystems.

Storage Subsystem Support

Data Manager currently supports monitoring and reporting on the following storage subsystems:

- ▶ Hitachi Data Systems
- ▶ HP StorageWorks
- ▶ IBM DS4000 series
- ▶ SAN Volume Controller
- ▶ ESS

File system support

Data Manager supports monitoring and reporting of the various file systems like FAT, NTFS, AIX_JFS, EXT2, EXT3, HP_HFS, REISERFS, etc.

Networked File system Support

Data Manager currently supports monitoring and reporting of the following networked file systems:

- ▶ IBM TotalStorage SAN File System 1.0 (Version 1 Release 1), from AIX V5.1 (32-bit) and Windows 2000 Server/Advanced Server clients
- ▶ IBM TotalStorage SAN File System 2.1, 2.2 from AIX V5.1 (32-bit), Windows 2000 Server/Advanced Server, Red Hat Enterprise Linux 3.0 Advanced Server, and SUN Solaris 9 clients
- ▶ General Parallel File System (GPFS) v2.1, v2.2

RDBMS Support

Data Manager currently supports the monitoring of the following Relational Database Management Systems (RDBMSs):

- ▶ Microsoft SQL Server 7.0, 2000
- ▶ Oracle 8i, 9i, 9i V2, 10G
- ▶ Sybase SQL Server 11.0.9 and higher
- ▶ DB2 Universal Database™ (UDB) 7.1, 7.2, 8.1, 8.2 (64-bit UDB/DB2 Instances are supported)

Note: For more details on supported systems, refer to the product and device support link on the following site:

<http://www.ibm.com/servers/storage/software/center/data/index.html>

16.4.3 Security considerations

IBM TotalStorage Productivity Center for Data has two security levels: non-administrative users and administrators:

- ▶ Non-administrator users can:
 - View the data collected by Data Manager, and
 - Create, generate, and save reports

- ▶ Administrators can:
 - Create, modify, and schedule pings, probes, and scans
 - Create, generate, and save reports
 - Perform administrative tasks and customize the Data Manager environment
 - Create groups, profiles, quotas, and constraints
 - Set alerts

16.5 Key features

Here we will discuss key features of IBM TotalStorage Productivity Center for Data.

16.5.1 Basic menu displays

Figure 16-3 shows the IBM TotalStorage Productivity Center for Data - Dashboard. This is the default right-hand pane display when you start the application and shows a quick summary of the overall health of the storage environment. It can quickly show you potential problem areas for further investigation.

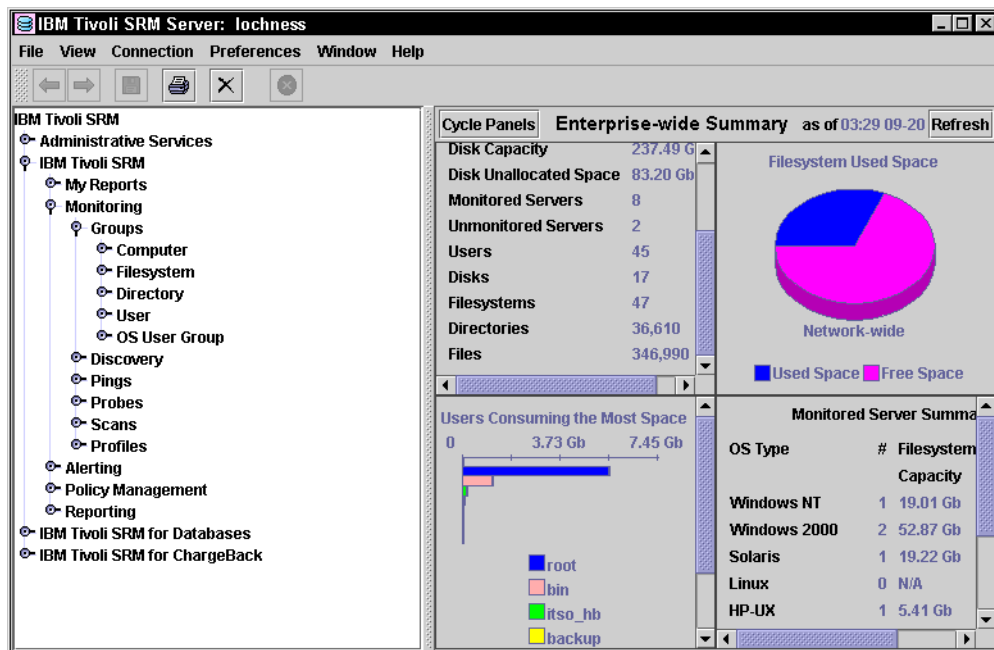


Figure 16-3 IBM TotalStorage Productivity Center for Data - dashboard

The dashboard contains four viewable areas, which cycle among seven predefined panels. To cycle, use the **Cycle Panels** button. Use the **Refresh** button to update the display.

Enterprise-wide summary

The Enterprise-wide Summary panel shows statistics accumulated from all the agents. The statistics are:

- ▶ Total file system capacity available
- ▶ Total file system capacity used
- ▶ Total file system free capacity
- ▶ Total allocated and unallocated disk space
- ▶ Total disk space unallocated to file systems
- ▶ Total number of monitored servers

- ▶ Total number of unmonitored servers
- ▶ Total number of users
- ▶ Total number of disks
- ▶ Total number of file systems
- ▶ Total number of directories
- ▶ Total number of files

File system used space

Displays a pie chart showing the distribution of used and free space in all file systems. Different chart types can be selected here. This provides a quick snapshot of your 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.

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 most full file systems, including the percent of space free, the total file system capacity and the file system mount point.

Users consuming the most space report

Shows the same information as the Users Consuming the Most Space panel, but in a table format.

Alerts pending

Shows active alerts which have been triggered but are still pending.

16.5.2 Discover and monitor information

IBM TotalStorage Productivity Center for Data 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 ping which checks registered agents for availability. 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
- ▶ Notification at login
- ▶ Entry in the Windows event log
- ▶ Run a script
- ▶ Send e-mail to a specified user(s)

Pings are used to generate Availability Reports, which lists the percentage of times a computer has responded to the ping. An example of an Availability Report for Ping is shown in Figure 16-4.

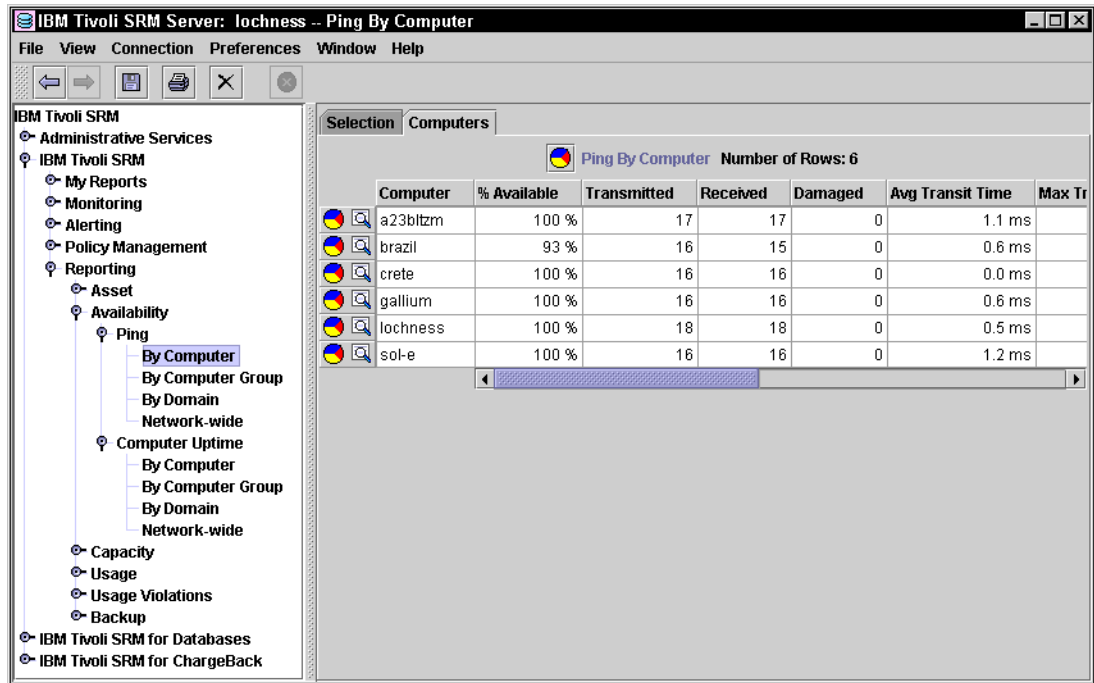


Figure 16-4 Availability Report - ping

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 16-5 shows an Asset Report for detected disks.

Computer	Path	Primary Defects	Manufacturer	Model
easter.almaden.ibm.com	/dev/rdisk/c2t2d0	752	SEAGATE	ST39204LC
easter.almaden.ibm.com	/dev/rdisk/c1t2d0	552	SEAGATE	ST39204LC
gallium	Disk 1	308	IBM-ESXS	ST318452LC
gallium	Disk 0	254	IBM-ESXS	ST318452LC
sol-e	/devices/pci@1f,4000/scsi@3/sd@0,0	116	IBM	DNES30917S

Figure 16-5 Asset report of discovered disks

Scans

The scan process is used to gather statistics about usage and trends of the server storage. Scans also gather information about the storage usage and trends within the monitored databases. Data collected by the scan jobs are tailored by *Profiles*. Results of Scan jobs are stored in the enterprise repository. This data supplies the data for the Capacity, Usage, Usage Violations, and Backup Reporting functions. These reports can be scheduled to run regularly, or they can be run ad hoc by the administrator.

Profiles limit the scanning according to the parameters specified in the profile. Profiles are used in scan jobs to specify what file patterns will be scanned, what attributes will be gathered, what summary view will be available in reports and the retention period for the statistics. IBM TotalStorage Productivity Center for Data supplies a number of default profiles which can be used, or additional profiles can be defined. Some of these include:

- ▶ **Largest files:** Gathers statistics on the largest files

- ▶ **Largest directories:** Gathers statistics on the largest directories
- ▶ **Most At risk:** Gathers statistics on the files that have been modified the longest time ago and have not been backed up since modified (Windows Agents only)

Figure 16-6 shows a sample of a report produced from data collected in scans.

Computer	Mount Point	Used Space	Percent Used Space	Free Space	Capacity
TOTAL =>		60.99 Gb	33 %	119.65 Gb	181.29 G
easter.almaden.ibm.com	/tmp	59.89 Mb	93 %	3.92 Mb	64.00 M
easter.almaden.ibm.com	/	115.95 Mb	83 %	22.65 Mb	140.00 M
brazil	/tivoli	837.14 Mb	80 %	202.86 Mb	1.02 G
sol-e	/usr	876.07 Mb	79 %	230.48 Mb	1.14 G
crete	/usr	390.09 Mb	76 %	121.91 Mb	512.00 M
sol-e	/var	189.29 Mb	75 %	60.92 Mb	278.01 M
vmwarew2ksrv1	C:/	1.45 Gb	72 %	553.54 Mb	1.99 G
easter.almaden.ibm.com	/var	747.38 Mb	70 %	312.46 Mb	1.05 G
NAS200	D:/	4.15 Gb	65 %	2.19 Gb	6.34 G
lochness	D:/	5.82 Gb	63 %	3.41 Gb	9.23 G
sol-e	/	274.94 Mb	63 %	158.96 Mb	482.12 M
NAS200	C:/	1.65 Gb	56 %	1.27 Gb	2.92 G
easter.almaden.ibm.com	/stand	38.83 Mb	52 %	34.76 Mb	81.77 M
brazil	/	522.14 Mb	50 %	517.86 Mb	1.02 G
lochness	C:/	4.79 Gb	49 %	4.98 Gb	9.77 G
sol-e	/tmp	45.35 Mb	49 %	46.49 Mb	102.04 M
crete	/tmp	255.70 Mb	49 %	256.30 Mb	512.00 M
brazil	/usr	555.03 Mb	46 %	628.97 Mb	1.16 G
easter.almaden.ibm.com	/usr	496.65 Mb	44 %	614.40 Mb	1.12 G
brazil	/mydata01	134.02 Mb	41 %	185.98 Mb	320.00 M
gallium	C:/	6.80 Gb	40 %	10.13 Gb	16.94 G
sol-e	/tivoli	1.11 Gb	39 %	1.70 Gb	2.87 G
NAS200	E:/	20.57 Gb	37 %	34.53 Gb	55.10 G
a23bitzm	D:/	1.37 Gb	34 %	2.63 Gb	4.00 G
sol-e	/fast101	308.24 Mb	34 %	595.88 Mb	961.83 M
brazil	/var	3.03 Mb	18 %	12.97 Mb	16.00 M

Figure 16-6 Summary view by file system, disk space used and disk space free

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.

16.5.3 Reporting

Reporting in IBM TotalStorage Productivity Center 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 or HTML 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 16-3 on page 262 shows a high-level summary report. Or, you can drill down to something very specific. Figure 16-7 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.

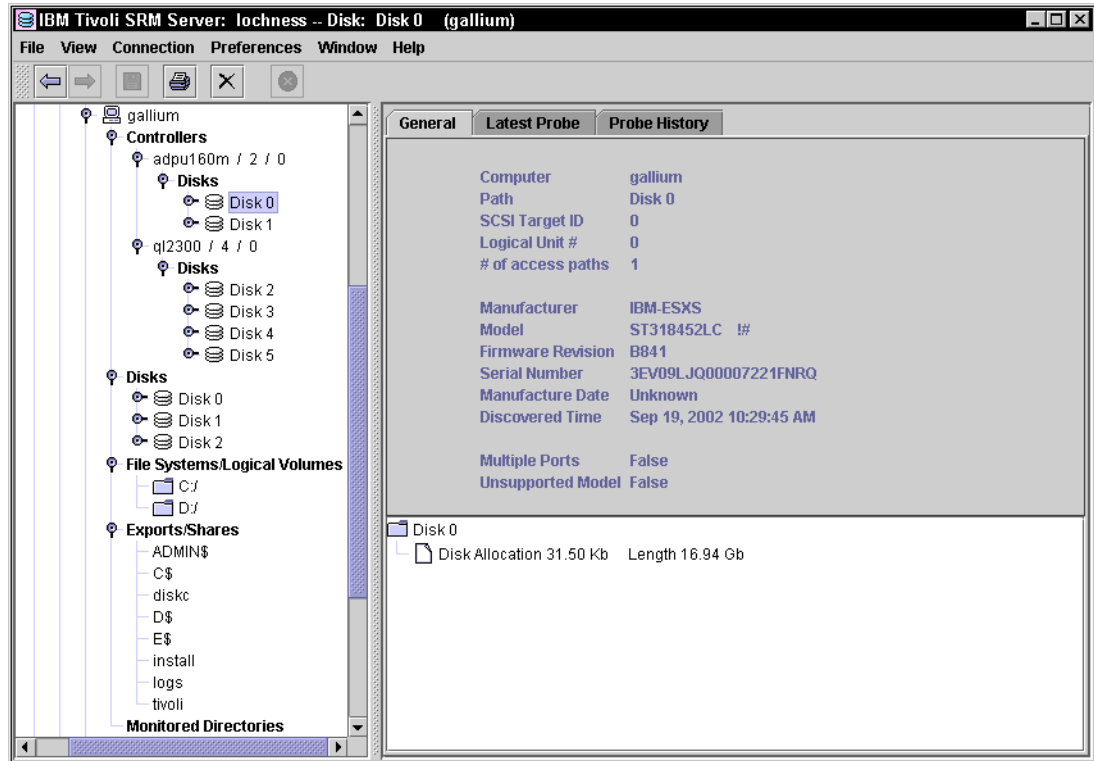


Figure 16-7 Asset reporting

Reports can be produced either system-wide or grouped into views, such as by computer, or operating system type.

IBM TotalStorage Productivity Center 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 IBM TotalStorage Productivity Center for Data. 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 files which are at risk because they have not been backed up.

Reporting on the Web

It is easy to customize IBM TotalStorage Productivity Center for Data to set up a reports Web site, so that anyone in the organization can view selected reports via their browser.

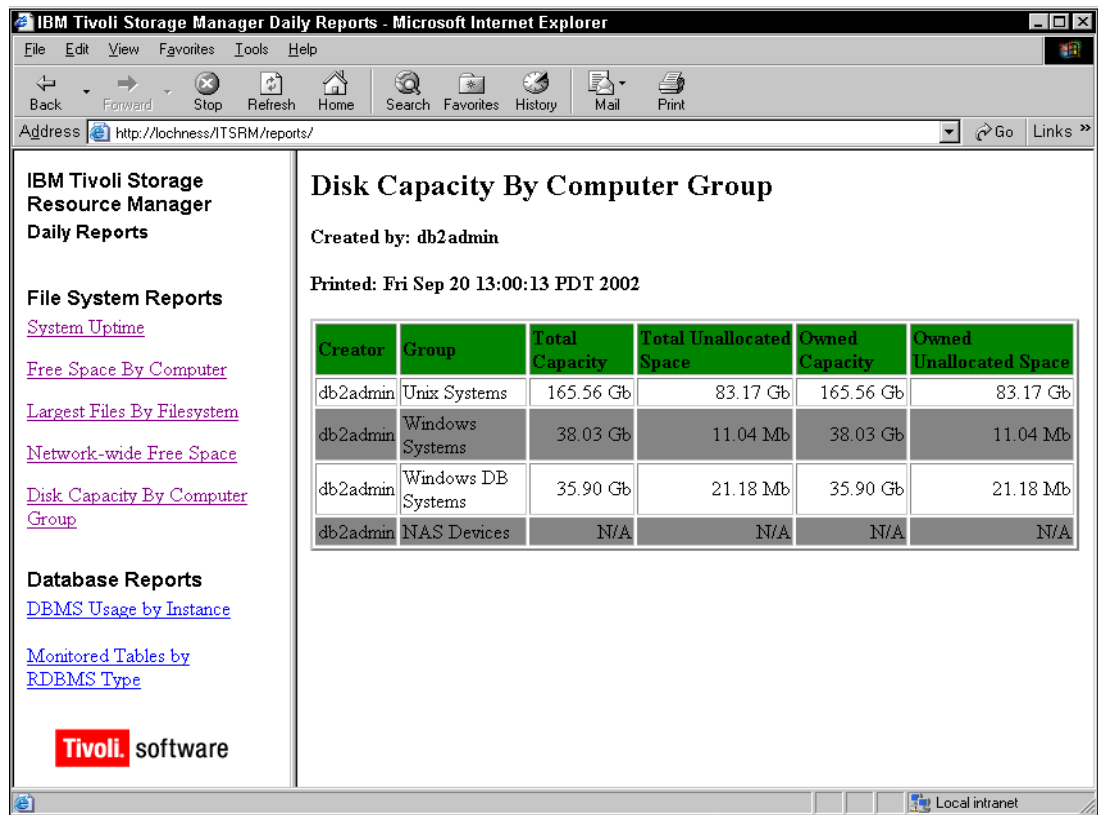


Figure 16-8 Reports on the Web

16.5.4 Alerts

An alert defines an action to be performed if a particular event occurs or condition is found. Alerts can be set on physical objects (computers and disks) and/or logical objects (file systems, directories, users, databases and operating system user groups). Alerts can tell you, for instance, if a disk has a lot of recent defects, or if a file system or database is approaching capacity.

Alerts on computers and disks come from the output of *probe* jobs and generate an alert for each object that meets the triggering condition. If you have specified a triggered action (running a script, sending an e-mail, and so on), then that action will happen if the condition is met.

Alerts on file systems, directories, users, and operating system user groups come from the combined output of a *probe* and a *scan*. Again, if you have specified an action, that action will be performed if the condition is met.

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 e-mail indicating the nature of the alert
- ▶ Run a specific script with relevant parameters supplied from the content of the Alert
- ▶ Make an entry into the Windows event log
- ▶ Popup next time the user logs in to IBM TotalStorage Productivity Center for Data
- ▶ Send an SNMP trap (Figure 16-9)

Figure 16-9 shows the Alert Log. The entries *Alert Log*, *All*, *Computer*, and *File system* are in red, signifying that an Alert threshold has been reached. Drilling down on *Computer*, you can see the details of the Alert. We can see it was caused by the system VMWARE2KSRV1 being *unreachable*.

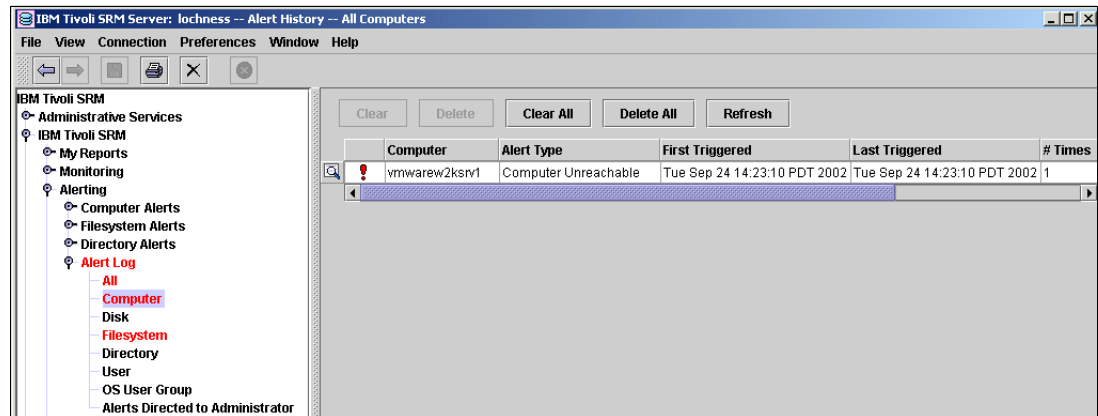


Figure 16-9 Alert Log and details

16.5.5 Chargeback: Charging for storage usage

Through the optional Tivoli Storage Resource Manager for Chargeback product, Tivoli Storage Resource Manager provides the ability to produce chargeback information for storage usage. The following items can have charges allocated against them:

- ▶ Operating system storage by user
- ▶ Operating system disk capacity by computer
- ▶ Storage usage by database user
- ▶ Total size by database-tablespace

Tivoli Storage Resource Manager can directly produce an invoice or create a file in CIMS format. CIMS is a set of resource accounting tools that allow you to track, manage, allocate and charge for IT resources and costs. For more information on CIMS, see the Web site:

<http://www.cims.com>

Chargeback is a very powerful tool for raising the awareness within the organization of the cost of storage and the need to have the appropriate tools and processes in place to manage storage effectively and efficiently.

Example 16-1 shows a chargeback report based on disk usage by user.

Example 16-1 Chargeback report based on disk usage by user

Tivoli SRM for ChargeBack
User Usage Invoice

page 1
24-Sep-02

Administrator.hb

Name	SPACE (MB)	COST 0.1042/MB
100	5	0.52
group total	5	0.52

Tivoli SRM for ChargeBack
User Usage Invoice

page 2
24-Sep-02

Tivoli.Default User Group

Name	SPACE (MB)	COST 0.1042/MB
[Supervisor]	524	54.60
1009	0	0.00
101	4	0.42
1010	1	0.10
1012	1	0.10
1084	1	0.10
111	1	0.10
1414	24	2.50
202	0	0.00
240	1	0.10
50	1	0.10
5115	10	1.04
8482	7	0.73
9727	0	0.00
adm	3	0.31
admin	523	54.50
Administrators	37,687	3,926.99
backup	27	2.81
bin	1,173	122.23
cbres	1	0.10
daemon	1	0.10
guest	1	0.10
imnadm	2	0.21
invscout	1	0.10
itso_hb	134	13.96
itso_usr	1	0.10
IUSR_LOCHNESS	1	0.10
IWAM_LOCHNESS	1	0.10
lotti	7	0.73
lp	1	0.10
nobody	1	0.10
oracle	5	0.52
root	5,857	610.30
sys	1	0.10
SYSTEM	3	0.31
uucp	2	0.21

group total	46,008	4,793.97
Tivoli SRM for ChargeBack Run Summary		page 3 24-Sep-02
User Usage Invoice	46,013 MB	4,794.49
run total		4,794

16.6 More information

IBM Tivoli Storage Resource Manager: A Practical Introduction, SG24-6886:

<http://www.redbooks.ibm.com/redbooks/pdfs/sg246886.pdf>

Exploring Storage Management Efficiencies and Provisioning, SG24-6373:

<http://www.redbooks.ibm.com/redbooks/pdfs/sg246373.pdf>

Refer to following links for other resources and related links:

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

<http://www-306.ibm.com/software/tivoli/products/totalstorage-data/>



IBM TotalStorage Productivity Center for Fabric

This chapter explains the functionality and the benefits of the IBM TotalStorage Productivity Center for Fabric (formerly known as Tivoli Storage Area Network Manager). 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 assists administrators with automatic discovery and visualization of multivendor SAN resources, active provisioning of SAN zones and zone sets, monitoring and alerting of SAN events, and intelligent and proactive prediction of future failures that may result in data unavailability.

17.1 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.

17.1.1 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 (SMIS) version 1.0. By early 2005, 17 vendors and more than 120 products had passed SMIS conformance tests. SMIS 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 SMIS 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 SMIS. 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 the IBM TotalStorage Enterprise Storage Server.

For more information on these evolving standards refer to Appendix A, "Introduction to SMIS and CIM/WBEM" on page 459 and the following Web sites:

- <http://www.snia.org>
- <http://www.ansi.org>
- <http://www.t11.org>

17.1.2 Discovery

The process of finding resources within an enterprise, including detection of network topology, is called *discovery*. IBM TotalStorage Productivity Center 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 TotalStorage Productivity Center 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. TotalStorage Productivity Center for Fabric gathers SNMP-collected information from out-of-band agents.

IBM TotalStorage Productivity Center 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.

Table 17-1 Differences in discovery capability

Function	Inband (uses fibre network)	Out-of-band (uses IP network)
Device discovery	X	X
Topology mapping	X	X
Topology monitoring	X	X
SAN identification		X
Element manager launch		X
Unit level events		X
Zone discovery		X
End point identification	X	
LUN identification	X	
Device status	X	

Function	Inband (uses fibre network)	Out-of-band (uses IP network)
Node and link level events	X	
End point port statistics	X	X
logical device-centric and host-centric views	X	

Each method has its advantages and usually a combination of both methods is used:

► In-band

- Agents are loaded onto monitored hosts.
- IBM TotalStorage Productivity Center for Fabric builds the topology map.
- Servers and disks correctly identified (name and type) on the map.
- Logical Views (LUNs, File System Capacity) of hosts are available in zones where a server has an agent loaded.
- In-band compliant devices can discover and report errors for adjoining devices.
- Agents can discover and manage the physical and logical connections from the switch to the fibre-attached disk.
- Agents can discover and manage fibre-attached hosts through contact with their HBAs.

► Out-of-band

- Discovery is not affected by zoning. Zoning limits in-band requests from management agents to discovering only those endpoints within the zone.
- MIB-enabled on switch.
- Servers and disks are shown as “Unidentified Devices” with WWN.

There are two types of discovery: a topology discovery and a full discovery. The topology discovery takes less time to perform than a full discovery. However, a topology discovery does not update the host-centric or device-centric views.

Data collection is driven by both schedule and events. A schedule triggers discovery and defaults to every 24 hours. When IBM TotalStorage Productivity Center for Fabric detects an event from a fabric switch, a new discovery begins. IBM TotalStorage Productivity Center for Fabric also supports iSCSI discovery. iSCSI discovery is performed independently from the discovery done by IBM TotalStorage Productivity Center for Fabric and requires that you enable the IP discovery of Tivoli NetView®. For details see 17.1.7, “iSCSI support” on page 276.

17.1.3 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. There are two types of zoning:

- Hardware zoning** In hardware, or port, zoning the members of a zone are the physical ports on the fabric switch.
- Software zoning** Software zoning uses the Simple Name Server (SNS) that runs inside the fabric switch. It is based on the node WWN or port WWN of the zone members to be included. Software zoning lets you create symbolic names for the zones and zone members.

The IBM TotalStorage Productivity Center for Fabric user interface lets you create, update, and delete zones and zone sets within a SAN. In addition, the interface lets you do the following:

- ▶ Assign zones to zone sets.
- ▶ Activate and deactivate zone sets.
- ▶ View zone membership.
- ▶ Add and remove zone members.

For information about viewing zones on your SAN and the zone control user interface, see the IBM TotalStorage Productivity Center for Fabric User's Guide.

17.1.4 Topology views

IBM TotalStorage Productivity Center for Fabric extracts fabric information from the database and uses Tivoli NetView to display the fabric topology. The following section describes the views available for SANs.

For each device discovered and displayed, you can open a properties dialog that shows the attributes and connections associated with that device.

Tivoli NetView displays one symbol for each SAN. Double-click an SAN symbol to display its submap. From a submap you can choose either of the following views:

- ▶ **Topology View:** Displays the entire SAN with two types of symbols, one for the SAN connection elements and another for each fabric segment.
- ▶ **Zone View:** Displays the SAN as a grouping of zones.
- ▶ **Host-centric:** Displays all host systems and their logical relationships to local and SAN-attached devices. It does not display the switches and other connection devices.
- ▶ **Device-centric:** Displays all the storage devices and their logical relationships to all the hosts. It does not display the switches and other connection devices.

17.1.5 SNMP events

IBM TotalStorage Productivity Center for Fabric can send events, which represent a change in the state of the fabric, to any event console within the enterprise that IBM TotalStorage Productivity Center for Fabric participates in. These events are generated in both SNMP and Tivoli Enterprise™ Console format.

17.1.6 Fault management

Fault management is also known as the *error detection and fault isolation* (EDFI) function. This is provided to help with problem determination on Fibre Channel SAN interconnect links. EDFI identifies faulty hardware by using a predictive failure analysis technique based on statistical modeling of the component failure distributions. EDFI looks at temporary error data across the SAN and analyzes it for patterns. Electrical and optical hardware often fails in such a way that intermittent conditions can be identified before they actually degrade into a permanent failure. Additionally, fault isolation is done to isolate the failing component. The

goal is to identify the failing optics and interconnect components and notify users before a hard failure actually occurs. It is important that this notification gets resolved promptly. The component with the elevated temporary error rates should be replaced before it actually becomes a permanent failure. EDFI does not analyze errors for internal components such as power supplies. Its function is to analyze counters that are associated with the transmission of data on SAN links. EDFI gathers temporary error counters for the entire SAN and analyzes the frequency and location of the errors by using statistical methods. The key to successfully using EDFI is to identify and remove the appropriate component before it fails permanently.

Some reasons to remove a component before it actually fails are as follows:

- ▶ A high temporary error rate causes performance problems. In some cases, these can be quite severe.
- ▶ To wait for a hard failure means unscheduled replacement of equipment when the failure actually occurs. With EDFI, equipment replacement can be scheduled and performed in a less disruptive way to the applications processing data on the affected link.
- ▶ Some failures (fractured cables are a good example) might never show a permanent error. This condition leaves the system with a degraded link indefinitely.
- ▶ Host systems, storage systems, and SAN component error recovery procedures in response to intermittent repeatable errors can cause unpredictable changes in the system I/O or Fibre Channel SAN configuration (or both). This can affect system and storage availability.

IBM TotalStorage Productivity Center for Fabric provides an EDFI user interface from the Tivoli NetView console that allows you to manage EDFI notifications and rule sets. EDFI generates standard Tivoli Enterprise Console® events that you can use for reporting purposes. EDFI can report on failing hardware components before they become permanent hardware failures. EDFI reports on faulty components in the EDFI Properties log and as an EDFI symbol on the device itself on the topology interface. Rule sets contain specific thresholds and policies that are used by the predictive failure analysis and associated fault isolation functions of IBM TotalStorage Productivity Center for Fabric. Hardware products change frequently, and rule sets allow IBM TotalStorage Productivity Center for Fabric to keep up with all the information needed to perform EDFI operations. Rule sets can be copied to the appropriate directory and then displayed through IBM TotalStorage Productivity Center for Fabric dynamically. For more information about EDFI, see “Configuring IBM TotalStorage Productivity Center for Fabric for EDFI” on page 129.

17.1.7 iSCSI support

IBM TotalStorage Productivity Center for Fabric provides basic support for iSCSI device discovery and monitoring through Tivoli NetView. iSCSI is a proposed industry-standard that allows SCSI block I/O protocols (commands, sequences and attributes) to be sent over a network using the TCP/IP protocol. The SCSI architecture is based on a client/server model. Internet SCSI takes this model into account when providing storage requests over TCP/IP networks. The client (initiator) is typically a host system such as a file server that issues read or write requests. The server (target) is a resource such as a disk array that responds to client requests. The client is the initiator and issues commands. The server is a target that fulfills the requests. The initiators and targets are identified by their world wide unique iSCSI names. The Internet Storage Name Service (iSNS) protocol is a complimentary IETF (Internet Engineering Task Force) storage management protocol for managing iSCSI devices. This protocol allows for storage devices and hosts to register themselves with an iSNS server. Subsequently, the hosts can either query the iSNS server or receive asynchronous updates from the iSNS server regarding the status of the storage devices. The iSNS protocol permits the iSNS server functions to be implemented in a number of devices: Switches, routers, storage controllers, and management console nodes. The IBM TotalStorage Productivity

Center for Fabric iSCSI support can be used either independently or in conjunction with the iSNS management framework. When iSNS is used in conjunction with IBM TotalStorage Productivity Center for Fabric, the loading of iSNS MIB files allows you to use the Tivoli NetView MIB browser. You can use the MIB browser to query the iSNS server and view the status of the iSCSI devices. For more information about iSCSI or iSNS, see the following Web site: <http://www.ietf.org> and refer to Chapter 10, “Introduction to storage networking” on page 161.

IBM TotalStorage Productivity Center for Fabric provides this support:

- ▶ All iSCSI devices discovered in the IP network can be placed in a unique iSCSI SmartSet. Also the user can create separate SmartSets for iSCSI initiator devices and target devices.
- ▶ The utility that supports the iSCSI discovery is called the `nvsniffer` program. The `nvsniffer` program is provided by Tivoli NetView.
- ▶ The iSCSI MIBs and Internet Storage Name Service (iSNS) MIBs are placed in a directory. You can then load the MIBs into NetView through the NetView MIB loading function.
- ▶ The iSCSI MIB trap definition files are used by Tivoli NetView event processing (event filtering, forwarding, pager, e-mail, and custom actions).

iSCSI discovery

The iSCSI discovery is performed independently from the discovery done by IBM TotalStorage Productivity Center for Fabric. The iSCSI discovery is done through the `nvsniffer` program and can be scheduled to refresh the iSCSI SmartSets at specified intervals. If you run `nvsniffer` regularly, the program notices when a service has been down for seven consecutive days, and will automatically delete the service object. You can start the iSCSI discovery through the Tivoli NetView menu. You can also start or stop the iSCSI discovery through the command-line interface. IBM TotalStorage Productivity Center for Fabric requires the following for iSCSI discovery:

- ▶ The device must have iSCSI MIB and SNMP support.
- ▶ The device should be configured so that the iSCSI MIB support is active. Some devices do not automatically activate iSCSI support (for example, iSCSI switch routers). Configuration can be done through the device’s configuration management interface.
- ▶ The iSCSI device must be discovered first as an IP device before `nvsniffer` can discover it as an iSCSI device. The iSCSI device discovery requires that you enable the IP discovery of Tivoli NetView.

Attention: The Tivoli NetView shipped with IBM TotalStorage Productivity Center for Fabric has disabled the IP discovery.

17.1.8 Tivoli Common Agent Services

The Tivoli Common Agent Services infrastructure provides an agent manager that provides authentication and authorization services and maintains a registry of configuration information about the managed computer systems in your environment. Management applications use the services of the agent manager to communicate securely with and to obtain information about the computer systems running the Tivoli common agent software. The Tivoli Common Agent Services infrastructure also provides common agents to act as a container to host product-specific code and common services. The common agent provides shared machine resources and secure connectivity. For more information about Tivoli Common Agent Services, see Chapter 3, “Tivoli Common Agent Services,” on page 19.

17.1.9 Interfaces to Fabric Manager

This section describes the type of interfaces available with IBM TotalStorage Productivity Center for Fabric.

▶ **IBM TotalStorage Productivity Center for Fabric console**

IBM TotalStorage Productivity Center for Fabric uses Tivoli NetView to display the fabric topology and view events in the event database.

▶ **Command-line interface**

You can use the command-line interface to issue logging commands, manager service commands, or when you create scripts that submit commands.

17.1.10 Integrating Fabric Manager into an SNMP management solution

Fabric Manager is designed to be part of an integrated SNMP management solution. If you are already implementing an SNMP solution for your storage network, Fabric Manager will integrate and add value to that solution. If you have not yet established an SNMP solution, Tivoli NetView can serve as your SNMP management console.

17.2 TPC for Fabric - overview

TPC for Fabric offers the following features:

- ▶ Zone control (Version 1.2 provided read-only access to zones) providing comprehensive SAN management from a single console.
- ▶ Performance management and capacity planning of the IBM virtualized environment.
- ▶ Provides an API to share SAN asset, topology information and to control zone configurations.
- ▶ Option of using the Cloudscape™ database which is suitable for smaller SAN installations.
- ▶ Requires single ID and password entry (the previous version required 4 entries) during install.
- ▶ Database snapshot of topology.
- ▶ Selectable launch providing the ability to launch element manager applications where an element has multiple management applications.
- ▶ Enhanced systems management integration providing an end-to-end IT solution for the enterprise.
- ▶ Providing an easy to use source side event filtering which can be done based on problem severity and entity type. This new source side event filtering can be applied to an SNMP as well as ITEC environment. This feature provides control over storage event notification.
- ▶ New level inserted in IBM TotalStorage Productivity Center for Fabric hierarchy to view VSAN; the Cisco MDS 9000 series of devices includes support for VSANs. Cisco switches offer the unique VSAN capability which allows the switch to be split into different logical SANs on the same switch hardware. This is useful where complete logical segregation of data is needed. TotalStorage Productivity Center for Fabric Version 2.2 supports this additional complexity by providing a new hierarchical view for each VSAN of its physical topology and zone views and is enabled by the SNMP discovery.
- ▶ Integration with Tivoli Enterprise Data Warehouse (TEDW).
- ▶ Integration with IBM Tivoli Service Level Advisor (ITSLA).

- ▶ Web console, providing SAN administrators access to management information from the Internet without the need to have a workstation with the IBM TotalStorage Productivity Center for Fabric console installed.

17.3 TPC for Fabric - architecture and components

Figure 17-1 shows the components of TotalStorage Productivity Center for Fabric that are used to manage storage resources. This includes devices that are connected to host systems through fibre-based hubs, bridges, switches, routers, gateways, and directors. TotalStorage Productivity Center for Fabric can also manage any host with an agent that is not connected to the SAN.

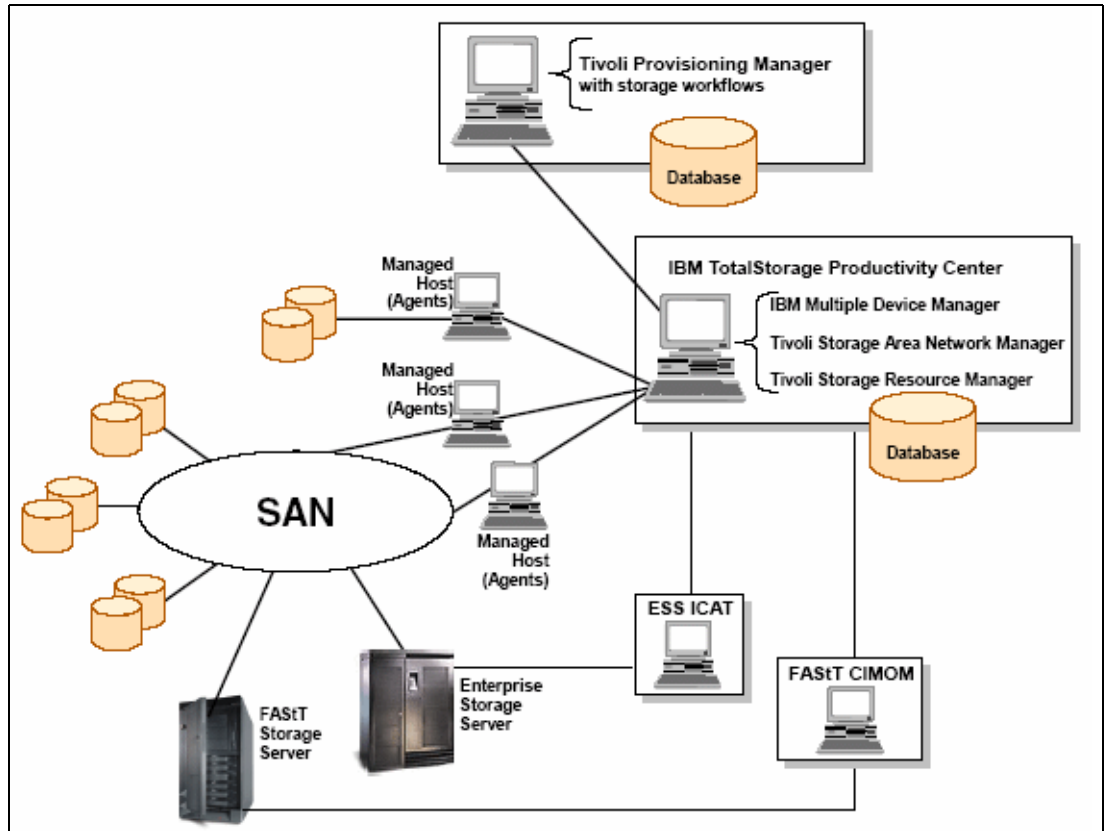


Figure 17-1 Components of TotalStorage Productivity Center for Fabric

17.3.1 Manager

With TotalStorage Productivity Center for Fabric, one system acts as the manager, and one or more other systems are the managed hosts:

The manager does the following:

- ▶ Gathers data from agents on managed hosts, such as descriptions of storage area networks (SANs), logical unit numbers (LUNs), and file system and host information.
- ▶ Provides graphical displays of SAN topology.
- ▶ Generates Simple Network Management Protocol (SNMP) events when a change is detected in the SAN fabric.
- ▶ Forwards events to the Tivoli Enterprise Console or an SNMP console.

The manager is actually composed of the *Fabric Manager* which is the manager component of IBM TotalStorage Productivity Center for Fabric and *Agent Manager* which provides authentication and authorization services for managed computer systems (and is usually installed on the same machine).

17.3.2 Agents

An agent resides on each managed host and do the following:

- ▶ Gather information about the SAN by querying switches and devices for attribute and topology information.
- ▶ Gather host-level information, such as file systems and mapping to logical units (LUNs).

17.3.3 DB2 and WebSphere®

WebSphere Express offers a cost-effective, out-of-the-box solution for managing simple yet dynamic Web sites with a simplified Web application server and a development environment based on WebSphere Studio.

DB2 Universal Database is the database management system that delivers a flexible and cost-effective database platform to build robust on demand business applications. DB2 UDB further leverages your resources with broad support for open standards and popular development platforms like J2EE and Microsoft.NET.

17.3.4 NetView

IBM Tivoli NetView discovers TCP/IP networks, displays network topologies, correlates and manages events and SNMP traps, monitors network health, and gathers performance data. Tivoli NetView meets the needs of managers of large networks by providing the scalability and flexibility to manage mission-critical environments.

The TotalStorage Productivity Center for Fabric Console performs the following functions:

- ▶ Graphically displays SAN topology — including physical and logical views
- ▶ Displays attributes of entities on the SAN
- ▶ Provides a GUI to configure and administer TotalStorage Productivity Center for Fabric
- ▶ Provides various reporting functions
- ▶ Provides a launching facility for SAN component management applications
- ▶ Displays EDFI status and logs.
- ▶ Displays status of discovered iSCSI devices

The TotalStorage Productivity Center for Fabric Console can be local - that is, installed on the manager system itself, and also remote - so that it is available on another system with NetView available. The installation process provides an option to install a remote NetView console. For more information, refer to *IBM TotalStorage Productivity Center Planning and Installation Guide*, GC32-1723:

<http://www-1.ibm.com/support/docview.wss?uid=pub1gc32172300>

17.3.5 Optional components

The Tivoli Enterprise Data Warehouse (TEDW) and IBM Tivoli Service Level Advisor (ITSLA) are optional components of TotalStorage Productivity Center for Fabric Version 2.2.

Tivoli Enterprise Data Warehouse (TEDW)

The TEDW is the central repository for enterprise systems management information where data from point Tivoli solutions as well as other vendors products can be combined to yield new value from new and integrated views. It collects, organizes, and makes data available for the purpose of analysis in order to give management the ability to access and analyze information about its business. IBM TotalStorage Productivity Center for Fabric provides ETL (Extract Translate and Load) code that allows TEDW to pull data from the TotalStorage Productivity Center for Fabric database.

The data includes basic asset and availability information on switches and port states managed by TotalStorage Productivity Center for Fabric. The state being tracked is the consensus state from various agents that see the switch or port. Each switch being tracked would be identified by manufacture, model number, WWN, IP address, for example.

IBM Tivoli Service Level Advisor (ITSLA)

TotalStorage Productivity Center for Fabric is integrated with ITSLA. Data sent by TotalStorage Productivity Center for Fabric to the TEDW is available to the ITSLA to easily add switch and port availability metrics into existing reports. TotalStorage Productivity Center for Fabric Version 2.2 provides a more complete picture of storage availability in service level reporting.

17.4 TPC for Fabric V2.2 - New features

The TotalStorage Productivity Center for Fabric Version 2.2 offers the following new features:

- ▶ Zone control
- ▶ API
- ▶ Option to use the (embedded) IBM Cloudscape database
- ▶ Database snapshot of topology
- ▶ Selectable launch

17.4.1 Zone control

To better enable comprehensive SAN management from a single console TotalStorage Productivity Center for Fabric Version 2.2 has added zone control (Figure 17-2). Existing zones can be discovered and the members that make up the zone can be viewed and modified by adding or deleting them. New zones can be created and existing zones can be deleted. TotalStorage Productivity Center for Fabric makes working with zones easy by providing an intuitive GUI that discovers SAN devices and makes them available to add or remove from zones as individual devices. Support for aliases (name assigned to a device so the name can be meaningful and familiar) is also provided. Also an alias can be a group of devices that can be managed together to make the zoning process easier.

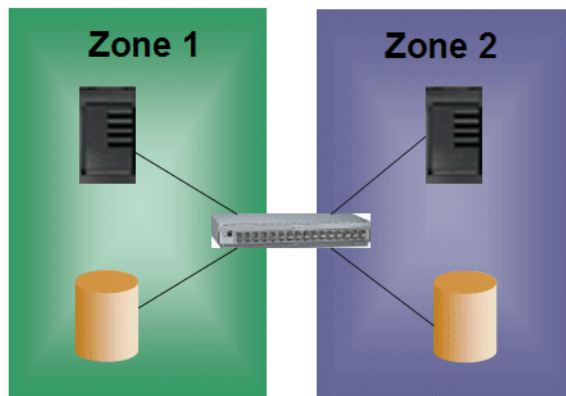


Figure 17-2 SAN management across zones

One or more zones make up a zone set. TotalStorage Productivity Center for Fabric displays the active zone set and the inactive set(s) and the zones that make up those zone sets. TotalStorage Productivity Center for Fabric can change which zone set is active.

17.4.2 IBM TotalStorage Productivity Center for Fabric API

IBM TotalStorage Productivity Center for Fabric simplifies the design, initial configuration, changes, operational monitoring, performance management and capacity planning of the IBM virtualized environment. IBM TotalStorage Productivity Center for Fabric has special recognition for the IBM SAN Volume Controller.

IBM TotalStorage Productivity Center for Fabric Version 2.2 provides an API to share SAN asset and topology information and to provide zone control. This function will query the Fabric manager database for various SAN asset and topology data and present this information through a programmable interface. The SAN data shared will include information on the SAN elements (switches, ports, hosts, disk LUNs, tape LUNs), how they are physically connected and their zoning. Access to this API is controlled through authentication.

This API will initially be exploited by the IBM virtualization family and allow the IBM TotalStorage Productivity Center for Fabric to help simplify the original configuration of the environment.

When additional physical storage is needed after the IBM Virtualization function creates LUNs on the new disk storage device it will then use the IBM TotalStorage Productivity Center for Fabric's awareness of the SAN topology to figure out the correct zone configuration that securely allows the SAN Volume Controller to have access to the additional physical storage and direct the TotalStorage Productivity Center for Fabric API to set the zones according to these instructions. Autozoning makes growing an IBM virtualized environment a more efficient process by reducing the workload of the storage administrator. Also the use of automation insures the configuration change is error free and helps meet service level objectives for application availability.

This TotalStorage Productivity Center for Fabric API is exploited only by the IBM Virtualization family at this time.

17.4.3 IBM Cloudscape database

The TotalStorage Productivity Center for Fabric Version 2.2 provides an option to use embedded Cloudscape database which is suitable for smaller SAN installations. This option

simplifies the installation and improves end user experience. During the install process the IBM Cloudscape database (V5.1.18) can be selected for use and substantially reduces the install time due to its lighter footprint and elimination of a separate install step for the database.

Small SAN installations can start out using Cloudscape to manage a smaller SAN (or in a proof of concept) and then migrate to DB2 with the assistance of a script when the SAN grows. This migration assist will maintain any manually entered data (such as nicknames for SAN devices) as well as the data used by the SAN Error Predictor (autonomic error detection and fault isolation that predicts fatal failures of optical transceivers based on data gathered over time).

TotalStorage Productivity Center for Fabric is built on proven enterprise software of NetView, WebSphere and DB2 or Cloudscape. The new installation process of TotalStorage Productivity Center for Fabric Version 2.2 reduces the time and complexity of the installation by reducing ID and password entry. It requires a single ID/password entry (the previous version required four entries).

17.4.4 Database snapshot of topology

IBM TotalStorage Productivity Center for Fabric provides a database snapshot of topology feature which is:

- ▶ An easy to use interface
 - Capture current view of SAN
 - Send current view of SAN to remote location
- ▶ For use by IBM SAN Central (level 2) when they need to understand the current SAN topology during troubleshooting
- ▶ Improving serviceability and usability

Snapshot support is provided to easily capture the current state of the SAN topology and share this information with others that may be located remotely and do not have a remote console view of the SAN.

This capability was developed to substantially reduce the amount of time needed to understand the SAN environment and speed problem solving.

This snapshot support is an automated snapshot capability that takes the needed database tables and logs on the TotalStorage Productivity Center for Fabric server, zips them and sends them to another location where they can be viewed by TotalStorage Productivity Center for Fabric. A snapshot can be taken in a Cloudscape environment and sent to a DB2 environment.

17.4.5 Selectable launch

Selectable launch is also new in IBM TotalStorage Productivity Center for Fabric Version. It provides the ability to launch element manager applications where an element has multiple management applications. The first exploitation of this is with the Cisco MDS 9000, but it not unique to Cisco and is planned for use with the IBM Virtualization family. Selectable launch providing launch options per device such as:

- ▶ VSAN configuration
- ▶ Switch-centric management
- ▶ Fabric-centric management
- ▶ Fault isolator

17.5 Functions

In this section, we illustrate with numerous screen captures the functions provided by the Total Storage Productivity center for Fabric and their usage.

17.5.1 Display a topology map of the SAN in physical and logical views

IBM TotalStorage Productivity Center for Fabric uses IBM Tivoli NetView to display topology information and views, and for monitoring devices. Tivoli NetView discovers, displays, and manages traditional TCP/IP networks. IBM TotalStorage Productivity Center for Fabric extends this function to the Storage Area Network by providing a new SAN pull-down menu item, as well as a SAN icon on the NetView top-level root map. These are shown in Figure 17-3.

Tivoli NetView, as customized with TotalStorage Productivity Center for Fabric, provides a single integrated platform to manage both traditional IP networks and SANs.

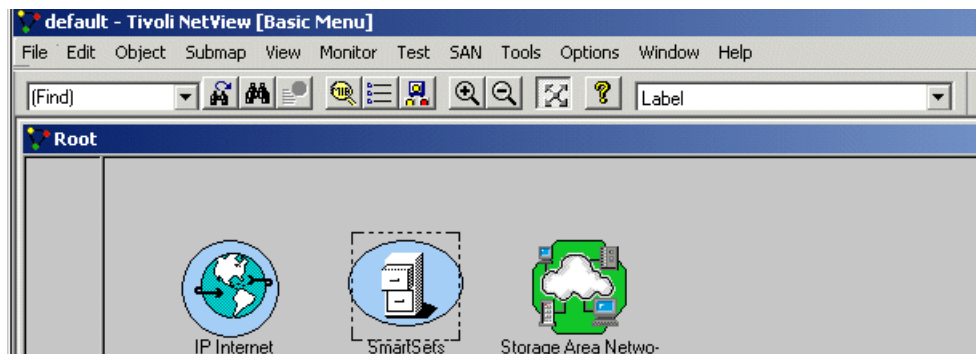


Figure 17-3 Root menu

Ways to display a topology

IBM TotalStorage Productivity Center for Fabric presents topology displays in two different ways - the icon display and the explorer display. Figure 17-3 is an example of an icon display. The explorer display is shown in Figure 17-4.

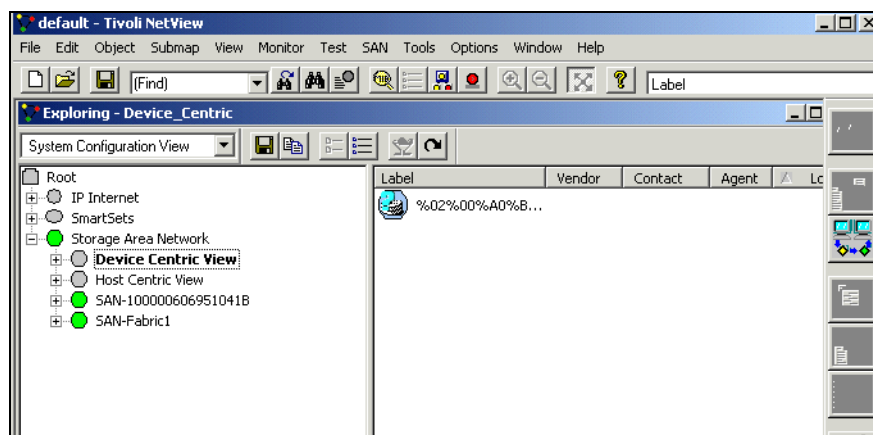


Figure 17-4 Explorer display

How is iSCSI topology information displayed

TotalStorage Productivity Center for Fabric displays iSCSI devices within the NetView SmartSet. Once IP discovery and the appropriate iSCSI Operation is selected, the iSCSI devices are discovered and a SmartSet is created.

In its recommended installation, TotalStorage Productivity Center for Fabric uses both in-band and out-of-band methods to discover and map the SAN topology. You can go from high-level views down to more specific views, focusing on different parts of the SAN. When you click the Storage Area Network icon shown in Figure 17-4, you will see the following submenu:

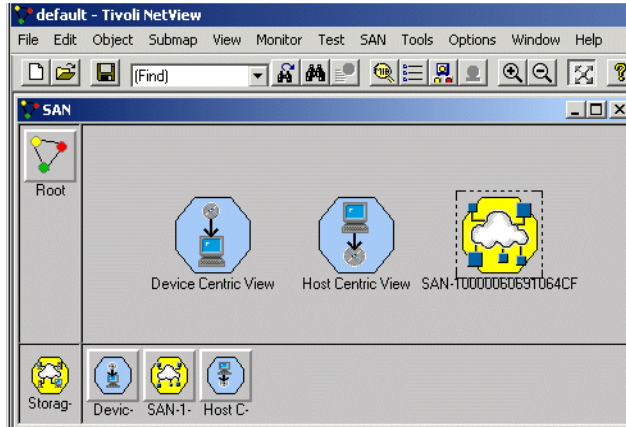


Figure 17-5 SAN submenu

The SAN icon on the right (highlighted) takes you down the physical topology displays, while the other two icons (Device Centric View and Host Centric View) provide access to the logical topology displays.

Physical topology display

A typical SAN physical topology display is shown in Figure 17-6.

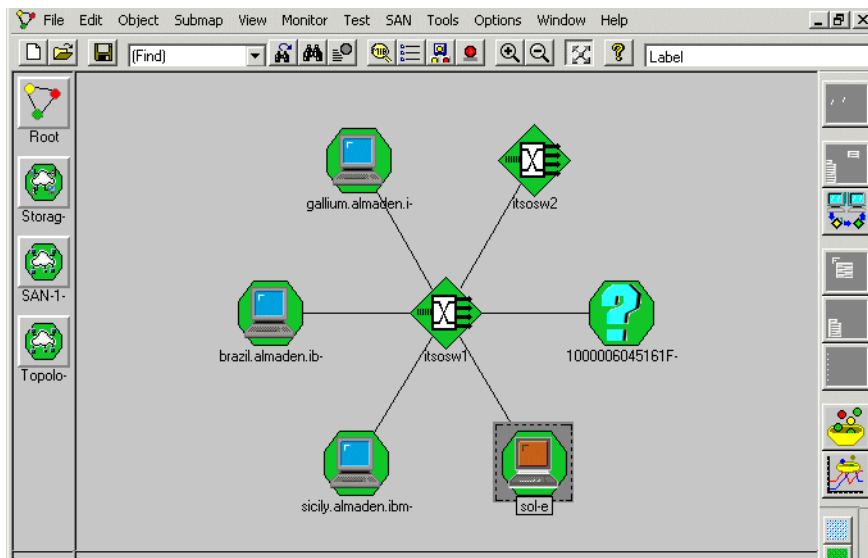


Figure 17-6 NetView physical topology display

In this case, one SAN switch is shown with the hosts and devices connected to it. All icons are shown in green, indicating they are active and available. Similarly, the connections are shown in black. NetView uses different colors for devices and connections to indicate their status. If something happens on the SAN (for example, a port on the switch fails), then the topology will be automatically updated to reflect that event. In this case an event is triggered that the port has failed, however the Server still communicates with the host Agent attached to that port. Therefore the connection line between the switch and host turns red, while the host system remains green.

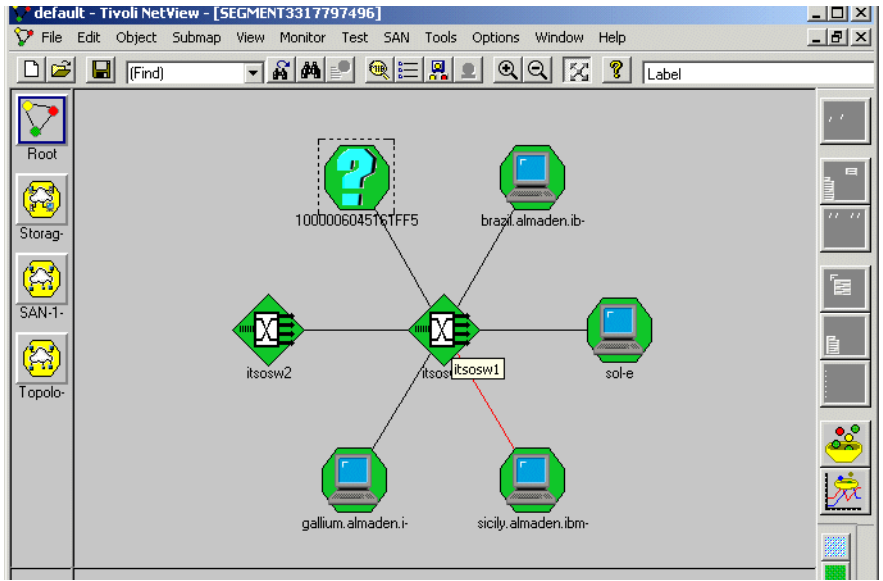


Figure 17-7 Map showing host connection lost

Zone view

TotalStorage Productivity Center for Fabric can also display switch zones, where supported by the switch API. Figure 17-8 shows two zones configured, FASTT and TSM.

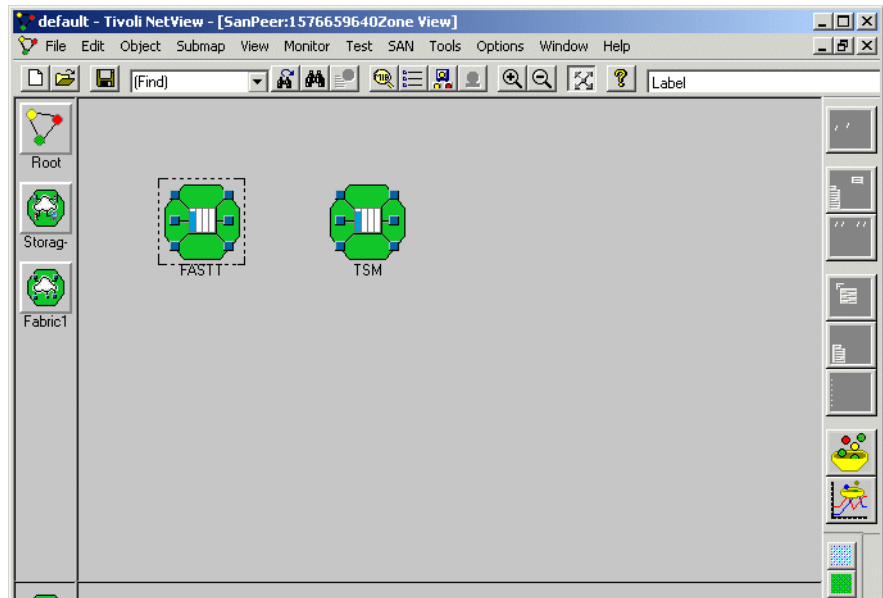


Figure 17-8 Zone view submap

If you click an individual zone, the members of that zone will be displayed. This is shown in Figure 17-9.

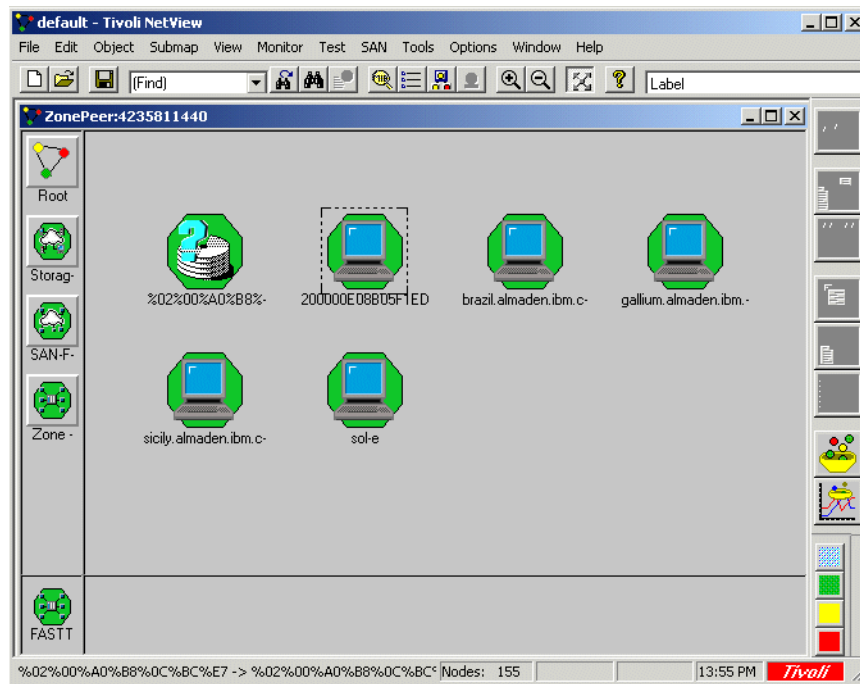


Figure 17-9 Zone members

Logical topology displays

The logical displays provided by TotalStorage Productivity Center for Fabric are the Device Centric View and the Host Centric View. Logical views, unlike the previously shown physical views, do not display the connection information between hosts and SAN fabric devices.

Device Centric View

The Device Centric View displays all the storage devices connected in the SAN, with their relationship to the hosts. The initial Device Centric View map is shown in Figure 17-10. It shows two disk systems, each with their serial numbers. The specific icons displayed depend on how your disk systems are supported by TotalStorage Productivity Center for Fabric.

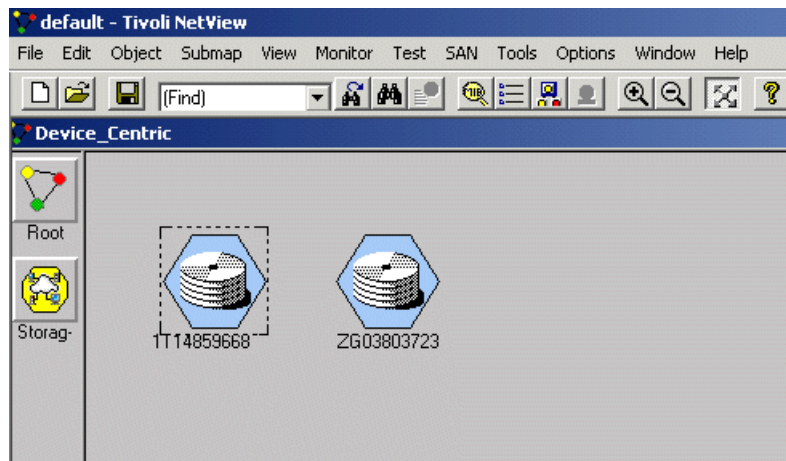


Figure 17-10 Device Centric View

You can drill down to individual devices, using the icon display, or display all the information in the explorer view. This is usually a more convenient way to display this information as it is more complete. If you select the **Explorer** icon, you will see the map shown in Figure 17-11. You can see that the LUNs for both storage systems displayed on the left hand panel. For each LUN, you can drill down to find out which host and operating system has assigned that LUN. In this example, disk system IT14859668 has five LUNs, and each LUN is associated with one or two named hosts. For example, the first LUN is associated with the hosts SENEGAL and DIOMEDE, running Windows 2000. You can drill down one step further from the operating system to display the filesystem installed on the LUN. There is one LUN discovered in the other disk system, which is used for the filesystem /mssfs on the AIX system CRETE.

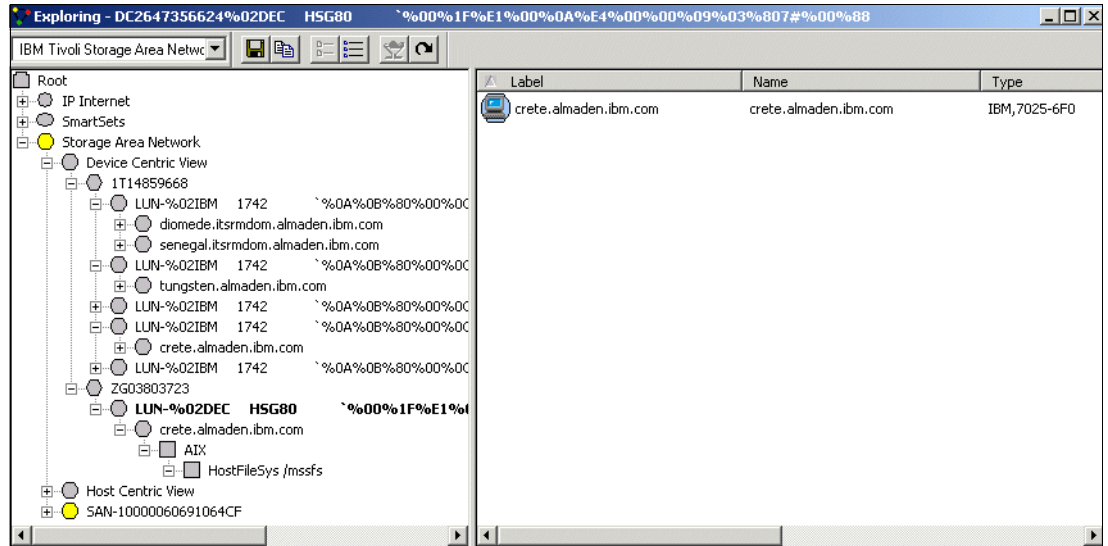


Figure 17-11 Device Centric View - explorer display

Host Centric View

The Host Centric View displays all the Managed Hosts, and their associated local and SAN-attached storage devices. The explorer Host Centric View is shown in Figure 17-12. You can see that each filesystem associated with each Managed Host or agent is displayed on the right-hand side. The left hand pane shows all the Managed Hosts.

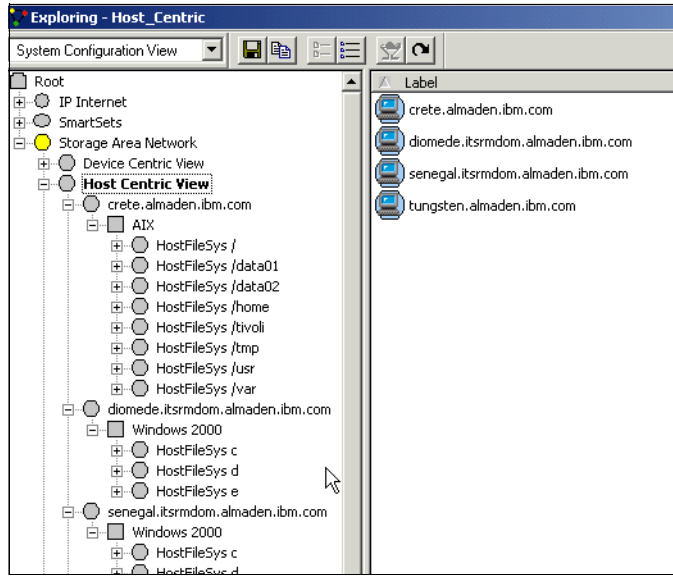


Figure 17-12 Host Centric View

You can drill down on the Host filesystem entries to show also the logical volume (and LUN if fibre-attached) associated with the filesystem, shown in Figure 17-13.

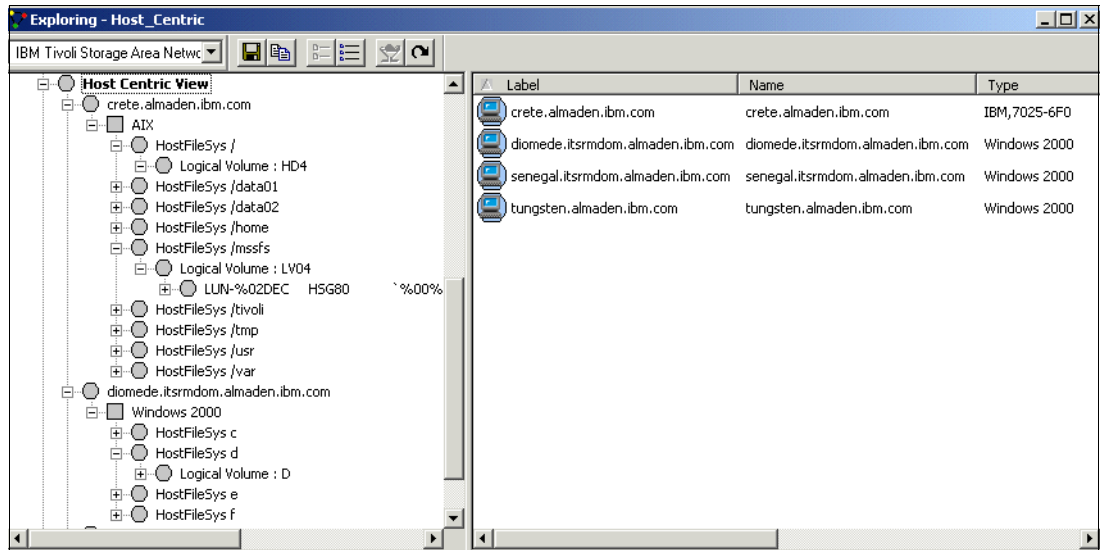


Figure 17-13 Host Centric View - logical volumes and LUN

Summary display

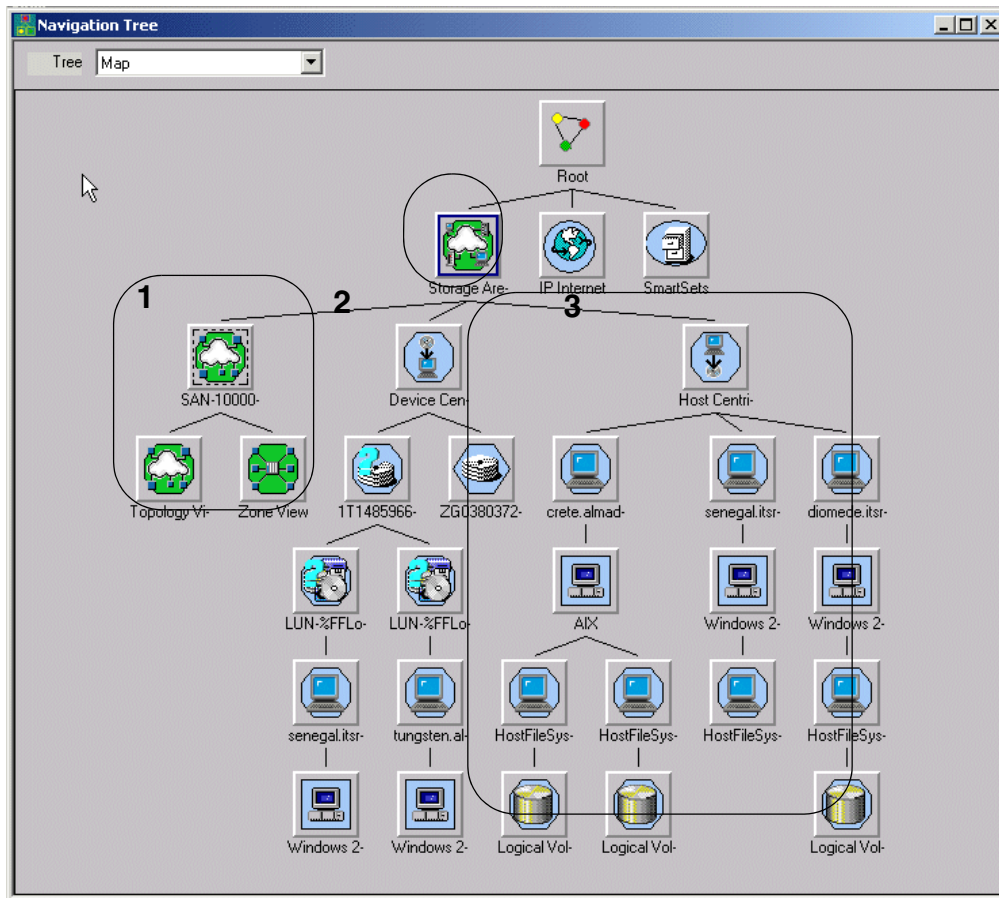


Figure 17-14 Navigation tree for IBM TotalStorage Productivity Center for Fabric

You can also see a summary or navigation display which holds a history of all the maps which you have navigated in TotalStorage Productivity Center for Fabric. In Figure 17-14, we have opened up all three views of TotalStorage Productivity Center for Fabric, and therefore can see a very comprehensive display. This is because we have drilled down a Device Centric View, drilled down a Host Centric View, and finally navigated the physical topology, and then opened the Navigation Tree.

The figure shows:

- ▶ (1) SAN View (third row, left, see #1)
 - Topology View
 - Zone View
- ▶ (2) Device Centric View (third row, middle, see #2)
- ▶ (3) Host Centric View (third row, right, see #3)

Object properties

If you click any device to select it, (from any map), then right-click and select **Object Properties** from the pop-up menu, this will bring up the specific properties of the object. In Figure 17-17, we selected a switch, and displayed the Properties window, which has seven different tabs. The figure shows the events tab, listing events which have been received for this switch.

17.5.2 Real-time monitoring and alerts

The primary business purpose for TotalStorage Productivity Center for Fabric is to keep the storage infrastructure running to support revenue-generating activities. We have already seen how the topology display will automatically update to reflect devices becoming available so that network administrator can quickly respond.

Another way to improve availability is to use monitoring and reports to anticipate devices that are beginning to fail.

Not all failures can be predicted - however, many devices may fail gradually over time, so reporting a history of problems can help anticipate this. For example, if you see a lot of transmission errors occurring over a period of time, you might anticipate a component failure, and schedule maintenance testing or pre-emptive replacement to minimize impact on generating revenue or other critical applications.

Reporting is provided by NetView and includes historical and real-time reporting, on an individual device, or on a defined group of devices.

With NetView you have a very flexible capability to build your own reports according to your specific needs. The reports we are interested in for TotalStorage Productivity Center for Fabric are reports against the objects in the MIB provided by the switch vendor.

The data elements in the MIB can report on status (device working, not working) and performance (x frames were transmitted over this switch port in Y seconds).

Historical reporting

With NetView you can display historical reports based on data collected. Figure 17-15 shows a report of data collected over 8 ports in a two minute interval. You can set up the data collection to look for thresholds on various MIB values, and send a trap when defined values are reached.

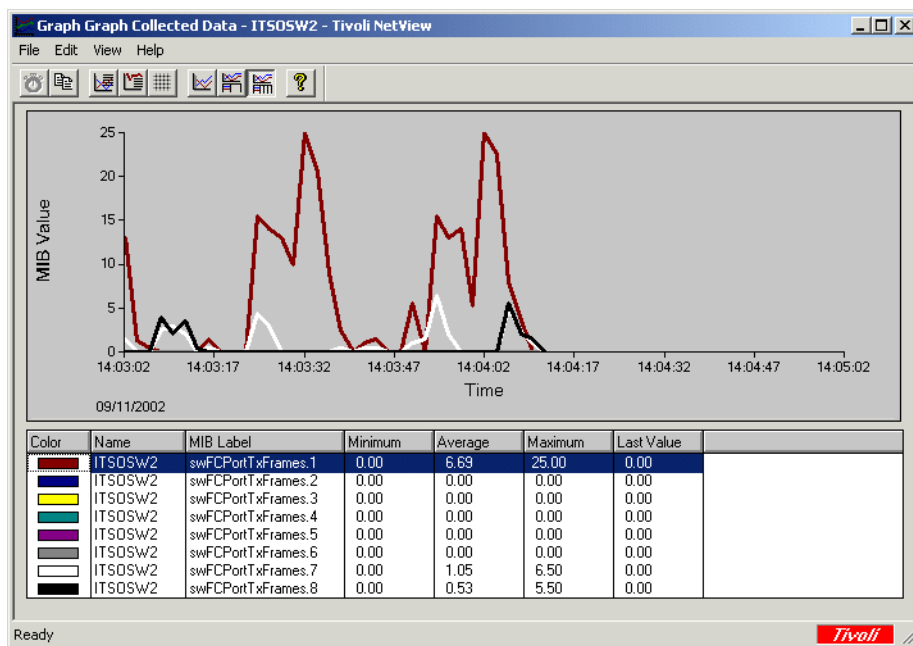


Figure 17-15 Graph of number frames transmitted over 8 ports in a 2-minute interval

Combining the MIB objects with the canned and customized reporting from NetView provides the storage administrator with the tools needed to help keep the SAN running all the time.

Real-time reporting

NetView can also track MIB values in real-time. Figure 17-16 shows real-time monitoring of traffic on switch ports. The graph shows the number of frames transmitted from a specific port on a particular switch over a specified time interval. You can set the polling interval to show how often the graph will update.

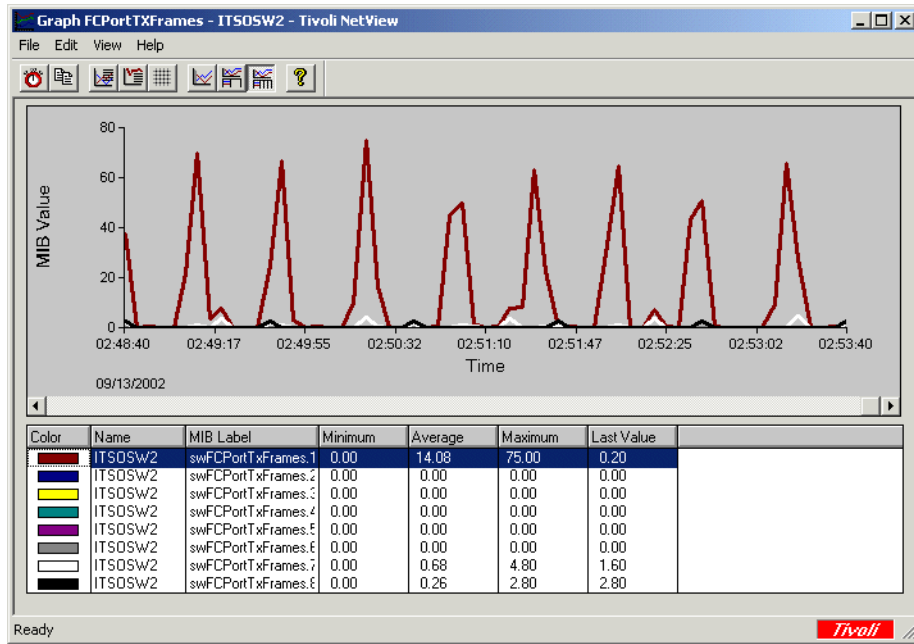


Figure 17-16 Number of frames transmitted over time

You can also create graphs from multiple devices using NetView SmartSets.

Log events

IBM TotalStorage Productivity Center for Fabric collects events received from devices. Figure 17-17 shows an example of events logged by one of the switches.

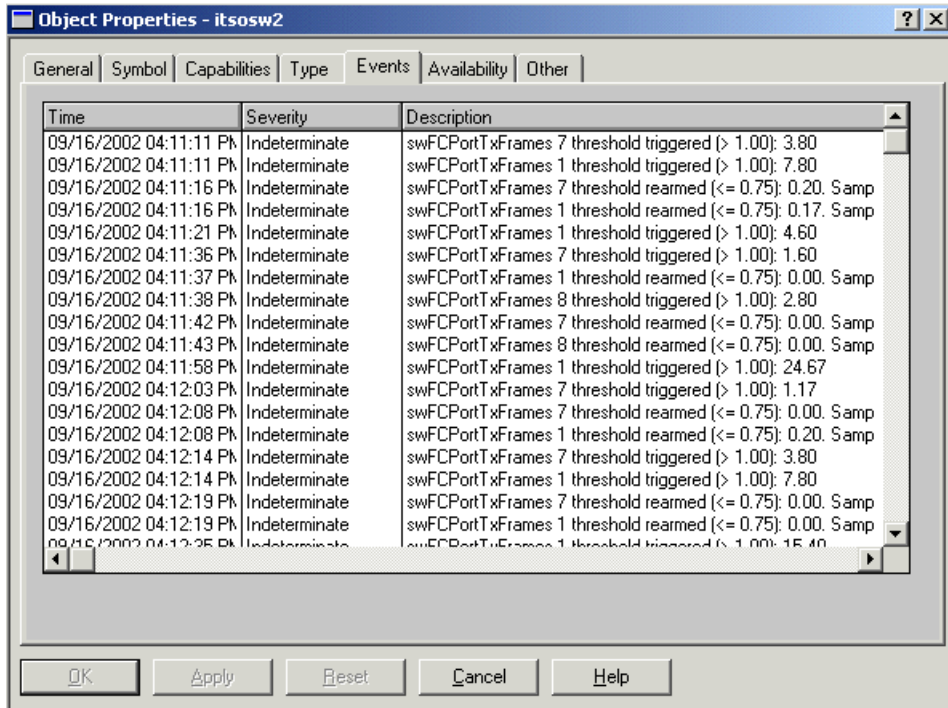


Figure 17-17 Switch events

17.5.3 Error Detection and Fault Isolation

Error Detection and Fault Isolation (EDFI - SAN Error Predictor) is implemented in TotalStorage Productivity Center for Fabric Version to provide a way to predict errors on the optical links that are used to connect SAN components (including HBA to switch, switch to switch, and switch to storage connections). EDFI is a standard function of IBM TotalStorage Productivity Center for Fabric and can be enabled or disabled from the standard menus as shown in Figure 17-18.

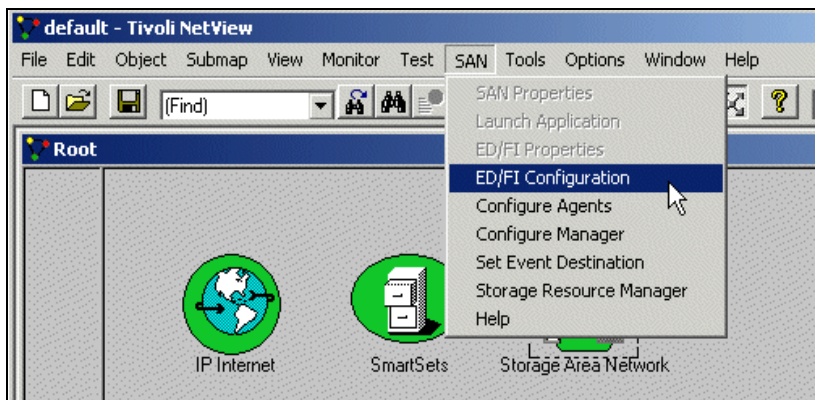


Figure 17-18 EDFI configuration

By using Predictive Failure Analysis (PFA), downtime of SAN components can be significantly decreased as it is possible to remove problematic components before failure. This can significantly reduce operational cost of SANs. The EDFI function collects data from TotalStorage Productivity Center for Fabric agents, out-of-band and/or inband as available. The polling interval is every 15 minutes. The data is stored in the TotalStorage Productivity Center for Fabric database. This data is then analyzed using various statistical methods and

from this future errors are predicted. The predicted errors are presented in the NetView interface by “adorning” the appropriate icons as shown in Figure 17-19.

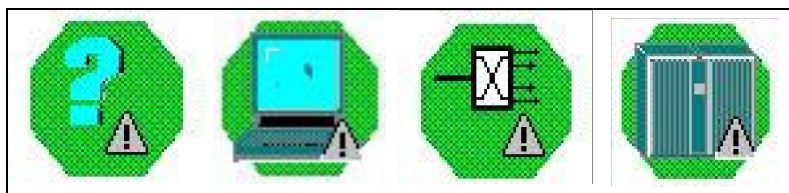


Figure 17-19 Adornment of icons in NetView

The adornment means that the exclamation point is superimposed on the icon representing the device where the error is predicted. A TEC event and SNMP trap are also generated. In a medium to large SAN, the topology map is complex and adornment icons may not be readily located, given its GUID or WWN. To make identification easier, you can use the NetView search function to find adorned devices.

Highlight faults

Events about SAN components are communicated to TotalStorage Productivity Center for Fabric by agents. TotalStorage Productivity Center for Fabric logs the events in a log that looks like Figure 17-17 above. TotalStorage Productivity Center for Fabric then evaluates the event, and decides if the event is (or is not) a fault.

If the event is a fault, TotalStorage Productivity Center for Fabric indicates the change in status by changing the color of the object in the topology map. Here is just one example, where we simulate the loss of the switch by powering it off. Other common faults include a Fibre Channel cable breaking, a host crash, and an unavailable storage frame.

Figure 17-20 shows the topology display once it is refreshed. The healthy devices are in green, and the unhealthy devices are in red. We have also circled the devices that turned red for clarity:

- ▶ Device #5 - the switch - is obvious. The power went out, so the switch no longer responds.
- ▶ All the links (connections) from the switch to the devices also went red.
- ▶ The Devices 1, 2, 3, and 4 went red. These devices have no agent on them, are known to TotalStorage Productivity Center for Fabric only through the switch nameserver, and are unavailable now that the switch is down. TotalStorage Productivity Center for Fabric's map displays this accurately.
- ▶ The hosts with agents on them (CRETE, TUNGSTEN, SENEGAL, DIOMEDE) remain green as they still communicate with the Server via TCP/IP. Only their connection link to the switch turns red.

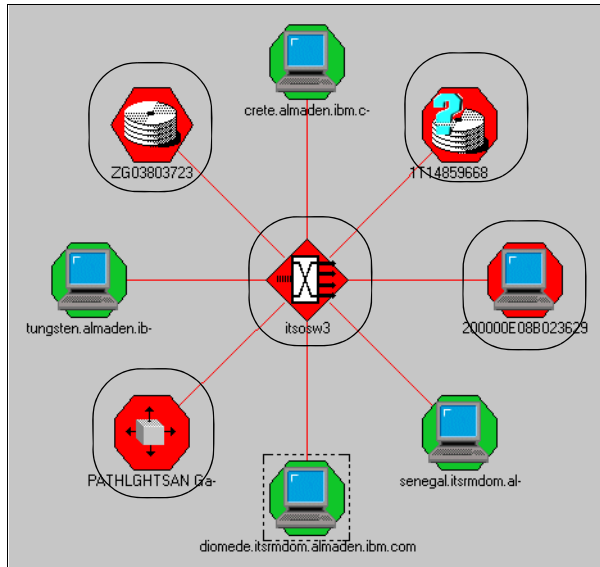


Figure 17-20 Map showing effects of switch losing power

When the switch is fixed, as it powers up, it will send an event back to TotalStorage Productivity Center for Fabric, who will re-query the switch (running both the topology scanner and the attribute scanner). The topology will be refreshed to reflect the switch being back online.

17.5.4 Launch vendor-provided applications to manage components

TotalStorage Productivity Center for Fabric provides a launch platform for many individual vendor management applications. In some cases these are automatically discovered, and in other cases they can be manually configured into NetView. These applications might be for a switch, hub, gateway or storage frame. Figure 17-21 shows an example of launching the automatically discovered switch management application.

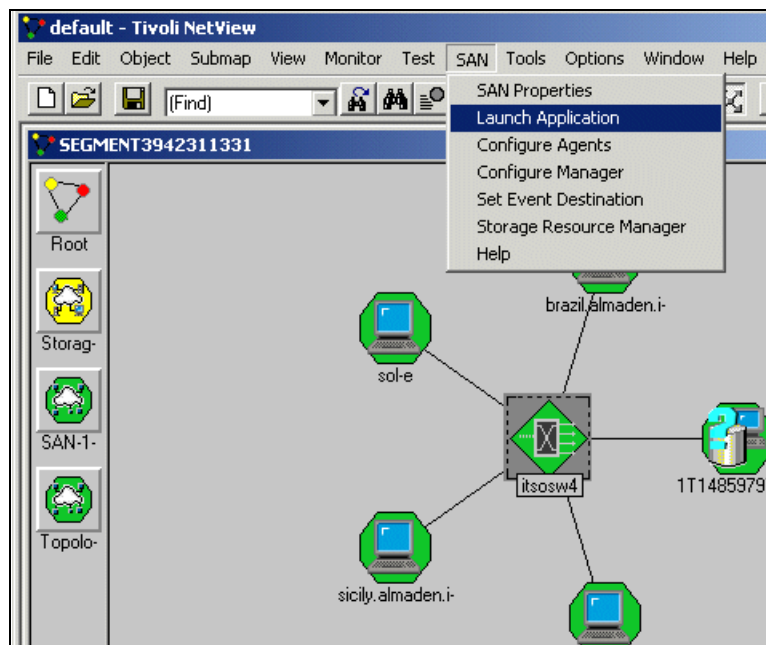


Figure 17-21 Vendor application launch

TotalStorage Productivity Center for Fabric provides three methods for locating such applications:

- ▶ **Native support:** For some devices, TotalStorage Productivity Center for Fabric will automatically discover and launch the device-related tool. IBM TotalStorage Productivity Center for Fabric has an internal set of rules (in XML format) by which it identifies the devices whose tool it can launch.
- ▶ **Web interface support:** Some devices are not discovered automatically, but have a Web interface. TotalStorage Productivity Center for Fabric can be configured with the URL, so that it can subsequently launch Web interface.

Non- Web interface support: Other applications have no Web interface. TotalStorage Productivity Center for Fabric offers you the ability to configure the toolbar menu to launch any locally-installed application from the TotalStorage Productivity Center for Fabric console. Note, these applications must be locally installed on the TotalStorage Productivity Center for Fabric Server.

17.5.5 System management

IBM TotalStorage Productivity Center for Fabric Version 2.2 provides enhanced system management integration to improve the end-to-end IT solution for the enterprise IBM TotalStorage Productivity Center for Fabric currently provides categories of events so an ITEC administrator can write rules to filter TotalStorage Productivity Center for Fabric generated events at the ITEC. In a very active network event storms can flood the ITEC, wasting resources sending alerts that are trivial or duplicates and even worse overwhelming the operator with useless data. ITEC integration is taken to a new level in TotalStorage Productivity Center for Fabric Version 2.2 by providing an easy to use source side event filtering. This filtering can be done based on problem severity and entity type. (Entity types are predefined categories for SAN resources such as HBA, LUN, zone set and so on). This new source side event filtering can be applied to an SNMP as well as ITEC environment.

17.6 More information

- ▶ For more information on SAN software, check the following Web site:
<http://www.storage.ibm.com/ibmsan/products/sansoftware.htm>
- ▶ For more information on ITSAMN, refer to this redbook and installation guide:
IBM Tivoli Storage Area Network Manager: A Practical Introduction, SG24-6848:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246848.html?Open>
- ▶ *IBM TotalStorage Productivity Center Planning and Installation Guide*, GC32-1723-00
<http://www-1.ibm.com/support/docview.wss?uid=pub1gc32172300>
- ▶ Product Manuals can be found at the following Web site:
<http://publib.boulder.ibm.com/tividd/td/StorageAreaNetworkManager1.3.html>
- ▶ Support for TotalStorage Productivity Center for Fabric can be found at:
<http://www.ibm.com/software/sysmgmt/products/support/IBMTotalStorageProductivityCenterforFabric.html>



TotalStorage Productivity Center for Disk

This chapter describes the functionality of the TotalStorage Productivity Center for Disk (formerly known as MDM Performance Manager).

TotalStorage Productivity Center 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

18.1 Overview

As a component of the IBM TotalStorage Productivity Center, TotalStorage Productivity Center for Disk (TPC for Disk) is designed to help reduce the complexity of managing storage devices by allowing administrators to configure, manage and monitor storage performance from a single console. The TotalStorage Productivity Center for Disk was formerly known as Multiple Device Manager (MDM) Performance Manager. MDM contained three separate subcomponents namely:

- ▶ Device Manager
- ▶ Performance Manager
- ▶ Replication Manager

These subcomponents have been separated for specific functionality into TotalStorage Productivity Center for Disk as well as TotalStorage Productivity Center for Replication. In this chapter we discuss TotalStorage Productivity Center for Disk which contains two subcomponents:

- ▶ **TotalStorage Productivity Center Common Base** - which is common with both TotalStorage Productivity Center for Disk as well as TotalStorage Productivity Center for Replication
- ▶ **TotalStorage Productivity Center for Disk** - Performance Manager Functionality

TotalStorage Productivity Center for Disk is designed to centralize management of networked storage devices that implement the Storage Management Interface Specification (SMI-S) established by the Storage Networking Industry Association (SNIA), including the IBM TotalStorage Enterprise Storage Server (ESS), IBM TotalStorage DS4000 series family (formerly known as FASTT family), and IBM TotalStorage SAN Volume Controller, and the devices that they manage. The TotalStorage Productivity Center for Disk is designed to:

- ▶ Help reduce storage management complexity and costs while improving data availability
- ▶ Centralize management of storage devices through open standards (SMI-S)
- ▶ Enhance storage administrator productivity
- ▶ Improve storage resource utilization
- ▶ Offer proactive management of storage devices

18.2 TotalStorage Productivity Center Common Base

The TotalStorage Productivity Center Common Base builds on the IBM Director technology. It uses Service Level Protocol (SLP) to discover supported storage systems on the SAN. Service Level Protocol enables the discovery and selection of generic services accessible through an IP network. The TotalStorage Productivity Center Common Base then creates managed objects to represent discovered devices. The discovered devices are displayed as individual icons in the Group Contents pane of the IBM Director console (See Figure 18-1).

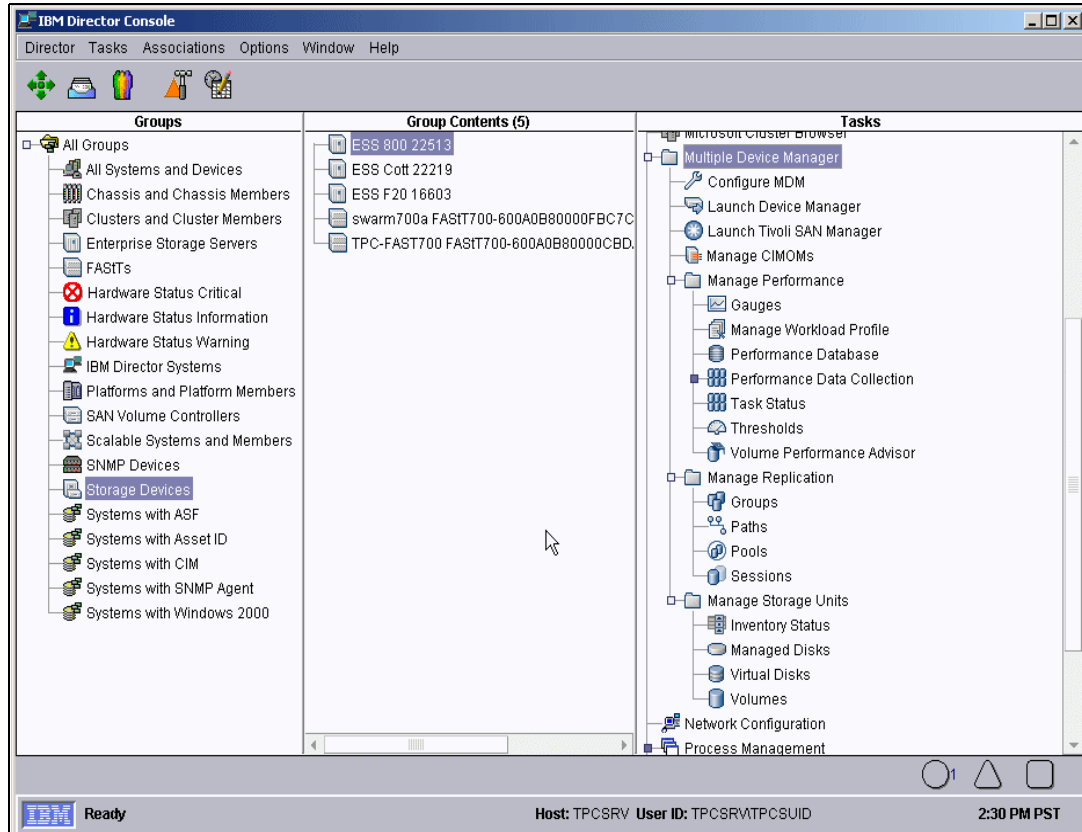


Figure 18-1 IBM Director Console

In summary, the TotalStorage Productivity Center Common Base is responsible for:

- ▶ Discovery of supported devices (SMI-S compliant)
- ▶ Data collection (asset, availability, configuration)
- ▶ Providing a topographical view of storage

The TotalStorage Productivity Center Common Base health monitoring keeps you aware of hardware status changes in the discovered storage devices. You can drill down to the status of the hardware device, if applicable. This enables you to understand which components of a device are malfunctioning and causing an error status for the device.

TotalStorage Productivity Center Common Base also provides a subset of configuration functions for the managed devices, primarily LUN allocation and assignment. Its function includes certain cross-device configuration, as well as the ability to show and traverse inter-device relationships. These services communicate with the CIM Agents that are associated with the particular devices, to perform any required configuration. Devices that are not SMI-S compliant are not supported. You can find more information on SMI specifications and CIM in Appendix A, "Introduction to SMIS and CIM/WBEM" on page 459.

TotalStorage Productivity Center for Disk supports the discovery and LUN provisioning of IBM TotalStorage DS4000 series storage systems, IBM TotalStorage SAN Volume Controller, and IBM TotalStorage Enterprise Storage Server. The user can view essential information about the storage, view the associations of the storage to other devices, and change the storage configuration. DS4000 and Enterprise Storage Servers, attached to the SAN or attached behind the SAN Volume Controller, can be managed by TotalStorage Productivity Center for Disk.

TotalStorage Productivity Center for Disk and IBM TotalStorage Productivity Center for Replication also include the IBM TotalStorage Productivity Center for Fabric Bonus Kit, which supports up to 64 Fibre Channel ports. With this function, the storage administrator can view Fibre Channel as well as SMI-S enabled storage devices, automate zoning of clients, LUNs, and storage devices for the TotalStorage SAN Volume Controller environment, and drill-down from storage devices to the underlying infrastructure. If more than 64 SAN ports are required in an enterprise, you can obtain additional capacity by ordering IBM TotalStorage Productivity Center for Fabric.

18.3 Performance monitoring and management

TotalStorage Productivity Center for Disk can provide performance monitoring of ESS and SVC storage devices, customization of thresholds based on your storage environment, and generation of events if thresholds are exceeded. In addition, TotalStorage Productivity Center for Disk is designed to help the IT administrators select the LUN for better performance. TotalStorage Productivity Center for Disk allows the IT administrator to specify both when, and how often the data should be collected. In addition, TotalStorage Productivity Center for Disk is designed to help support high availability for critical applications by providing customization of threshold settings and generating alerts when these thresholds are exceeded. TotalStorage Productivity Center for Disk can also provide gauges to track real-time performance. TotalStorage Productivity Center for Disk is designed to enable the IT administrator to:

- ▶ Monitor performance metrics across storage subsystems from a single console
- ▶ Receive timely alerts to enable event action based on customer policies
- ▶ Focus on storage optimization through identification of best LUN

The TotalStorage Productivity Center for Disk provides the raw capabilities of initiating and scheduling performance data collection on the supported devices, of storing the received performance statistics into database tables for later use, and of analyzing the stored data and generating reports for various metrics of the monitored devices. In conjunction with data collection, the TotalStorage Productivity Center for Disk is responsible for managing and monitoring the performance of the supported storage devices. This includes the ability to configure performance thresholds for the devices based on performance metrics, the generation of alerts when these thresholds are exceeded, the collection and maintenance of historical performance data, and the creation of gauges, or performance reports, for the various metrics to display the collected historical data to the end user. The TotalStorage Productivity Center for Disk enables you to perform sophisticated performance analysis for the supported storage devices.

Figure 18-2 on page 301 describes TotalStorage Productivity Center for Disk performance management functions.

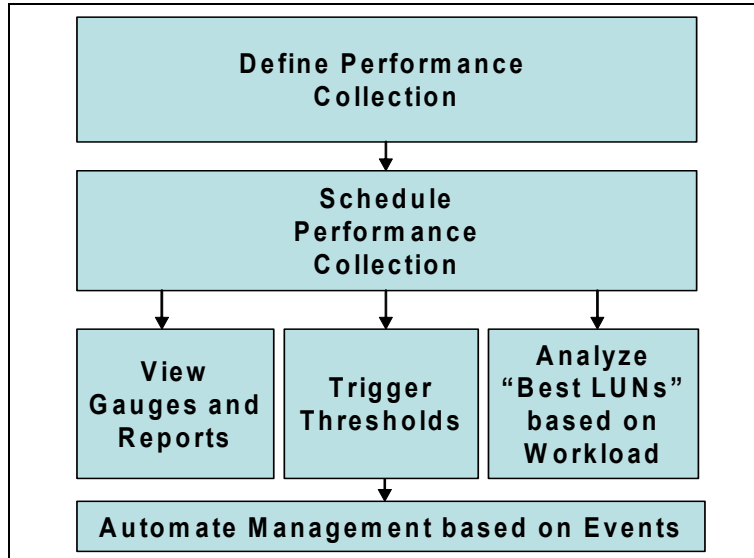


Figure 18-2 Performance management functions

Collecting data from devices

The TotalStorage Productivity Center for Disk collects data from the IBM TotalStorage Enterprise Storage Server (ESS), SAN Volume Controller (SVC), and SMI-S enabled devices. Each data collector collects performance data from one or more storage groups, all of the same device type (for example, ESS or SAN Volume Controller). Each data collector has a start time, a stop time, and a sampling frequency. The performance sample data is stored in DB2 database tables.

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 to 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.

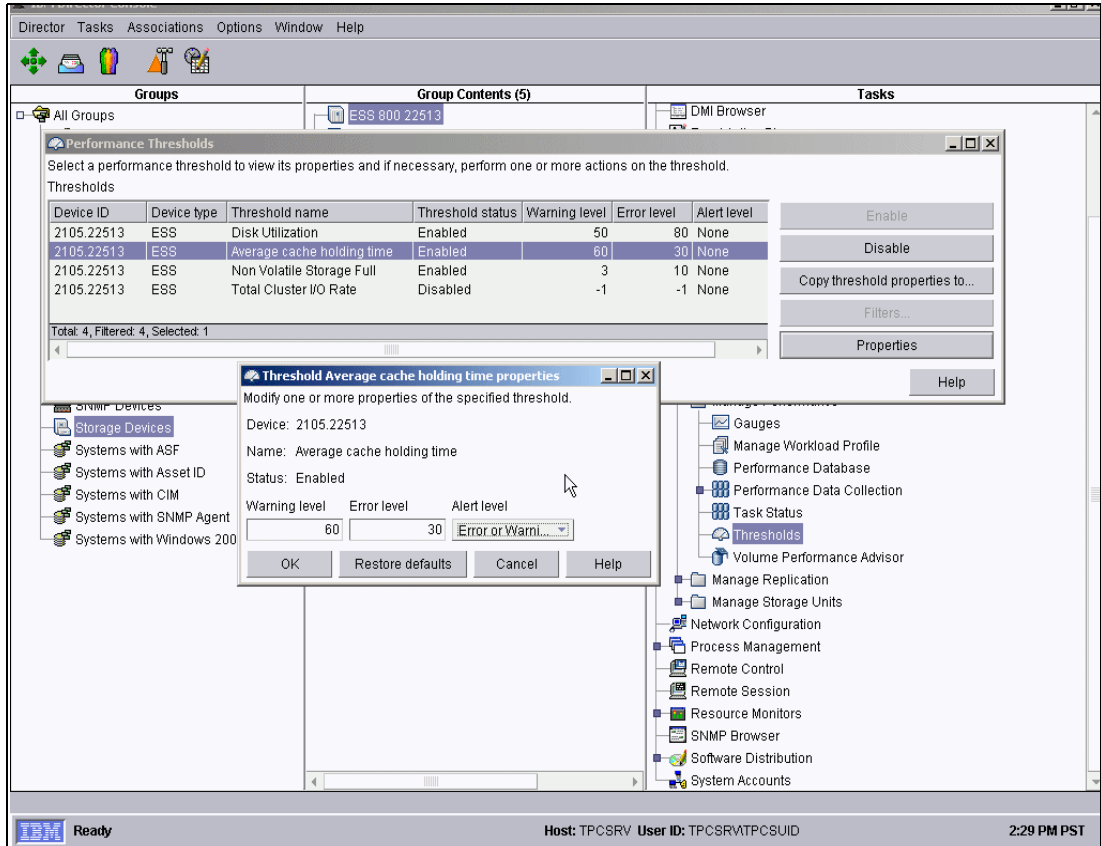


Figure 18-3 Performance threshold properties

The eligible metrics for threshold checking are fixed for each storage device. If the threshold metrics are modified by the user, the modifications are accepted immediately and applied to checking being performed by active data collectors. Examples of threshold metrics include:

- ▶ Disk utilization value
- ▶ Average cache hold time
- ▶ Percent of sequential I/Os
- ▶ I/O rate
- ▶ NVS full value
- ▶ Virtual disk I/O rate
- ▶ Managed disk I/O rate

There is a user interface that supports threshold settings, enabling a user to:

- ▶ Modify a threshold property for a set of devices of like type.
- ▶ Modify a threshold property for a single device.
 - Reset a threshold property to the IBM-recommended value (if defined) for a set of devices of like type.

IBM-recommended critical and warning values will be provided for all thresholds known to indicate potential performance problems for IBM storage devices.
 - Reset a threshold property to the IBM-recommended value (if defined) for a single device.
- ▶ Show a summary of threshold properties for all of the devices of like type.
- ▶ View performance data from the TotalStorage Productivity Center for Disk database.

Gauges

The TotalStorage Productivity Center for Disk supports a performance-type gauge. The performance-type gauge presents sample-level performance data. The frequency at which performance data is sampled on a device depends on the sampling frequency that you specify when you define the performance collection task. The maximum and minimum values of the sampling frequency depend on the device type. The static display presents historical data over time. The refreshable display presents near real-time data from a device that is currently collecting performance data.

The TotalStorage Productivity Center for Disk enables a Multiple Device Manager user to access recent performance data in terms of a series of values of one or more metrics associated with a finite set of components per device. Only recent performance data is available for gauges. Data that has been purged from the database cannot be viewed. You can define one or more gauges by selecting certain gauge properties and saving them for later referral. Each gauge is identified through a user-specified name and, when defined, a gauge can be started, which means that it is then displayed in a separate window of the Multiple Device Manager GUI. You can have multiple gauges active at the same time.

Gauge definition is accomplished through a wizard to aid in entering a valid set of gauge properties. Gauges are saved in the TotalStorage Productivity Center for Disk database and retrieved upon request. When you request data pertaining to a defined gauge, the TotalStorage Productivity Center for Disk builds a query to the database, retrieves and formats the data, and returns it to you. When started, a gauge is displayed in its own window, and it displays all available performance data for the specified initial date/time range. The date/time range can be changed after the initial gauge window is displayed.



Figure 18-4 Gauges

For performance-type gauges, if a metric selected for display is associated with a threshold enabled for checking, the current threshold properties are also displayed in the gauge window and are updated each time the gauge data is refreshed.

Database services for managing the collected performance data

The performance data collected from the supported devices is stored in a DB2 database. Database services are provided that enable you to manage the potential volumes of performance data.

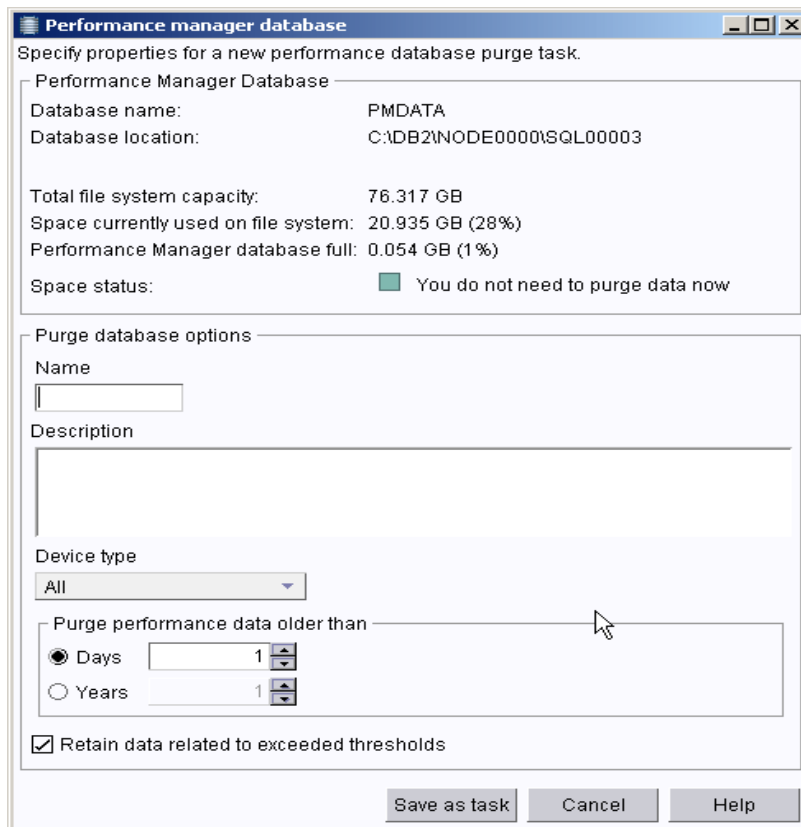


Figure 18-5 TotalStorage Productivity Center for Disk database purge task

– Database purge function

A database purge function deletes older performance data samples and, optionally, the associated exception data. Flexibility is built into the purge function, and it enables you to specify the data to purge, allowing important data to be maintained for trend purposes. You can specify to purge all of the sample data from all types of devices older than a specified number of days. You can specify to purge the data associated with a particular type of device. If threshold checking was enabled at the time of data collection, you can exclude data that exceeded at least one threshold value from being purged. You can specify the number of days that data is to remain in the database before being purged. Sample data and, optionally, exception data older than the specified number of days will be purged.

A reorg function is performed on the database tables after the sample data is deleted from the respective database tables.

– Database information function

Due to the amount of data collected by the TotalStorage Productivity Center for Disk, the database should be monitored to prevent it from running out of space. The

database information function returns the database percentage full. This function is invoked from either the Web user interface or the CLI.

Volume Performance Advisor

The advanced performance analysis provided by Multiple Device Manager is intended to address the challenge of allocating more storage in a storage system so that the users of the newly allocated storage achieve the best possible performance.

The Volume Performance Advisor is an automated tool that helps the storage administrator pick the best possible placement of a new LUN to be allocated (that is, the best placement from a performance perspective). It also uses the historical performance statistics collected from the supported devices to locate unused storage capacity on the SAN that exhibits the best (estimated) performance characteristics. Allocation optimization involves several variables that are user-controlled, such as required performance level and the time of day/week/month of prevalent access. This function is fully integrated with the Device Manager function so that, for example, when a new LUN is added to the ESS, the Device Manager can seamlessly select the best possible LUN.

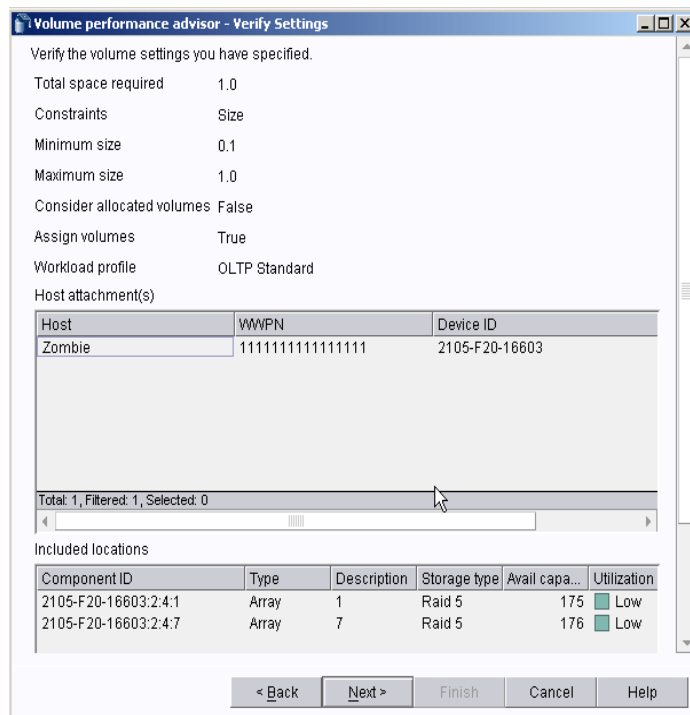


Figure 18-6 Volume Performance Advisor

Workload profiles

You can use the TotalStorage Productivity Center for Disk to select a predefined workload profile or to create a new workload profile that is based on historical performance data or on an existing workload profile. The TotalStorage Productivity Center for Disk uses these profiles to create a performance recommendation for volume allocation on a storage server.

The following five predefined Workload Profiles provided with TotalStorage Productivity Center for Disk:

- ▶ Standard OLTP
- ▶ High OLTP
- ▶ Data Warehouse

- ▶ Sequential Batch
- ▶ Document Archival

18.3.1 IBM TotalStorage Expert

The IBM TotalStorage Expert is an existing software tool that gives storage administrators powerful yet flexible storage asset, capacity, and performance-management capabilities to centrally manage ESSs located anywhere in the enterprise.

While the TotalStorage Productivity Center for Disk and the ESS Expert can coexist, the TotalStorage Productivity Center for Disk enhances several of the ESS Expert capabilities.

For more information on IBM TotalStorage Expert refer to Chapter 23, “TotalStorage Expert” on page 387.

18.4 More information

For more information on 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: Getting Started*, SG24-6490-00, which can be found at:

<http://publib-b.boulder.ibm.com/redpieces/abstracts/sg246490.html>

- ▶ *Exploring Storage Management Efficiencies and Provisioning - Understanding IBM TotalStorage Productivity Center and IBM TotalStorage Productivity Center with Advanced Provisioning*, SG24-6373-00, which can be found at:

<http://publib-b.boulder.ibm.com/abstracts/sg246373.html>



TotalStorage Productivity Center for Replication

In this chapter we describe the functionality as well as advantages of the TotalStorage Productivity Center for Replication (formerly known as Multiple Device Manager for Replication).

The function of the TotalStorage Productivity Center is designed to give customers the ability to monitor configure and manage storage devices and subsystems from a central console.

19.1 Overview

TotalStorage Productivity Center for Replication (TPC for Replication) is designed to help simplify copy services management for ESS. TotalStorage Productivity Center for Replication is also designed to provide configuration and management of the FlashCopy and Synchronous Metro Mirror capabilities of the ESS. TotalStorage Productivity Center for Replication is designed to control and monitor copy services operations in storage environments, it also provides advanced copy services functions for supported storage subsystems on the SAN.

TotalStorage Productivity Center for Replication was formerly known as Multiple Device Manager Replication Manager. MDM contained three separate subcomponents, namely:

- ▶ Device Manager
- ▶ Performance Manager
- ▶ Replication Manager

These subcomponents have been separated for specific functionality into TotalStorage Productivity Center for Replication as well as TotalStorage Productivity Center for Disk.

In this chapter we discuss TotalStorage Productivity Center for Replication which is made up of two subcomponents:

- ▶ **TotalStorage Productivity Center Common Base** - which is common with both TotalStorage Productivity Center for Replication as well as TotalStorage Productivity Center for Disk
- ▶ **TotalStorage Productivity Center for Replication** - Replication Manager Functionality

Replication Manager functionality is core to the TotalStorage Productivity Center for Replication, and we discuss Replication Manager functionality.

19.2 TotalStorage Productivity Center Common Base

The TotalStorage Productivity Center Common Base builds on the IBM Director technology. It uses Service Level Protocol (SLP) to discover supported storage systems on the SAN. Service Level Protocol enables the discovery and selection of generic services accessible through an IP network. The TotalStorage Productivity Center Common Base then uses managed objects to manage these devices. TotalStorage Productivity Center Common Base also provides a subset of configuration functions for the managed devices, primarily LUN allocation and assignment. A sample IBM Director console is shown in Figure 19-1.

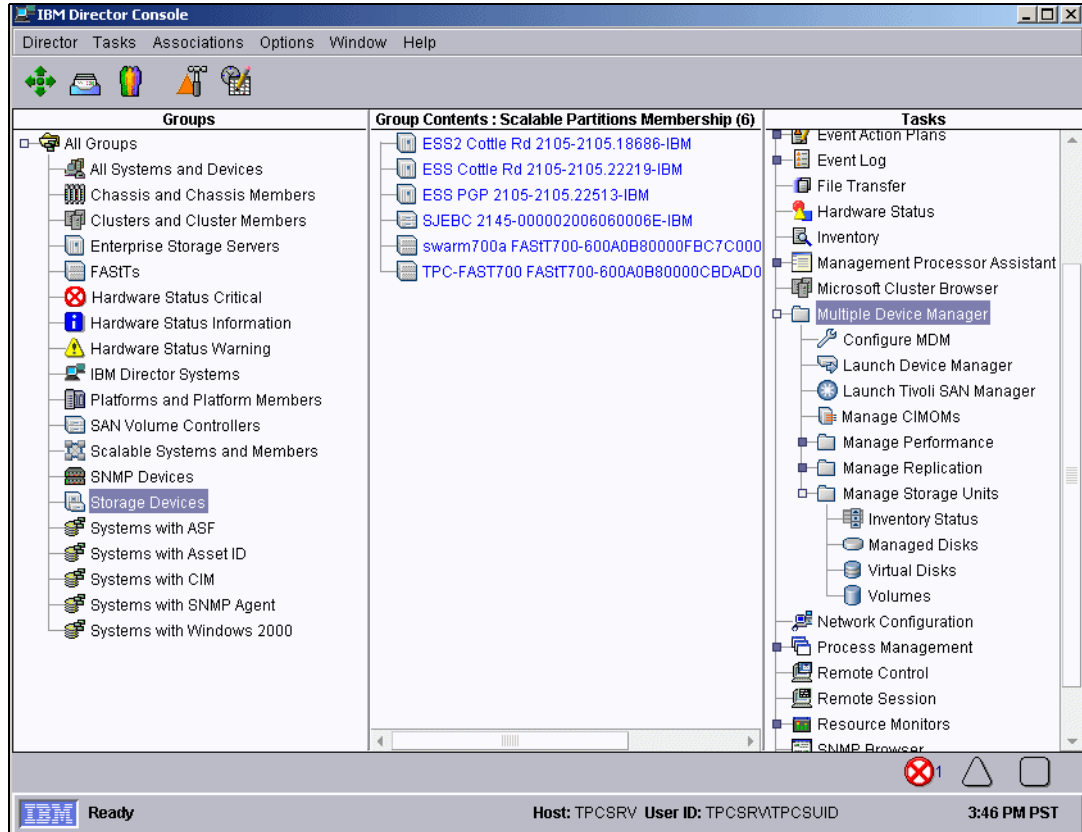


Figure 19-1 IBM Director Console

In summary, the TotalStorage Productivity Center Common Base is responsible for:

- ▶ Discovery of supported devices (SMI-S compliant)
- ▶ Data collection (asset, availability, configuration)
- ▶ Providing a topographical view of storage

The TotalStorage Productivity Center Common Base health monitoring keeps track of changes to the status of devices that have been “discovered” — meaning those within its knowledge.

19.3 Replication Manager functionality

IBM TotalStorage Productivity Center for Replication provides integrated administration, optimization, and replication features for interacting SAN devices. This provides an integrated view of the system so that administrators can easily perform complex configuration tasks and more productively manage the SAN infrastructure.

User-defined volume groups can be configured and stored in IBM TotalStorage Productivity Center for Replication without the need for separate scripts. Pairing of source and target volumes for either FlashCopy or Metro Mirror can be:

- ▶ Automatically paired by IBM TotalStorage Productivity Center for Replication
- ▶ Manually created by the administrator through GUI
- ▶ Created programmatically via a command line interface

IBM TotalStorage Productivity Center for Replication provides a central point for monitoring and managing FlashCopy and Metro Mirror operations for the IBM TotalStorage Enterprise Storage Server solution, through its SMI-S enabled API. Replication status is displayed in a GUI or returned through a command line interface request. Replication commands are issued to operate on the volume groups without the need for scripts. IBM TotalStorage Productivity Center for Replication monitors the copy services execution of the storage device.

IBM TotalStorage Productivity Center for Replication is designed to maintain data consistency across ESS Metro Mirror volumes during link outages and other failures. A Productivity Center for Replication copy session ensures that data on multiple, related, heterogeneous volumes is kept consistent by managing the volume pairs in the session as a consistent unit. TPC for Replication provides an interface for creating, maintaining, and using volume groups and for scheduling copy tasks. TPC for Replication populates the lists of volumes using the interface.

The TotalStorage Productivity Center for Replication helps perform the following tasks:

- ▶ Create replication groups
- ▶ Set up a group for replication
- ▶ Create, save and name a replication task
- ▶ Create a replication session using the following interface panels:
 - Create a session wizard
 - Select a source group
 - Select a copy type
 - Select a target pool
 - Save a session
- ▶ Manage a replication session

Note: Additional platform support

Support for Windows 2003 Enterprise Edition is added in IBM TotalStorage Productivity Center for Replication V2.1.

Important: TotalStorage Productivity Center for Replication replaces MDM for Replication, when you have installed TotalStorage Productivity Center for Disk or Replication common base (which is Device Manager), you will need to install TotalStorage Productivity Center for Replication. Figure 19-2 below gives an example of the how the new interface looks once the installation of the TotalStorage Productivity Center for Disk or Replication common base (known as Multiple Device Manager on the screen) is installed as well as TotalStorage Productivity Center for Replication as highlighted.

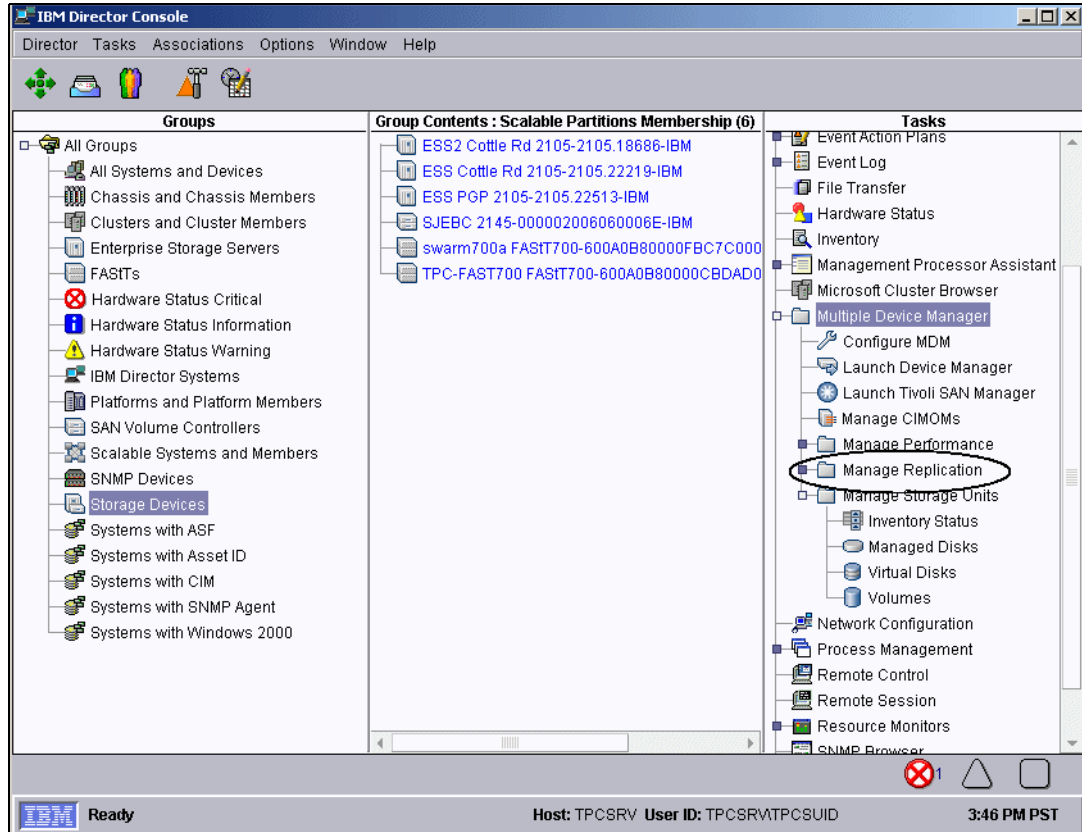


Figure 19-2 TotalStorage Productivity Center for Replication

19.3.1 Replication session

A replication session is a set of copy relationships which are maintained as a unit in a manner which provides consistency; especially across box or other hardware boundaries. The replication session, then, associates a pool with a group and gives them a particular copy relationship - either a continuous synchronous remote copy or a point-in-time copy

Replication Manager supports the session concept in which multiple pairs are handled as a consistent unit. You can create and manage copy relationships between source and target volume pairs or source volume groups, and among target pools through a Replication Manager copy session.

The Replication Manager Sessions panel shows sessions and their associated status. The status indicates if the volume is a source, target, or both; and it shows the copy mode of the volume. You can also use this panel to assess if current replication activities are proceeding normally or abnormally.

When you are creating a Replication session, you can select source and target volume pairs or volume groups, then establish a continuous synchronous remote copy (remote copy) or point-in-time copy (PIT copy) relationship between them. The Sessions panel includes the following options:

- ▶ **Create** invokes the Create Session wizard, which you can use to create copy relationships for a new session. You select copy types (remote copy or PIT copy), source volume groups, and target pools during the creation of the new session.
- ▶ **Delete** deletes an existing session.

- ▶ **Start** starts an existing session.
- ▶ **Increment** makes another PIT copy of the data for an existing session; the previous copy is lost.
- ▶ **Suspend (consistent)** freezes an existing session, which results in a consistent target copy if there are no errors.
- ▶ **Suspend (immediate)** stops an existing session with no guarantee of consistency.
- ▶ **Terminate** stops an existing session and withdraws the relationships.
- ▶ **Status** provides an at-a-glance status view of an existing session.
- ▶ **Properties** displays the Session Properties panel for an existing session.

Replication Manager session monitoring

Replication Manager session monitoring provides information about the status of replication activities.

You can view the following information about a Replication Manager session:

- ▶ Replication session properties, you can view the replication properties of a session.
- ▶ Storage group properties, you can view the storage group properties of a session.
- ▶ Storage pool properties, you can view the storage pool properties of a session.

19.3.2 Storage group

A storage group is a collection of storage units that jointly contain all the data for a specified set of storage units, such as volumes. The storage units in a group must be from storage devices of the same type. Groups can be created to identify sets of volumes that need to be managed as a consistent unit. A general purpose group can be used as a container for volumes that share some association, for example a group of volumes that are all associated with a specific application. A group can contain any set of compatible managed elements, for example managed disks (MDisks).

After a storage group is created, you can perform the following tasks:

- ▶ Add volumes or other managed elements
- ▶ Delete volumes or other managed elements
- ▶ Change the description of the group

A storage group can be managed by a Replication Manager session, so that copy tasks can be performed between storage groups.

19.3.3 Storage pools

A storage pool is an aggregation of storage resources on a storage area network (SAN) that you have set aside for a particular purpose. For example, you could use a storage pool for targets of copy operations that a collection of storage devices on the SAN can use.

The pool can be split into virtual disks (vDisks) that are visible to the storage devices that use them. The storage devices can be from different vendors but must be a type that TPC for Replication supports.

19.3.4 Relationship of group, pool, and session

This section illustrates the interdependency between a replication group, pool, and session in the context of the TPC for Replication.

To review, the definitions are:

- ▶ **Group:** A set of volumes containing related data, which are managed as a unit – they are managed concurrently.
- ▶ **Pool:** Volumes set aside for copy services targets – these must not be in use by any other application.
- ▶ **Session:** A set of copy relationships which are maintained as a unit to provide consistency, across storage and server hardware boundaries.

TPC for Replication provides the ability to copy a group to a pool, in which case it creates the valid mappings for source and target volumes and optionally presents them to the user for verification that the mapping is acceptable. Sessions are a set of multiple pairs that are managed as a consistent unit from which freeze and run functions can be performed when errors occur. The session can also be viewed as a consistency group.

The following chart (Figure 19-3) graphically depicts the interactions of groups, pools and a session. It shows one group of related volumes on a source ESS (volumes A1 and A2) that we want to mirror to another target pool of volumes (the B volumes). Once we've identified and created the source volumes in the group and the target volumes in a pool, we can then establish the mirroring relationship.

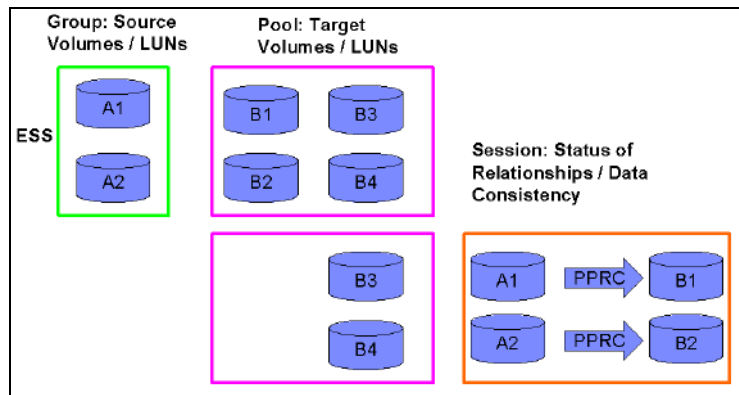


Figure 19-3 Relationships of groups, pools and sessions

Our example session shows that A1 is associated with B1, similarly A2 with B2. The B1 and B2 volumes are now persistently bound to the relationship – whereas B3 and B4 are still available for use.

Replication Manager can automatically create the source to target relationship on your behalf. Once created, these volumes are now part of a session or consistency group. This means again that any error on any of the volumes in this session could trigger a suspend across all the volumes to ensure data consistency. Events such as loss of access to a source subsystem or the loss of the PPRC links could be examples of such conditions to trigger a freeze event.

19.3.5 Copyset and sequence concepts

A copyset incorporates all the volumes that make up an instance of a given copy type. In other words, it comprises the; source volume, target volume, and the copy relationships. With TPC for Replication you can manually select target volumes, or have the Replication Manager select them for you.

Subsequent releases of Replication Manager will enhance the number of volumes in a copy set, and a session will be able to manage one to thousands of copy sets.

A sequence includes the set of all copy relationships at any given stage of a copy operation. For PPRC and FlashCopy there is only one sequence. A sequence will share the same pool criteria policy.

Using Figure 19-4, in a PPRC relationship, A1, A2, and A3 are each members of a copyset. You can visualize this as three copysets, along with their target volumes. When you create the PPRC session, MDM automatically maps the disk in each copyset to appropriate available disks, or you can choose the targets manually if you want to. The target disk must be larger than the source disk on ESS. When you perform the PPRC, the sequence dictates that each copyset will kick off more or less concurrently.

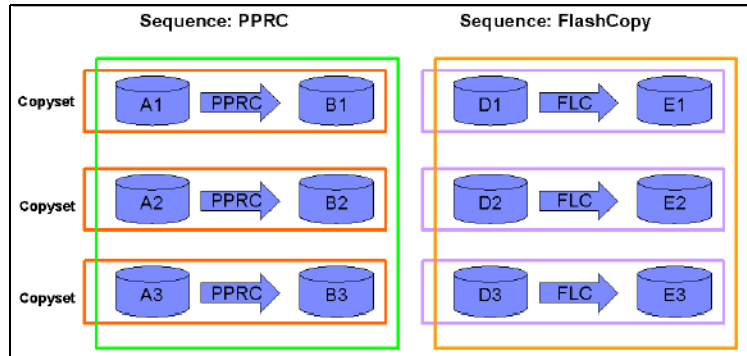


Figure 19-4 RM PPRC sequence relationship example

Note: The same process as described above also applies to FlashCopy.

19.4 More Information

For more information on TotalStorage Productivity Center for Replication, refer to:

<http://www.redbooks.ibm.com/redbooks/pdfs/sg246373.pdf>

<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.

20.1 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.

20.1.1 IBM Tivoli Storage Manager family of products

The IBM Tivoli Storage Manager is available as either a basic or extended version:

- ▶ The 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
- ▶ The 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

20.1.2 Licensing

Licensing in IBM Tivoli Storage Manager Version 5.3.0 server has changed from previous versions. Before upgrading from a previous version, delete or rename the nodelock file (for all operating systems except z/OS) or remove existing LICENSE options from the server's option file (z/OS). After installing the Version 5.3.0 server, you must register new licenses.

The REGISTER LICENSE command has changed with this version, and its function is currently limited:

- ▶ You can register licenses for server components. This includes Tivoli Storage Manager (base), Tivoli Storage Manager Extended Edition, and Tivoli Data Retention Protection.

- ▶ You cannot register licenses for components that are licensed on the basis of processors (for example, Tivoli Storage Manager for Mail, Tivoli Storage Manager for Databases, Tivoli Storage Manager for ERP, Tivoli Storage Manager for Hardware, Tivoli Storage Manager for Space Management).

Note: Your license agreement determines what you are licensed to use, even if you cannot use the REGISTER LICENSE command to register all components. You are expected to comply with the license agreement and use only what you have purchased. Use of the REGISTER LICENSE command implies that you agree to and accept the license terms specified in your license agreement.

20.1.3 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.

Note: In the latest version of Tivoli Storage Manager which is version 5.3, the retention periods for archive and backup sets have been increased to 30000 days which is 82 years instead of the 9999 days in the previous versions.

- ▶ **Data protection for business-critical applications:** The 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.

The 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 Hardware
- 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 R/3, mySAP.com, 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), and DB2 Universal Database (UDB) from IBM integrates 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 and EMC *intelligent disk subsystem* advanced copy functions such as FlashCopy and TimeFinder. 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, Microsoft SQL Server, and Informix®. 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 SAP R/3 Servers. It provides automated data protection, reduces the CPU performance impact of data backups and restores on the R/3 server, and greatly reduces the administrator workload necessary to meet data protection requirements.

IBM Tivoli Storage Manager for Hardware - improves the data protection of business-critical databases and ERP applications that require 24x365 availability. 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 Enterprise Storage Server (ESS) with Storage Manager and its database protection capabilities.

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 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

- ▶ **Disaster Recovery:** *IBM Tivoli Storage Manager Disaster Recovery Manager* - Part of Tivoli Storage Manager Extended edition, Tivoli Storage Manager's Disaster Recovery Manager is designed to get your critical business systems up and minimize recovery time from any type of disaster that might affect your on-site storage. It creates a disaster recovery plan and facilitates the tracking of off-site volumes. The recovery plan is integrated with and maintained on the IBM Tivoli Storage Manager server. It contains detailed recovery steps and automated computer scripts. It also tracks where your off-site copies of volumes are.
- ▶ **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.
- ▶ **Hierarchical Space Management:** HSM and SAN software extensions to Tivoli Storage Manager are also available.

IBM Tivoli Storage Manager for Space Management — Helps optimize usage of online storage and can automatically migrate infrequently or unused data files to less expensive offline storage and still make those files appear to be online for easy retrieval.

This process 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.

Core Functions:

- 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.
- ▶ **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.

20.1.4 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 TSM 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 20-1 shows many of the client and server platforms supported by Tivoli Storage Manager.

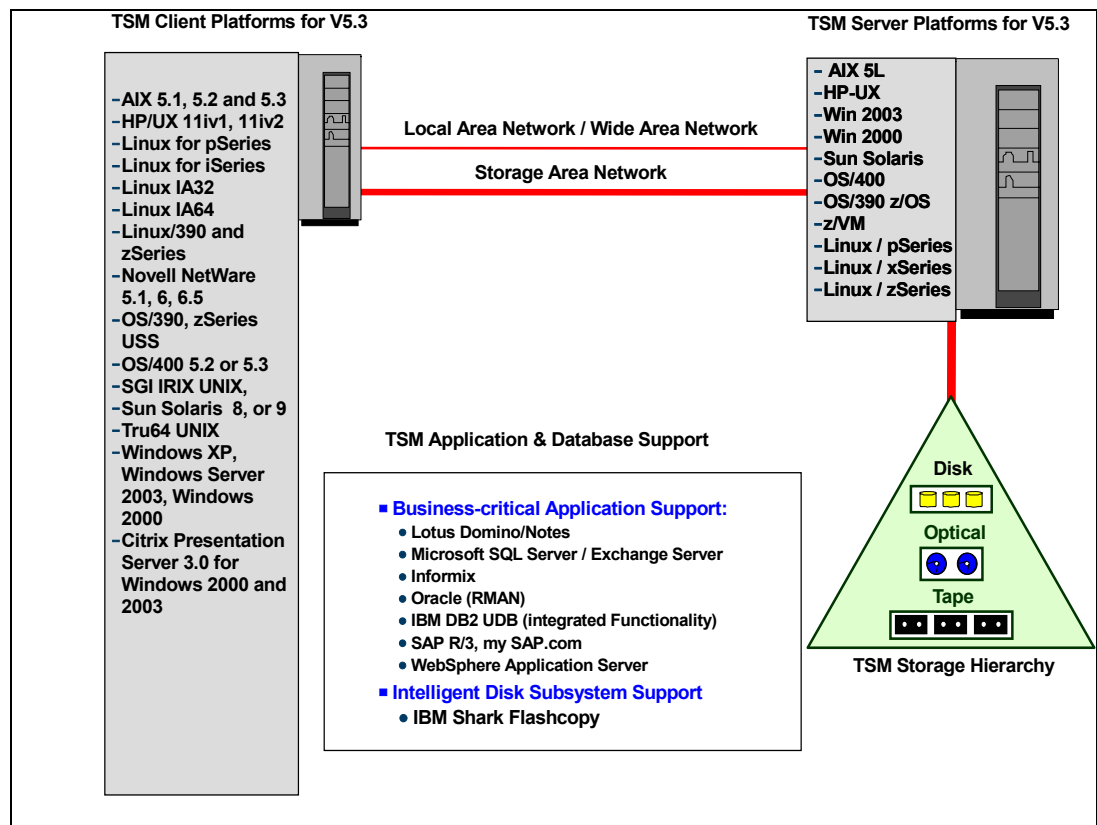


Figure 20-1 Tivoli Storage Manager client, server, application, and database support

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.

For newly defined storage pools, the default storage pool collocation setting is now **GROUP**.

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

Note: Operational reporting is included as part of the Tivoli Storage Manager for Windows server and is also available as a stand-alone package on a Windows server.

Powered by the relational database inside Tivoli Storage Manager, the granular policy engine can automate capabilities down to the individual file level. 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

20.2 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 20-2 on page 324 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.

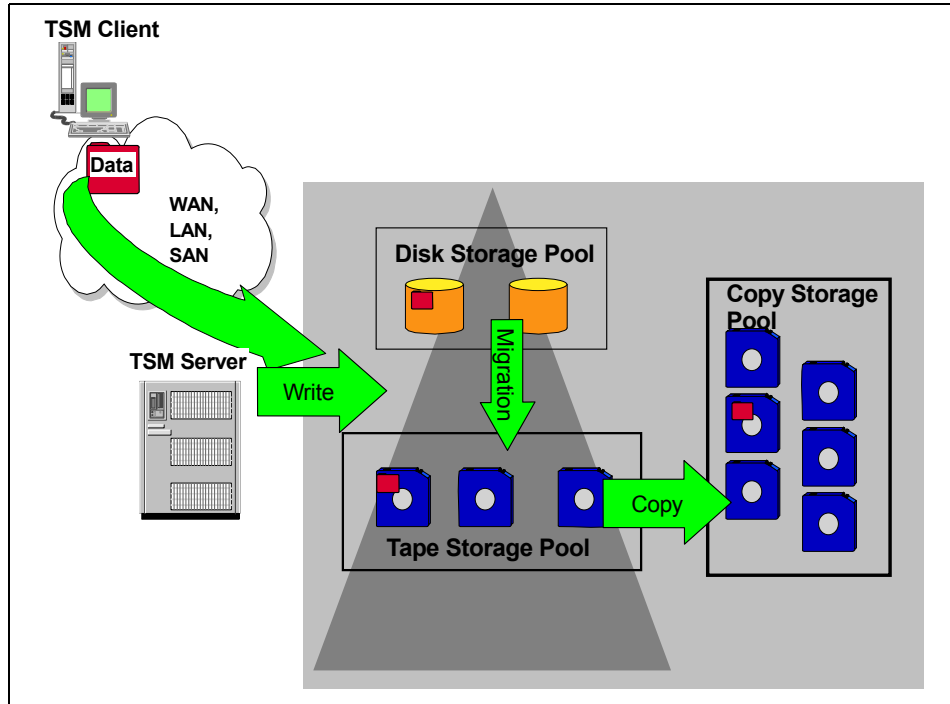


Figure 20-2 Data movement with Tivoli Storage Manager and the storage hierarchy

20.2.1 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.

20.2.2 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.

Note: The Tivoli Storage Manager Server and the Administrations Center can be installed on the same machine. The Administration Center requires at a minimum 512 MB RAM in addition to the RAM required for the Tivoli Storage Manager Server.

For the latest recommendation on the Administration Center installation, use keyword TSMADMINCENTER when you visit:

<http://www.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageManager.html>

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 20-3 is an example of how the Welcome page looks after installing ISC and the Administration Center.

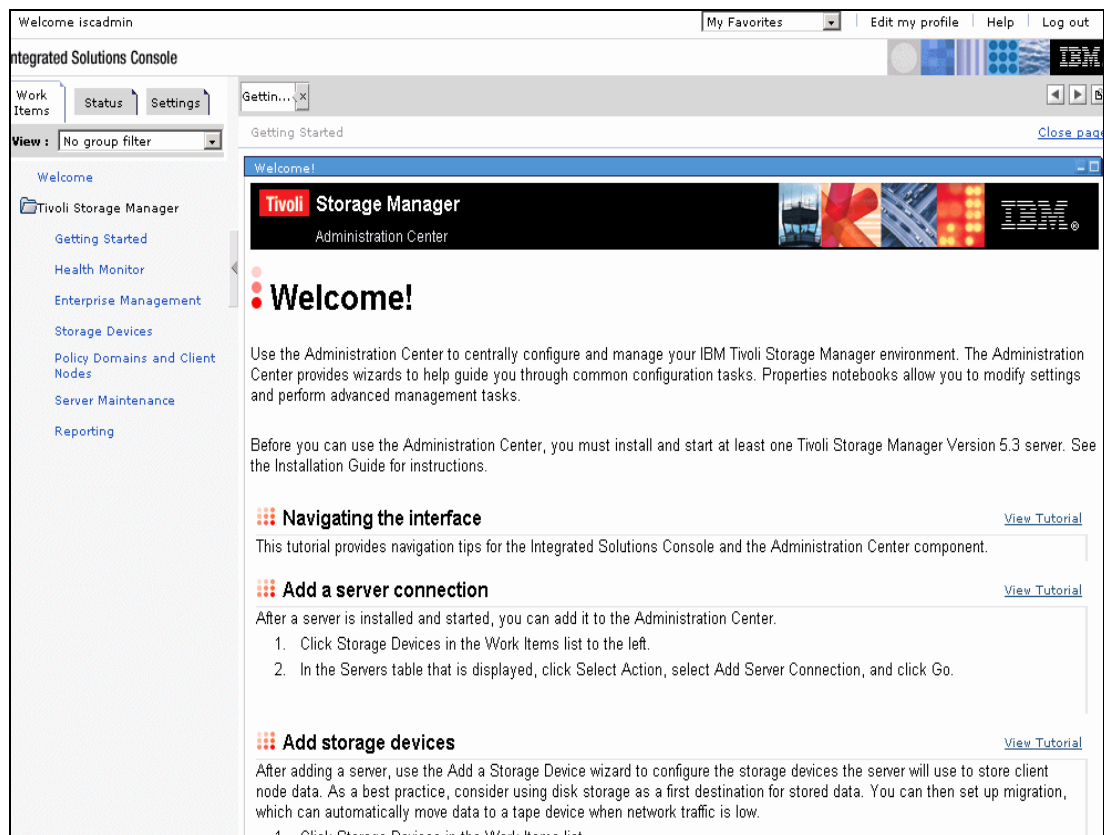


Figure 20-3 Integrated Solutions Console Administration Center

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 wizards

Listed below are the wizards available within the Administration Center these are divided into four area's namely *Enterprise Management*, *Storage Devices*, *Policy Domains* and *Client nodes* and lastly *Server Maintenance*.

Enterprise Management

Figure 20-4 shows the enterprise console.

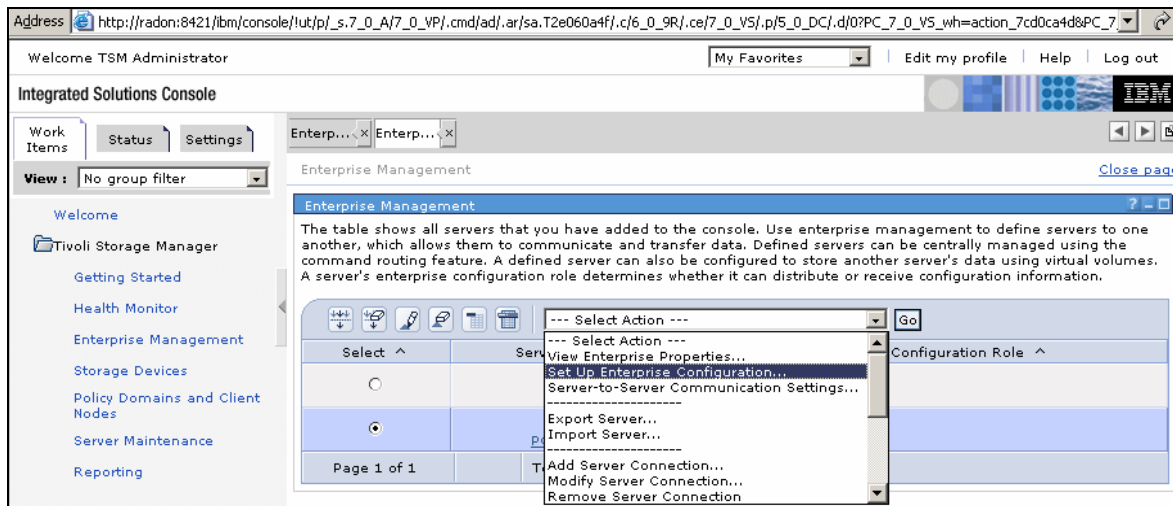


Figure 20-4 Enterprise management

Set up enterprise configuration

Used to set up a configuration manager and managed servers. You must set up the configuration manager server before setting up its managed servers.

- ▶ Define a server
 - Used to enable server-to-server communications.

- ▶ Create a server group
 - Used to create a group of defined servers, which can be used to facilitate command routing.
- ▶ Create a profile
 - Used to create new profiles for configuration manager servers.
- ▶ Create a subscription
 - Used to create new profile subscriptions for managed servers.
- ▶ Import or Export a server
 - Used to import or export server definitions.
- ▶ Create administrative schedule
 - Used to setup a schedule to perform administrative tasks.
- ▶ Import or Export Administrator
 - Used to import or export administrator definitions.

Storage devices

Figure 20-5 shows the Storage devices console.

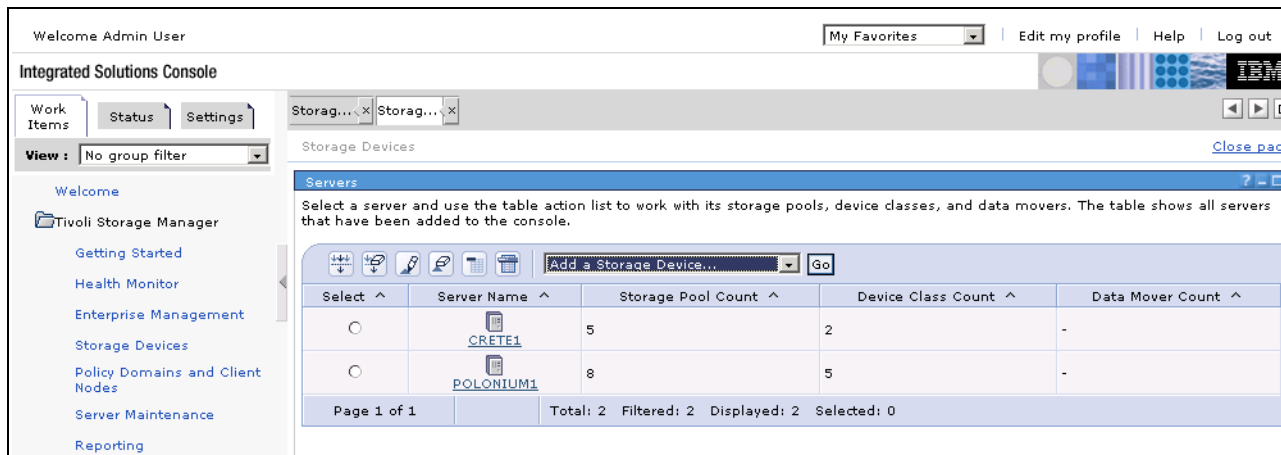


Figure 20-5 Storage devices

- ▶ Add storage device
 - Used to create a library, drives, paths, a device class, and storage pools for a storage device.
- ▶ Create a library
 - Provides a fast way to define a library and its drives.
- ▶ Create a device class
 - Used to create a device class, which represents a set of storage devices with similar characteristics.
- ▶ Create a storage pool
 - Used to set up a primary or copy storage pool for use.

- ▶ Protect an NAS file server
 - Used to create a complete configuration so that the server can back up a network-attached storage (NAS) file server using network data management protocol (NDMP).
- ▶ Create a collocation group
 - Used to create a collocation group to store data for selected client nodes on as few volumes as possible.
- ▶ Add volumes
 - Used to check volumes into the library inventory, and label them if necessary. Also used to label volumes without checkin them in.
- ▶ Restore volumes
 - Used to restore all files on damaged volumes in a primary storage pool that was backed up to a copy storage pool.
- ▶ Delete volume history
 - Used to delete volume history file records that are no longer needed (for example, records for obsolete database backup volumes).

Policy domains and client nodes

Figure 20-6 shows the policy domain and client nodes console.

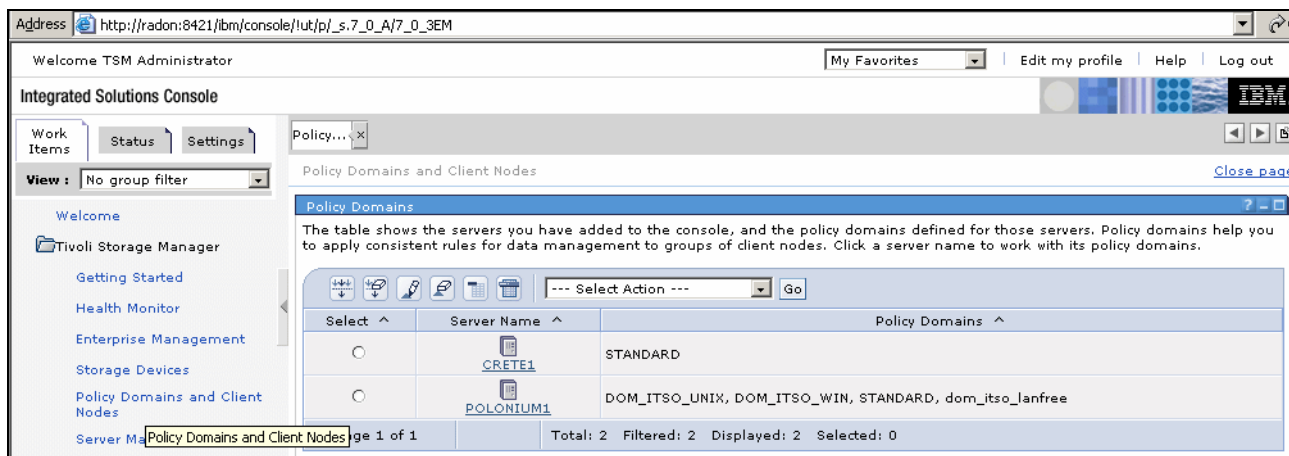


Figure 20-6 Policy Domains and client nodes

- ▶ Create a client node
 - Used to create a client node.
- ▶ Create client node schedule
 - Used to set up a schedule for automating client node operations.
- ▶ Create a management class
 - Used to create a new management class.
- ▶ Create an option set
 - Used to create an option set, which allows you to centrally manage many client node processing options.
- ▶ Create a policy domain
 - Used to create a policy domain with a default management class.

- ▶ Enable LAN-free data movement
 - Used to set up a storage agent so that it can move data on behalf of client nodes over a SAN directly to storage devices.
- ▶ Generate a backup set
 - Used to create a point in time copy of a client node's data. The copy is created directly from server storage, so the data does not have to be transferred over the network to create the backup set.
- ▶ Import or export client node
 - Used to import or export client node definitions.
- ▶ Import or export policy domain
 - Used to import or export policy domain definitions.
- ▶ Restore by file name(NAS)
 - Used to restore files for an NAS file server by file name when a table of contents was created at the time of the backup, or when you know the names of individual files to restore.

Server maintenance

Figure 20-7 shows the server maintenance console.

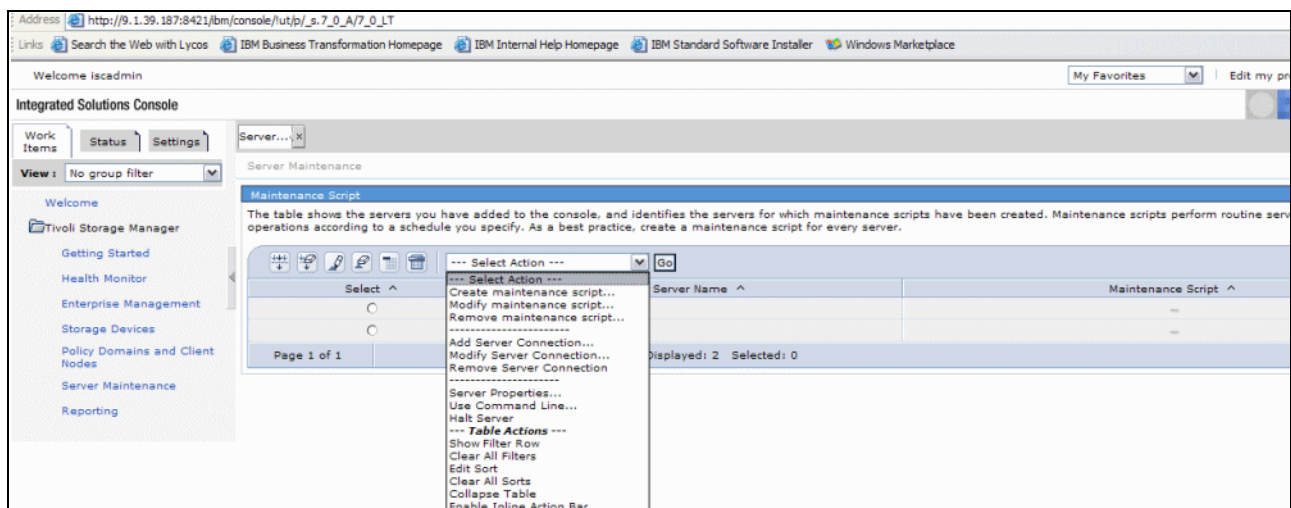


Figure 20-7 Server maintenance

- ▶ Create maintenance script
 - Used to create a maintenance script to automatically run the essential processes that protect the server database and storage pools. The script helps keep the server running well.

Tip: For more information and an in depth review of the Tivoli Storage Manager Administration Center, see the redbook *Tivoli Storage Manager V5.3*, SG24-6638.

20.2.3 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 20-8 shows the Windows Tivoli Storage Manager Client GUI.

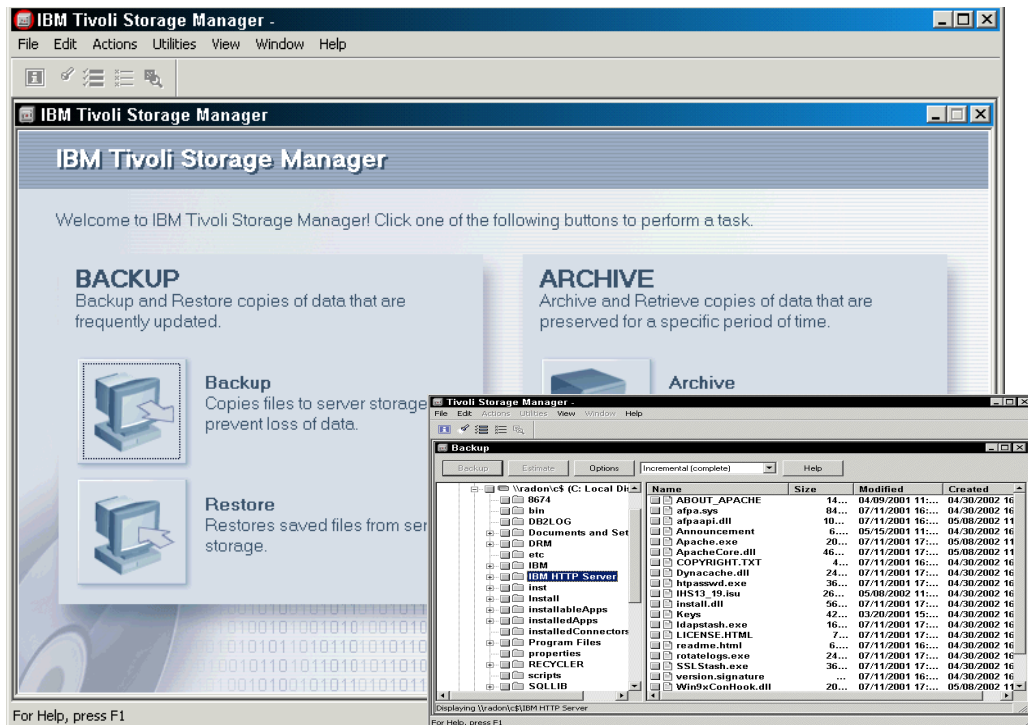


Figure 20-8 The Tivoli Storage Manager Version 5 client interface

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.

20.2.4 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, R/3 and Lotus Domino.

20.2.5 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.

20.2.6 Tivoli Storage Manager API

The Tivoli Storage Manager APIs are used for Tivoli's own Disaster Protection products but 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 TotalStorage DR550 is a storage solution. For an overview of the DR550, refer to Chapter 24, "IBM TotalStorage DR550" on page 401 or alternatively go to the link below:

<http://www.redbooks.ibm.com/redbooks/pdfs/sg247091.pdf>

Figure 20-9 on page 332 is an example of how the API fits into the DR550 solution.

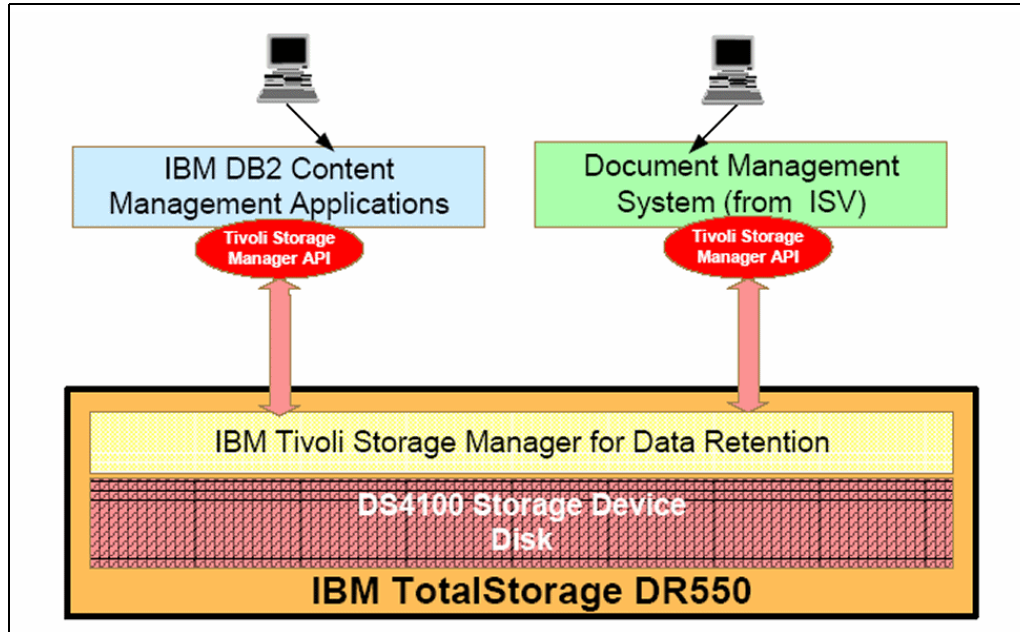


Figure 20-9 How the API fits into the DR550 solution

20.3 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. Table 20-1 on page 332 describes the various client backup and restore operations supported with Tivoli Storage Manager and provides a description, usage, and restore options. The standard backup method that Tivoli Storage Manager uses is called progressive incremental backup (see 20.3.7, “Progressive incremental backups” on page 340). It is a unique and efficient method for backup.

Table 20-1 Summary of client backup and restore operations

Type	Description	Usage	Restore options
Progressive Incremental backup	The standard method of backup used by the Tivoli Storage Manager backup client. Only files that are new or have changed since the last backup are backed up. There is no dependency on a full backup, as each file is tracked individually.	Helps ensure complete, effective, policy-based backup of data. Eliminates the need to retransmit backup data that has not been changed during successive backup operations.	The user can restore just the version of the file that is needed (depending on the retention parameters). Tivoli Storage Manager does not need to restore a base file followed by incremental backups. This means reduced time and fewer tape mounts, as well as less data transmitted over the network.

Type	Description	Usage	Restore options
Selective backup	Backup of files that are selected by the user, regardless of whether the files have changed since the last backup.	Allows users to protect a subset of their data independent of the normal incremental backup process.	The user can restore just the version of the file that is needed. Tivoli Storage Manager does not need to restore a base file followed by incremental one. This means reduced time, fewer tape mounts, and less data over the network.
Adaptive subfile backup	A method that backs up only the parts of a file that have changed since the last backup. The server stores the base file and subsequent subfiles (the changed parts) that depend on the base file. The process works with either the standard progressive incremental backup or with selective backup.	Maintains backups of data while minimizing connect time and data transmission for the backup of mobile and remote users. Applicable to clients on Windows systems.	The base file plus a maximum of one subfile is restored to the client.
Journal-based backup	Aids all types of backups (progressive incremental backup, selective backup, adaptive subfile backup) by basing the backups on a list of changed files. The list is maintained on the client by the journal engine service of the Tivoli Storage Manager backup/archive client.	Reduces the amount of time required for backup. The files eligible for backup are known before the backup operation begins. Applicable to clients on Windows NT and Windows 2000 systems.	Journal-based backup has no effect on how files are restored; this depends on the type of backup performed.
Image backup	Full volume backup. Nondisruptive, online backup is possible for Windows 2000 and Linux clients by using the Tivoli Storage Manager snapshot function.	Allows backup of an entire file system or raw volume as a single object. Can be selected by backup- archive clients on UNIX and Windows systems. Used by Windows clients that are using server-free data movement.	The entire image is restored.
Image backup with differential backups	Full volume backup, which can be followed by subsequent differential backups.	Used only for the image backups of NAS file servers, performed by using Tivoli Data Protection for NDMP.	The full image backup plus a maximum of one differential backup are restored.

Type	Description	Usage	Restore options
Backup using hardware snapshot capabilities	A method of backup that exploits the capabilities of IBM Enterprise Storage Server FlashCopy and EMC TimeFinder to make copies of volumes used by database servers. The Tivoli Data Protection for applications and databases use the volume copies to back up the database volumes.	Implements high-efficiency backup and recovery of business-critical applications while virtually eliminating backup-related downtime or user disruption on the database server.	See 20.3.4, "Split-mirror/point-in-time copy backup using SAN" on page 337 for details.
Group backup	<p>A method that backs up files that you specify as a named group. The files can be from one or more file spaces. The backup can be a full or a differential backup.</p> <p>Applicable to clients on Unix and Linux systems.</p>	<p>Creates a consistent point in time backup of a group of related files. The files can reside in different files spaces on the client. All objects in the group are assigned to the same management class. The server manages the group as a single logical entity, and stores the files in a virtual file space in server storage.</p> <p>A group can be included in a backup set.</p>	The user can select to restore the entire group or just selected members of the group. The user can restore just the version off the file that is needed.
Archive	The process creates a copy of files and stores them for a specific time.	Use for maintaining copies of vital records for legal or historical purposes. If you frequently create archives for the same data, consider using instant archive (backup sets) instead. Frequent archive operations can create a large amount of metadata in the server database resulting in increased database growth and decreased performance of expiration server operations.	The selected version of the file is retrieved on request.

Type	Description	Usage	Restore options
Instant archive	The process creates a backup set of the most recent versions of the files for the client, using files already in server storage from earlier backup operations.	Use when portability of the recovery media or rapid recovery of a backup- archive client is important. Also use for efficient archiving.	The files are restored directly from the backup set. The backup set resides on media that can be mounted on the client system, for example, CD, tape drive, file system. The Tivoli Storage Manager server does not have to be contacted for the restore process, so the network and Tivoli Storage Manager server are not used.

20.3.1 Traditional LAN and WAN backup topology

In a traditional LAN and WAN environment the Tivoli Storage Manager backup and archive client or application would read data from locally attached disks and send it over the LAN to the Tivoli Storage Manager backup server as shown in Figure 20-10. The server receives the data 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.

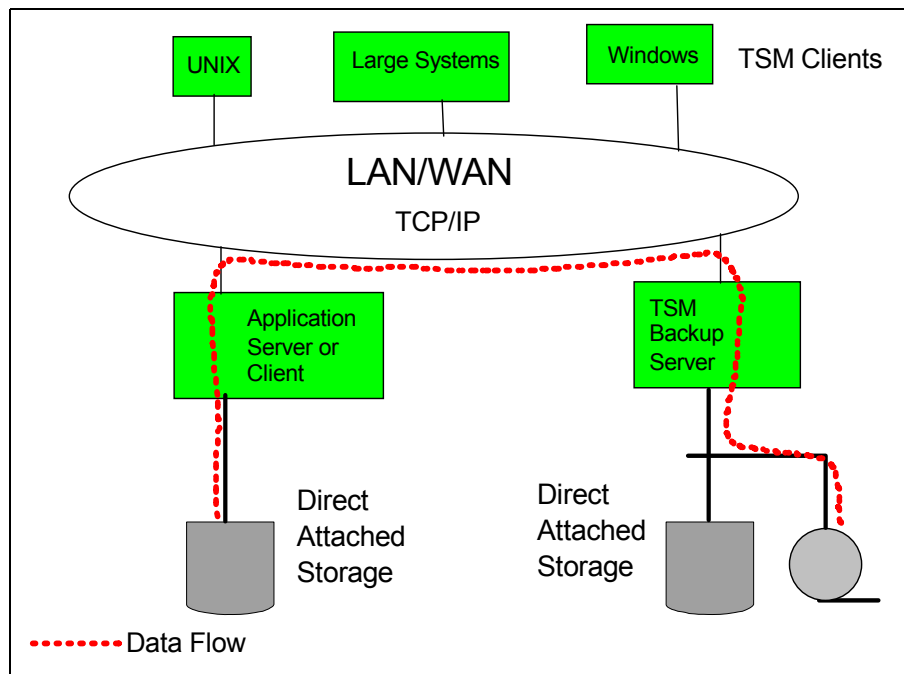


Figure 20-10 Tivoli Storage Manager LAN and WAN backup

20.3.2 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.

Note: In IBM Tivoli Storage Manager V5.3 enhancements have been made to allow you to quickly determine if your LAN-free environment has been configured correctly. You can request validation for one client node and one storage agent. When validation is requested, a detailed report is generated explaining why the storage pool is or is not LAN-free capable. As a result, you can determine if there is a setting or configuration issue on the server preventing LAN-free data movement. The new command is: **validate lanfree**

A LAN-free backup environment is shown in Figure 20-11.

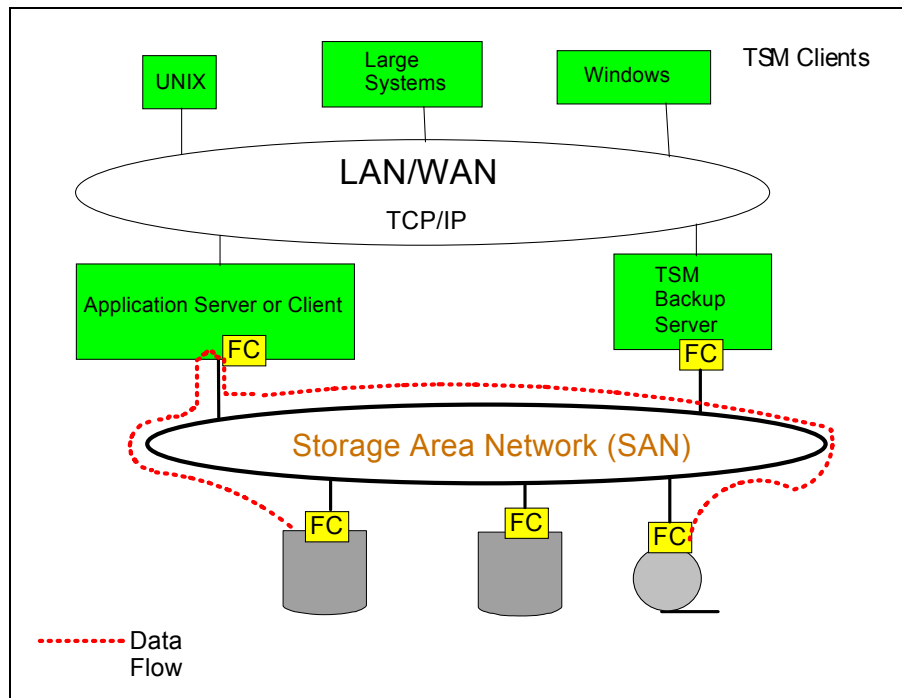


Figure 20-11 Tivoli Storage Manager LAN-free backup

20.3.3 Server-free backup

Server-free backup/restore capability is introduced into Tivoli Storage Manager (Tivoli Storage Manager) in Version 5. In a server-free backup environment, data is copied directly from the SAN attached Tivoli Storage Manager client disk to the SAN attached tape drive via the SAN Data Gateway data mover as shown in Figure 20-12. The Storage Agent used in LAN-free backups is not used. The data movement is actually done by a SAN Data Gateway (SDG) or similar device on the SAN. Therefore both Tivoli Storage Manager client and server machines do not have to read and write the data at all. The Tivoli Storage Manager server sends commands to the SDG device to tell it which blocks to move from which SAN attached disk to which SAN attached tape device. The data is actually copied rather than moved from one location to another. This provides a way to back up and restore large volumes of data

between client-owned disks and storage devices using a method that considerably reduces overhead on the Tivoli Storage Manager server and the client. Only volume images and not individual files can be moved by server-free data movement. The data is transferred block by block rather than by doing file I/O. Both raw and NTFS volumes can be backed up using server-free data movement.

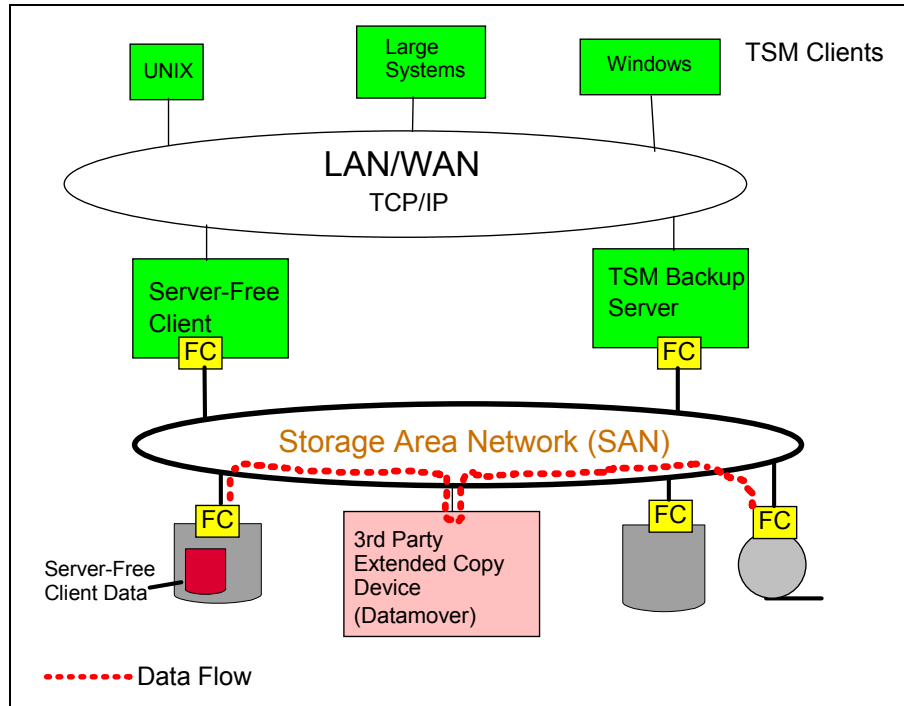


Figure 20-12 Tivoli Storage Manager server-free backup

Data that has been backed up using server-free data movement can be restored over a server-free path, over a LAN-free path, or over the LAN itself. The impact on application servers is now minimized with Server-Free Data Movement. It reduces both Tivoli Storage Manager client and server CPU utilization. The use of a SCSI-3 extended copy command causes data to be transferred directly between devices over the SAN or SCSI bus. The data mover device must support the SCSI-3 EXTENDED COPY command, which conforms to the ANSI T10 SPC-2 standard. The data mover device can be anywhere in the SAN, but it has to be able to address the LUNs for both the disk and tape devices it is moving data between.

20.3.4 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 20-13. 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 Hardware component by integrating IBM's Enterprise Storage Server FlashCopy function with Tivoli Storage Manager and its database protection capabilities for Oracle, R/3, 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.

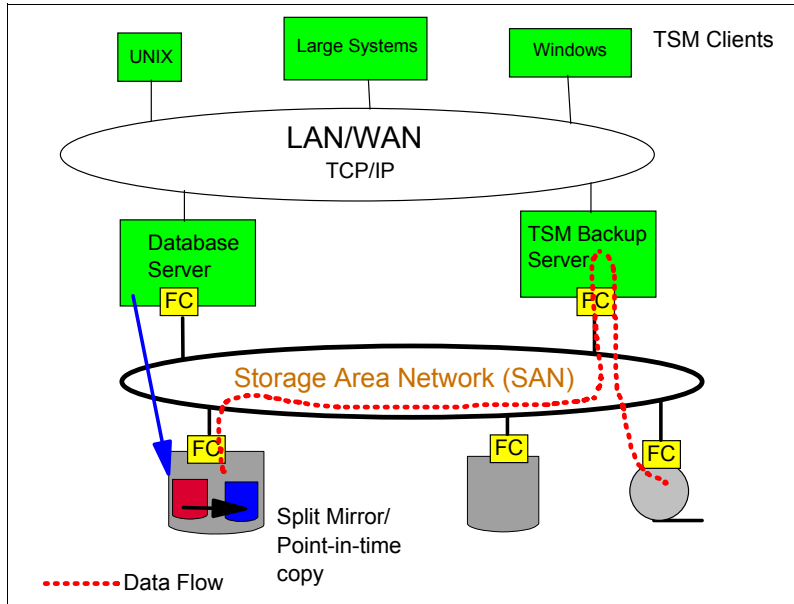


Figure 20-13 Tivoli Storage Manager split-mirror/point-in-time copy backup

20.3.5 NAS backup and restore

The IBM Network Attached Storage products have been pre-installed with the Tivoli Storage Manager client which enables an existing or planned Tivoli Storage Manager server to back up the data in the NAS system. This backup client is designed to provide file level and sub-file level backup and restore functionality. A Tivoli Storage Manager environment integrated with NAS systems can be used to manage a broad range of data storage, recovery, and availability functions across the entire computing infrastructure. Based on the Tivoli Storage Manager server's configuration, the final destination of the NAS appliance's backup may either be located in the Tivoli Storage Manager server's disk storage or an attached tape subsystem. Just like a standard Tivoli Storage Manager client an *option* file is created and stored on the NAS system. For more information on backing up IBM NAS appliances, see *IBM TotalStorage NAS Backup and Recovery Solutions*, SG24-6831.

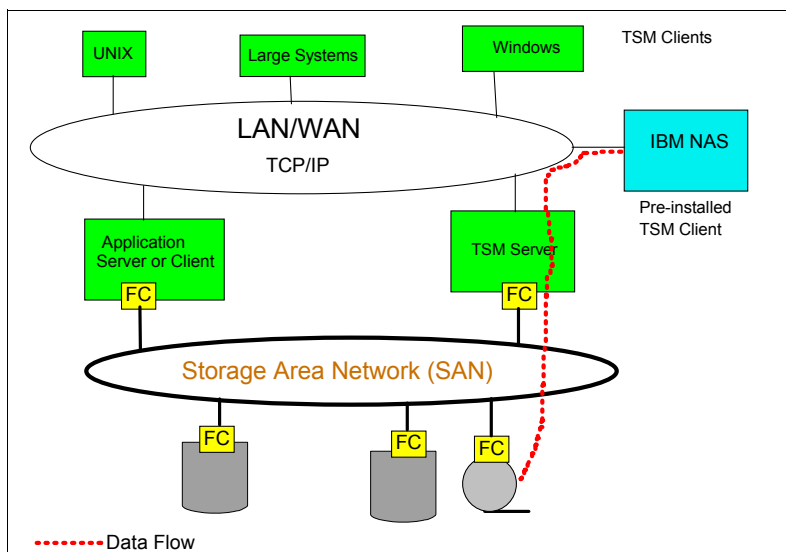


Figure 20-14 IBM TotalStorage NAS backup

NAS appliances supporting NDMP can be backed up with Tivoli Storage Manager Extended Edition. It provides backup and recovery support on Tivoli Storage Manager servers for NAS file servers from various vendors. NAS file servers often require a unique approach to providing backup and recovery services, because these file servers are not typically intended to run third-party software. The NAS file server does not require installation of Tivoli Storage Manager 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, as shown in Figure 20-15. Network Data management Protocol (NDMP) is used to communicate with and provide 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 higher-performance backup to tape devices without moving the data over the LAN.

Note: NDMP operations for backup of NAS file servers have been enhanced to support the following functions:

- ▶ Directory-level backup of NAS data, which enables the division of a file system backup operation among several NDMP backup operations as well as several tape drives. This enhancement will reduce backup and restore times
- ▶ NDMP Directory Level Backup will enable Tivoli Storage Manager to backup user created snapshots that are stored as sub-directories, specifically Network Appliance snapshots.

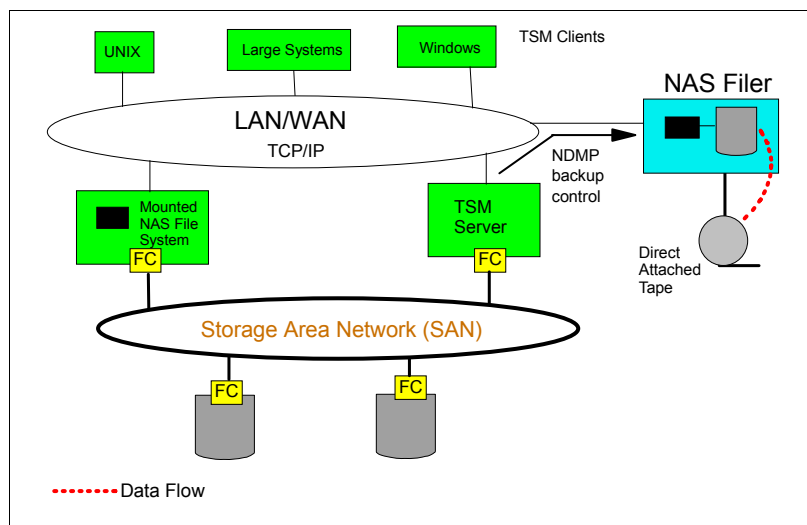


Figure 20-15 Tivoli Storage Manager and NDMP backup

20.3.6 Image backup

An image backup is a block-by-block copy of data from the Tivoli Storage Manager client to the Tivoli Storage Manager backup server. One important function of an image restore is to accelerate recovery in a Disaster Recovery scenario. 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 administration documents found at:

http://www.tivoli.com/support/public/Prodman/public_manuals/td/TD_PROD_LIST.html

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 will be applied to the file system image as a whole. An image backup provides the following benefits.

- ▶ Image backup 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 20-16 illustrates the process for image backup with Tivoli Storage Manager.

On the Windows 2000 client platform a Logical Volume Storage Agent (LVSA) has been introduced which is capable of taking a 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 is able to send a consistent image of the volume as it was at the start of the snapshot process to the Tivoli Storage Manager server.

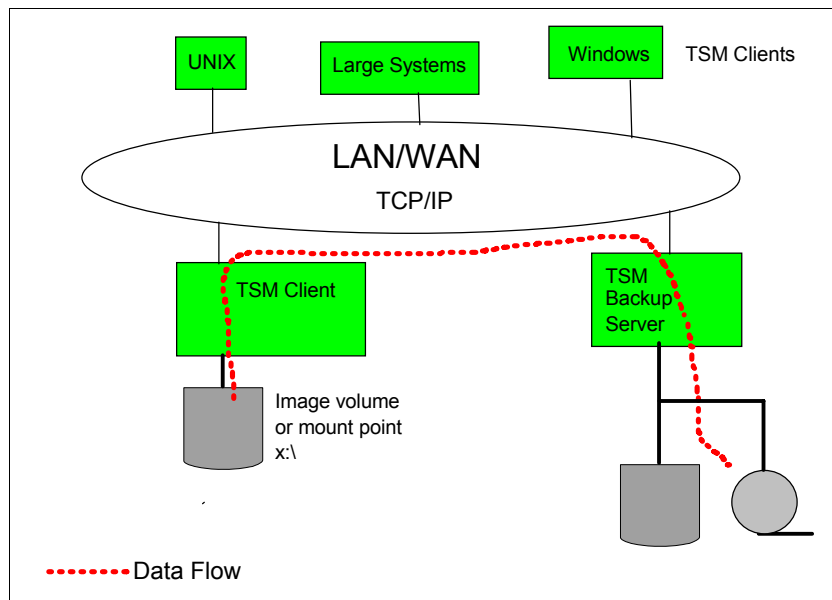


Figure 20-16 Tivoli Storage Manager image backup

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.

20.3.7 Progressive incremental backups

One of the many advantages of Tivoli Storage Manager 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.

20.4 IBM Tivoli Storage Manager for Data Retention

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 Tivoli Storage Manager for Data Retention V5.2 implements data retention policies that help meet regulatory requirements.

20.4.1 Overview

Tivoli Storage Manager for Data Retention provides enhancements to Tivoli Storage Manager Extended Edition to help meet additional requirements defined by the regulatory agencies for retention and disposition 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 Tivoli Storage Manager for Data Retention extends IBM Tivoli Storage Manager Extended Edition functionality to help meet regulatory retention requirements.

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.

Your content management and archive applications can utilize the IBM Tivoli Storage Manager API to apply business policy management for ultimate deletion of archived data at the appropriate time. The product is available for the IBM Tivoli Storage Manager Extended Edition of all current server platforms.

Sold as a separately licensed software product integrated into Tivoli Storage Manager Extended Edition server software, Tivoli Storage Manager for Data Retention requires a stand-alone Tivoli Storage Manager-Extended Edition server be dedicated for its use.

Attention: For more information on data retention, refer to Chapter 24, "IBM TotalStorage DR550" on page 401.

Device support enhancements

IBM Tivoli Storage Manager device support has enhanced support for:

- ▶ WORM Tape (IBM Tivoli Storage Manager 5.3 now supports the 3592 WORM tape device)
- ▶ DVD

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.

Tivoli Storage Manager for Data Retention supports selected WORM Tape devices and will extend that support, as more devices are available.

Tivoli Storage Manager for Data Retention also includes support for 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.

20.5 IBM Tivoli Storage Manager Disaster Recovery

Disaster preparation, planning and recovery (formerly Tivoli Disaster Recovery Manager) is a component of IBM Tivoli Storage Manager Version 5 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 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 the basic Information Technology recovery strategy, the steps to rebuild core systems, as well as the critical machines that must be recovered.

One of the key features of Tivoli Storage Manager and its 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.

Tivoli disaster recovery functions helps maintain business continuance by:

- ▶ Establishing and helping to automate a thorough *server* Disaster Recovery Plan — 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 recover ability of the server. The DRP can be easily recreated daily so that it is up to date. The main functions are illustrated in Figure 20-17.

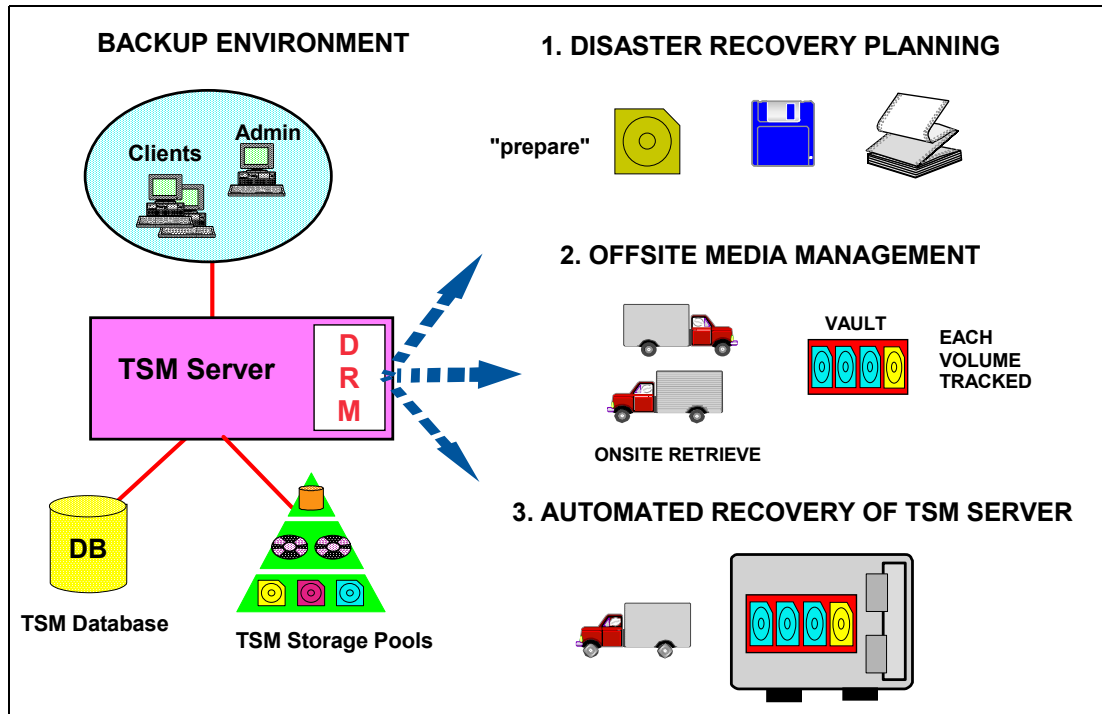


Figure 20-17 Tivoli Disaster Recovery Management (DRM) functions

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

DRM uses the **PREPARE** command to generate a plan file that will contain critical information needed for recovery. Information in the plan is arranged in stanzas, these can be considered to be somewhat like headers. For example, the stanza PLANFILE.DESCRPTION shown in Example 20-1 provides summary information about the plan file as a whole.

Example 20-1 DMR Plan File Example Stanza

```
begin PLANFILE.DESCRPTION

Recovery Plan for Server TESTSERVER

Created by DRM PREPARE on 03/14/2005 17:19:32

DRM PLANPREFIX C:\PROGRAM FILES\TIVOLI\TSM\SERVER1\

Storage Management Server for Windows - Version 5, Release 3, Level 0.0

end
PLANFILE.DESCRPTION
```

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.

20.5.1 DRM and Tivoli Storage Manager clients

DRM is designed to automate the recovery of your Tivoli Storage Manager server(s), rather than automate client recovery. The recovery of the Tivoli Storage Manager server(s) is most critical because it is from the server that clients can continue to backup their data and recover if required. For the clients, DRM allows the machine information needed to help recover the Tivoli Storage Manager clients to be stored in the Tivoli Storage Manager server database. The type of information stored includes:

- ▶ Tivoli Storage Manager client machine location, machine characteristics, and recovery instructions
- ▶ Business priorities associated with the Tivoli Storage Manager client machines
- ▶ Description, location, and volume/diskette/CD labels of Tivoli Storage Manager client boot media

With this information stored in the DRM plan file, you can then use appropriate operating system, CDs, or tape images to perform a bare metal recovery (BMR) of your client system. Finally, the Tivoli Storage Manager client on that machine can be reinstalled, which allows for restoration of client data from the Tivoli Storage Manager server.

20.6 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 recover ability. 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

This section provides an overview of server-to-server communication and virtual volumes.

20.6.1 Server-to-server communication

The server-to-server communication function in Tivoli Storage Manager provides the ability to enable enterprise configuration, enterprise event logging and monitoring, command routing (Figure 20-18).

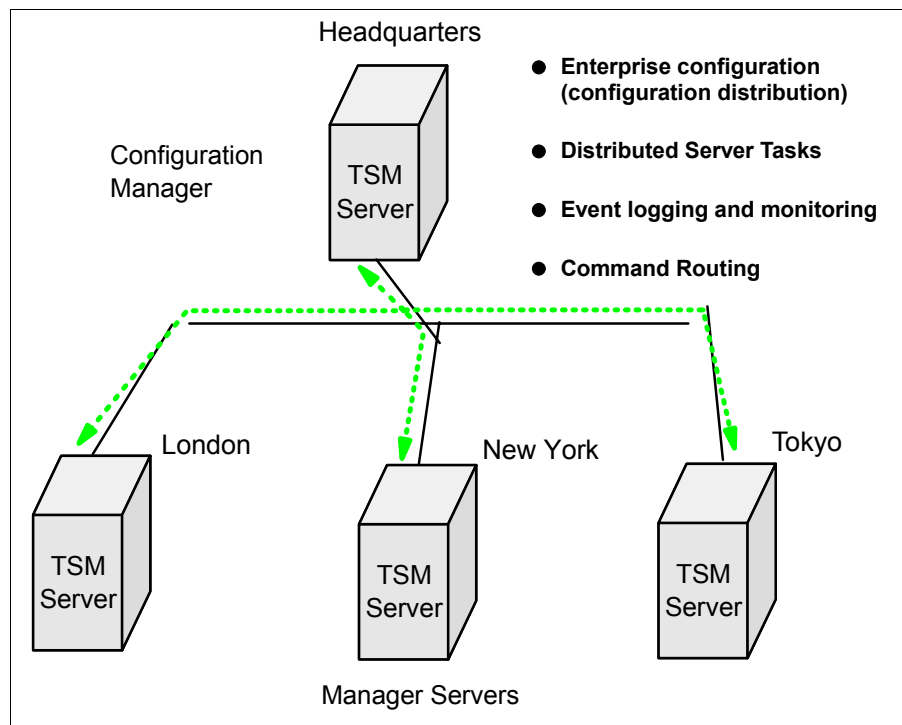


Figure 20-18 Tivoli Storage Manager server-to-server communications

The enterprise configuration functions of the Tivoli Storage Manager make it easier to consistently set up and manage a network of Tivoli Storage Manager servers. You set up configurations on one server and distribute the configurations to the other servers. You can

make changes to configurations and have the changes automatically distributed. To use enterprise configuration, you first select the Tivoli Storage Manager server that is to act as the *configuration manager*. You may want to dedicate a new server for this purpose. At the configuration manager, you define the details of the server configurations that you want to distribute.

For example:

- ▶ You set up backup and archive policies and client option sets.
- ▶ You designate one or more administrators to have access to the servers, and control their authority levels.
- ▶ You define the servers that you want the configuration manager to manage or communicate with, and you set up communications among the servers.

On each server that is to receive the configuration information, you identify the server as a *managed server* by defining a *subscription* to one or more profiles owned by the configuration manager.

When you connect to the configuration manager via a Web browser, you are presented with the enterprise console. From the enterprise console you can perform tasks on the configuration manager and on one or more of the managed servers. You can also connect to another server to perform tasks directly on that server. As long as you are registered with the same administrator ID and password, you can do this work on many servers without having to log on each time.

From the command line of the administrative Web interface or from the command-line administrative client, you can also route commands to other servers. The other servers must be defined to the server to which you are connected. You must also be registered on the other servers as an administrator with the administrative authority that is required for the command. Command routing enables an administrator to send commands for processing to one or more servers at the same time. The output is collected and displayed at the server that issued the routed commands. A system administrator can configure and monitor many different servers from a central server by using command routing. To make routing commands easier, you can define a server group that has servers as members. Commands that you route to a server group are sent to all servers in the group.

One or more servers can send server events and events from their own clients to another server for logging. The sending server receives the enabled events and routes them to a designated event server. This is done by a receiver that Tivoli Storage Manager provides. At the event server, an administrator can enable one or more receivers for the events being routed from other servers.

20.6.2 Server-to-server virtual volumes

Tivoli Storage Manager server-to-server virtual volume support lets a server (a *source* server) use another Tivoli Storage Manager server (a *target* server) as the destination device class for certain Tivoli Storage Manager entities. The data is stored as *virtual volumes*, which are actually archive files on a target server. The target server has no knowledge of the contents or type of objects which are stored in the virtual volumes. The only detailed information about the virtual volume contents is contained in the source server's database. Virtual volumes can be any of the following:

- ▶ Database backups
- ▶ Storage pool backups (copy storage pools)
- ▶ Data that is backed, archived or space managed from client nodes (primary storage pools)
- ▶ Any data that can be moved by the **EXPORT** and **IMPORT** commands

- ▶ DRM plan files (output of the DRM **PREPARE** command)

The source server is a client of the target server, and the data for the source server is managed only by the source server. In other words, the source server controls the expiration and deletion of the files that comprise the virtual volumes on the target server. At the target server, the virtual volumes from the source server are seen as archive data. The relationship between the source and target Tivoli Storage Manager servers is illustrated in Figure 20-19.

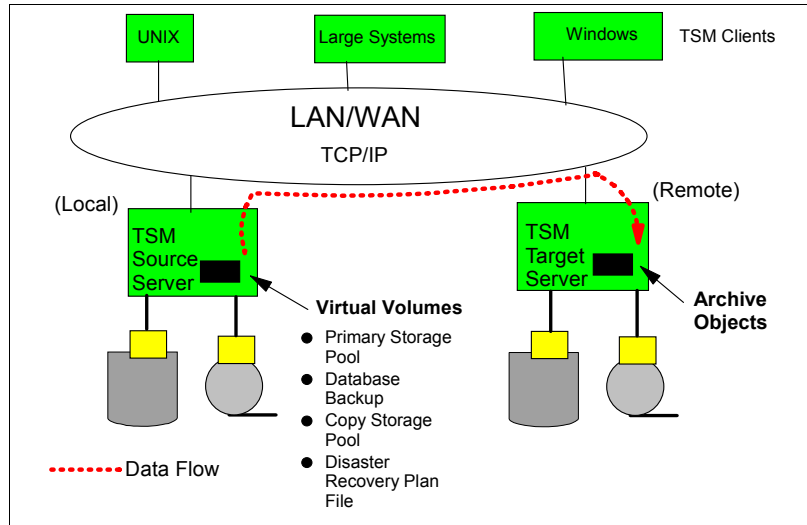


Figure 20-19 Server-to-server virtual volumes

The source server is registered as a client node (of TYPE=SERVER) at the target server and is assigned to a policy domain. The archive copy group of the default management class of that domain specifies the storage pool for the data from the source server. TCP/IP is the only communications method that can be used for server-to-server virtual volumes.

To use virtual volumes, the source server needs to define one or more device classes of TYPE=SERVER, using the **DEFINE DEVCLASS** command. The device class definition indicates on which remote or target server the volumes will be created. Having defined the device class, it can be used for primary or copy storage pools, database backups and other virtual volume functions.

All data destined for virtual volumes is sent to the target server, using virtual volumes rather than direct attached storage devices. For example, if a client is backing up data which is bound to a backup copy group which is using a virtual volume primary storage pool, this data will be sent to the target server. If a client needs to restore the data, the source server gets the data back from the target server. A Tivoli Storage Manager client can always perform the same granularity of restore, retrieve or recall operation, whether the data is stored on a local Tivoli Storage Manager server, or on a target server using server-to-server communication. That is, remote storage pools (using server-to-server communication) are transparent to the client. The only requirement is that the TCP/IP communication link between the source and target server must be working correctly.

Note: Source server objects such as database and storage pool backups are stored on the target server as *archived* data. Therefore, the target server cannot directly restore these objects in the event of a disaster at the source server site. In the event of a disaster at the source server site, the source server should be re-installed (likely at an alternate location) and then objects originally stored on the target server can be restored over the network, using the same server-to-server communication.

Using virtual volumes can benefit you in the following ways:

- ▶ The source server can use the target server as an electronic vault. This removes the need to physically ship media off site, or retrieve it physically in a disaster.
- ▶ Smaller Tivoli Storage Manager source servers can use the storage pools and tape devices of larger Tivoli Storage Manager servers.
- ▶ For incremental database backups, it can decrease wasted space on volumes and under use of high-end tape drives.

20.6.3 Using server-to-server virtual volumes for disaster recovery

Server-to-server virtual volumes can be used as part of strategies 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 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). To display plan file information on the target server you can issue the **QUERY RPFIL** command, specifying the source server node name that prepared the plan. You can easily display a list of all recovery plan files that have been saved on a target server.

20.6.4 Considerations for server-to-server virtual volumes

When using server-to-server virtual volumes consideration must be given to network bandwidth, the location of Tivoli Storage Manager database and storage pools, and data flow duplication. As mentioned earlier, TCP/IP is the only supported communications protocol. The following should be considered when implementing server-to-server virtual volumes.

- ▶ If you use virtual volumes for database backups, you might have the following situation: SERVER_A backs up its database to SERVER_B in the same location, and SERVER_B backs up its database to SERVER_A. If this is the only way databases are backed up, and a disaster strikes at that location, you may have no backups available from which to restore your databases.
- ▶ Moving large amounts of data between the servers may slow down your communications significantly, depending on the network bandwidth and availability. In general a Tivoli Storage Manager server-to-server communications operation will achieve slower throughput than a typical client backup operation. You should carefully test to make sure sufficient performance is provided.
- ▶ Storage space limitations on the target server affect the amount of data that you can store on that server. Make sure that you include virtual volume requirements in your planning for each server.
- ▶ Duplicate data flow over the network can occur when one storage pool is backed up to another storage pool and both primary and copy storage pools are using virtual volumes on the same target server. In this situation, data will be moved from a primary storage pool volume (on the target server), back to the source server, then to a copy storage pool

volume on the target server. This is clearly more time-consuming than a simple copy storage pool operation without using server-to-server communication.

20.7 Tivoli Storage Manager and high availability clustering

What exactly is a cluster? In simple terms, a cluster is a group of computers that, together, provide a set of resources to a client. A cluster consists of a minimum of two and in theory, up to any number, of systems. The key point of clustering is that the client has no knowledge of the underlying physical hardware of the cluster. This means that the 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. For example, a client can request the use of an application without being concerned about either where the application resides or which physical server is processing the request. The user simply gains access to the application in a timely and reliable manner.

Another benefit is scalability. If you need to add users or applications to your system and want performance to be maintained at existing levels, additional systems can be incorporated into the cluster. A typical example would be a Web site that shows rapid growth in the number of demands for Web pages from browser clients. Running the site on a cluster would allow the growth in demand to be easily accommodated by adding servers to the cluster as needed.

Whether you will configure your system to include clusters depends on your business needs. A cluster can provide system level high availability to ensure a Tivoli Storage Manager server or client can continue normal backup and restore processes without significant disruption for users and administrators. In addition to assuring the right type of hardware and the applicable software, varying failover patterns between cluster nodes exist and play different roles, for example, hot-standby versus concurrent cluster operation. This section overviews Tivoli Storage Manager server and client clustering support and configuration using IBM High Availability Cluster Multi-Processing (HACMP) and Microsoft Cluster Server (MSCS). It should be noted that Tivoli Storage Manager also supports High Availability on Novell NetWare clients and with Domino, R/3 and Exchange databases.

20.7.1 IBM High Availability Cluster Multi-Processing (HACMP) and TSM

A Tivoli Storage Manager server can use IBM's High Availability Cluster Multiprocessing (HACMP) software for high availability. HACMP provides the leading AIX-based clustering solution, which allows automatic system recovery on system failure detection. By using HACMP together with Tivoli Storage Manager, you can ensure 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 loss of end-user time. You can set up a Tivoli Storage Manager server on a system in an HACMP cluster so that, if the system fails, the Tivoli Storage Manager server will be brought back up on another system in the cluster. In both failover and fallback, it appears that the Tivoli Storage Manager server has crashed or halted and was then restarted. Any transactions that were in progress at the time of the failover or fallback are rolled back, and all

completed transactions are still complete. Tivoli Storage Manager clients see this as a communications failure and try to re-establish their connections as shown in Figure 20-20.

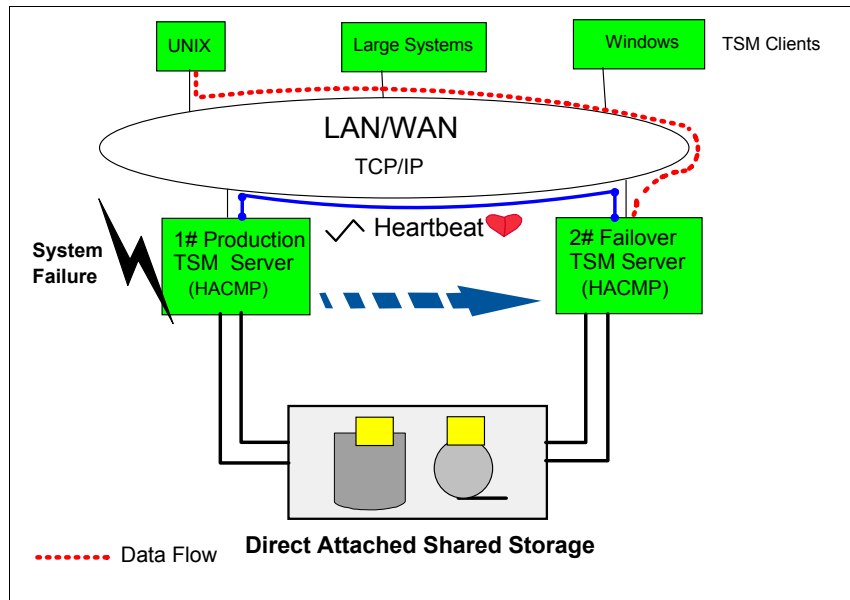


Figure 20-20 HACMP and Tivoli Storage Manager server high availability configuration

HACMP support for the Tivoli Storage Manager server on AIX has been officially supported since V4.2. Several HACMP scripts are provided with the Tivoli Storage Manager server files. These scripts can then be customized to suit the local environment. When failover occurs, HACMP calls the Tivoli Storage Manager *startserver* script on the standby node. The script verifies the devices, breaking the SCSI reserves, and starts the server. On fallback, the *stopserver* script runs on the standby node, which causes the Tivoli Storage Manager server to halt. Then the *startserver* script runs on the production node. HACMP handles taking over the TCP/IP address and mounting the shared file systems on the standby node or production node, as appropriate. By default, the *startserver* script will not start the Tivoli Storage Manager server unless all the devices in the VerifyDevice statements can be made available. However, you can modify the *startserver* script to start the server even if no devices can be made available.

Both failover and fallback act as though a Tivoli Storage Manager server has crashed or halted and was then restarted. Any transactions that were in-flight at the time are rolled back, and all completed transactions are still complete. Tivoli Storage Manager clients see this as a communications failure and try to re-establish connection based on their COMMRESTARTDURATION and COMMRESTARTINTERVAL settings. The backup-archive client can usually restart from the last committed transaction. The clients and agents will behave as they normally do if the server was halted and restarted while they were connected. The only difference is that the server is physically restarted on different hardware.

Note: For a detailed discussion on supported environments, prerequisites, install, setup, and testing of a HACMP and Tivoli Storage Manager server failover environment, see *Tivoli Storage Manager for AIX Quick Start Version 5.1*, GC32-0770, and *IBM Tivoli Storage Manager version 5.3 Technical Guide*, SG24-6638.

20.7.2 TSM backup/archive and HSM client support with HACMP

As of Tivoli Storage Manager Version 5.1, the backup-archive client itself (including the administrator, backup/archive, HSM and API pieces) is supported for use in an HACMP 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. See Figure 20-21 for an illustration of how this works.

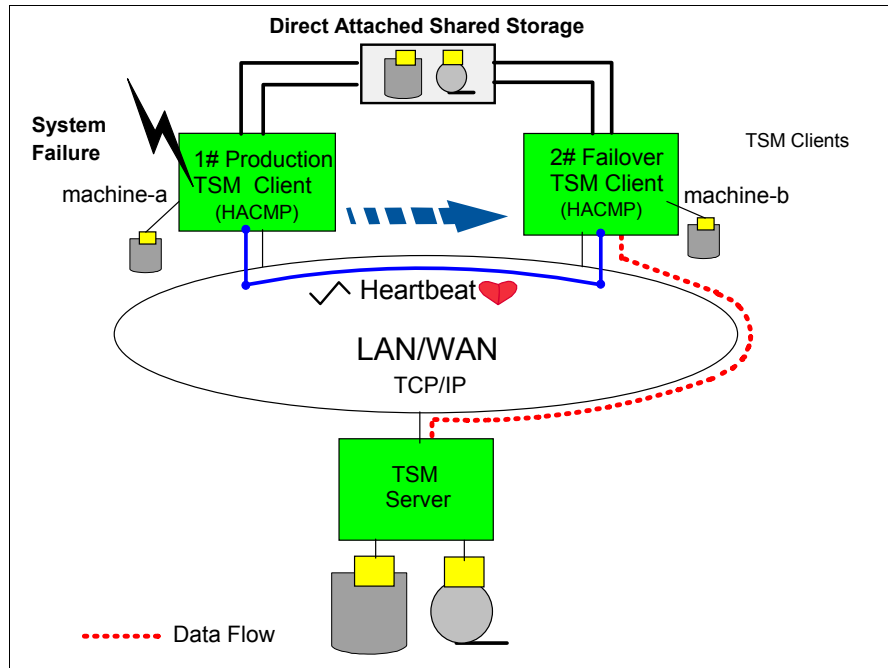


Figure 20-21 HACMP and Tivoli Storage Manager client high availability configuration

The CLUSTERNODE option in the AIX client `dsm.sys` file determines if you want the Tivoli Storage Manager client to back up cluster resources and participate in cluster failover for high availability.

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 filesystems can continue to operate in the case of an HACMP node failover and fallback.

20.7.3 Tivoli Storage Manager and Microsoft Cluster Server (MSCS)

Tivoli Storage Manager is a cluster-aware application and can be configured in a MSCS high availability environment. The administrator uses the MSCS Cluster Administrator interface and Tivoli Storage Manager to designate cluster arrangements and define the failover pattern. The 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. Clustering allows you to join two Windows servers, or nodes, using a shared disk subsystem. This provides the nodes with the ability to share data, which provides high server availability.

For example, in the MSCS failover environment shown in Figure 20-22, a clustered Tivoli Storage Manager server called Tivoli Storage ManagerSERVER1 runs on node A and a clustered Tivoli Storage Manager server called Tivoli Storage ManagerSERVER2 runs on node B. Clients connect to the Tivoli Storage Manager server Tivoli Storage ManagerSERVER1 and the Tivoli Storage Manager server Tivoli Storage ManagerSERVER2 without knowing which node currently hosts their server. The MSCS concept of a virtual server ensures that the server's location is transparent to client applications. To the client, it appears that the Tivoli Storage Manager server is running on a virtual server called Tivoli Storage ManagerSERVER1.

When 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 remaining node. The remaining node takes over the Tivoli Storage Manager server resource group, restarts the Tivoli Storage Manager service, and provides access to administrators and clients. If node A fails, node B assumes the role of running Tivoli Storage ManagerSERVER1. To a client, it is exactly as though node A were turned off and immediately turned back on again.

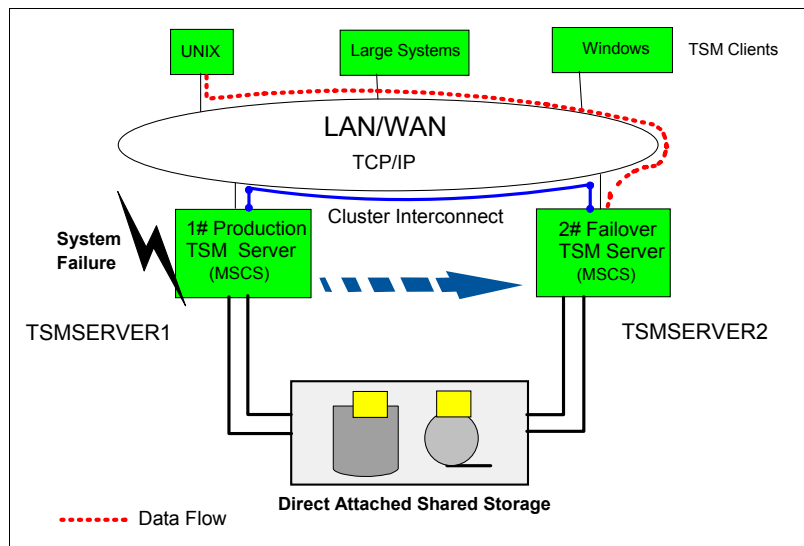


Figure 20-22 MSCS and Tivoli Storage Manager server high availability configuration

Clients experience the loss of all connections to Tivoli Storage ManagerSERVER1 and all active transactions are rolled back to the client. Clients must reconnect to Tivoli Storage ManagerSERVER1 after this occurs, which is normally handled as an automatic attempt to reconnect by the Tivoli Storage Manager client. The location of Tivoli Storage ManagerSERVER1 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.

Two failover configurations are supported with MSCS and Tivoli Storage Manager: *active/passive* and *active/active*. 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. In the *active/active* configuration the cluster runs two independent instances of a Tivoli Storage Manager server, one on each server. In the event of a system failure, the server on the failed instance transfers to the surviving instance, so that it is running both instances. Even if both instances are running on the same physical server, users believe they are accessing a separate server.

Clusters consist of many components such as nodes, cluster objects, Microsoft Cluster Server (MSCS) virtual servers, and even the hardware and software. If any one of these components is missing, the cluster cannot work. MSCS requires each Tivoli Storage Manager server instance to have a private set of disk resources. Although nodes can share disk resources, only one node can actively control a disk at a time. TCP/IP is used as the communications method in a MSCS environment with Tivoli Storage Manager.

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 Version 5.2*, GC32-0782, and *Tivoli Storage Manager Version 3.7.3 and 4.1: Technical Guide*, SG24-6116.

20.7.4 Tivoli Storage Manager backup/archive client support with MSCS

The Tivoli Storage Manager client is supported within a 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 20-23.

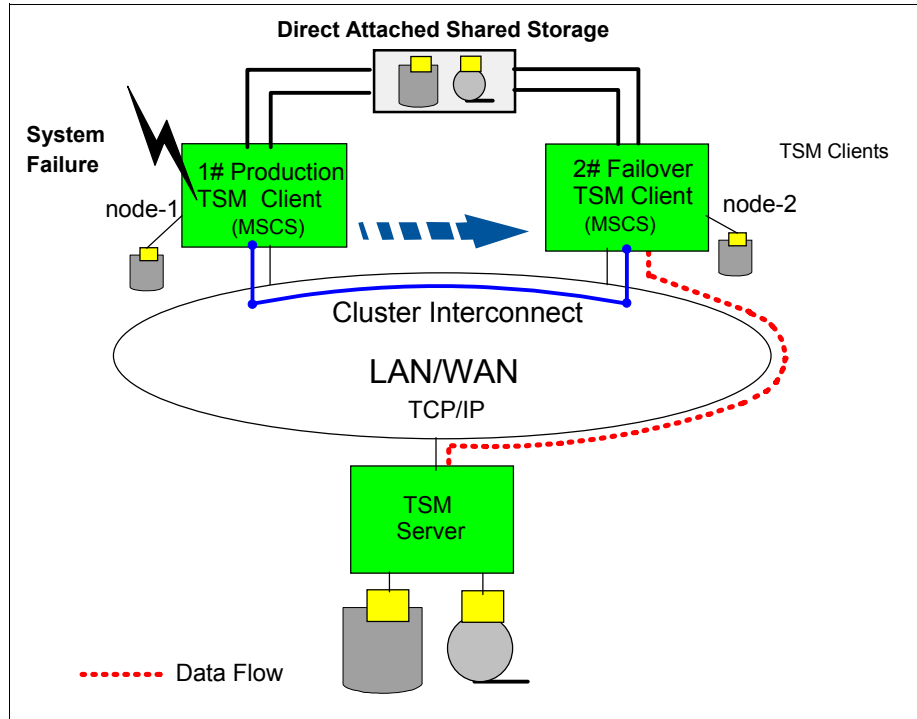


Figure 20-23 MSCS and Tivoli Storage Manager client high availability configuration

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 should be installed for each node: node-1, node-2, and physical disk resources. 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 5.1*, GC32-0788.

20.8 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. DRM helps you to 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.

20.8.1 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 20-24.

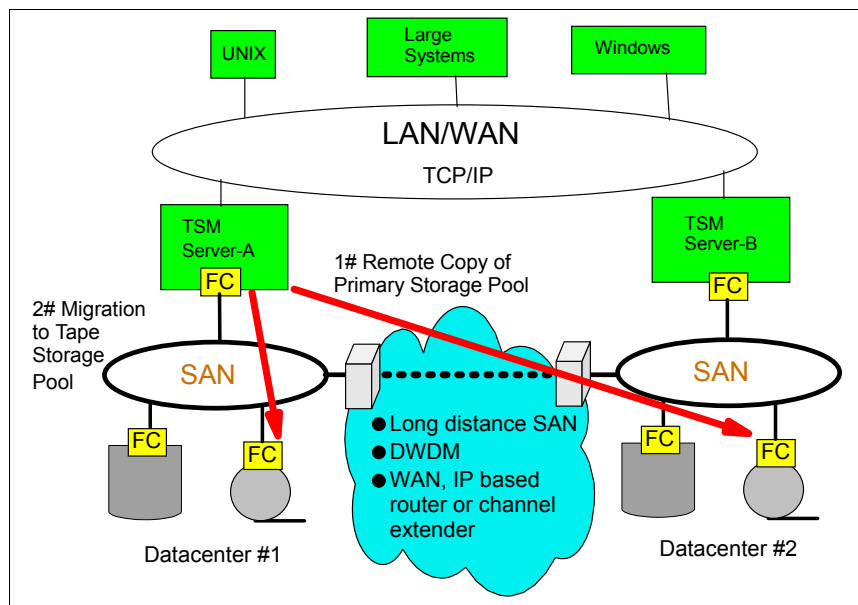


Figure 20-24 Tivoli Storage Manager and electronic tape vaulting

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.

20.9 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 20-2 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 TotalStorage Enterprise Storage Server uses PPRC to achieve data replication. PPRC is supported via ESCON links, which can be further extended via DWDM or WAN channel extension.

Table 20-2 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) INRANGE (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.

20.9.1 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.

20.9.2 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.

20.10 More information

- ▶ More information on Tivoli Storage Manager can be found on the following Web site:
<http://www.tivoli.com/TivoliStorageManager>
- ▶ Refer to this link for details of all Tivoli Application Protection solutions:
<http://www.ibm.com/software/tivoli/solutions/storage/application/>
- ▶ These Redbooks give a good insight into Tivoli Storage Manager as well:
Getting Started with Tivoli Storage Manager Implementation Guide, SG24-5416:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg245416.html?Open>
Disaster Recovery Strategies with Tivoli Storage Management, SG24-6844:
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246844.html?Open>
TSM 5.3 Windows Backup-Archive Client Installation and User's Guide, GC32-0788
TSM V5.3 for Windows Administrator's Guide, GC32-0782
IBM Tivoli Storage Manager version 5.3 Technical Guide, SG24-6638



Part 5

z/OS storage management

Part 5 describes these products:

- ▶ DFSMS
- ▶ DFSORT
- ▶ ESS and ETL Expert
- ▶ z/OS storage management tools



DFSMS

DFSMS is a software suite that automatically manages data from creation to expiration. DFSMS, formerly called DFSMS/MVS® is a stand-alone product, and is now an integral part of the z/OS operating system. DFSMS provides allocation control for availability and performance; backup/recovery and disaster recovery services; space management; tape management; and reporting and simulation for performance and configuration tuning.

These are the components of DFSMS:

- ▶ **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.

In addition:

- ▶ **z/OS NFS** (Network File System), formerly called DFSMS/MVS NFS: Provides a way to optimize performance and efficiency in a distributed network while still capitalizing on the capacity, security, and integrity of z/OS. It does not support tape and optical drives.
- ▶ **DFSORT**: Separate complementary product that provides high performance, sorting, merging, copying, reporting, and analysis of data.

21.1 DFSMSdfp (Data Facility Product)

DFSMSdfp and the BASE Control Program (BCP) form the foundation of the z/OS operating system where DFSMSdfp performs the essential data, storage, and device management functions of the system: it provides the logical and physical input and output for z/OS storage, it keeps track of all data and programs managed within z/OS, and it provides data access both for native z/OS applications and other platforms such as AIX/UNIX, the Windows family, or OS/400.

The other elements of DFSMS: DFSMSdss, DFSMSHsm, DFSMSRmm, and DFSMSStvs complement DFSMSdfp to provide a fully-integrated approach to data and storage management. In a system-managed storage environment, DFSMS automates and centralizes storage management based on the policies that your installation defines for availability, performance, space and security. With these optional features enabled, you can take full advantage of all the functions that DFSMS offers.

21.1.1 Overview

The major features and characteristics of DFSMSdfp are:

- ▶ **Storage management:** DFSMSdfp lets you define and maintain policies (using the ISMF interface) to manage your storage resources. These policies help to improve the use of storage devices and to increase levels of service for user data, with minimal effort required from users. DFSMSdfp provides the Storage management subsystem (SMS) which manages these policies for the operating system. You can also use the NaviQuest tool under ISMF to help you migrate to SMS, maintain your SMS configuration, and perform many testing, implementation, and reporting tasks in batch.
- ▶ **Tape mount management:** SMS provides a means for implementing tape mount management, a methodology for improving tape usage and reducing tape costs. This methodology involves intercepting selected tape data set allocations through the SMS automatic class selection (ACS) process and redirecting them to a DASD buffer. Once on DASD, these data sets can be migrated to a single tape or small set of tapes, thereby reducing the overhead associated with multiple tape mounts.
- ▶ **Data management:** DFSMSdfp helps you store and catalog information on DASD, optical, and tape resources so that it can be quickly identified and retrieved from the system. You can use the Catalog Search Interface, now part of DFSMSdfp, to quickly access catalog information.
- ▶ **Device Management:** DFSMSdfp is involved in defining your input and output devices to the system and in controlling the operation of those devices in the z/OS environment.
- ▶ **Distributed Data Access:** Distributed data access allows all authorized systems and users in a network to exploit the powerful features of system-managed storage, or automated storage management provided by DFSMS. DFSMSdfp uses the Distributed FileManager (DFM) to support remote access of z/OS data and storage resources from workstations, personal computers, or any other system on a SNA LU 6.2 network.
- ▶ **Advanced Copy services:** Advanced Copy Services includes remote and point-in-time copy functions that provide backup and recovery of data. When used before a disaster occurs, Advanced Copy Services provides rapid backup of critical data with minimal impact to business applications. If a disaster occurs to your data center, Advanced Copy Services provides rapid recovery of critical data.
- ▶ **Object Management on Multiple Media Types:** DFSMSdfp includes the Object Access Method (OAM) component. DFSMSdfp OAM is a functional component of the DFSMS family that manages a class of data referred to as objects for the z/OS environment. Separately it is also used in the support of tape libraries.

OAM manages large objects, like compressed scanned images or coded data, in their entirety and contains no restrictions on the data in an object. Objects may be stored on disk, tape, or optical platters, and freely staged up and down the storage hierarchy based on SMS control parameters.

Secondly, in the support of tape libraries, OAM uses system-managed storage concepts within the SMS component of DFSMS to provide the management facilities for the physical movement and tracking of tape volumes used within tape libraries.

21.1.2 ISMF

The components of DFSMS automate and centralize storage management, based on policies your installation defines for availability, performance, space, and security. The Interactive Storage Management Facility (ISMF) provides the user interface for defining and maintaining these policies, which the Storage Management Subsystem (SMS) governs for the system.

ISMF is an Interactive System Productivity Facility (ISPF) application. ISMF works with the following products, which you should be familiar with:

- ▶ Interactive System Productivity Facility/Program Development Facility (ISPF/PDF), which provides the edit, browse, Data Set, and Library utility functions
- ▶ TSO/Extensions (TSO/E), TSO CLISTs and commands
- ▶ Data Facility SORT (DFSORT), which provides the record-level functions
- ▶ RACF®, a component of the Security Server for z/OS, which provides the access control function for data and services
- ▶ Device Support Facilities (ICKDSF), which provide the storage device support and analysis functions

ISMF provides access to the functions of the DFSMS family of products, as represented in Figure 21-1.

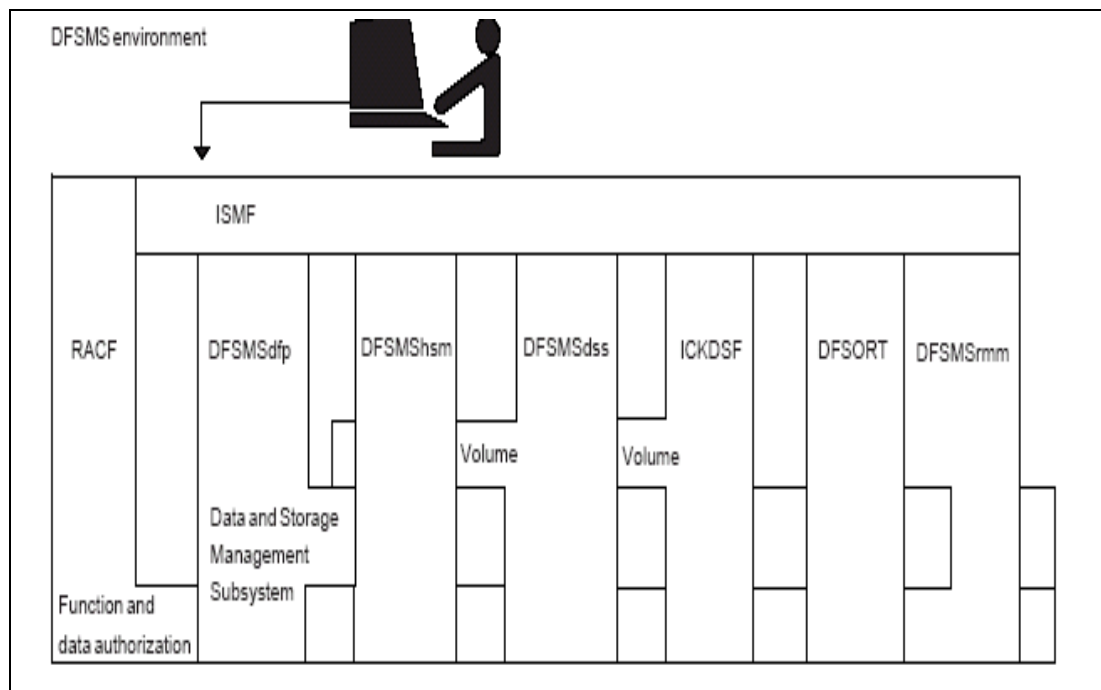


Figure 21-1 DFSMS environment

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 datasets 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 21-2 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 21-2 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

You cannot allocate data sets from ISMF. Data sets are allocated from ISPF, from TSO, or with job control language (JCL) commands. ISMF provides the DSUTIL command, which enables users to get to ISPF and toggle back to ISMF.

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.

21.1.3 System Managed Storage

Using the ISMF menu driven utility, DFSMSdfp allows the storage administrator to assign attributes to data sets and objects through the definition of data class, storage class, management class, and storage group constructs.

The following paragraphs illustrate some of the possibilities and facilities available to storage administrators through DFSMSdfp:

Data Classes

A storage administrator can define data classes (a collection of allocation and space attributes) to provide users with the following benefits (users need not specify attributes because the data class attributes can be automatically assigned, based on the name of the data set for instance):

- ▶ Simplify JCL specification
- ▶ Use allocation defaults, specifying allocation values that make efficient use of storage
- ▶ Specify when a data set should be allocated in extended format
- ▶ Allocate VSAM data sets using JCL, dynamic allocation, or the TSO ALLOCATE command
- ▶ Retry data set allocations on new volumes when allocation fails due to space constraints. Allocations are also retried when extending multivolume data sets to new volumes
- ▶ Let the system determine how many buffers and which buffering algorithms to use when a VSAM data set allocated in extended format is accessed using nonshared resources (NSR) and batch programs

Storage Class

The storage administrator can also define a storage class (a collection of performance goals and device availability requirements). SMS uses it to select a device that can best meet those goals and requirements in terms of the performance of the device, the amount of space available on the volume, and how available a data set or object can be on that device.

Storage class availability and accessibility attributes are used with devices such as the Enterprise Storage Server to provide a required level of service. When you define a storage class with a specific service level, SMS attempts to ensure that a data set associated with that storage class is allocated to the required hardware and that the caching abilities are maximized to deliver the required service level.

By using the storage class attributes with ACS routines, your installation can provide users with the following benefits:

- ▶ Specify performance requirements at the data set or object level.
- ▶ Use SMS to make optimal use of storage, matching performance and availability characteristics with user-specified needs.
- ▶ 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.
- ▶ Direct allocations to fault-tolerant devices, to keep critical DASD data sets continuously available.

- ▶ Use the concurrent copy function and DFSMSdss to keep DASD data sets almost continuously accessible while consistent backups or copies are made.
- ▶ Use the sequential data striping function and request the sustained data rate that is required. DFSMS determines the number of devices to meet the requirement. Data for one data set is distributed and transferred concurrently for improved I/O performance.
- ▶ Designate tape data sets for placement in tape libraries.

Management Class

A *management class* is a collection of management attributes defined by the storage administrator. 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. You can also use management classes to define how your aggregate groups should be managed. DFSMSHsm acts on many management class attributes.

By using the management class attributes with ACS routines, your installation can provide users with the following benefits:

- ▶ Provide migration, backup, and deletion at the data set or object level. This enables the storage administrator to manage DASD, optical, and tape storage more effectively and to satisfy user requests for space management and backup.
- ▶ Specify whether a DASD data set must be copied concurrently using concurrent copy, should be copied concurrently if possible, or should be copied using normal backup processing techniques.
- ▶ Specify whether backup copies of an aggregate group of data sets must be backed up concurrently using concurrent copy, should be backed up concurrently if possible, or should be backed up using normal backup processing techniques.
- ▶ Reduce the effort users and the storage administrator require for storage management tasks.

Storage Group

A storage group is a collection of storage volumes and attributes defined by the storage administrator. The storage administrator groups the volumes to meet a specific service or business strategy. Allocating data to storage groups can exploit the advantages of pooled volumes, improving your use of personnel and storage resources.

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.

By using the storage group attributes with ACS routines, your installation can provide users with the following benefits:

- ▶ Simplify creation of objects and segregate objects into groups according to customer usage and location within the storage hierarchy.
- ▶ 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.
- ▶ Direct SMS to select volumes from storage groups for new data. For pool storage groups, the storage thresholds of each DASD volume are taken into account during the selection process so that space usage can be balanced across volumes within storage groups and DASD space can be used more efficiently. For tape storage groups, the number of

- available scratch tape volumes in each library is also taken into account during the selection process, so that the storage resources can be balanced across storage groups.
- ▶ Simplify device installation and hardware configuration management. An administrator can add or delete devices through the ISMF Storage Group Application. An initial program load (IPL) is not necessary when making changes to storage groups.
 - ▶ Simplify data set allocation for users. When SMS uses ACS routines to automatically direct new data sets to appropriate storage groups, users do not have to provide unit and volume information during allocation.
 - ▶ Improve the balance of I/O activity across devices.

21.1.4 Managing data with DFSMSdfp

DFSMSdfp provides several access methods for formatting and accessing data. An access method defines the organization of the data in a data set and the technique by which the data is stored and retrieved. DFSMSdfp access methods have their own data set structures to organize data, macro instructions to process data sets, and utility programs to manipulate data sets.

Access Methods

- ▶ 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.
- ▶ Virtual Storage Access Method (VSAM)

VSAM arranges records by an index key, by relative byte address, or by relative record number. Data organized by VSAM is cataloged for easy retrieval and is stored in one of four types of data sets:

 - Key-sequenced data set (KSDS)
 - Entry-sequenced data set (ESDS)
 - Relative-record data set (RRDS)
 - Linear data set (LDS)
- ▶ Basic Sequential Access Method (BSAM)

BSAM arranges records sequentially in the order in which they are entered. Records are stored in physical blocks and retrieved as requested. This is called basic access. You can use BSAM with physical, sequential data sets or members of partitioned data sets or partitioned data sets extended (PDSE). BSAM data sets can be compressed and striped.
- ▶ Queued Sequential Access Method (QSAM)

QSAM arranges records sequentially in the order in which they are entered. QSAM collects records into blocks. This is called blocking. QSAM anticipates the need for records based on their order, and, to improve performance, reads these records into storage before they are requested.

This is called queued access. You can use QSAM with physical sequential data sets, or with members of partitioned data sets (PDSs) or partitioned data set extended (PDSEs). QSAM data sets can be compressed and striped.
- ▶ Basic Partitioned Access Method (BPAM)

BPAM arranges records as members of partitioned data sets or PDSEs. A partitioned data set or PDSE includes a directory that relates member names to locations on the DASD volume, in order to retrieve individual members.

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), 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.

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.
- ▶ **IEBPTPCH** 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 (VTOCs).

Managing data sets and volumes

To manage the storage and placement of data sets, DFSMSdftp uses catalogs and volume tables of contents (VTOCs).

A catalog describes data set attributes and records the location of a data set so that the data set can be retrieved without requiring the user to specify the data set's location. Multiple user catalogs contain information about user data sets, and a single master catalog contains entries for system data sets and user catalogs.

Tape libraries and the volumes they contain must be cataloged in a volume catalog. A volume catalog is user catalog that can only contain entries for tape volumes and tape libraries. You can use access method services to define and maintain a volume catalog the same way you define and maintain any catalog, except that you cannot define aliases to a volume catalog.

DFSMSdftp provides numerous facilities to manage catalogs, that is:

- ▶ Create catalogs
- ▶ Define catalog entries

- ▶ Alter catalog characteristics
- ▶ Diagnose catalog errors
- ▶ Delete catalog entries
- ▶ Copy or merge catalog entries
- ▶ Back up catalogs
- ▶ Lock catalogs
- ▶ Recover catalogs
- ▶ List the entries in a catalog
- ▶ Examine the structural consistency of a catalog.

21.2 DFSMSdss (Data Set Services)

DFSMSdss is a reliable utility to quickly move, copy and backup data.

This high-speed component of DFSMS is a fast, efficient data mover that is particularly valuable to customers moving many data sets from old to new disk subsystems.

DFSMSdss is used to move data from disk to disk (copy) or disk to tape (dump) very quickly and efficiently. It performs both logical dumps (to copy data) and physical dumps (to copy track images).

21.2.1 Overview

DFSMSdss is a component of DFSMS, and is used to:

- ▶ **Move and Replicate data:** DFSMSdss offers powerful, user-friendly functions that let you move or copy data between volumes of like and unlike device types. It can also copy data that has been backed up.
- ▶ **Manage storage space efficiently:** DFSMSdss can increase performance by reducing or eliminating DASD free-space fragmentation.
- ▶ **Backup and Recover data:** DFSMSdss provides you with host system backup and recovery functions at both the data set and volume levels. It also includes an enhanced Stand-alone Restore Program that lets you restore vital system packs during disaster recovery - without a host operating system.
- ▶ **Convert data sets and volumes:** DFSMSdss can convert your data sets and volumes to system-managed storage, or return your data to a non-system-managed state as part of a recovery procedure - all without data movement.

21.2.2 Interface to other services

DFSMSdss is an external interface to Concurrent Copy and FlashCopy. It is also the file manager for DFSMSHsm and fully supports ABARS.

DFSMS Advanced Copy Services

Advanced Copy Services, running on the IBM Enterprise Storage Server (ESS), consists of the following individual functions and enhancements:

- ▶ **Remote copy**, which consists of the following:
 - Extended Remote Copy (XRC): XRC is a combined hardware and software solution that offers the highest levels of continuous data availability in a disaster recovery and workload movement environment. Offered on the ESS, XRC provides an asynchronous remote copy solution, of both system-managed and non-system-managed data to a second, remote location.

- Coupled extended remote copy (CXRC) and Multiple Extended Remote Copy (MXRC): CXRC and MXRC are designed to support large environments that have an expanded number of primary storage controls and DASD volumes, in excess of those supported by a single data mover configuration.
- Peer-to-Peer Remote Copy (PPRC): PPRC is a hardware solution to rapid and accurate disaster recovery as well as a solution to workload and DASD migration. Updates made on the primary DASD volumes are synchronously shadowed to the secondary DASD volumes.
- ▶ **FlashCopy:** FlashCopy is a point-in-time copy services function that can quickly copy data from a source location to a target location.
- ▶ **Concurrent Copy (CC):** Concurrent copy is a copy services function that generates a copy of data while applications are updating that data.

Many of these functions rely on the System Data Mover (SDM). The SDM is the high-speed data transfer engine behind the Extended Remote Copy, SnapShot, and Concurrent Copy functions. Peer-to-Peer Remote Copy does not rely on the SDM for data movement.

DFSMSdss Stand-alone Services

The Stand-Alone Services function is available to all supported releases of DFSMSdss. The Stand-Alone Services function completely replaces and makes obsolete the previous stand-alone restore function, which is no longer available as part of DFSMSdss.

This version of IBM's Stand-Alone restore function is a single-purpose program designed to allow the system programmer to restore vital system packs during disaster recovery without needing to rely on a z/OS environment. Stand-Alone Services runs independently of a system environment either as a "true" stand-alone system or under a VM system.

Stand-Alone Services can perform either a full-volume or a tracks restore from dump tapes produced by DFSMSdss or DFDSS, and offers the following benefits when compared to the previous DFSMSdss stand-alone functions:

- ▶ Provides user-friendly commands to replace the previous control statements
- ▶ Supports IBM 3494 and 3495 Tape Libraries, and 3590 Tape Subsystems
- ▶ Supports IPL-ing from a DASD volume, in addition to tape and card readers
- ▶ Allows you to predefine the operator console to be used during Stand-Alone Services processing

21.3 DFSMShsm (Hierarchical Storage Manager)

DFSMShsm is a functional component of DFSMS that provides facilities for managing storage devices.

21.3.1 Overview

DFSMShsm provides automatic management of low activity and inactive data, and automatic backup and recovery of active data in both system-managed and non-system-managed environments. DFSMShsm accomplishes this by providing the following types of management:

- ▶ Automatic Storage Management

DFSMSHsm is a tool that improves productivity by effectively managing storage. DFSMSHsm uses a hierarchy of storage devices in its automatic management of data that relieves users from manual storage management tasks and improves DASD utilization.

- ▶ Space Management

DFSMSHsm automatically manages DASD space by enabling active data sets to be kept on fast access storage devices. DFSMSHsm frees available space on user volumes by releasing over allocated space, deleting eligible data sets, and by moving low activity data sets to lower cost-per-byte devices. By using DFSMSHsm in conjunction with your automatic class selection (ACS) routines to implement the tape mount management methodology, you can write multiple output data sets to a single tape, with a single tape mount. This not only improves tape use, but also greatly reduces the number of tape mounts required by your installation.

- ▶ Availability Management

DFSMSHsm makes data available by automatically copying new and changed data sets to backup volumes. If the original data sets are damaged or accidentally deleted, having backup versions ensures that the data can be made available. DFSMSHsm provides the ability to back up aggregate groups of critical data sets and programs that can be taken to remote locations and used for disaster recovery or other business needs. DFSMSHsm creates and manages backup versions of storage groups defined within copy pools. DFSMSHsm recovers a fast replication backup version from the volume or copy pool level. DFSMSHsm also provides the secondary host promotion function to improve data availability in instances where a DFSMSHsm host processor experiences a failure in a multisystem environment. Using the secondary host promotion function, with a cross-system coupling facility (XCF) in either a basic or Parallel Sysplex environment, a secondary host can take over the unique functions that were performed by a failed primary host. The secondary host promotion function also enables other DFSMSHsm hosts to take over secondary space management functions from either a primary or secondary host that has failed. If a promoted host also fails, any remaining host that is eligible for promotion will take over.

DFSMSHsm provides a number of functions to support space and availability management. Full exploitation of DFSMSHsm services in DFSMS environments requires the use of DFSMSdss for certain functions.

21.3.2 Data Hierarchy

DFSMSHsm-owned volumes contain low activity and inactive user data sets that were moved (migrated data) or copied (dumped and backed up data) into the DFSMSHsm-owned portion of the storage hierarchy. The data is stored in a space-saving format. The data stored in a space-saving format is not directly accessible by users, but must be returned to a level 0 volume in user format before users can access it. Information about DFSMSHsm-owned data is recorded in the inventory maintained in the DFSMSHsm control data sets.

- ▶ Migration level 1 and migration level 2 volumes

Low activity data is stored on migration level 1 (ML1) and migration level 2 (ML2) volumes. The DFSMSHsm automatic management of ML1 and ML2 volumes can be supplemented by DFSMSHsm commands to manage them manually.

- ▶ Backup tapes

Backup tapes contain the backup copies of data sets placed there on a daily basis.

- ▶ Spill backup tapes

Spill backup tapes contain data moved from backup volumes as a result of recycle processing. The data is written in a DFSMSShsm space saving format. The recycle function is a tape consolidation process performed on ML2 and backup tapes that become sparsely populated with valid data.

- ▶ **Dump tapes**

Dump tapes contain all the data of a physical volume and are written in a DFSMSDss space saving format. Several dumps from different DASD volumes can be automatically stacked onto a single tape.

- ▶ **Aggregate data tapes**

Aggregate data tapes contain user-defined groups of data sets that were copied by aggregate backup for recovery at a remote site. Aggregate control tapes can contain control information, data from level 0 DASD, data from ML2 and user tape, and instruction and log data.

- ▶ **Alternate tapes**

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 provide site disaster protection, supplementing the ABARS (aggregate backup and recovery support) aggregates and protecting against media damage.

- ▶ **Fast replication target volumes**

Fast replication target volumes contain the fast replication backup copies created during DFSMSShsm fast replication processing. Fast replication target volumes are defined within copy pool backup storage groups.

Tape environment

DFSMSShsm works with all major tape management systems. DFSMSShsm provides an internal interface to IBM's tape management system, DFSMSRmm, and provides a general exit as an interface to other tape management systems.

DFSMSShsm is generally used with a global or public scratch pool. That is, tapes enter DFSMSShsm's usage by a non-specific tape mount and then when all data has been invalidated, the tape is returned to the scratch pool. DFSMSShsm is also able to work with a private scratch pool when tapes are specifically identified for the private use of DFSMSShsm.

Tapes can be protected by RACF or expiration dates and are expected to be included in a tape management product of the customer's choice.

DFSMSShsm 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. DFSMSShsm uses the capacity of the cartridges and provides ways to handle contentions for different data sets residing on the same cartridge.

DFSMSShsm provides tape duplication for migration or backup tapes by providing a second or alternate tape copy. This duplicate tape can be made by concurrent creation (duplex) or by copying when the tape becomes full (TAPECOPY).

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.

21.4 DFSMSrmm (Removable Media Manager)

DFSMSrmm is a functional component of DFSMS for managing removable media resources including automatic libraries (such as IBM's Virtual Tape Server), for the z/OS environment.

Introduction

In an enterprise, protection of tape media is a primary concern and priority. DFSMSrmm incorporates many functions, such as using the VSAM access method for containing inventory data of tape information, to provide superior media protection and availability. DFSMSrmm was designed to use KSDS VSAM and has eliminated the need for pointers within the database. Therefore, no database maintenance (for example, to correct chain or pointer errors) is required.

DFSMSrmm is a functional component of DFSMS for managing tape resources for the z/OS environment. DFSMSrmm automatically records all system tape activity and, manages traditional, as well as, system-managed tape libraries (including IBM's Virtual Tape Server).

DFSMSrmm supports removable media in many types of media libraries. In addition to a traditional tape library, a room with tapes, tape racks, and tape drives, several automated, manual and virtual tape libraries may exist. In addition to these on-site libraries, off-site storage locations, also known as vaults or stores are supported by DFSMSrmm.

Tape volume and tape data set retention requirements are processed by DFSMSrmm retention policies. These policies may be implemented by utilizing one or more techniques: Some of these are: DFSMS ACS routines and management classes, DFSMSrmm filters, or existing JCL expiration values of 99nn.

Many different types of reporting are available with the DFSMSrmm reporting feature. Sample reports are shipped with DFSMSrmm.

DFSMSrmm automatically validates volumes, ensuring that only valid scratch volumes are mounted for nonspecific mount requests, and that the right volume is mounted for a specific mount request using full 44 character data set names. This eliminates unintentionally overwriting a valid master or a volume retained for disaster recovery or vital record management.

Volume movement and retention are fully automated based on organization directed policies capable of supporting a variety of criteria including cycle, last used and catalog control. Off-site Bin assignment for volume slotting may be specified for any defined off-site location. Security protected operator controls exist allowing for manual intervention of volume movement and retention when necessary.

Conversion from other products

Shipped with DFSMSrmm are conversion utilities, for most common tape managers, that are capable of building a fully functional DFSMSrmm tape system from the existing tape manager. These utilities when used with DFSMSrmm parallel run capability allow migrations to be performed in a non-intrusive, risk-managed environment.

Product integration

DFSMSrmm is an integral part of DFSMS and is shipped as part of DFSMS with the z/OS operating system.

DFSMSrmm tape label support

DFSMSrmm supports the following tape label types:

- ▶ IBM Standard labels (SL)
- ▶ ISO/ANSI/FIPS labels (SL)
- ▶ Both IBM standard and user header or trailer labels (SUL)
- ▶ Both ISO/ANSI/FIPS and user header or trailer labels (AUL)
- ▶ No labels (NL)

Infrastructure benefits

These are the infrastructure benefits for the DFSMSrmm.

- ▶ **Integration:** DFSMSrmm is an integral part of DFSMS and is shipped as part of DFSMS with the z/OS operating system.
- ▶ **Compatibility:** DFSMSrmm is functionally compatible with existing tape management systems and runs in parallel during conversion. The expiration date formats of 99nnn and 98nnn are supported by DFSMSrmm.

Application benefits

DFSMSrmm volume retention and movement are specified interactively with ISPF panels. This allows authorized application owners to alter existing values without contacting the tape librarian.

21.4.1 Overview

DFSMS Removable Media Manager is a full function tape management system and is available as DFSMSrmm, a functional component of DFSMS.

DFSMSrmm enables you to manage your removable media as an enterprise library across systems that can share DASD. It can manage all of your tape volumes and the data sets on those volumes. It protects tape data sets from being accidentally overwritten, manages the movement of tape volumes between libraries and vaults over the life of the tape data sets, and expire and scratch tapes, all according to policies that you define. DFSMSrmm also manages other removable media that you define to it; for example, it can record the shelf location for optical disks and track their vital record status. DFSMSrmm does not yet automatically record information for optical volumes.

21.4.2 Volume management

Most of the volumes that DFSMSrmm manages in your installation are probably scratch volumes—volumes that are used repeatedly by different users. Each time a volume is used it is retained and managed by policies that you define 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. It also can manage the RACF TAPEVOL profiles for tape volumes and uncatalog data sets.

In addition to managing scratch volumes, DFSMSrmm can manage any volume you define to it, including those that you treat as foreign volumes under your existing system. DFSMSrmm supports any volume serial (VOLSER) and allows you to use a rack number that is different from the VOLSER. Therefore you can manage all volumes except those that are duplicates of volumes already defined to DFSMSrmm. You can still process volumes as foreign to DFSMSrmm by either not defining them to DFSMSrmm or using the DFSMSrmm ignore support through the EDGUX100 exit.

21.4.3 Validation

DFSMSrmm automatically validates volumes, ensuring that only valid scratch volumes are mounted for nonspecific mount requests and that the right volume is mounted for a specific mount request. This validation eliminates the unintentional overwriting of a valid master volume or a volume retained for disaster recovery or vital record management.

When a data set on a volume is opened and closed, DFSMSrmm automatically:

- ▶ Changes the volume status from scratch to master for nonspecific mount requests.
- ▶ Sets an expiration date for the volume and ensures that the maximum expiration date is not exceeded (repeated at close time).
- ▶ Records information about data sets on the volume (the data set name is recorded at open time; all other information is recorded at close time).
- ▶ Counts the number of times a volume is used.
- ▶ Counts the number of temporary and permanent errors encountered.
- ▶ Sets a security classification based on the data sets that reside on the volume.
- ▶ Prevents reading of data on a volume in scratch status when DFSMSrmm is running in protect mode.

All of the above actions occur at open time, except where noted.

21.4.4 Library and storage location management

You decide where to store your removable media according to how often the media are accessed and for what purpose they are retained. For example, you might keep volumes that are frequently accessed in an automated tape library data server, and you probably use at least one storage location or vault to retain volumes for disaster recovery and audit purposes. You might also have locations where volumes are sent for further processing. Such locations would be data centers within your company or customer and vendor sites.

DFSMSrmm manages:

- ▶ **Removable media library**, which contains all tape and optical volumes that are available for immediate use and includes the shelves where they reside. A removable media library usually includes other libraries, such as:
 - System-managed tape libraries; for example, the IBM 3494 and 3495 Automated Tape Library Dataserver models, and the manual IBM 3495 Tape Library Dataserver.
 - Non-system-managed tape libraries, or traditional tape libraries.
- ▶ **Storage locations**, which are not part of the removable media library because their volumes are not generally available for immediate use. Storage locations are typically used to store removable media that are kept for disaster recovery or vital records.

Removable media library

In the removable media library, you store your volumes in shelves, where each volume occupies a single shelf location. In the RMM TSO subcommands and DFSMSrmm Interactive System Productivity Facility (ISPF) dialog, a shelf location in the library is called a rack number. A rack number matches the volume's external label. DFSMSrmm uses the external VOLSER to assign a rack number when adding a volume, unless you specify otherwise. The rack number is the external VOLSER. Therefore both must be six alphanumeric characters. The internal VOLSER can be one to six alphanumeric characters. There are no restrictions on the format of the six character volume serial numbers you can define to DFSMSrmm.

- ▶ **System-managed tape library:** A system-managed tape library is a collection of tape volumes and tape devices, defined in the tape configuration database (TCDB).

The TCDB is an ICF catalog of type VOLCAT. It is a new SMS control data set containing tape library and tape volume records.

A system-managed tape library can be either automated or manual. You can have several automated or manual tape library dataservers; you use an installation-defined library name to define each automated tape library datserver or manual tape library datserver to the system, and DFSMSrmm treats each system-managed tape library as a separate location or destination.

An automated tape library datserver is a device consisting of robotics components, cartridge storage areas (or shelves), tape subsystems, and controlling hardware and software, together with the set of tape volumes that reside in the library and can be mounted on the library tape drives. IBM's automated tape libraries are the IBM 3494 and 3495 Automated Tape Library Dataservers.

A manual tape library datserver is a set of tape drives and the set of system-managed volumes the operator can mount on those drives. IBM's manual tape library is the IBM 3495 Tape Library Datserver Model M10.

- ▶ **Non-system-managed tape library:** A non-system-managed tape library is all the volumes, shelves, and drives not in an automated or manual tape library datserver. You might know this library as the traditional *tape library*. DFSMSrmm provides complete tape management functions for the volumes and shelves in this traditional tape library. DFSMSrmm defines volumes in a non-system-managed library as being shelf resident.

All tape media and drives that the operating system supports are supported in a non-system-managed tape library environment. Using DFSMSrmm, you can fully manage all types of tapes in a non-system-managed tape library, including 3420 reels and 3480 and 3590 cartridge system tapes.

Storage location

A storage location comprises shelf locations that you define to DFSMSrmm. A shelf location in a storage location is identified by a bin number. DFSMSrmm manages two types of storage locations: installation-defined storage locations and DFSMSrmm built-in storage locations.

You can define an unlimited number of installation-defined storage locations. You can use any eight-character alphanumeric name for a storage location and define the type or shape of the media in the location and the bin numbers that DFSMSrmm assigns to the shelf locations in the location. You can request DFSMSrmm shelf management when you want DFSMSrmm to assign a specific shelf location to a volume in the location.

You can also use the DFSMSrmm built-in storage locations, LOCAL, DISTANT, and REMOTE. Although the names of these locations imply their purpose, they do not mandate their actual location. All volumes can be in the same or separate physical locations. For example, an installation could have the LOCAL storage location onsite, as a vault in the computer room, the DISTANT storage location could be a vault in an adjacent building, and the REMOTE storage location could be a secure facility across town. An installation defined storage location of SHELF could be onsite, and used for volume movement within a library to make better use of the library. Because DFSMSrmm provides shelf management for storage locations, storage locations can be managed at the shelf location level. Optionally you can define storage locations that are not shelf-managed.

Although DFSMSrmm automatically shelf-manages built-in storage locations, you must first define the bins you want to use to DFSMSrmm. For bin numbers in built-in storage locations, the numbers are fixed in range, starting at bin number 000001.

21.4.5 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. The retention and movement policies you define to DFSMSrmm are known as vital record specifications (VRSs). You use them to indicate how long and where you want to keep data sets or volumes. You also use them to define how volumes are to be moved among the libraries that DFSMSrmm supports and the storage locations defined for vital records and disaster recovery purposes.

You can create a vital record specification chain to cause a sequence of volume moves. The first vital record specification in the chain is a data set or volume vital record specification. These specify one location in which the data set or volume being retained is to be stored. Add one name vital record specification for each additional location or retention value that applies to the volume or data set.

DFSMSrmm records the starting locations for volume when the volume is initially defined to DFSMSrmm, or when volume information is changed. This starting location is known to DFSMSrmm as a home location. Home is where volumes start from and are returned to when the identified retention and movement actions have been completed.

21.4.6 Environment

All data set, volume, and policy information is kept in the control data set (CDS). The CDS is a VSAM key-sequenced data set (KSDS) that contains all inventory information.

All DFSMSrmm interactions take place under control of the subsystem. Utilities are provided to start housekeeping functions so that you can schedule DFSMSrmm work through your production scheduling systems. One benefit of running as a subsystem is evident in the support for system-managed tape: DFSMSrmm uses a subsystem-to-subsystem interface to provide as much support as possible without involving the operator.

DFSMSrmm supplies ISPF panels as well as TSO commands. Therefore almost anything you can do online you can also do in batch through TSO command execution.

DFSMSrmm provides utilities to manage your inventory, create reports, maintain the DFSMSrmm CDS, and erase and initialize volumes.

Use the EDGHSKP utility, with the DFSMSrmm subsystem active, to run inventory management activities, which include:

- ▶ Vital record processing, to determine which data sets to retain and which volume moves are required, based on VRSs
- ▶ A trial-run of inventory management vital record processing without making any changes to the control data set or journal
- ▶ Expiration processing, to identify volumes ready to be released and returned to scratch
- ▶ Storage location management processing, to assign shelf locations to volumes being moved to storage locations
- ▶ Backing up the CDS and the journal, and clearing the journal
- ▶ Creating an extract data set for report generation

Use the EDGUTIL utility to create, update, and verify the CDS. Use the EDGBKUP utility to back up and recover the CDS and journal. EDGUTIL and EDGBKUP can be executed independently of the subsystem.

21.5 More information

- ▶ For more information on DFSMS, refer to this Web site:

<http://www.storage.ibm.com/software/sms/index.html>

- ▶ These Redbooks also provide more information:

See: <http://www.redbooks.ibm.com>

DFSMSshsm ABARS and Mainstar Solutions, SG24-5089

<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg245089.html?Open>

IBM TotalStorage Virtual Tape Server Planning, Implementing and Monitoring, SG24-2229

<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/SG242229.html>

z/OS V1R3 DFSMS Technical Guide, SG24-6569

<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/sg246569.html?Open>

Z/OS V1R3 and V1R5 DFSMS Technical Guide SG24-6979

<http://www.redbooks.ibm.com/abstracts/sg246979.html?Open>



DFSORT

DFSORT is IBM's high-performance sort, merge, copy, analysis, and reporting product. It is an optional component of the z/OS and is available for the MVS and VSE environments.

22.1 Overview

The DFSORT licensed program is a high-performance data arranger developed by IBM for z/OS (and OS/390).

DFSORT is available for MVS and VSE.

DFSORT manipulates data stored in data sets (files). For most of the processing done by DFSORT, the whole data set (file) is affected. However, some forms of DFSORT processing involve only certain individual records in that data set.

DFSORT provides tools to sort, merge, and copy data sets. It can perform simple tasks such as alphabetizing a list of names, or complex tasks such as taking inventory or running a billing system. It can also perform data management tasks using its record-level editing capability.

DFSORT is designed to optimize the efficiency and speed with which operations are completed through synergy with processor, device, and system features (for example, Hiperspace™, data space, striping, compression, extended addressing, DASD and tape device architecture, processor memory, processor cache, and so on) and other products (for example, the SAS System, COBOL, PL/I, IDCAMS BLDINDEX, and so on).

Some of the DFSORT features (common to MVS and VSE) are:

- ▶ Sort, merge or copy data sets while including, excluding or reformatting records
- ▶ Analyze data and produce detailed reports using the ICETOOL utility
- ▶ Use symbols for fields and constants in DFSORT and ICETOOL statements
- ▶ Join and match records from different data sets in various ways
- ▶ Convert FB data sets to VB data sets, and VB data sets to FB data sets
- ▶ Add timestamps to records, include or omit records based on today's date, and convert SMF date and time values to readable form
- ▶ Collate and compare data using DFSORT's National Language Support

While sorting, merging, or copying data sets, you can also:

- ▶ Select a subset of records from an input data set
- ▶ Reformat records, add or delete fields, and insert blanks, constants, or binary zeros
- ▶ Sum the values in selected records while sorting or merging (but not while copying)
- ▶ Create multiple output data sets and reports from a single pass over an input data set
- ▶ Sort, merge, include or omit records according to the collating rules defined in a selected locale
- ▶ Alter the collating sequence when sorting or merging records (but not while copying)
- ▶ Sort, merge, or copy Japanese data if the IBM Double Byte Character Set Ordering Support (DBCS Ordering) (5665-360 Licensed Program, Release 2.0 or an equivalent product) is used with DFSORT to process the records

22.1.1 Other DFSORT utilities and features

This section gives an overview of additional utilities or tools included with DFSORT. There are some differences (as indicated) between MVS and VSE, in the name or availability of specific utilities.

ICEGENER (MVS)

ICEGENER is a DFSORT feature that provides excellent performance improvements as a replacement for IEBGENER. It can be installed as a direct replacement for IEBGENER without changing existing IEBGENER jobs. ICEGENER uses DFSORT to process IEBGENER jobs when possible and transfers control to IEBGENER when DFSORT can't be used.

ICETOOL (MVS and VSE)

ICETOOL is a versatile file processing and reporting utility that provides an easy-to-use batch front-end for DFSORT/VSE. It performs multiple operations on one or more files in a single pass over the data. It uses the capabilities of DFSORT/VSE to perform the required operations, calling DFSORT/VSE for each operation with the particular DFSORT/VSE control statements and options required. It combines new features with previously available DFSORT/VSE features to perform complex sorting, copying, reporting and analytical tasks using multiple files in a single pass over the data. It provides the following features:

- ▶ Copy one or more input files to one or more output files.
- ▶ Print a message containing the count of records.
- ▶ Print the DFSORT/VSE installation defaults.
- ▶ Specify input or output file characteristics.
- ▶ Print the values or characters of specified numeric or character fields. Simple, tailored or sectioned reports can be produced.
- ▶ Set/reset scanning and error actions.
- ▶ Print each unique value for specified character or numeric fields and how many times it occurs.
- ▶ Print a message containing the count of values in a specified range for a specified numeric field.
- ▶ Select records from files for inclusion in an output file based on meeting criteria for the number of times specified field values occur.
- ▶ Sort one or more input files to one or more output files.
- ▶ Print messages containing the minimum, maximum, average, and total for specified numeric fields.
- ▶ Print a message containing the count of unique values for a specified numeric or character field.
- ▶ Examine specified decimal fields in one or more input files and prints a message identifying each invalid value found for each field.

DFSORT symbols (MVS)

A symbol is a name 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 or the value of a constant. It is a powerful, simple and flexible way to create symbol mappings for frequently used data. In addition, IBM-created symbol mappings and sample jobs can be used.

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.

OUTFIL (MVS) or control statements (VSE)

OUTFIL is a very versatile DFSORT/MVS control statement. It provides a facility to create one or more output data sets for a sort, copy or merge application from a single pass over one or more input data sets, which can result in significant performance improvements. It provides the following rich set of features that can be used with one output data set or many output data sets. You can achieve similar functions with DFSORT/VSE control statements.

- ▶ Select the records to be included in each output data set. SAVE can be used to select the records that are not selected for any other subset, eliminating the need to specify complex conditions.
- ▶ Reformat the records for output.
- ▶ Generate constants for the current date and time in a variety of character, zoned decimal and packed decimal formats, providing an easy way to add timestamps to output records.
- ▶ Generate constants for the current date and time.
- ▶ Display any field in a record in hexadecimal.
- ▶ Translate characters.
- ▶ Provides sophisticated editing capabilities for controlling how numeric fields, including SMF date and time fields and two-digit year date fields, are presented with respect to length. Twenty-seven predefined editing masks are available for commonly used numeric patterns, encompassing many of the numeric notations used throughout the world. In addition, a virtually unlimited number of numeric editing patterns are available via user-defined editing masks.
- ▶ Convert a field in one numeric format, including SMF date and time fields and two-digit year date fields, to another numeric format.
- ▶ Generate sequence numbers in the output records.
- ▶ Create many output records from each input record.
- ▶ Substitute meaningful words or phrases for cryptic values using lookup table for selection of a character or hexadecimal string for output.
- ▶ Provides short record padding (for variable-length records).
- ▶ Remove trailing bytes of the specified type from the end of variable-length records.
- ▶ convert fixed-length to variable-length output records and vice versa.
- ▶ Select a range of records to be included in each output data set.
- ▶ Split the records in rotation among the output data sets.
- ▶ Create highly detailed three-level (report, page and section) reports containing a variety of report elements.

Specific benefits of using DFSORT/VSE

DFSORT is a higher performance data arranger than Sort/Merge V2R5 for the VSE/ESA™ operating system: better performance for fixed-length record (FLR) and variable-length record (VLR) for copy and merge jobs with DASD and tape input/output, for dataspace sorting and GETVIS sorting.

Dataspace sorting

Dataspace sorting is a DFSORT/VSE capability that uses data space available with VSE/ESA systems (ESA supervisor mode) in place of 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 dataspace sorting, data space is used as an extension of main storage for storing records during a sort. Dataspace sorting reduces CPU time and elapsed time by using data space in place of DASD work files.

You can use the DSPSIZE installation or run-time option to enable dataspace sorting.

Dataspace sorting can be used with both fixed-length records (FLR) and variable-length records (VLR). If you currently use partition program area sorting for your applications, using dataspace sorting should provide significant CPU and elapsed time performance improvements.

Getvis sorting

Getvis sorting is a DFSORT/VSE capability that uses partition virtual storage for sort applications. Getvis sorting is similar to dataspace sorting except that it uses the partition GETVIS area for virtual storage instead of dataspace.

You can use the GVSIZE installation or run-time option to enable getvis sorting.

If you currently use partition program sorting for your applications, getvis sorting should provide significant CPU and elapsed time performance improvements. Also, using the partition GETVIS area exploits the available partition virtual storage above 16M virtual.

22.2 More information

- ▶ For more information, see:
<http://www.storage.ibm.com/software/sort/>



TotalStorage Expert

IBM TotalStorage Enterprise Storage Server (ESS) Expert and IBM TotalStorage Enterprise Tape Library (ETL) Expert components of Expert are packaged together, run on the same server, share a common database, and offer seamless integration with a common look and feel.

TotalStorage Expert is a tool that gives administrators powerful yet flexible storage asset, capacity, and performance management capabilities to centrally manage Enterprise Storage Servers and 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 IBM ESS disk and tape subsystems.

23.1 Overview

IBM TotalStorage Expert is an application in the IBM TotalStorage software family which helps you manage your Enterprise Storage Server (ESS) and Enterprise Tape Library (ETL) using a Web browser based user interface.

IBM TotalStorage Expert has two available features, the ESS feature which supports ESS and the ETL feature which supports Enterprise tape library products (see Chapter 8, “IBM Enterprise Tape” on page 131).

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 performance tools such as Resource Management Facility (RMF™), DFSMS Optimizer, AIX Performance Toolkit and similar host-based performance monitors. While these tools provide performance statistics from the host system's perspective, the TotalStorage Expert provides statistics from the ESS and ETL system perspective.

TotalStorage Expert gathers and presents information which provides a complete management solution for storage monitoring and administration. TotalStorage Expert helps storage administrators by increasing the productivity of storage resources.

The TotalStorage ESS Expert is packaged with the TotalStorage ETL Expert. The ETL Expert provides performance, asset, and capacity management for IBM's three ETL solutions; IBM TotalStorage Enterprise Automated Tape Library, IBM TotalStorage Virtual Tape Server, and IBM TotalStorage Peer-to-Peer Virtual Tape Server.

Both tools can run on the same server, share a common database, efficiently monitor storage resources from any location within the enterprise, and provide a similar look and feel through a Web browser user interface. Together they provide a complete solution that helps optimize the potential of IBM disk and tape subsystems (Figure 23-1).

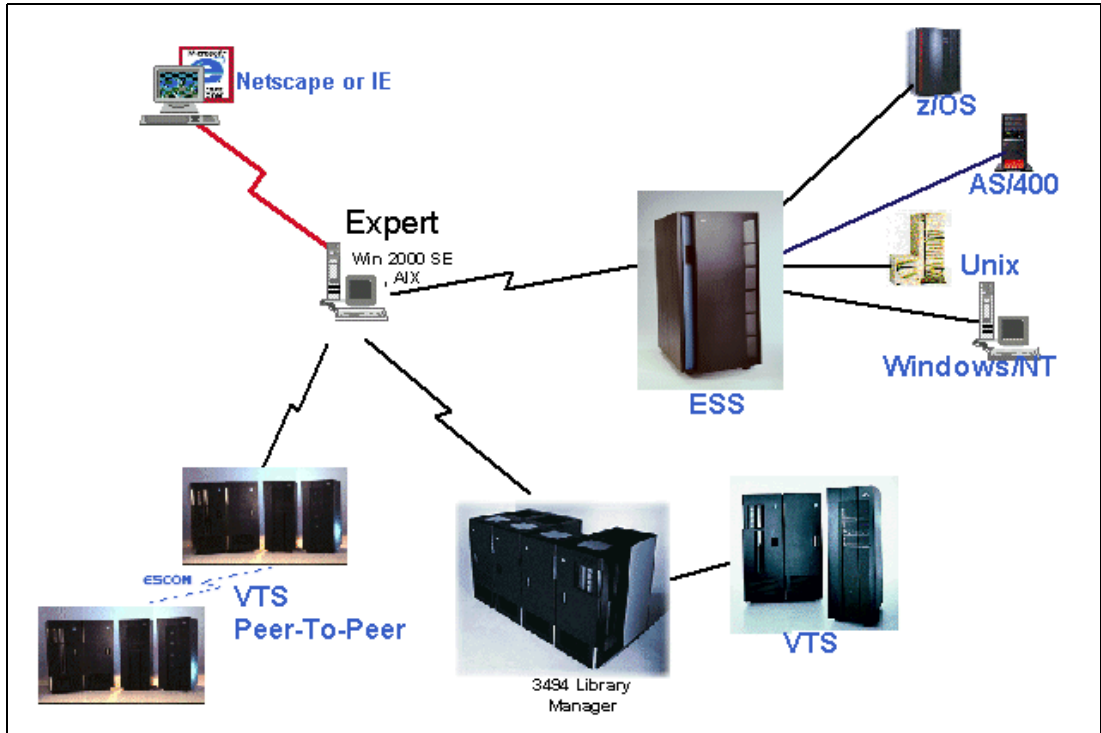


Figure 23-1 IBM TotalStorage Expert

The ESS feature and the ETL feature both provide 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.

The infrastructure benefits provided by the TotalStorage Expert include:

- ▶ Keeping track of the total capacity, assigned capacity and free capacity of the disk storage on all the ESSs in the enterprise.
- ▶ ESS and ETL components of Expert are packaged together, run on the same server, share a common (DB2) database, and offer seamless integration with a common look and feel.
- ▶ The two features can be separately licensed to accommodate users who may only be interested in one component.

The TotalStorage Expert helps you do the following tasks in addition to providing several powerful new features:

- ▶ Performance management

Capability to customize and enable threshold events relating to ESS performance metrics including disk utilization.

- ▶ Asset management

Hosts disk mapping through a Logical Unit Number (LUN), providing complete information on logical volumes including physical disks and adapters from the host perspective. Host Mapping for AIX 4.3.3 and 5L, Windows NT, Windows 2000 Server and Advanced Server, HP UX, and Sun Solaris.

- ▶ Capacity management
 - Storage capacity, including storage that is assigned to application server hosts, formatted into volumes but not yet assigned to an application server host and free space.
- ▶ Use of Simple Network Management Protocol (SNMP) alerts for ESS exception events that exceed customized threshold values.
- ▶ Support for Windows 2000 Server and Advanced Server operating systems.
- ▶ Enhanced usability characteristics:
 - A drill-down view of volumes per host, including the volumes accessible to a particular host, for each ESS
 - A drill-down view of capacity assigned to each application server and shared between application servers, as well as the names and types of each application server that can access the ESS
 - A volume drill-down capability, which provides information about a particular volume and identifies all the application server hosts that can access it.

23.2 Key features

This section summarizes the key features of the IBM TotalStorage Expert.

23.2.1 Performance management for ESS and ETL

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.

ESS performance management

By using tools which are available on the host processors, you can investigate how many I/Os are issued, or how much data is being transferred during a certain period of time. However, suppose that you need to know how well your DASD subsystems are working? In addition to hard disk drives, today's high-end DASD subsystems have many complex features, such as large cache memory and multiple host bus adapter connectivity, so they can provide broad bandwidth to satisfy performance requirements. This multiplicity of features is especially true with the ESS.

Just looking at the physical specifications of the disk drive modules (DDMs) that are loaded on your ESS will not help you analyze the ESS's performance — you may also need detailed performance data. One of the unique features of the ESS is its ability to maintain various performance statistics on a logical volume basis. For example, it records the total number of I/O requests, as well as the number of I/O requests that completed within cache memory. This is referred to as *cache hit*.

A new feature of the TotalStorage Expert provides the ability to customize and enable threshold events relating to ESS performance and utilization. In addition, default settings may be used which are derived from the IBM recommended thresholds. You can modify the Default threshold settings to create User-defined General ESS thresholds. The User-defined General ESS threshold settings are applied to all ESSs unless you set specific thresholds for an ESS (User-defined specific ESS). This feature will enable you to fine-tune your ESS storage specifically to monitor application, cache, and disk performance from a host perspective.

The following parameters can be user-defined:

- ▶ Disk utilization
 - Customize this threshold to define an utilization percentage that indicates there is a bottleneck for a disk group.
 - Customize this threshold to prevent it from applying to disks with a particular percent of sequential I/O activity. This will prevent the application from identifying instances of high disk utilization where most of the I/Os are sequential as bottlenecks.
- ▶ Cache holding time
 - Customize this threshold to monitor the time data remains in cache. Data should remain in cache at least 30 seconds. If it does not, this indicates a possible bottleneck.
- ▶ NVS cache full
 - Customize this threshold to indicate the percent of space temporarily unavailable in NVS memory.

Another new TotalStorage Expert performance monitoring feature is the Activate Simple Network Management Protocol (SNMP) function panel to configure SNMP managers. This enables their ability to receive alerts and notify you of thresholds that have been exceeded and exception events that have occurred. You also have the option to view the alert logs and filter alert messages to meet your monitoring requirements. This will provide you the administrative ability to react promptly and precisely to your enterprise performance demands.

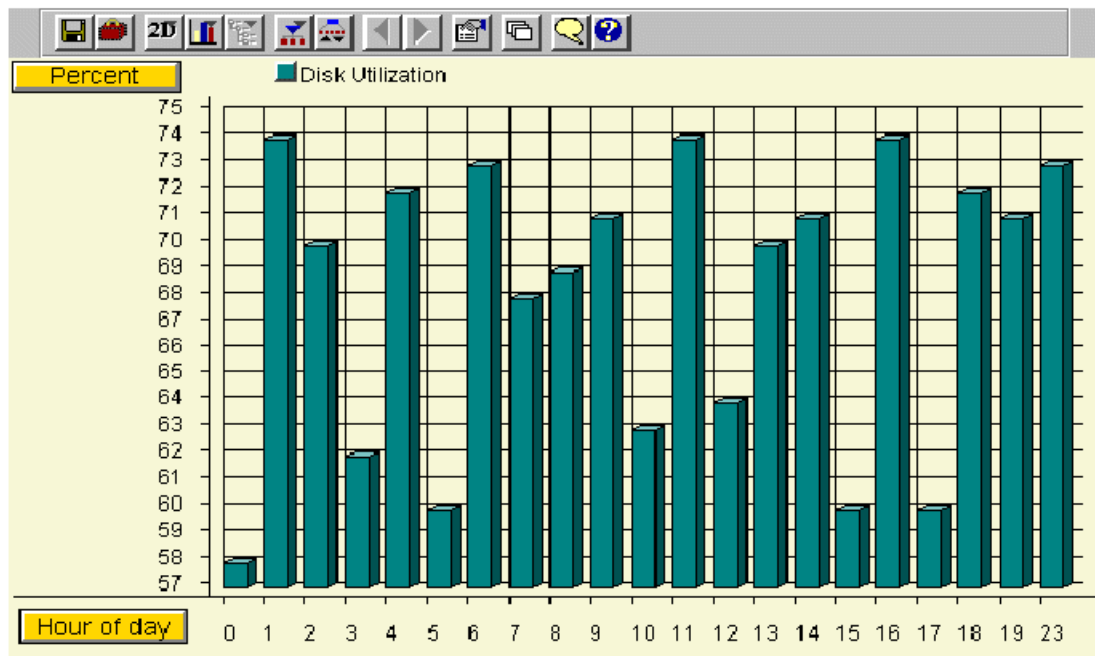


Figure 23-2 IBM TotalStorage Disk Utilization Report

ETL performance management

What about tape subsystems? The preceding description also applies to tape subsystems. The Enterprise Tape Library (ETL) is no longer a bunch of physical tape drives. This is particularly true with the Virtual Tape Server (VTS) or the Peer-to-Peer VTS. 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 3494 native library, the VTS, and the Peer-to-Peer VTS. This is especially useful for the VTS. Since the TotalStorage Expert shows you the VTS activity along with the back end native drives activity, you can correlate this information to determine how well your VTS is working and to help perform capacity planning.

Like the ESS, 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 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.

23.2.2 Asset management for your ESS and ETL

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.

In this section, we provide examples of how the TotalStorage Expert helps you manage your ESS and ETL assets.

ESS asset management

Three main features comprise the TotalStorage Asset management toolkit:

► **Storage Server Asset Inventory - Summary by Storage Server**

In a multiple ESS environment, you may need to determine host connectivity requirements for a clustered Windows 2000 implementation. You want to incorporate the benefits provided by Fibre Channel technology and restrict application access discretely within the cluster (LUN Masking). To quickly develop a picture of the future configuration, you need to know which ESS host adapters have Fibre Channel capabilities and which will need to be modified.

Within the summary report, the Access Mode feature shows the access mode of storage servers with Fibre Channel capability. The access mode identifies the current setting for restricting access from Fibre Channel attached hosts to logical volumes in the storage server. Depending on your requirements, you may use the access mode definition to allow unconfigured Fibre Channel attached hosts to access any, all, or none of the logical volumes defined in the ESS.

► **Storage Server Profile**

The Storage Server Profile feature enables you to acquire asset information on your attached ESSs. A snapshot of your discrete ESS information is provided whenever you require it for ESS code upgrades, planning, reporting, budgeting, and connectivity. Regularly scheduled or point-in-time information can be derived such as; the Server World-Wide Node-Name (of any fibre attached hosts), the most recent date of capacity data collected, server Expansion Features, Storage Cluster X (1 or 2) information including Code Level Details, This information is invaluable and is easily accessed when you require it for specific needs.

From this panel, you can also launch the IBM TotalStorage ESS Specialist to view or change configuration and add user names.

► **Storage Server Asset Inventory - Summary of LIC Code Levels**

As time goes by, you may need to upgrade the licensed internal code (LIC) for your ESS. For example, assume that you want enhanced XRC function so you can get the extensive benefits of unplanned outage support. The enhanced XRC function requires not only a certain level of software, but also requires a certain level of ESS LIC.

Therefore, you would need to check if the LIC of your ESS meets this requirement. Perhaps you have a configuration map which shows all of your assets' machine types, serial numbers, and latest LIC level. But how can you make sure that your map is up to date — what if it says *as of* six months before? The more ESSs you have, the harder it is for you to keep this information current. The TotalStorage Expert feature, *Viewing the Recent Data Report for ESS Assets*, allows current and historical comparative analysis of storage assets to be done quickly and accurately. The user-defined tasks which collect this data can be customized to accommodate the most diverse enterprise implementations.

As an administrative tool, the ESS Expert can automatically probe all of your ESS assets through your network and can provide you all the information you need for a particular purpose. Through the user-defined asset tasks, historical and current information can be retrieved and stored. This will provide the administrator and management with the capability to determine appropriate budgetary and storage allocation decisions.

ETL asset management

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.

23.2.3 Capacity management for ESS and ETL

In this section we provide an overview of capacity management considerations for the ESS and ETL.

ESS capacity management

With the LUN to host disk mapping, you can quickly ascertain resource usage and determine requirements for all of your ESS attached hosts.

You might have both z/OS CKD volumes and SCSI logical unit number (LUN) volumes on the same ESS. In the event that you need to balance storage allocation, you can administer the complete distribution of logical volumes across your ESS. This will also be much more efficient than maintaining a hardcopy version of storage allocation maps, or investigating every host processor's configuration serially.

ETL-VTS 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.

23.3 How IBM TotalStorage Expert works

Figure 23-3 on page 395 shows how the TotalStorage Expert works with your ESS and ETL. The TotalStorage Expert communicates with your ESS and ETL through the TCP/IP network. Therefore, you can gather information on any ESS and ETL around the world as long as you have a communication path between the TotalStorage Expert and the ESS and ETL through your intranet or the World Wide Web (Internet).

You will need to provide a LAN link between the ESS and the IBM TotalStorage Expert, so that the performance information, in particular, can be sent from the ESS to the IBM TotalStorage Expert. If you put the ESS on a private network, (for example, one that has only the ESS and the ESSNET configuration console), then you will need to provide a communication capability for the TotalStorage Expert with that private network, or attach the private network to the your existing intranet through a LAN router which does address translation.

Note: TotalStorage Expert requires more memory than is available on the ESSNET configuration console, so that you will need to install it on a separate workstation or server. 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 interface. In other words, you can operate the ESS Expert through Netscape Navigator or Microsoft Internet Explorer from any machine which has network access to the machine where the ESS Expert is installed.

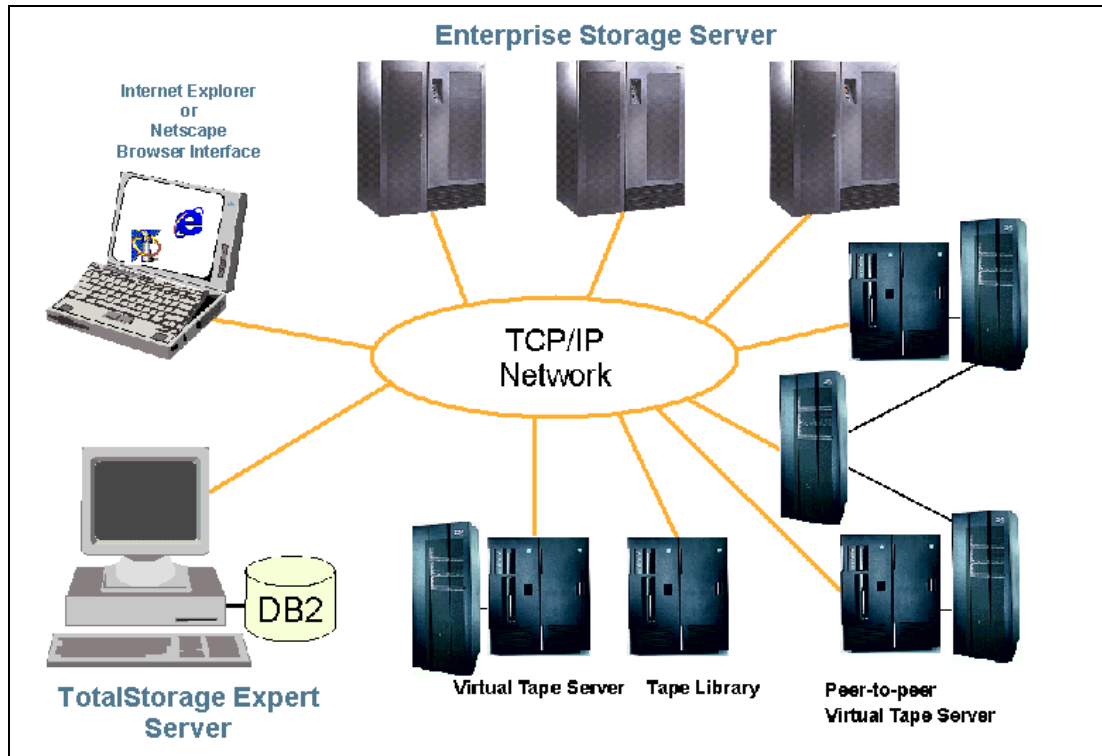


Figure 23-3 IBM TotalStorage Expert sample configuration

The TotalStorage Expert solicits your ESS and ETL 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 Version 2.1 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.

23.4 Planning and implementation considerations

This section provides information on planning and implementing your TotalStorage Expert environment including network, operating system co-requisite software requirements.

Network planning

Since the TotalStorage Expert communicates with your ESS and ETL through TCP/IP network, you need to carefully plan how the TotalStorage Expert is going to attach to your network.

You will need to provide a LAN link between the ESS and the IBM TotalStorage Expert, of course, so that the information can be sent from the ESS to the IBM TotalStorage Expert. If you put the ESS on a private network, (for example, one that has only the ESS and the ESSNET configuration console), then you will need to provide a communication capability for IBM TotalStorage Expert with that private network, or attach the private network to the your existing intranet via a LAN router (see Figure 23-4 on page 396) which does address

translation. IBM TotalStorage Expert requires more memory than is available on the ESSNET configuration console, so that you will need to install it on a separate workstation or server.

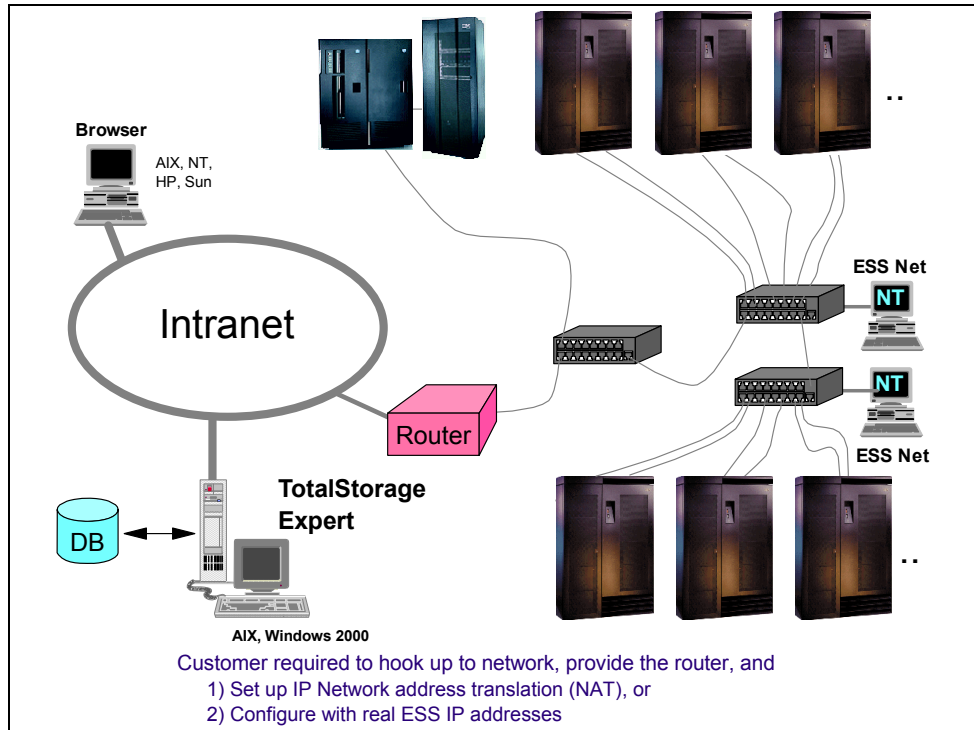


Figure 23-4 TotalStorage Expert with ESS or ETL on a private network

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 Version 1 Release 2:

- ▶ 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.

23.5 More information

For more information on 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 StorWatch Expert Hands-On Usage Guide*, SG24-6102
<http://www.redbooks.ibm.com/redbooks/pdfs/sg246102.pdf>
- ▶ *IBM TotalStorage Expert Reporting: How to Produce Built-In and Customized Reports*, SG24-7016
<http://www.redbooks.ibm.com/redbooks/pdfs/sg247016.pdf>



Part 6

Storage solutions

Part 6 presents some of the storage solutions that IBM offers using and combining storage products that we with have described in previous parts of this redbook:

- ▶ Data protection solutions that are important for business continuance of companies and organizations
- ▶ Disaster Recovery solutions that provide continuance to the companies
- ▶ Solutions for e-mails that are the hearth of the communication in the companies
- ▶ Solutions to keep the all data in one place, centralization
- ▶ Solutions in the entertainment environment, where the data increases exponentially because of high capacity visual and audible data
- ▶ Solutions in life sciences, where laboratories need high storage capacity, such as medical imaging



IBM TotalStorage DR550

This chapter discusses the IBM TotalStorage DR550 (DR550) 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 56TB of disk storage. 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 DR550 is built with IBM's latest server technology (POWER5 processors), storage technology (IBM TotalStorage DS4100 subsystems), and data protection technology (Tivoli Storage Manager for Data Retention).

24.1 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.

24.1.1 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 life cycle:** 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.

24.1.2 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's unique experience 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 the IBM Tivoli Storage Manager, IBM DB2 Content Manager, along with any needed disk and tape-based storage.

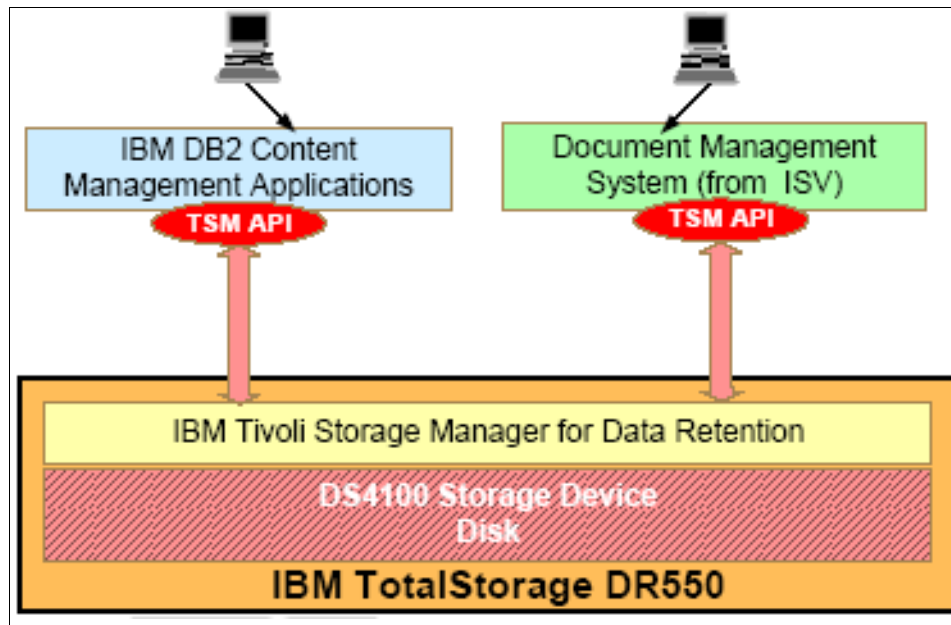


Figure 24-1 Representation of components of the DR550

The main focus of the IBM TotalStorage 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 TotalStorage 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.

24.2 Positioning of the DR550

The IBM 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 (the IBM Tivoli Storage Manager in particular) 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 Tivoli Storage Manager API.

24.3 DR550 overview

The DR550 brings together off-the-shelf IBM hardware and software products.

24.3.1 Hardware

Note: The hardware comes pre-mounted in a secure rack; the software is pre-installed and to a large extent pre-configured. There is very little the customer has to do to deploy this offering.

The 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 (1-way), one IBM TotalStorage DS4100 Storage Server with 3.5 TB of raw disk capacity, one IBM TotalStorage SAN Switch H08 that connect the p5 520 servers to the storage, and a convenient Hardware Management Console, everything installed in a 7014 rack. The second single node configuration comes with one additional EXP100 expansion enclosure, increasing the total disk storage capacity to 7 TB.

Dual node configurations consist of two IBM @server p5 520 (1-way) and two Fibre Channel switches 2005-H08 and one or two DS4100 Storage Servers. Depending on the number of installed EXP100 expansion enclosures, the dual node configurations scale from 3.5 TB over 7 TB, 14 TB and 28 TB up to a maximum of 56 TB total disk storage capacity. The 56 TB configuration features a second 7014 rack to accommodate the second DS4100 Storage Server and the additional seven EXP100 expansion units.

The customer also has the choice between a so-called disk-only version and a disk-and-tape-ready version, that provides additional hardware required to attach tape devices. IBM recommends the IBM TotalStorage 3592 Enterprise Tape Drive in combination with the IBM TotalStorage 3592 WORM media, to extend the DR550 characteristics for non-erasable and non-rewritable data also to the tape storage pool.

24.3.2 Software

The software bundle includes the IBM AIX v5.2 operating system, HACMP cluster management software (dual node only), the IBM Tivoli Storage Manager for Data Retention, and the IBM TotalStorage 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 on the DR550 is via the IBM Tivoli Storage Manager Client API. As such, document management applications need to be capable of communicating with this API in order to archive data on the DR550.

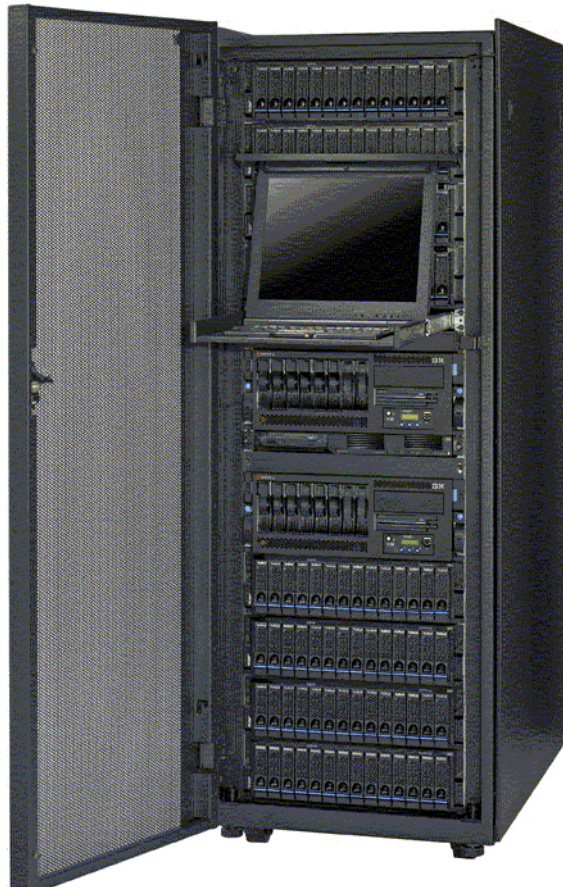


Figure 24-2 DR550 with management console

Tape device attachment

While the DR550 does not come with a tape drive in its standard configuration, the IBM Tivoli Storage 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 TSM database

Like the actual data, the TSM database needs protection from the very same scenarios described above.

The *Understanding the IBM TotalStorage DR550*, SG24-7091 documents two examples showing how to attach and use tape in the context of the DR550 archiving solution. The first example explains the basics of attaching a 3494 Enterprise Tape Library with 3592 Enterprise Tape Drive technology. The second example, based on a 3584 IBM UltraScalable Tape Library with 3592 Enterprise Tape Drive technology, and using 3592 WORM (write once read many) media.

24.4 Data retention features

The key software component of the DR550 is an IBM Tivoli Storage Manager for Data Retention server. The IBM Tivoli Storage Manager for Data Retention is what gives the DR550 its unique data retention features.

Note: The IBM TotalStorage DR550 is an archive solution focused on providing a data retention repository for applications that utilize the Tivoli API to archive data to the IBM Tivoli Storage Manager for Data Retention server.

It is beyond the scope of this document to explain IBM Tivoli Storage Manager in every detail. For a detailed overview of IBM Tivoli Storage Manager and its complementary products, refer to *IBM Tivoli Storage Manager Administrator's Guide Version 5.2*, GC32-0768-02, *IBM Tivoli Storage Manager Concepts*, SG24-4877, and *IBM Tivoli Storage Manager Implementation Guide*, SG24-5416 and Chapter 20, "IBM Tivoli Storage Manager" on page 315.

In this section we discuss the IBM Tivoli Storage 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 Tivoli Storage 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 Tivoli Storage 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 version 5.2.2 to make it possible for applications utilizing the API to further control the retention period (RETVER) for archive objects.

24.4.1 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.

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 24-3 shows a simplified view of a chronological retention policy. With `RETINIT=creation` and `RETVER=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 Tivoli Storage 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 (not enabled for Data Retention protection) through the normal backup/archive client (not applicable for DR550).

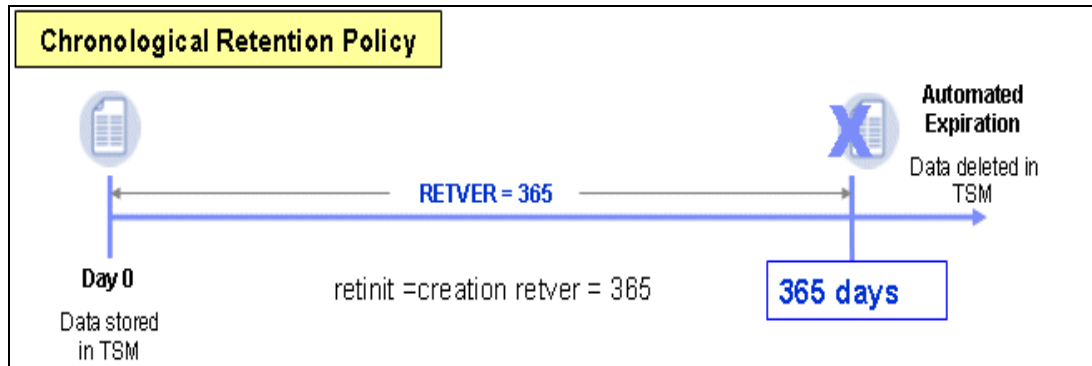


Figure 24-3 Chronological retention policy

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 24-4 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 (RETVER) is activated through an API retention event, IBM Tivoli Storage Manager assigns an expiration date for this object. The expiration date that IBM Tivoli Storage 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 RETVER 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 Tivoli Storage Manager database making the data irretrievable from IBM Tivoli Storage 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.

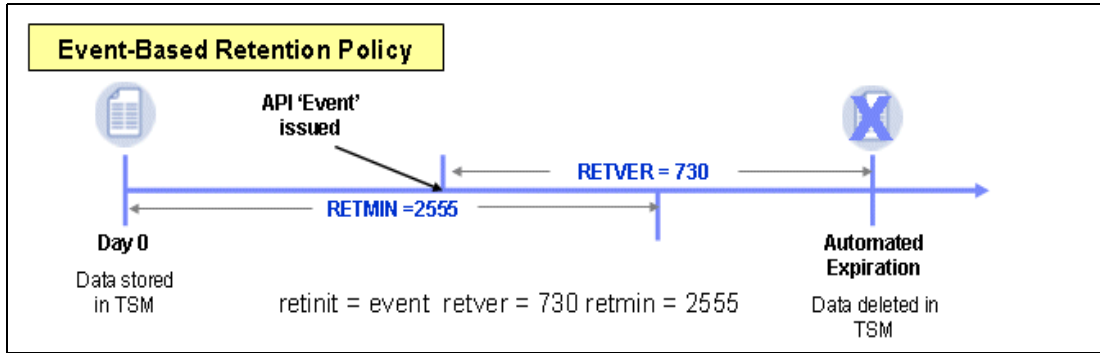


Figure 24-4 Event based retention policy

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 24-5 shows a timeline depicting deletion hold and release.

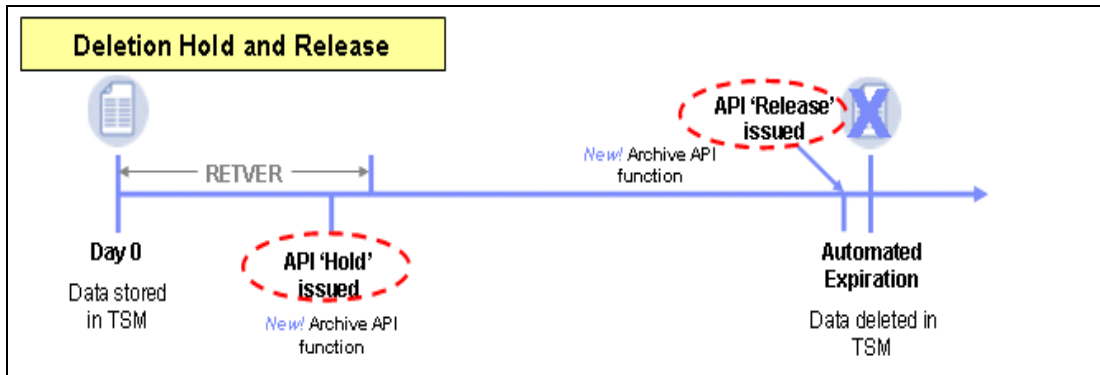


Figure 24-5 Deletion hold and release

24.5 Document Management Systems

Any Document Management System (DMS) that can use the archive and retention functions offered through the Tivoli Storage 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 Tivoli Storage Manager API Client may be found at:

<http://publib.boulder.ibm.com/tividd/td/IBMTivoliStorageManagerClient5.2.2.html>

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. These vendors include:

- ▶ Ceyoniq
- ▶ Compendium
- ▶ Easy Software
- ▶ Hummingbird® Ltd
- ▶ Hyperwave
- ▶ iLumin Software Services
- ▶ IRIS Software
- ▶ IXOS Software
- ▶ KVS, Inc. (a business unit of VERITAS software)
- ▶ Neeb and Partner
- ▶ OuterBay Technologies
- ▶ Princeton Softech
- ▶ SAPERION
- ▶ SER Solutions
- ▶ Waters

Additional information on qualified ISVs may be found at:

<http://www.ibm.com/storage/dr550>

You can also find information on ISV support for the IBM Tivoli Storage Manager API on the Tivoli Web pages:

<http://www.ibm.com/software/tivoli/partners/public.jsp?tab=connect&content=rft-validation-tsm>

Additional information on IBM DB2 Content Manager can be found in *Understanding the IBM TotalStorage DR550*, SG24-7091.

24.6 More information

For a more detailed understanding of the DR550 refer to these sources:

- ▶ *Understanding the IBM TotalStorage DR550*, SG24-7091
- ▶ *IBM TotalStorage DR550 Installation, Setup, Operations, and Problem Determination Guide*:

<http://www-1.ibm.com/servers/storage/disk/dr/pdf/DR550-Users-Guide.pdf>

- ▶ IBM TotalStorage DR550 Web page:

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



Business Continuity solutions

In this chapter, we discuss the concepts and the solutions for Business Continuity as it relates to IBM TotalStorage. In particular, we discuss:

- ▶ Business Continuity Planning overview and terminology
- ▶ Business Continuity Planning processes and procedures
- ▶ Examples of solutions

25.1 Evolving definitions of disaster and recovery

Before we proceed, we need to clarify the terms Business Continuity and Disaster Recovery. These terms are sometimes used interchangeably, as are the similar terms Business Resumption and Contingency Planning.

Business Continuity can be defined as the ability of an organization to ensure continuity of service and support for its customers and to maintain its viability before, during and after an event. Each business or industry is unique. A total Business Continuity Plan has a broad focus and includes items, such as a crisis management plan, business impact analysis, human resources management, business recovery plan procedure, documentation and so forth.

Disaster Recovery can be defined as the processes that are required to maintain Business Continuity when a catastrophic outage occurs. Disaster recovery plan refers to an *IT-focused* plan. In practical terms this means Disaster Recovery is the ability to recover a datacenter at a different site if a disaster destroys the primary site; it is only one component of an overall Business Continuity plan.

The forces that are currently impacting the recovery strategies that organizations must implement include:

- ▶ The exponential growth in the amount of data that must be stored has created the need for a comprehensive program to manage that growth.
- ▶ The rate of technological change complicates the recovery planning process. Rapid change and adjustments further complicate staying online or being able to come back online.
- ▶ Closely related to the above point, the mixture of disparate platforms, operating systems, and communication protocols found within most organizations intensifies the already complex task of recovering/preserving data.
- ▶ Customers, suppliers, management, and others expect that data be available at anytime from anywhere.
- ▶ The concept of what is a threat has also changed and now must include: terrorism; vandalism from outside as well as from within the organization.
- ▶ New regulatory requirements have forced many to rethink their approach to data survivability.
- ▶ Finally, globalization of a firm creates new opportunities for some firms, but it can also create additional recovery and availability concerns that did not exist in the centralized computing model.

Historically, Disaster Recovery referred to the recovery of an organization's central processing facility. Recovery efforts were almost always triggered by a fire, flood, storm, or other physical devastation. These hazards still exist, of course; however, a disaster in today's frame of reference can mean the corruption of files or any event that makes the organization's data *unavailable, even for relatively short periods of time*. Business' tolerance level for a system outage is getting leaner and leaner. Data disruptions for some industries, such as financial institutions, can result in staggering financial losses. Inability to cope or plan for such disruptions normally cause businesses huge amounts of financial loss.

The methods available to assist recovery efforts are further hampered by application architectures such as distributed applications, distributed processing, distributed data, and hybrid computing environments.

25.2 Types of contingency plans

In this section, we discuss different types of Contingency Planning and what the objectives of these plans should be.

25.2.1 Business Continuity Plan (BCP)

The *Business Continuity Plan* confronts the likelihood of a disaster, how the disaster interrupts the business process, and how the business can continue in operation.

Business Continuity Planning relates to considerations that affect the whole of the business process. The BCP focuses on sustaining an organization's business functions during and after a disruption. An example of a business function may be an organization's payroll process or customer support process. A BCP may be written for a specific business process or may address all key business processes. Information systems are considered in the BCP only in terms of their support to the larger business processes. In some cases, the BCP may not address long-term recovery of processes and return to normal operations, but only cover interim Business Continuity requirements.

25.2.2 Business Recovery Plan (BRP)

The Business Recovery Plan (BRP) addresses the restoration of business processes after an emergency. The BRP is similar to the BCP, but unlike that plan, the BRP typically lacks procedures to ensure continuity of critical processes throughout an emergency or disruption.

25.2.3 Continuity of Operations Plan (COOP)

The COOP focuses on restoring an organization's essential functions at an alternate site and performing those functions for an extended period of time before returning to normal operations. This term traditionally is used by the Federal Government and its supporting agencies to describe activities otherwise known as Disaster Recovery.

25.2.4 Incident Response Plan (IRP)

The Incident Response Plan establishes procedures to address cyber attacks against an organization's IT server and workstation systems. These procedures are designed to enable security personnel to identify, mitigate, and recover from malicious computer incidents, such as unauthorized access to a system or data, denial of service, or unauthorized changes to system hardware or a virus.

25.2.5 Occupant Emergency Plan (OEP)

The OEP provides the Response Procedures for occupants of a facility in the event of a situation posing a potential threat to the health and safety of personnel, the environment, or property. Such events would include a fire, hurricane, criminal attack, or a medical emergency. OEPs are developed at the facility level, specific to the geographic location and structural design of the building.

25.2.6 Disaster recovery plan (DRP)

As suggested by its name, the disaster recovery plan applies to major events that deny access to the normal facility for an extended period. The disaster recovery plan scope may overlap that of an IT Contingency Plan; however, the disaster recovery plan is narrower in scope and does not address minor disruptions that do not require relocation such as a denial-of-service attack. The Figure 25-1 depicts how the various plans relate to each other.

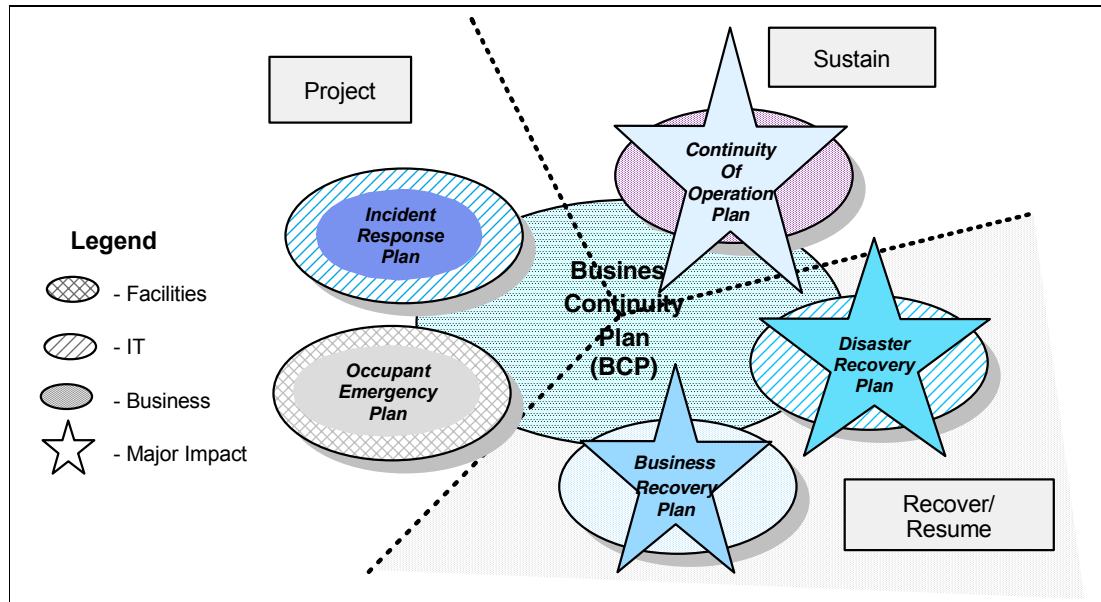


Figure 25-1 Interrelationship of emergency preparedness plans

25.3 Disaster recovery planning phases

Disaster recovery planning is typically undertaken as a multi phase project. Figure 25-2 on page 415 is a diagram that illustrates the phases of a typical disaster recovery planning project.

The central *data storage* presented receives the information from the Collection Phase (and from subsequent testing and change management activity) creating centralized data storage that can be used in the other phases of planning.

Here is a brief description of each phase:

1. Project Initiation and Team Selection

Project Initiation involves the formalization of management approval, the selection of a project coordinator and team selection, the standardization of information collection methods, and the scheduling of resources for the project.

2. Data Collection and Critical Needs

The Collection Phase entails the collection of information about business processes, their technology infrastructure supports, potential outage costs, disaster exposures, and techniques/strategies used by other companies to mitigate risk.

3. Risk Analysis

The Risk Analysis Phase entails the processing of collected data to arrive at disaster recovery planning targets and objectives. Basically, the outcome of analysis is an understanding of what is at risk and what resources are required for recovery within an acceptable time frame.

4. Data Protection

Data protection is a key component of disaster recovery planning. Measures to classify and safeguard information assets (including records, microfilm, and electronic and optical data) need to be formalized and articulated.

5. Recovery Plans

The Recovery Plans Phase involves the formulation of strategies for replacing systems, networks, and end users in the wake of an unplanned interruption event.

6. Training and Testing

Training And Plan Testing validate the disaster recovery strategies that have been developed and provide a vehicle for identifying plan deficits so they can be resolved.

7. Change Management

Change management provides a mechanism for keeping the plan up-to-date with changing business and IT parameters. Change management provides feedback that is used to keep the disaster recovery capability current at all times.

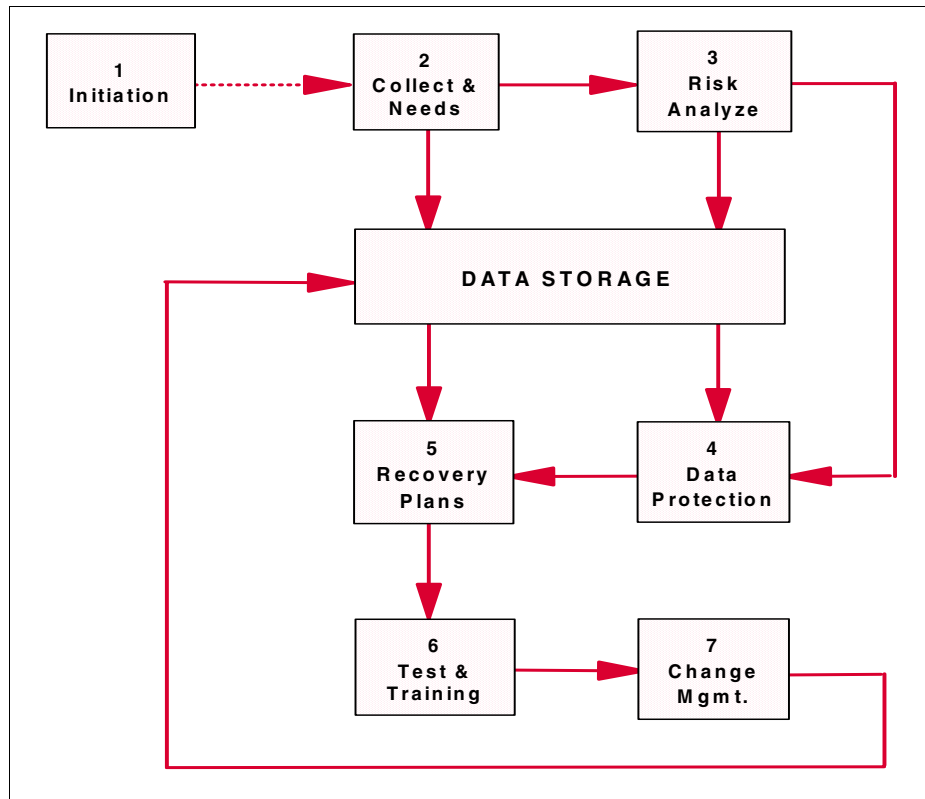


Figure 25-2 Diagram of disaster recovery planning phases

25.4 Project phases in detail

In this section we introduce some details about each phase. We are covering all phases as separate topics, but the discussion is from the entire project point of view.

25.4.1 Phase 1 - Project initiation and team selection

Following are the major aspects of this project initiation and team selection phase.

► **Top management commitment**

Top management must support and be involved in developing, coordinating, and ensuring the disaster recovery effectiveness within the organization.

► **Establish a planning committee**

A planning committee should be appointed to oversee the development and implementation of the plan. The planning committee should include representatives from all functional areas of the organization. Key committee members should include the operations manager and the data processing manager. The committee also should define the scope of the plan.

► **Scope**

The plan should have a broad scope if it is to effectively address the many disaster scenarios that could affect the organization.

► **Assumptions**

A worst-case scenario should be the basis for developing the plan. The worst-case scenario is the destruction of the main or primary facility.

The assumptions can often be identified by asking the following questions:

- What equipment/facilities have been destroyed?
- What is the timing of the disruption?
- What records, files, and materials were protected from destruction?
- What resources are available following the disaster:

► **Teams and responsibilities**

The disaster recovery planning should be structured using a team approach. Personnel should be chosen to staff these teams based on their skills and leadership. Ideally, teams would be staffed with the personnel responsible for the same or similar operation under normal conditions. Team members must understand not only the contingency plan purpose, but also the procedures necessary for executing the recovery strategy. Teams should be sufficient in size to remain viable even if some members are unavailable. Each team is led by a leader. The most important team in disaster recovery planning is the Management Team, which is necessary for providing overall guidance following a major system disruption or emergency.

25.4.2 Phase 2 - Data collection and critical needs

To determine the critical needs of the organization, each department should document all the functions performed within that department.

An analysis over a period of two weeks to one month can indicate the principle functions performed inside and outside the department, and assist in identifying the necessary data requirements for the department to conduct its daily operations satisfactorily.

On top of that, there are some recommended data gathering materials and documentation to be included: backup position listing, critical telephone numbers, communications inventory, distribution register, documentation inventory, equipment inventory, forms inventory, insurance policy inventory, main computer hardware inventory, master call list, master vendor list, microcomputer hardware and software inventory, notification checklist, office supply inventory, off-site storage location inventory, software and data files backup/retention schedules, telephone inventory, temporary location specifications, other materials, and documentation.

Important: Each business unit should declare the maximum time of recovery of their application. There are two pieces of information that should be delivered: The *Recovery Time Objective (RTO)*, how long should it take to recover to the point for an end user to be online again; and the *Recovery Point Objective (RPO)*, how much (or how little) data the user can afford to recreate after the disaster. Usually department personnel are biased in setting higher than needed figures for recovery. The other key piece of information is the *Network Recovery Objective (NRO)*. This is the time to recover the network and move the users to the disaster site. There is no point in recovering the applications in 10 minutes if it takes several hours to connect the users to the new site.

Once the critical needs have been documented, management can set priorities within departments for the overall recovery of the organization:

- ▶ Essential activities (vital):
- ▶ Recommended activities (critical):
- ▶ Nonessential activities (non-critical)

Although the classification for RTO and RPO will vary from company to company, some general guidelines can be shown to help qualifying the application:

- ▶ Vital: RTO < 8 hours after the outage occurred and RPO within 15 minutes of the time of outage
- ▶ Critical: RTO < 72 hours after the outage occurred and RPO start of the day of the outage
- ▶ Non-critical: RTO < 168 hours (1 week) after the outage occurred and RPO within 48 hours of the time of the outage

25.4.3 Phase 3 - Risk analysis

The planning teams should prepare a risk management process and at the same time the Business Impact Analysis (BIA) that includes a range of possible disasters, including natural, technical, and human threats.

Each functional area of the organization should be analyzed to determine the potential consequence and impact associated with several disaster scenarios. The Risk Analysis Phase should also evaluate the safety of critical documents and vital records.

Risk management process

Risk management encompasses a broad range of activities to identify, control, and mitigate risks to an IT system. Risk management activities may be considered to have two primary functions:

- ▶ Prevent or reduce the likelihood of damaging incidents by reducing or eliminating risks.
- ▶ Encompass actions to reduce or limit the consequences of threats in the event that they successfully disrupt a system.

Figure 25-3 illustrates the relationship between preemptive security controls and post-event disaster recovery plan implementation.



Figure 25-3 Disaster recovery plan as an element of risk management implementation

Risks result from a variety of factors, although typically they are classified into three types:

- ▶ Natural: Hurricane, tornado, flood, and fire
- ▶ Human: Operator error, sabotage, and malicious code
- ▶ Technological: Equipment failure, software error, telecommunications network outage, and electric power failure

Not all risks are present with respect to a given IT department (or site). A thorough risk assessment should identify all potential risks and attempt to determine the probability of the risk actually occurring.

Ideally, all identified risks would be eliminated completely. However, rarely is this possible or cost-effective. Rather, an attempt is made to reduce risks to an acceptable level and remain aware of residual risks. The scope of the Contingency Plan may be reduced to address only these residual risks. As a result, the Contingency Plan can be more narrowly focused, conserving company resources while ensuring an effective system recovery capability.

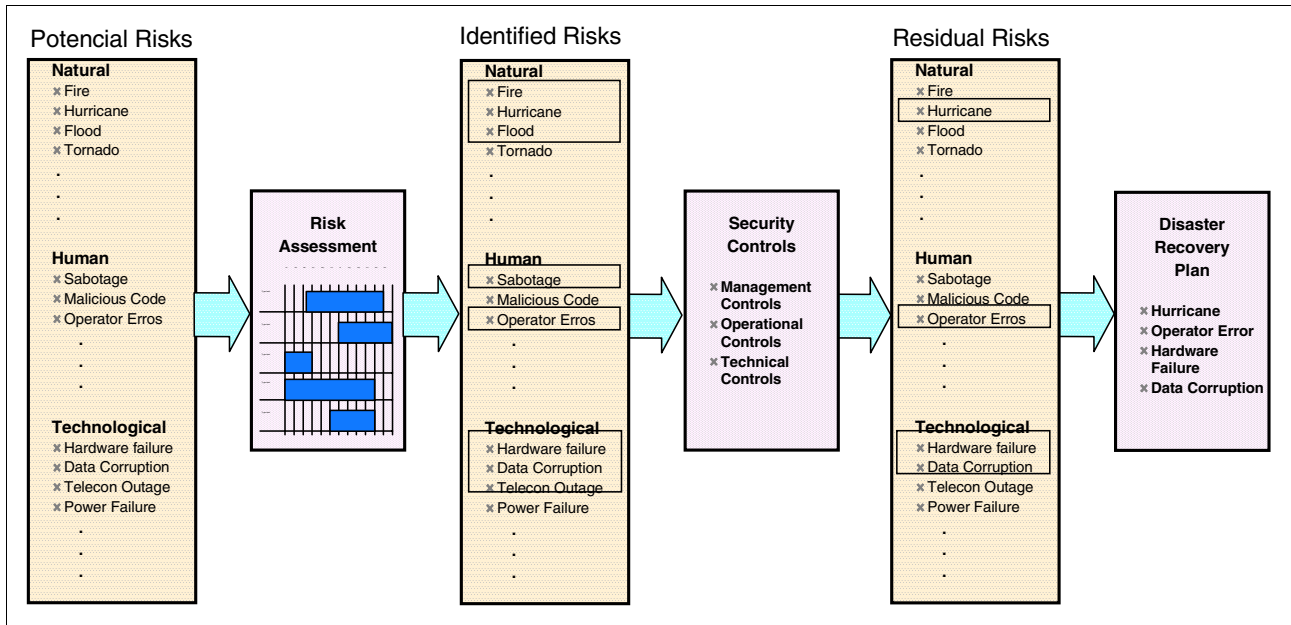


Figure 25-4 Critical risk assessment - disaster recovery plan relationship

The disaster recovery plan should also analyze the costs related to minimizing the potential exposures.

Business Impact Analysis (BIA)

The BIA enables the disaster recovery group to characterize fully the system requirements, processes, and interdependencies, and use this information to determine contingency requirements and priorities.

The sample BIA process, illustrated in Figure 25-5, helps the disaster recovery group streamline and focus their plan development activities to achieve a more effective plan:

1. Identify critical resources.

The first BIA step evaluates the IT system to determine the critical functions performed by the system and to identify the specific system resources (hardware, network, software, data) required to perform them.

2. Identify disruption impacts and allowable outage times.

In this step you should analyze the critical resources identified in the previous step and determine the impact(s) on IT operations if a given resource were disrupted or damaged.

3. Develop recovery priorities.

The outage impact(s) and allowable outage times, characterized in the previous step, allow recovery strategies to be prioritized and developed to be implemented during disaster recovery plan activation. For example, if the outage impacts step determines that the system must be recovered within four hours, the disaster recovery plan will adopt measures to meet that need.

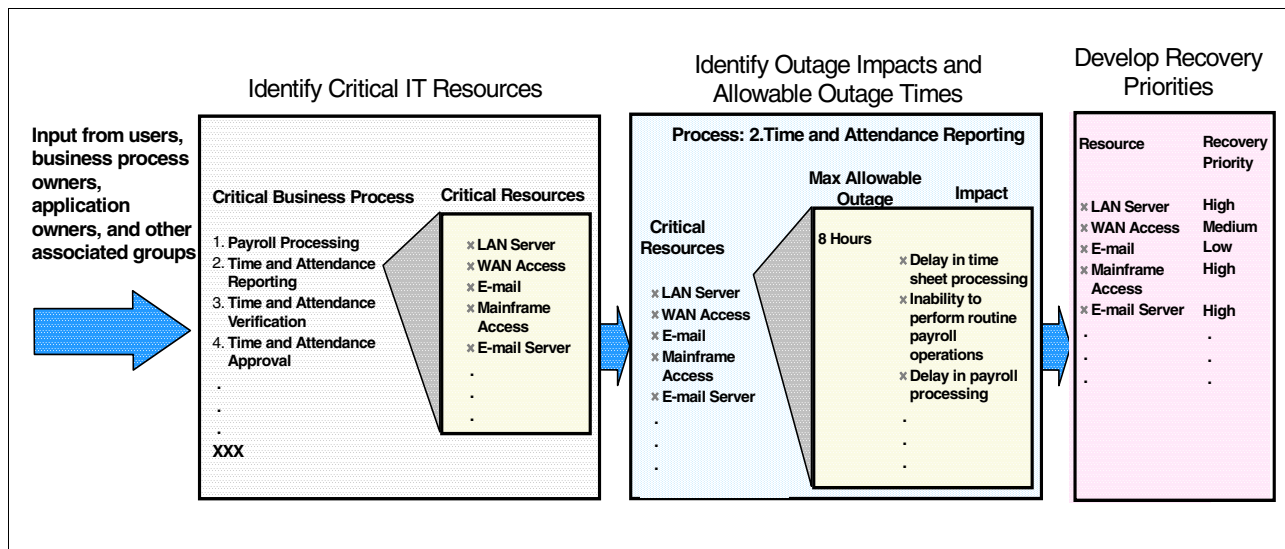


Figure 25-5 Example of a business impact analysis process

Creating an avoidance system

Because a goal of Business Recovery Planning is to ensure the safety of personnel and assets during and following a disaster, a critical aspect of the risk analysis process is to identify the preparedness and preventive measures in place at any point-in-time. Once the potential areas of high exposure to the organization are identified, additional preventative measures can be considered for implementation.

Disaster prevention techniques include two categories: Procedural prevention and physical prevention.

► **Procedural prevention**

Procedural prevention relates to activities performed on a day-to-day, month-to-month, or annual basis, relating to security and recovery. Procedural prevention begins with assigning responsibility for overall security of the organization to an individual with adequate competence and authority to meet the challenges.

► **Physical prevention**

Preparedness for disaster begins when a site is constructed. It includes special requirements for building construction, as well as fire protection for various equipment components.

25.4.4 Phase 4 - Data protection

Data Protection refers to all the measures adopted within an organization to safeguard assets, ensure the accuracy and reliability of records, and encourage operational efficiency and adherence to prescribed procedures.

Certain Control Procedures are necessary for a proper execution of other Control Procedures. This interdependence of Control Procedures may be significant because certain control objectives that appear to have been achieved may, in fact, not have been achieved because of weaknesses in other Control Procedures upon which they depend.

Security is an increasing concern because computer systems are increasingly complex. Particular security concerns result from the proliferation of PCs, local area networking, and online systems that allow more access to the mainframe and departmental computers.

Computer internal controls are especially important because computer processing can circumvent traditional security and control techniques. Important areas of concern related to general computer internal controls include: organization controls, systems development and maintenance controls, documentation controls, access controls, data and procedural controls, physical security, password security systems, and communications security.

Insurance coverage

Adequate insurance coverage is a key consideration when developing a Business Recovery Plan and performing a risk analysis. Having a Disaster Recovery Plan and testing it regularly may not, in itself, lower insurance rates in all circumstances. However, a good plan can reduce risks and address many concerns of the underwriter, in addition to affecting the cost or availability of the insurance.

Types of insurance coverages to be considered include: computer hardware replacement, extra expense coverage, business interruption coverage, valuable paper and records coverage, errors and omissions coverage, fidelity coverage, and media transportation coverage.

Records classification

Critical records can be obtained or reproduced at considerable expense and only after considerable delay. Non-critical records would cause inconvenience if lost, but can be replaced without considerable expense.

Vital and critical records should be duplicated and stored in an area protected from fire or its effects.

Records kept in the computer room should be minimized and should be stored in closed metal files or cabinets. Records stored outside the computer room should be in fire-resistant file cabinets with a fire resistance of at least two hours.

Records should not be retained only as proof of financial transactions, but also to verify compliance with legal and statutory requirements. In addition, businesses must satisfy retention requirements as an organization and employer. These records are used for independent examination and verification of sound business practices.

Federal and state requirements for record retention must be analyzed. Each organization should have its' legal counsel approve its own retention schedule. As well as retaining records, the organization should be aware of the specific Record Salvage Procedures to follow for different types of media after a disaster.

25.4.5 Phase 5 - Recovery Plan

The Recovery Plan phase deals with recovery strategies that provide a mean to restore IT operations quickly and effectively following a service disruption. The strategies should address residual risks identified in the Business Impact Analysis (BIA). Several alternatives should be considered when developing the strategy, including cost, allowable outage time, security, and integration with larger, organization-level disaster recovery plans.

The recovery strategy selected should address the potential impacts identified in the BIA and should be integrated into the system architecture during the Design and Implementation Phases of the system life cycle.

For this phase automation is a key feature. It is too demanding for personnel staff to be able to execute all the needed functions at the verge of a disaster. However, all the automation policies and people interaction must be rehearsed to exhaustion to reach success.

The topics that can be considered when developing the Recovery Plans are:

- ▶ Systems recovery

A Systems Recovery Plan should address mission-critical application hosts, both centralized and distributed.

- ▶ Network recovery

Network Recovery Plans are developed to provide for the restoration of:

- Internal LAN and peripheral network supports for mission-critical business processes
- External WANs and telecommunications services
- Communications between recovered systems and end users

- ▶ Activation of disaster recovery teams

The disaster recovery Management Team must coordinate all activities of all emergency recovery teams. Quick and valid decisions are key in an exceptional situation. A successful DR plan will take this into account to ensure that the business can be recovered and maintained even if key people are not available.

- ▶ End user recovery

End User Recovery Plans, often ignored in traditional disaster recovery planning, consist of strategies to provide end users with a mechanism for interacting with restored systems and networks to perform useful work.

Alternative sites

Although major disruptions with long-term effects may be rare, they should be accounted for in the disaster recovery plan. Thus, the plan must include a strategy to recover and perform system operations at an alternate facility for an extended period. In general, three types of alternate sites are available:

- ▶ Dedicated site owned or operated by the company
- ▶ Reciprocal agreement with an internal or external entity
- ▶ Commercially leased facility

Regardless of the type of alternate site chosen, the facility must be able to support system operations as defined in the Contingency Plan. The three alternate site types also may be categorized in terms of their operational readiness. Based on this factor, sites may be classified as cold sites, warm sites, hot sites (see Figure 25-6), mobile sites, and mirrored sites. Progressing from basic to advanced, the sites are described below.

Cold sites

Cold sites typically consist of a facility with adequate space and infrastructure (electric power, telecommunications connections, and environmental controls) to support the IT system. The space may have raised floors and other attributes suited for IT operations. The site does not contain IT equipment and usually does not contain office automation equipment. The organization using the cold site is responsible for providing and installing the necessary equipment and telecommunications capabilities.

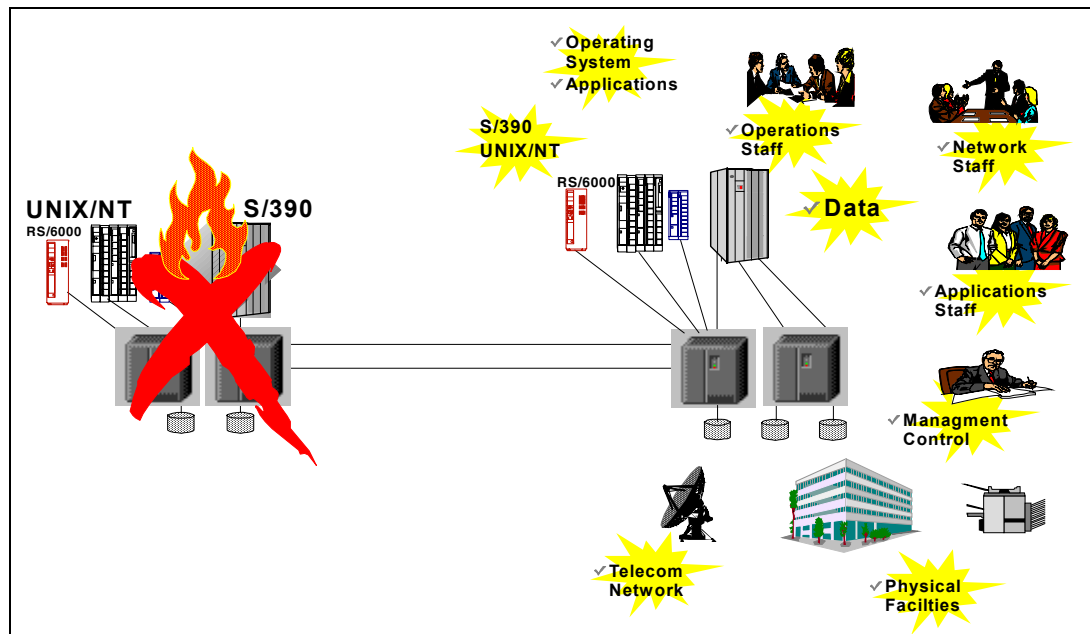


Figure 25-6 Hot site alternative

Warm sites

Warm sites are partially equipped office spaces that contain some or all of the system hardware, software, telecommunications, and power sources. The warm site is maintained in an operational status ready to receive the relocated system. A warm site may serve as a normal operational facility for another system or function, and in the event of Contingency Plan activation, the normal activities are displaced temporarily to accommodate the disrupted system.

Hot sites

Hot sites are office spaces appropriately sized to support system requirements and configured with the necessary system hardware, supporting infrastructure, and support personnel. Hot sites are typically staffed 24 hours a day, 7 days a week. Hot site personnel begin to prepare for the system arrival as soon as they are notified that the Contingency Plan has been activated.

Mobile sites

Mobile sites are self-contained, transportable shells custom-fitted with specific telecommunications and IT equipment necessary to meet system requirements. These are available for lease through commercial vendors. The facility often is contained in a tractor-trailer and may be driven to and set up at the desired alternate location.

Mirrored sites

Mirrored sites are fully redundant facilities with full, real-time information mirroring. Mirrored sites are identical to the primary site in all technical respects. These sites provide the highest degree of availability because the data is processed and stored at the primary and alternate site simultaneously. These sites typically are designed, built, operated, and maintained by the organization.

A requirement for these mirrored sites is the availability of synchronous remote copy among the I/O controllers. Usually companies implementing mirrored sites use both sites to produce work to decrease the expenses - not just a hot stand-by site. Depending on the distance, it is also possible to install a GDPS (geographically dispersed parallel sysplex) between the two sites. In this case we may have workload distribution and continuous availability on top of the disaster recovery solution.

Phase 6 - Testing and training

Plan testing is a critical element of a viable contingency capability. Testing enables plan deficiencies to be identified and addressed. Testing also helps evaluate the ability of recovery staff to implement the plan quickly and effectively.

To derive the most value from the test, explicit test objectives and success criteria should be identified. The use of test objectives and success criteria enable the effectiveness of each plan element and the overall plan to be assessed. Test results and lessons learned should be documented and reviewed by test participants and other personnel as appropriate.

Training for personnel with disaster recovery plan responsibilities should complement testing. Training should be provided at least annually. Ultimately, Contingency Plan personnel should be trained to the extent that they are able to execute their respective Recovery Procedures without aid of the actual document. Recovery personnel should be trained on the following plan elements:

- ▶ Purpose of the plan
- ▶ Cross-team coordination and communication
- ▶ Reporting Procedures
- ▶ Security requirements
- ▶ Team-specific processes (Activation/Notification, Recovery, and Reconstitution Phases)
- ▶ Individual responsibilities (Activation/Notification, Recovery, and Reconstitution Phases)

25.4.6 Phase 7 - Maintenance and changes

To be effective, the plan must be maintained in a ready state that accurately reflects system requirements, procedures, and policies. IT systems undergo frequent changes because of shifting business needs, technology upgrades, or new internal or external policies. Therefore, it is essential that the disaster recovery plan be reviewed and updated regularly to ensure new information is documented and contingency measures are revised if required.

As a general rule, the plan should be reviewed for accuracy and completeness at least annually or whenever significant changes occur to any element of the plan. Certain elements will require more frequent reviews, such as contact lists. Based on the system type and criticality, it may be reasonable to evaluate plan contents and procedures more frequently. At a minimum, plan reviews should focus on the following elements:

- ▶ Operational requirements
- ▶ Security requirements
- ▶ Technical Procedures
- ▶ Hardware, software, and other equipment (types, specifications, and amount)
- ▶ Names and contact information of team members
- ▶ Names and contact information of vendors, including alternate and off-site POCs (Points of Contact)
- ▶ Alternate and off-site facility requirements
- ▶ Vital records (electronic and hardcopy)

Although some changes may be quite visible, others will require additional analysis. The BIA should be reviewed periodically and updated with new information to identify new contingency requirements or priorities. As new technologies become available, preventive controls may be enhanced and recovery strategies may be modified.

This phase must be totally connected with the IT Change Management Team, which is in charge of informing the disaster recovery team of eventual updates in the system, which cause modification in the Recovery Plans.

25.5 Notification and activation procedures

Notification/Activation Procedures which are initial actions taken once a system disruption or emergency has been detected or appears to be imminent. This phase includes activities to notify recovery personnel, assess system damage, and implement the plan. At the completion of this phase, recovery staff will be prepared to perform contingency measures to restore system functions on a temporary basis.

25.5.1 Notification procedures

An event may occur with or without prior notice. For example, advanced notice is often given that a hurricane will affect an area or that a computer virus is expected on a certain date. However, there may be no notice of equipment failure or a criminal act. Notification procedures should be documented in the plan for either type of situation. The procedures should describe the methods used to notify recovery personnel during business and non-business hours.

Prompt notification is important for reducing the effects on the IT system; in some cases, it may provide enough time to allow system personnel to shut down the system gracefully to avoid a hard crash. Following the disaster event, notification should be sent to the team responsible for the damage assessment, so that it may determine the status of the situation and the appropriate next steps. When damage assessment is complete, the appropriate recovery and support teams should be notified.

Notifications can be accomplished through a variety of methods, including telephone, pager, work or personal electronic mail (e-mail), or cell phone. Notification tools that are effective during widespread disasters include radio and television announcements and Web sites. The notification strategy should define procedures to be followed in the event that certain personnel cannot be contacted. Notification procedures should be documented clearly in the disaster recovery plan.

A common notification method is a call tree. This technique involves assigning notification duties to specific individuals, who in turn are responsible for notifying other recovery personnel. The call tree should account for primary and alternate contact methods and should discuss procedures to be followed if an individual cannot be contacted. Figure 25-7 on page 426 presents a sample call tree.

Personnel to be notified should be clearly identified in the contact list appended to the plan. This list should identify personnel by their team position, name, and contact information, including home, work, and pager numbers, e-mail addresses, and home addresses. Additionally the primary personnel may not be available due to the disaster, so you would need to have contracted for alternate personnel at the recovery site to execute the disaster recovery plan.

Note: Be aware, that with some incidents of a physical site disaster, telephone lines can be impacted as well, or the phone lines can be heavily overloaded for hours or even days! A second alternative plan for notification should be considered.

Notification should also be sent to POCs of external organizations or interconnected system partners that may be adversely affected if they are unaware of the situation. Dependent on the type of disruption, the POC may have recovery responsibilities. Therefore, for each system interconnection with an external organization, a POC should be identified to the extent that the organizations will assist each other and the terms under which the assistance will be provided. These POCs should also be listed in an appendix to the plan.

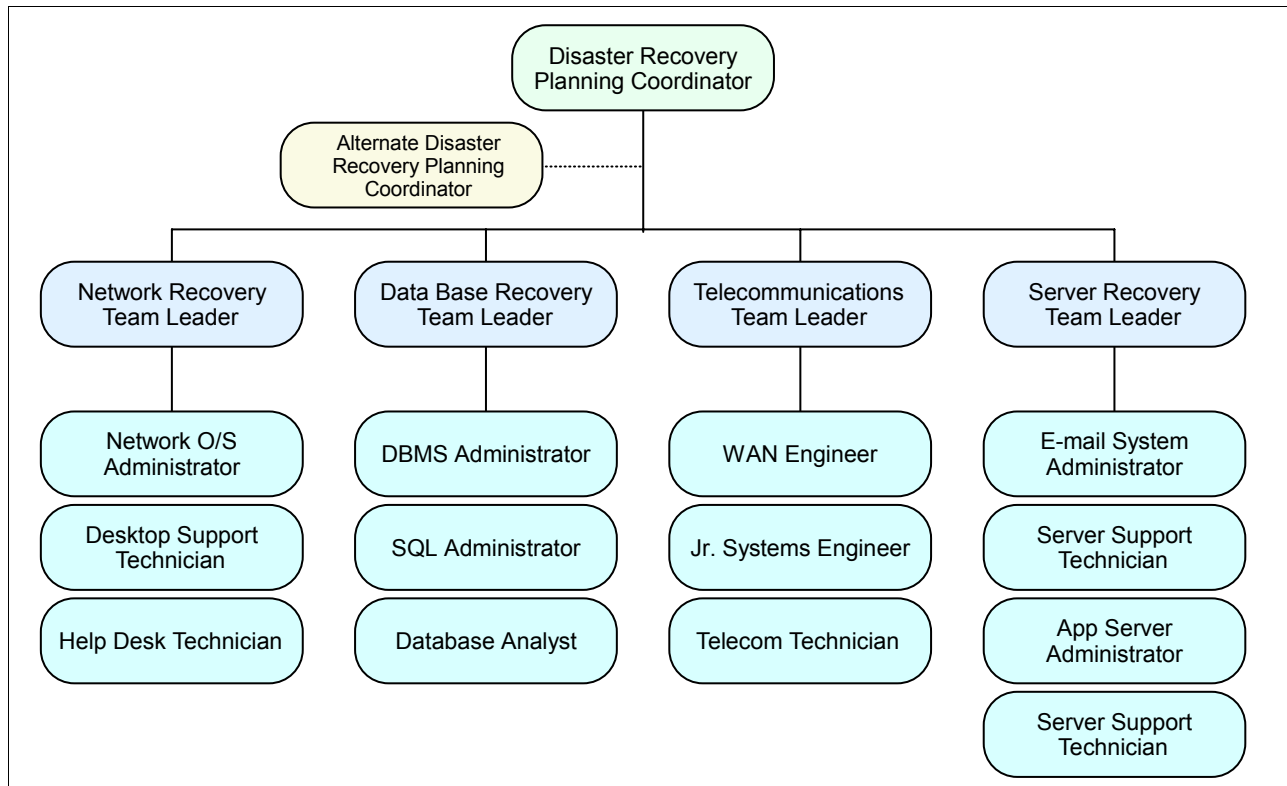


Figure 25-7 Sample call tree

The type of information to be relayed to those being notified should be documented in the plan. The amount and detail of information relayed may depend on the specific team being notified.

25.5.2 Activation

The disaster recovery plan should be activated only when the damage assessment indicates that some of the activation criteria for that system are met. If an activation criterion is met, the disaster recovery planning coordinator should activate the plan. Activation criteria for events are unique for each organization and should be stated in the disaster recovery planning policy statement and may be based on:

- ▶ Safety of personnel and/or extent of damage to the facility
- ▶ Extent of damage to system (physical, operational, or cost)
- ▶ Criticality of the system to the organization's mission (such as, critical infrastructure protection asset)
- ▶ Anticipated duration of disruption

Once the system damage has been characterized, the disaster recovery planning coordinator may select the appropriate recovery strategy, and the associated recovery teams may be notified.

25.6 Recovery procedures

Recovery operations begin after the disaster recovery plan has been activated, damage assessment has been completed (if possible), personnel have been notified, and appropriate teams have been mobilized. Recovery Phase activities focus on contingency measures to restore temporary IT processing capabilities, whereas activities executed during the Reconstitution Phase in 25.7, “Reconstitution phase” on page 428 are directed to repair damage to the original system and restore operational capabilities at the original or new facility. At the completion of the recovery, the system will be operational and performing the functions designated in the plan. Depending on the recovery strategies defined in the plan, these functions could include temporary manual processing, recovery and operation on an alternate system, or relocation and recovery at an alternate site. Teams with recovery responsibilities should understand and be able to perform these recovery strategies well enough that if the paper plan is unavailable during an event, they can still perform the necessary activities.

25.6.1 Sequence of recovery activities

When recovering a complex system, such as a WAN involving multiple independent components, Recovery procedures should reflect system priorities identified in the BIA. The sequence of activities should reflect the system's allowable outage time to avoid significant impacts to related systems and their application. Procedures should be written in a sequential format so system components may be restored in a logical manner. For example, if a LAN is being recovered after a disruption, the most critical servers should be recovered before other, less critical devices, such as printers. Similarly, to recover an application server, procedures should first address operating system restoration and verification before the application and its data are recovered. The procedures should also include instructions to coordinate with other teams when certain situations occur, such as:

- ▶ An action is not completed within the expected time frame
- ▶ A key step has been completed
- ▶ Items must be procured
- ▶ Other system-specific concerns

If conditions require the system to be recovered at an alternate site, certain materials will need to be transferred or procured. These items may include shipment of data backup tapes from off-site storage, hardware, copies of the Recovery Plan, and software programs.

Procedures should designate the appropriate team or team members to coordinate shipment of equipment, data, and vital records. References to applicable appendixes, such as equipment lists or vendor contact information, should be made in the plan where necessary. Procedures should clearly describe requirements to package, transport, and purchase materials required to recover the system.

25.6.2 Recovery procedure highlights

To facilitate the Recovery Phase operations, the disaster recovery plan should provide detailed procedures to restore the system or system components. Given the extensive variety of system types, configurations, and applications, this planning guide does not provide specific recovery procedures.

Recovery procedures should be written in a straightforward, step-by-step style. To prevent difficulty or confusion in an emergency, no procedural steps should be assumed or omitted. A checklist format is useful for documenting the sequential Recovery Procedures. It also is useful for troubleshooting problems if the system cannot be recovered properly.

25.7 Reconstitution phase

In the Reconstitution Phase, recovery activities are terminated and normal operations are transferred back to the organization's facility. During the Recovery Phase, as the contingency activities are performed, reconstitution of the original site should be under way. If the original facility is unrecoverable, the activities in this phase can also be applied to preparing a new facility to support system processing requirements.

Once the original or new site is restored to the level that it can support the IT system and its normal processes, the system may be restored back to the original or to the new site. Until the primary system is restored and tested, the contingency system should continue to be operated.

The Reconstitution Phase should specify teams responsible for restoring or replacing both the site and the IT system. The following major activities occur in this phase:

1. Ensuring adequate infrastructure support, such as electric power, water, telecommunications, security, environmental controls, office equipment, and supplies.
2. Installing system hardware, software, and firmware. This activity should include detailed Restoration Procedures similar to those followed in the Recovery Phase.
3. Establishing connectivity and interfaces with network components and external systems.
4. Testing system operations to ensure full functionality.
5. Backing up operational data on the contingency system and uploading to the restored system.
6. Shutting down the contingency system.
7. Terminating contingency operations.
8. Removing and/or relocating all sensitive materials at the contingency site.
9. Arranging for recovery personnel to return to the original facility.

These teams should understand and be able to perform their required functions without a paper plan in the event such documentation is unavailable.

Note: The IBM Global Service Method provides a highly sophisticated method to develop and tailor phases, activities and tasks for any kind of projects.

25.8 Tiers of Business Continuity

The goal of any disaster protection planning is to protect the most business critical processes and minimize unplanned downtime. Keep in mind that all planning for any type of a disaster tolerant solution is always subject to balancing the solution versus the downtime versus the cost.

In 1992, the SHARE user group in the United States, in combination with IBM, defined a set of Business Continuity tier levels. This was done to address the need to properly describe and quantify various different methodologies for successful mission-critical computer systems Business Continuity implementations. Accordingly, within the IT Business Continuance industry, the tier concept continues to be used, and it is very useful for describing today's Business Continuity capabilities. They need only to be updated for today's specific Business Continuity technologies and associated RTO/RPO.

The Seven Tiers of Business Continuity solutions (Figure 25-8) offer a simple methodology of how to define your current service level, the current risk, and the target service level and target environment.

This section describes the seven tiers of Business Continuity and gives a more detailed breakdown of the possibilities that exist under each tier.

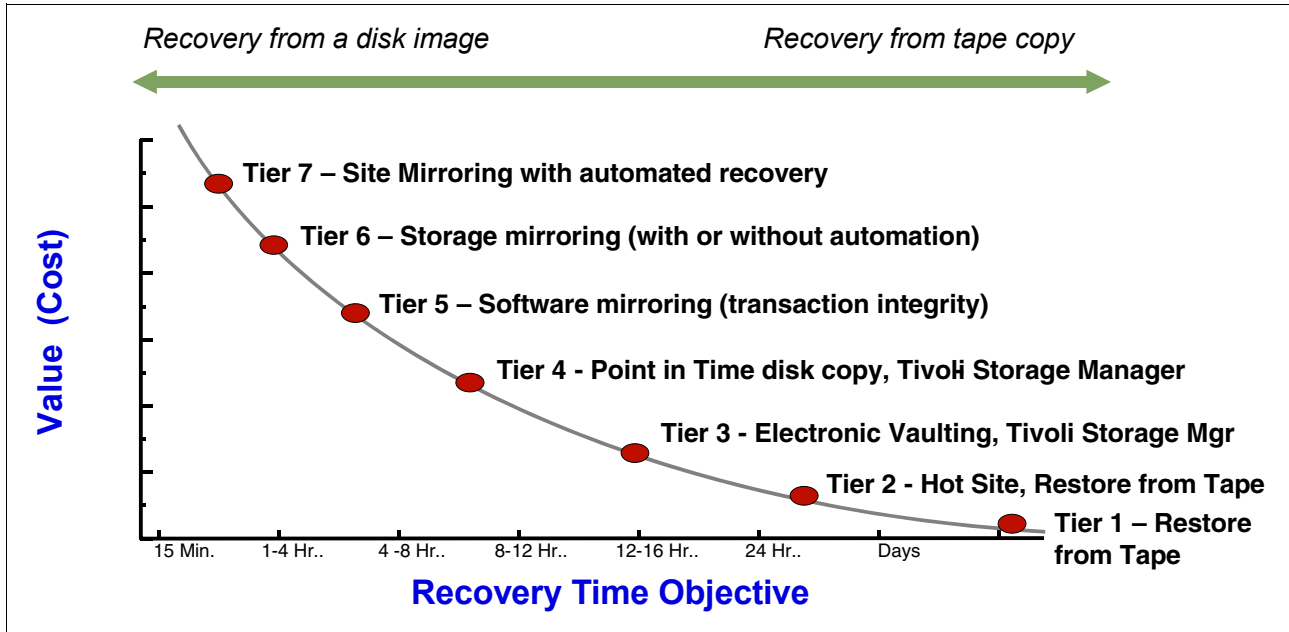


Figure 25-8 Seven tiers of Business Continuity

25.8.1 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.
- ▶ Typical recovery time: The length of recovery time in this instance is unpredictable. In fact, it may not be possible to recover at all.

25.8.2 Tier 1: Data backup with no Hot Site

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.

Example 25-1 Examples of Tier 1 Business Continuity solutions

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

25.8.3 Tier 2: Data backup with a Hot Site

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 hot site) 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 less unpredictable in recovery time.

Example 25-2 Tier 2 Business Continuity Solutions

PTAM with Hot-site available, IBM Tivoli Storage Manager

25.8.4 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 PTAM. As a result there is less data recreation or loss after a disaster occurs.

Example 25-3 Tier 3 Business Continuity Solutions

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

25.8.5 Tier 4: Point-in-time copies

Tier 4 solutions are used by businesses who 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 more disk based solutions. Several hours of data loss is still possible, but it is easier to make such point-in-time (PIT) copies with greater frequency than data can be replicated through tape based solutions.

Example 25-4 Tier 4 Business Continuity Solutions

Batch/Online Database Shadowing and Journaling, Global Copy, FlashCopy, FlashCopy Manager, Peer-to-Peer Virtual Tape Server, Metro/Global Copy, IBM Tivoli Storage Manager - Disaster Recovery Manager, FlashCopy Backup/Restore for SAP Databases, DS4000 Integrated Backup for Databases, Rapid Data Recovery for UNIX and Windows (eRCMF)

25.8.6 Tier 5: Transaction integrity

Tier 5 solutions are used by businesses with a requirement for consistency of data between production and recovery data centers. There is little to no data loss in such solutions, however the presence of this functionality is entirely dependent on the application in use.

Example 25-5 Tier 5 Business Continuity Solutions

Software, two-phase commit, such as DB2 remote replication, Oracle Data-Guard, etc.

25.8.7 Tier 6: Zero or little data loss

Tier 6 Business Continuity solutions maintain the highest levels of data currency. They are used by businesses with little or no tolerance for data loss and who need to restore data to applications rapidly. These solutions have no dependence on the applications to provide data consistency.

Example 25-6 Tier 6 Business Continuity Solutions

Metro Mirror, Global Mirror, z/OS Global Mirror, GDPS HyperSwap Manager, Peer-to-Peer VTS with synchronous write, Metro/Global Mirror, PPRC Migration Manager, eRCMF, HACMP/XD with Logical Volume Mirroring, HACMP/XD with Metro Mirror

25.8.8 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. This allows a Tier 7 solution to ensure consistency of data above that which is granted by Tier 6 solutions. Additionally, recovery of the applications is automated, allowing for restoration of systems and applications much faster and more reliably than would be possible through manual Business Continuity procedures.

Example 25-7 Tier 7 Business Continuity Solutions

GDPS/PPRC with or without HyperSwap™, GDPS/XRC, AIX HACMP/XD with Metro Mirror, TotalStorage Continuous Availability for Windows, iSeries High Availability Business Partner software

25.9 Blending tiers into an optimized solution

To use the tiers to select the best Business Continuity solution, a blend of Business Continuity tiers is the practical method for cost-optimizing the entire solution. The most common result, from an enterprise standpoint, is a strategic architecture of three tiers. Three tiers generally appear as an optimum number, because at the enterprise level, two tiers generally are insufficiently optimized (in other words, overkill at some point and underkill at others), and four tiers are more complex but usually do not provide enough additional strategic benefit.

To use the tiers to derive a blended, optimized Business Continuity architecture, we suggest:

1. Categorize the business' entire set of applications into three bands: Low Tolerance to Outage, Somewhat Tolerant to Outage, and Very Tolerant to Outage. Of course, while some applications that are not in and of themselves critical, they do feed the critical applications. Those applications would need to be included in the higher tier.
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.

25.9.1 Three Business Continuity solution segments

IBM's Business Continuity solutions in the TotalStorage Resiliency Family have been segmented for you into these three bands; that segmentation is shown in Figure 25-9:

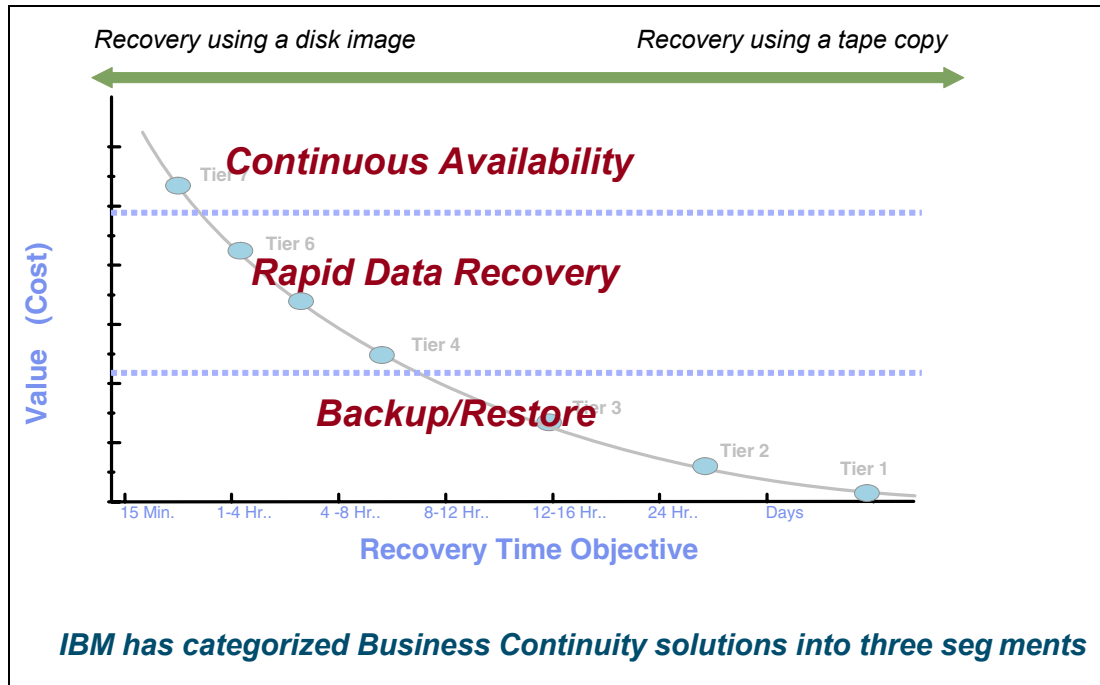


Figure 25-9 Three Business Continuity solution segments

25.10 Select a solution by segment

Once you have decided upon your desired Recovery Time Objective, find the appropriate solution segment band in the chart above.

IBM TotalStorage Business Continuity solutions in the *Continuous Availability* solution segment are Tier 7 solutions. They include (but are not limited to):

- ▶ **Windows:** TotalStorage Continuous Availability for Windows
- ▶ **pSeries:** TotalStorage Continuous Availability for pSeries (AIX HACMP/XD)
- ▶ **zSeries:** TotalStorage Continuous Availability for zSeries (GDPS)

IBM TotalStorage Business Continuity solutions in the *Rapid Data Recovery* solution segment are Tier 4 to 6 solutions. They include (but are not limited to):

- ▶ **Heterogeneous high end servers:** enterprise Remote Copy Management Facility (eRCMF)
- ▶ **Heterogeneous open system disk vendor mirroring:** IBM SAN Volume Controller Metro Mirror
- ▶ **zSeries:** GDPS HyperSwap Manager

IBM TotalStorage Business Continuity solutions in the *Backup/Restore* solution segment are Tiers 4 to 1. They include (but are not limited to):

- ▶ **Disk to Disk Backup for DB2:** ADS4000-based DB2 integrated backup/restore solution exploiting integration with Tivoli Storage Manager
- ▶ **FlashCopy Backup / Restore for SAP:** An ESS-based integrated backup/restore solution exploiting SAP integration with Tivoli Storage Manager for ERP

- ▶ **Centralized Backup / Restore:** An enterprise class Tivoli Storage Manager planning and implementation offering

In the redbook *IBM TotalStorage Business Continuity Solutions Guide*, SG24-6547, all solutions are discussed in more detail. The above set of solutions will be briefly reviewed in subsequent sections.

25.11 Continuous Availability solutions

This section briefly describes Continuous Availability solutions in the following environments:

- ▶ Continuous availability for zSeries (GDPS)
- ▶ Continuous availability for open systems (GDOC)
- ▶ Continuous availability for AIX (HACMP-XD)
- ▶ Continuous availability for Windows (GDS for MSCS)

25.11.1 Geographically Dispersed Parallel Sysplex (GDPS)

GDPS is a family of IBM Global Services offerings for single site or a multi-site application availability solution, designed to provide 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 zSeries 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 zSeries Business Continuity solutions consists of two major offering categories, with sub-offerings in each category. They are:

- ▶ GDPS/PPRC solutions, based on IBM TotalStorage Metro Mirror (Metro Mirror was formerly known as Peer-to-Peer Remote Copy (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 TotalStorage z/OS Global Mirror (z/OS Global Mirror was formerly known as Extended Remote Copy (XRC)), including:
 - GDPS/XRC, a Tier 7 solution
 - RCMF/XRC, a remote copy management solution for XRC

GDPS/PPRC overview

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

The physical topology of a GDPS/PPRC, referring to Figure 25-10, consists of a zSeries base or Parallel Sysplex cluster spread across two sites separated by up to 100 kilometers (km) of fiber – approximately 62 miles – with one or more z/OS systems at each site.

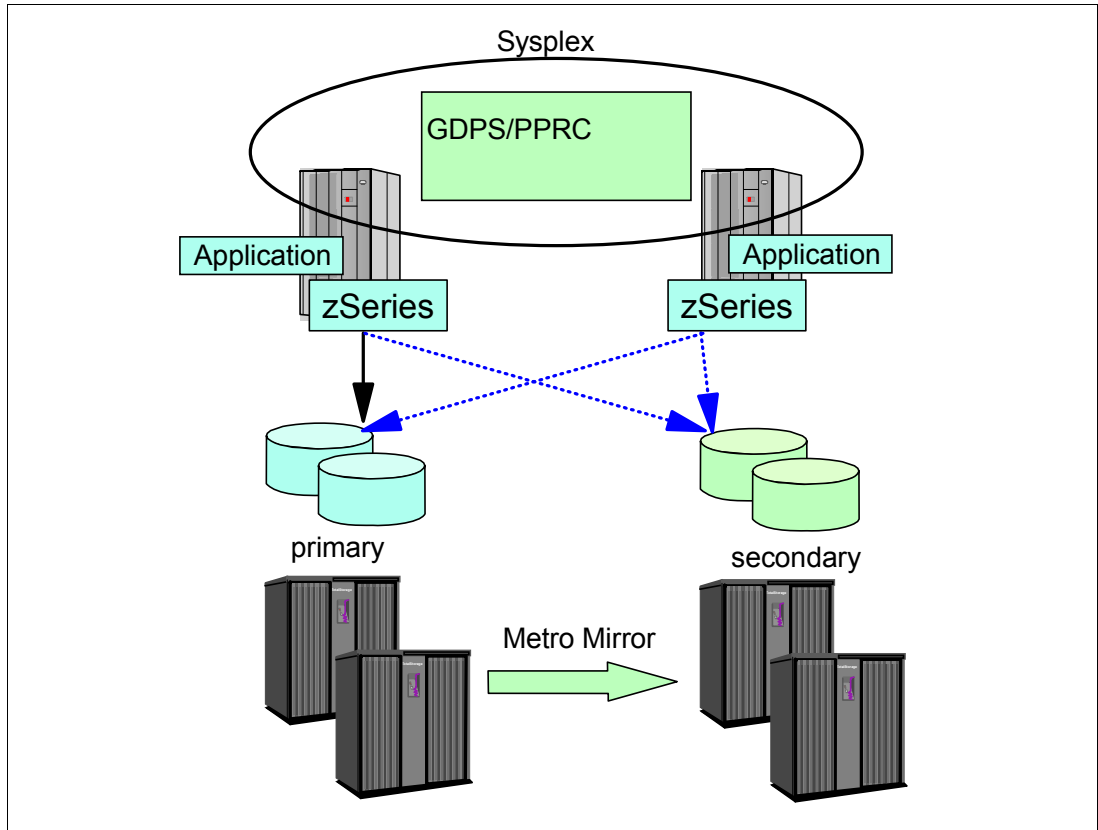


Figure 25-10 GDPS/PPRC topology

GDPS/PPRC is designed to provide the ability to 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 Data Base Management System (DBMS) restart – not DBMS recovery – in the second site. GDPS/PPRC is application independent, and therefore can cover the client's complete application environment.

Near Continuous Availability of data with HyperSwap

GDPS in the PPRC 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 so doing, HyperSwap is designed to broaden 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 later in this section. With this sub-offering, Parallel Sysplex availability can now be extended to disk subsystems, even if multiple sites are not available and the Parallel Sysplex is configured in 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 production operating systems; these include 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 PPRC, 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 for both zSeries 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 will provide 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 is designed to 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 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 using 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/PPRC HyperSwap Manager is designed to provide both rapid data recovery for enterprise data, and a lower initial price and cost-effective initial entry point to the GDPS family of GDPS offerings.

GDPS/XRC overview

GDPS/XRC has the attributes of a *disaster recovery* solution. z/OS Global Mirror (XRC) is a combined hardware 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 DFSMSdfp™ component called System Data Mover (SDM), typically running in the recovery site, asynchronously offloads data from the primary storage subsystem's cache and updates the secondary disk volumes.

GDPS/XRC is capable of 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, as shown in Figure 25-11, consists of production system(s) in one site. The production systems could be a single system, multiple systems sharing disk, or a base or Parallel Sysplex cluster1. The recovery site can be located at a virtually unlimited distance from the production site.

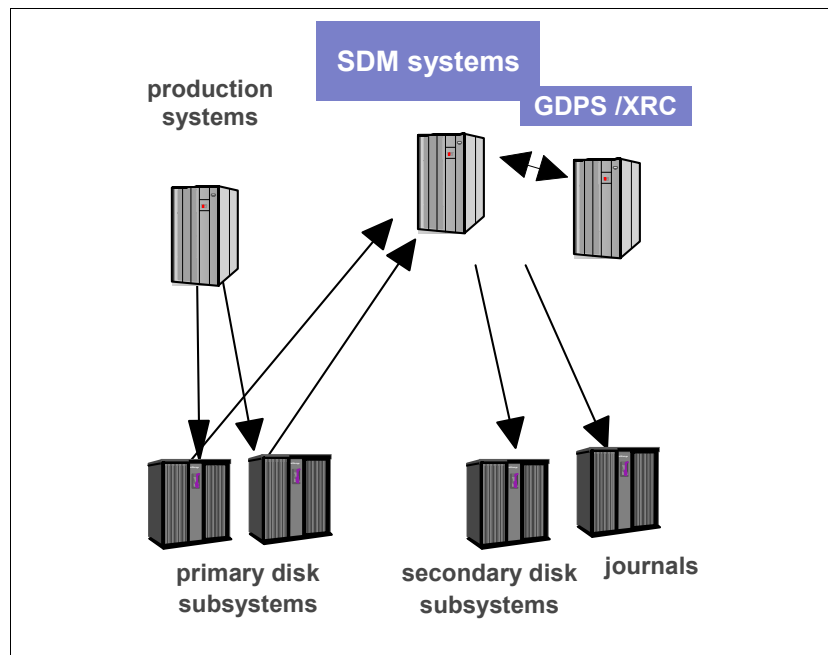


Figure 25-11 GDPS/XRC topology

In GDPS/XRC, the production system(s) located in the production site can be a single system, multiple systems 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) to allow a business to attain "near transparent" Disaster Recovery with minimal data loss. GDPS/XRC is designed to provide the ability to perform a controlled site switch for an unplanned site outage, maintaining data integrity across multiple volumes and storage subsystems and the ability to perform a normal Data Base Management System (DBMS) restart – not DBMS recovery – in the recovery site. 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, you can refer to the following Web sites and manuals or Redbooks:

GDPS home page:

<http://www.ibm.com/servers/eserver/zseries/gdps/>

zSeries Business Resiliency Web site:

<http://www.ibm.com/servers/eserver/zseries/resiliency/features.html>

For an overview of how planning, implementation, and proper exploitation of zSeries Parallel Sysplex clustering technology can enable your zSeries IT infrastructure to achieve near Continuous Availability, refer to "Five Nines / Five Minutes — Achieving Near Continuous Availability", found at:

<http://www.ibm.com/servers/eserver/zseries/pso/>

Also, refer to the following IBM manuals and Redbooks:

- ▶ *IBM @server zSeries: Device Support Facilities User's Guide and Reference R17, GC35-0033*
- ▶ *IBM @server zSeries: z/OS V1R6.0 DFSMS Advanced Copy Services, SC35-0428*
- ▶ *IBM TotalStorage Enterprise Storage Server: Implementing Copy Services in an Open Environment, SG24-5757*
- ▶ *IBM TotalStorage Enterprise Storage Server: Implementing ESS Copy Services with IBM @server zSeries, SG24-5680*

25.11.2 Geographically Dispersed Open Clusters

Geographically Dispersed Open Clusters (GDOC) is a multivendor solution designed to protect the availability of critical applications that run on UNIX, Windows, or Linux servers. GDOC is based on an Open Systems Cluster architecture spread across two or more sites with data mirrored between sites to provide high availability and Disaster Recovery. In this section, we provide a concise overview of the GDOC function, its benefits, and applications.

This is a Tier 7 solution. It is composed of both software automation and IGS services to configure and deploy the solution.

The GDOC solution is based on VERITAS Foundation suite and VERITAS Cluster software products. Because of this the GDOC solution is available only on operating system platforms supported by these VERITAS products. The operating system platforms that are currently supported are:

- ▶ IBM AIX
- ▶ Linux
- ▶ Microsoft Windows
- ▶ SUN Solaris

► HP UX

A GDOC solution controls resource and application availability and initiates application failover to alternative servers when the need arises. GDOC is based on VERITAS products, such as VERITAS Storage Foundation for volume and file management. Replication can be performed either at the server level using VERITAS Volume Replicator for software replication, or at the hardware level using storage subsystem remote replication functions. Figure 25-12 shows a diagram of the GDOC functionality.

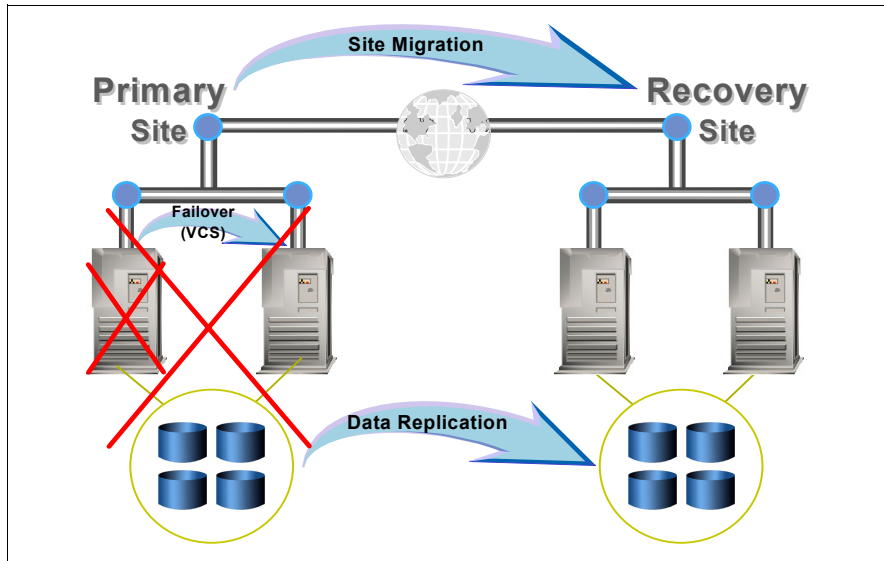


Figure 25-12 GDOC functional overview

Application availability and failover are controlled using VERITAS Cluster Server (VCS) and specialized modules for applications, such as SAP, DB2, and others. Additional cluster modules are available.

Normally applications run on the primary site. Application data is contained in an external storage subsystem. This allows the application to fail over to a secondary server at the primary site under VCS control. The application accesses the data in the primary site's storage subsystem.

Data is replicated continuously between the storage subsystems located at the primary and secondary sites using data replication functions such as ESS and DS6000/DS8000 Metro Mirror or VERITAS Volume Replicator. VERITAS Volume Replicator maintains dependent write consistency, and write order fidelity using VERITAS terminology, at the individual server level. If you have a requirement to maintain dependent write consistency across multiple servers, then you must consolidate the storage in one or more external storage subsystems and use ESS or DS6000/DS8000 Metro Mirror or Global Mirror solutions.

The VERITAS CommandCentral Availability, the follow-on product to VERITAS Global Cluster Manager, controls failover of applications to the secondary site. This is shown in Figure 25-12 on page 438 as a site migration. A site migration will typically require operator intervention; all that is required is an operator confirmation to perform the site failover.

This solution extends the local High Availability (HA) model to many sites. Dispersed clusters and sites are linked by public carrier over a wide area network or SAN. Each site is aware of the configuration and state of all of the sites (global cluster management). Complete site failover occurs in the event of a catastrophe, and the basis for this failover is the replicated

data. This solution allows you to perform global availability management from a single console.

25.11.3 TotalStorage Continuous Availability for AIX

TotalStorage Continuous Availability for AIX is designed to allow cluster resources, that is applications and data, to fail over between two geographically dispersed sites. It offers both High Availability and Disaster Recovery across geographically-dispersed HACMP clusters, protecting business-critical applications and data against disasters that affect an entire data center.

This is a Tier 7 solution when High Availability Cluster Multiprocessing (HACMP) XD (eXtended Distance) is used in combination with the Metro Mirror (formerly Synchronous PPRC) feature. It can be considered an automated Tier 4 solution when HACMP-XD is used in combination with Global Copy (previously known as PPRC-XD).

Solution description

For over 10 years, HACMP has been managing high availability for AIX clusters located within a single computer room or small campus. HACMP has been providing reliable, high availability services, monitoring capabilities and dependable detection of application failures. HACMP manages the fallover of business application environments to backup servers. With the introduction of the new optional package High Availability Cluster Multiprocessing (HACMP) XD (eXtended Distance), IBM addresses the need for high availability and long distance mirroring by offering a world-class solution.

TotalStorage Continuous Availability for AIX offers both High Availability and Disaster Recovery across geographically-dispersed HACMP clusters, protecting business-critical applications and data against disasters that affect an entire data center. The solution integrates HACMP with the disk remote replication feature to provide a robust fully automated disaster tolerant solution for mission-critical applications. Resources automatically fail over to the surviving node, the applications restart, no manual steps are required and traffic is re-routed to the surviving node, and it manages the resynchronization of data during site recovery too.

HACMP/XD: Remote Copy increases data availability for IBM TotalStorage Enterprise Storage Server (ESS) volumes that use Peer-to-Peer Remote Copy (PPRC) to copy data to a remote site for disaster recovery purposes. This feature allows HACMP clusters to support automatic fallover of disks that are PPRC pairs and creates a powerful solution for clients using the ESS with PPRC. By automating the management of PPRC, the recovery time is minimized after an outage, regardless of whether the clustered environment is local or geographically dispersed. HACMP/XD in combination with PPRC manages a clustered environment to automatically start the application at the remote site using the PPRC mirrored data. It also ensures that the PPRC mirrored data is synchronized prior to allowing a fallover. So this option provides a well integrated hardware-based solution.

A Disaster Recovery Plan that includes geographically dispersed Enterprise Storage Servers (ESS), Peer-to-Peer Remote Copy (PPRC), and IBM HACMP/XD will automate the real-time replication of your applications and data, while monitoring and maintaining the availability of your system.

These three products work together as an integrated system to provide:

- ▶ Automatic backup and recovery after failures: The recovery of your business-critical applications and data after a wide range of system failures is not dependent on the availability of any one component.

- ▶ Automated control of data mirrors: The complicated tasks of establishing, suspending, reversing, and resynchronizing data mirrors are automatic, thus reducing the chance of data loss or corruption due to user error.
- ▶ Easier execution of planned system outages: Tools for user-controlled operations help you gracefully bring individual system components offline for scheduled maintenance while minimizing the downtime experienced by the end users.

Solution components

IBM TotalStorage Continuous Availability for AIX is a solution based on a combination of hardware and software products together with a service engagement, where IBM Global Services or an IBM Business Partner will work with your IT department to configure the solution to match your unique business environment.

Hardware requirements

The hardware requirements for using TotalStorage Continuous Availability for AIX include:

- ▶ Two supported storage subsystems linked by a minimum of two Peer-to-Peer Remote Copy links (to enable logical PPRC paths in both directions). At the time of the writing of this redbook, only the ESS is supported, but there are plans to support SAN Volume Controller in the next months.
- ▶ pSeries, Cluster 1600, and RS/6000® models connected to the storage subsystem using one or more of the supported host bus adapters (HBAs).
- ▶ LAN interconnect as required by HACMP.
- ▶ We recommend using non-IP point-to-point networks for the heartbeat to connect the cluster nodes and avoid the TCP/IP stack as a single point of failure.

You should also check the supported microcode versions for your storage subsystems.

If you need channel extenders, you can find a list of supported devices in the interoperability matrixes located on the Web:

- ▶ IBM TotalStorage ESS Model 800
<http://www-1.ibm.com/servers/storage/disk/ess/ess800/interop-matrix.html>
- ▶ SAN Volume Controller
http://www-1.ibm.com/support/docview.wss?rs=591&uid=ssg1S1002442#_Inter_Cluster

Software requirements

The software requirements for using TotalStorage Continuous Availability for AIX include:

- ▶ High Availability Cluster Multi-Processing for AIX 5L™
- ▶ High Availability Cluster Multi-Processing Extended Distance (HACMP/XD) for ESS PPRC
- ▶ AIX 5L Version 5.1, Version 5.2, or Version 5.3 with the suggested Recommended Maintenance packages and APARs
- ▶ PSSP Version 3.5 is also required on RS/6000 SP (9076) systems
- ▶ AIX 5L Java™ Runtime Environment Version 1.3.0.15, or later
- ▶ The Command Line Interface (CLI) needed to automate the Copy Services tasks on your storage subsystem
- ▶ IBM TotalStorage Subsystem Device Driver (SDD) Service Version 1.3.3.6, or later
- ▶ HACMP/XD for eRCMF requires eRCMF V2.0

Service and ordering

For the TotalStorage Continuous Availability for AIX solution, the storage subsystems, servers, and network components have to be ordered through the normal process. Then you engage IBM Global Services to install the ESS and eventually implement Metro Mirror/Global Copy and HACMP/XD for ESS PPRC. IBM Business Partners can also wrap implementation services around HACMP/XD. These specialists will work with your IT department to configure the solution to match the needs of your business environment as well as possible.

Additional information

You can find further information about the IBM Web site:

http://www-1.ibm.com/servers/eserver/pseries/ha/disaster_tech.html

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

For more information, contact your IBM sales representative.

25.11.4 TotalStorage Continuous Availability for Windows

TotalStorage Continuous Availability for Windows is designed to allow cluster resources, that is applications and data, to fail over between two geographically dispersed sites. By combining the high availability of Microsoft Cluster Server (MSCS) services for application cluster failover with PPRC mirroring for Disaster Recovery, TotalStorage Continuous Availability for Windows provides a continuous availability solution to protect critical cluster-aware applications and data in a Microsoft Cluster environment.

This is a Tier 7 solution. As a Tier 7 Disaster Recovery solution, the failover of applications and data and recovery on the secondary site will complete in a matter of minutes.

Currently IBM TotalStorage offers support for IBM TotalStorage Continuous Availability for Windows (formerly named IBM TotalStorage Geographically Dispersed Clusters for Microsoft Cluster Server, or GDS offering) on the Enterprise Storage Server (ESS) and SAN Volume Controller. Support for the recently announced DS6000 and DS8000 storage subsystems is forthcoming. This services offering, available by RPQ through IBM Systems Group, enables high availability and Disaster Recovery in the Microsoft environment by incorporating the benefits of Metro Mirror (previously known as Synchronous PPRC).

TotalStorage Continuous Availability for Windows is the IBM implementation for disk subsystems within the Microsoft Geographically Dispersed Sites for clusters framework. It will provide clustering failover support for any MSCS cluster-aware applications such as DBMS, messaging and file sharing. Any MSCS cluster-aware applications will run in a TotalStorage Continuous Availability for Windows configuration.

TotalStorage Continuous Availability for Windows is a disaster recovery solution for Microsoft Cluster server applications using automated cluster failover services combined with IBM storage subsystems using PPRC for remote site mirroring and automated failover of data. It can allow for zero or little data loss and little or no downtime and it provides for complete site server application and data disaster recovery across two geographically dispersed sites.

It monitors the health states of the applications, network and storage PPRC mirroring to determine automatic actions for failover to the secondary site.

Additional information

Additional information can be found on the IBM Web site:

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

IBM personnel can also access the following Web site to get additional information about this offering:

http://web.mainz.de.ibm.com/e_dir/esccw3.nsf/ContentDocsByTitle/IBM+TotalStorage+ATS+Services+EMEA

For more information, contact your IBM sales representative.

25.12 Rapid Data Recovery

Rapid Data Recovery is based on maintaining a second disk-resident 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 distinguished from Continuous Availability by the fact that the automation required to be a Tier 7 solution is not present.

This section describes Rapid Data Recovery in the following environments:

- ▶ zSeries and mixed z + Open
- ▶ UNIX and Windows (eRCMF)
- ▶ UNIX and Windows (SVC)
- ▶ iSeries

25.12.1 zSeries and mixed z+Open platforms

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 falls short of being a Tier 7 solution due to the lack of recovery automation provided by a full GDPS/PPRC implementation.

Rapid Data Recovery for zSeries is provided by a new IGS service offering, GDPS/PPRC HyperSwap Manager (GDPS/PPRC HM), in the GDPS suite of offerings. It is based on IBM TotalStorage Metro Mirror, previously known as Synchronous PPRC, a hardware-based mirroring and remote copying solution for the IBM TotalStorage Enterprise Storage Server (ESS), DS6000, and DS8000.

25.12.2 TotalStorage Rapid Data Recovery for UNIX and Windows

In a disaster recovery solution based on disk remote mirroring we want to restart a database application following an outage without having to restore the database. This process has to be consistent, repeatable and fast, measurable in minutes. A Database Recovery where you have to restore the last set of Image Copy tapes and apply the log changes to bring the database up to the point of failure is a process that can be measured in hours or even days. And this is what you want to avoid in a Tier 4 or higher solution. However in a real disaster (fire, explosion, earthquake) you can never expect your complex to fail all at the same moment. 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 caused during a Rolling Disaster.

The TotalStorage Rapid Data Recovery for UNIX and Windows solution is based on *enterprise Remote Copy Management Facility (eRCMF)*, which is a service offering from IBM Global Services. Enterprise Remote Copy Management Facility is a multi-site disaster recovery solution which manages data on volumes for Fixed Block (open systems hosts). The solution does not depend on the operating system as no client installation is required. This is a Tier 4-6 solution.

This solution simplifies the process of repairing inconsistent PPRC pairs after an error with the copy process. eRCMF assists the user in automating the PPRC and FlashCopy processes. eRCMF is a scalable and flexible solution that protects business data and maintains consistent restartable data. It can be used to manage copy relationships for both:

- ▶ Planned outages such as hardware and software upgrades, or disaster practice
- ▶ Unplanned outages such as an actual disaster

Further information can be found on the IBM Web site at:

<http://www-1.ibm.com/services/us/index.wss/so/its/a1000110>

You can also refer to the following papers:

- ▶ *enterprise Remote Copy Management Facility eRCMF V3 Implementation Guide*, S7001075
- ▶ *enterprise Remote Copy Management Facility eRCMF V3 User Guide*, S7001074

that you can find at:

<http://www-1.ibm.com/servers/storage/support/software/opensystems.html>

25.12.3 IBM TotalStorage SAN Volume Controller

IBM TotalStorage SAN Volume Controller brings your devices together in a virtual pool to make all your storage appear 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 IBM TotalStorage SAN Volume Controller addresses the increasing costs and complexity in data storage management. It addresses this increased complexity by shifting storage management intelligence from individual SAN controllers into the network and by using virtualization.

The SAN Volume Controller falls under Tier 4 for FlashCopy and Tier 6 for Metro Mirror.

25.12.4 Global Mirror Utility

The Global Mirror Utility (GMU) is a tool that can be used as a Tier 6 solution. It can be used only to manage Global Mirror and not other Copy Services. It is distributed on a CD together with installation and user guide documentation with every ESS running Licensed Code (LIC) 2.4.0 or higher. It is available for ESS, DS6000 and DS8000. The GMU tool can be used for free in a single Global Mirror copy services domain (up to 8 storage subsystems), it is a service for bigger configurations or if clients need specific customizations.

25.13 Backup and Restore

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.

This section describes Backup and Restore in the following solutions:

- ▶ Volume Shadow Copy Services (VSS)
- ▶ Space efficient FlashCopy with SFS
- ▶ Tivoli Storage Manager
- ▶ IBM TotalStorage DS4000 Integrated Backup for Data
- ▶ TotalStorage Data Manager for SQL 2000 (TDM) and SAP on SQL Server

25.13.1 Microsoft Volume Shadow Copy Service (VSS)

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

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

VSS is used for a number of purposes such as:

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

Volume Shadow Copy service for ESS

The ESS Application Programming Interface (API) enables users to perform ESS FlashCopy Services through Microsoft Volume Shadow Copy Service. The ESS API is a non-proprietary storage management client application that supports and enables Copy Services such as FlashCopy with support for Microsoft VSS.

Volume Shadow Copy Service for DS8000/DS6000/DS4000

The DS8000/DS6000/DS4000 storage servers provide an integration with Microsoft Volume Shadow Copy service to produce consistent shadow copies. This enables you to create full volume copies of data in a storage unit.

25.13.2 SAN File System FlashCopy

IBM Total Storage SAN File System (SAN File System) FlashCopy offers a rapid method of protecting data. This is a Tier 4 solution. The SAN File System has a FlashCopy function, which creates a point-in-time copy or image of a fileset. The created image is a read-only, space-efficient image of the contents of a SAN File System fileset at the time it was taken. You can use backup applications or utilities on SAN File System clients to back up the contents of FlashCopy images, rather than the actual fileset. Doing this avoids any issues with open files that might cause problems when backing up live data.

FlashCopy images are file-based. Therefore the SAN File System clients can see all the files and directories in the image. The administrator can control access to FlashCopy images using file system security. They can use this image for quick restore of parts of the fileset if required by simply copying the required files and folders back to the actual fileset. Finally, the entire fileset can be quickly reverted from a FlashCopy image by the SAN File System administrator.

25.13.3 IBM Tivoli Storage Manager

As part of the IBM TotalStorage Open Software Family, 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 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

25.13.4 IBM TotalStorage DS4000 Integrated Backup for Data

The storage servers within the IBM TotalStorage DS4000 series are redundant arrays of independent disks (RAID) storage subsystems and the disk drive enclosures. The DS4000 series provides high availability through the use of hot-swappable and redundant components.

IBM TotalStorage DS4000 falls under Tier 6 for Enhanced Remote Mirror, Metro Mirror and Tier 4 for Volume Copy and FlashCopy. The DS4000 Storage Manager software (Storage Manager 9.12) is used to configure, manage and troubleshoot the IBM TotalStorage DS4000 storage server. Enhanced Remote Mirroring provides the ability to create remote logical drive mirror pairs with two new features: asynchronous write mode, and the write consistency group option. Asynchronous mirroring with the consistency group option is referred to as Global Mirroring, and asynchronous mirroring without the consistency group option is referred to as Global Copy. Integrated Backup for Databases offering, a new feature for IBM TotalStorage DS4500, DS4400 and DS4300 systems, provides the utilities and documentation necessary to install an integrated data protection backup solution for DB2 databases.

25.13.5 TotalStorage Data Manager for SQL 2000

TotalStorage Data Manager for SQL 2000 (TDM) is a FlashCopy backup and recovery solution for Microsoft SQL Server databases that have their data in IBM TotalStorage Enterprise Storage Server (ESS) subsystems. This solution can also be used to backup SAP database servers that use MS SQL Server as the database engine. This is a Tier 4 services based solution. TDM creates online snapshot backups and offline snap-back restores of Microsoft SQL Server 2000 databases on selected IBM storage subsystems. TDM uses the Microsoft Virtual Device Interface (VDI) to suspend writes on the database during the FlashCopy backup operation. It can be a complex and time-consuming task for installations to develop their own scripts to perform online snapshot backups of Microsoft SQL Server databases on the ESS or DS4000.

25.14 More information

For a more detailed understanding of Business Continuity Solutions, refer to these tips, papers, Redbooks, and Web sites:

- ▶ *IBM TotalStorage Business Continuity Solutions Guide*, SG24-6547
- ▶ *IBM TotalStorage Business Continuity Solutions Overview*, SG24-6684
- ▶ *Disaster Recovery Strategies with Tivoli Storage Management*, SG24-6844
- ▶ *Seven Tiers of Disaster Recovery*, TIPS-0340
- ▶ *A Disaster Recovery Solution Selection Methodology*, REDP-3847
- ▶ *Planning for Disaster Recovery in a Heterogeneous Environment*, REDP-3848
- ▶ The Disaster Recovery Journal Web site at: <http://www.drj.com>



Digital media solutions

Digital media—video, audio, text, images, games and graphics captured and stored in digital format—has become an important tool in many industries. An effective digital media system requires powerful servers and storage, proven database management systems, capable middleware and other software. Of all the major parts of a digital media system, none figures more prominently than storage.

IBM Digital Content Creation is a complete solution for the media and entertainment industries. This offering provides a comprehensive suite of IBM workstations, software, storage and servers, all designed to support animators, special effects wizards and digital media producers as their work transitions from proprietary workstations and servers to Linux and Intel platforms.

26.1 The magic of digital

For creative professionals, the power and magic of digital content creation is opening up fantastic new worlds to explore. This same technology used by special effects creators is being adapted to the business world, where educational and corporate messages can now be brought to life in ways never before imagined.

As a technology innovator, IBM has been at the forefront of this exciting breakthrough in human creativity.

An effective digital media system requires powerful servers and storage, proven database management systems, capable middleware and other software. Of all the major parts of a digital media system, perhaps none figures more prominently than storage.

Digital media applications require a large amount of fast storage to enable on demand viewing of content. Whether digital media is stored on network attached storage (NAS) or contained within a Storage Area Network (SAN), the hard disk drives and tape subsystems form the very heart of the digital media delivery system.

26.1.1 Digital Media Center

IBM has introduced a new digital video storage system based on Storage Area Network (SAN) technology. The storage solution allows broadcasters to escape from their traditional reliance on videotape as the shared video storage medium of choice. The system helps transform broadcasting production environments from analog to digital, and eventually tapeless environments.

Comprised of industry-standard IBM hardware and software, Digital Media Center solutions provide broadcasters with a pool of shared storage that is not dependent on any specific machine or vendor.

On demand

The Digital Media Center can help you better leverage your media assets to improve system availability and efficiency. Furthermore, you will be able to take advantage of on demand technology from across IBM. Digital Media Center enables a completely new view of broadcast production infrastructure, one which embodies these key on demand qualities:

- ▶ **Integrated:** Shares hardware and software resources across the enterprise
- ▶ **Open:** Shares data and manages diverse systems
- ▶ **Virtualized:** Hides complexity to permit better utilization
- ▶ **Autonomic:** Incorporates IBM self-managing, self-healing technology

Paired with IBM digital media solutions, these storage offerings are designed to provide the high performance, scalability and efficiency required for digital media systems today.

Digital Media Center components

Digital Media Center is based on IBM Storage Area Network (SAN) technology. SANs can provide exceptional configuration flexibility with any-to-any device connectivity, and can help transform a storage infrastructure into an easy-to manage, highly scalable system. The principal components of the Digital Media Center network are:

- ▶ **IBM TotalStorage DS Family:** The IBM TotalStorage DS Family offers a broad array of storage solutions, giving you the freedom to select the storage enclosure, or combination of enclosures, that best matches the size and value of digital content you intend to store. The IBM TotalStorage DS Family can help provide the levels of performance, availability and expandability you need to create your Digital Media Center.

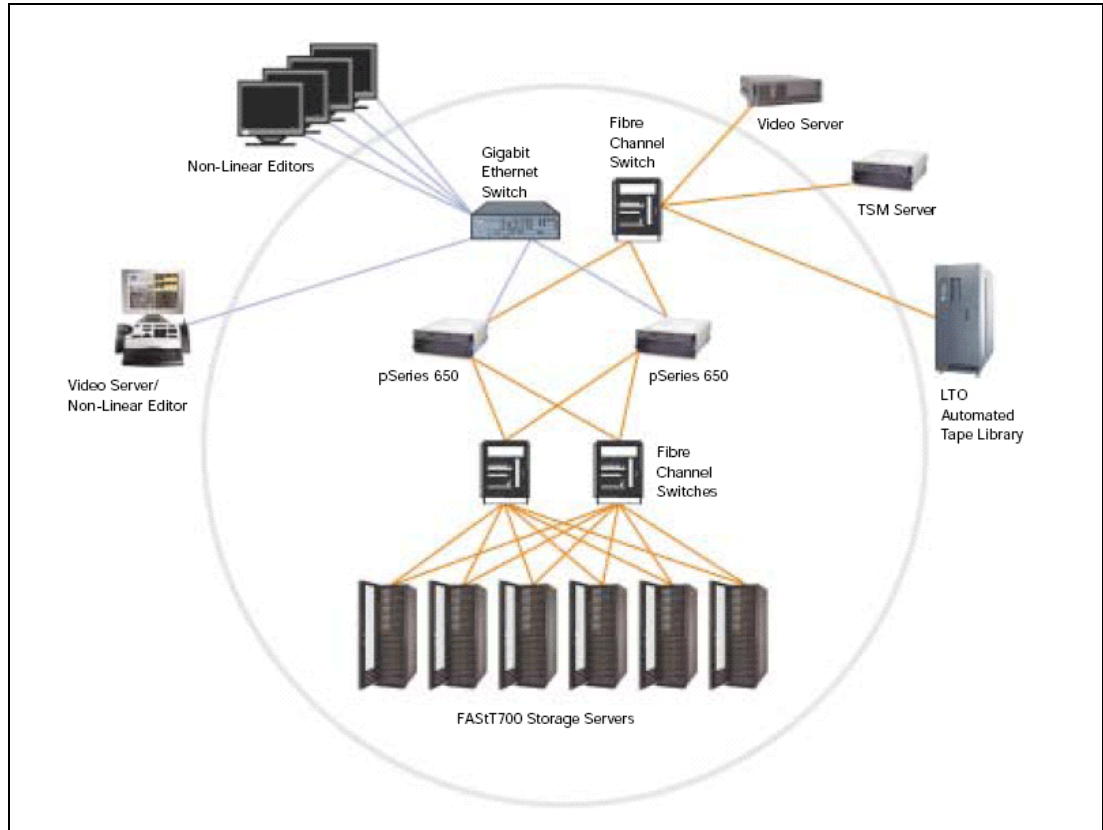


Figure 26-1 Digital Media Center components

- ▶ **pSeries servers:** pSeries servers offer an excellent platform for deploying business-critical applications requiring performance, reliability, scalability and manageability.
- ▶ **xSeries servers:** xSeries systems help provide critical Intel processor-based server solutions. xSeries servers give you outstanding reliability through mainframe-inspired technologies and intelligent management tools, available in systems ranging from affordable, edge-of-network servers to enterprise-class machines. These rack-optimized servers help solve the datacenter space dilemma by delivering outstanding rack performance and manageability in a slim chassis, and can deliver outstanding levels of availability, upward scalability and transaction performance.
- ▶ **GPFS:** GPFS (General Parallel File System) is a high-performance, shared-disk file system available on pSeries and xSeries servers with AIX and Linux, operating systems. Clustered applications impose stout requirements on the file system - parallel applications, for example, need access to the same data from multiple nodes. GPFS provides this access and meets other requirements through features such as parallel data and metadata access, reliability and fault tolerance, online system management, uniform access across the cluster and high throughput.
- ▶ **IBM Tivoli Storage Manager:** IBM Tivoli Storage Manager represents a comprehensive storage management solution suite to help protect data integrity, minimize costs and maximize productivity. It provides Hierarchical Storage Management (HSM) functions for Digital Media Center by providing the intelligence to move files along a hierarchy of storage devices that are ranked in terms of cost per megabyte of storage, speed of storage and retrieval, and overall capacity limits. Based upon established policies, files are migrated along the hierarchy to less expensive forms of storage based on rules tied to

the frequency of data access, file size or both. File migration and retrieval remains transparent to users.

- ▶ **IBM Ultrium 2:** Low cost, high-capacity tape systems are essential for backing up digital media files. The new IBM TotalStorage Ultrium 2 Tape Drive will double the capacity and more than double the performance compared to the first-generation IBM Ultrium offerings.
- ▶ **Tivoli Storage Resource Manager:** The policy-based Tivoli Storage Resource Manager can help detect potential problems and automatically make adjustments—without human intervention—increasing productivity and saving costs.

26.1.2 Digital content creation

The digital content creation solution for media and entertainment from IBM is flexible and can be configured for a one- or two-workstation shop, as well as for an entertainment enterprise. It is part of the IBM Digital Media Factory, a framework of IBM and Business Partner technologies that employs open systems and modular components to simplify complex implementations.

This solution can:

- ▶ Fuel artists' and animators' creativity with leading-edge tools for 2D/3D, animation, effects, paint, high-end graphics, images and rendering, and non-linear editing capabilities.
- ▶ Shorten content creation cycle times, allowing greater output and creativity.
- ▶ Decrease the per unit cost of rendering images.
- ▶ Deliver more value from your content assets through quicker, universal distribution and management that is more effective.
- ▶ Eliminate the time spent troubleshooting among different technology vendors—we coordinate all of the elements of the solution, which are integrated and supported by IBM experts.
- ▶ Facilitate the transition of studio production to a Linux platform

In addition IBM can provide project management support to coordinate all aspects of delivery and installation, including hardware and software setup services. Warranty service is offered on all components as well.

26.1.3 Digital Video Surveillance and Security

Digital Video Surveillance and Security solution provides a fully automated, end-to-end video asset management and monitoring solution for rapid access and retrieval of digital video content. The Digital Video Surveillance solution consists of the following components:

- ▶ IBM DB2 Content Manager
- ▶ WebSphere Application Server
- ▶ Tivoli Storage Manager
- ▶ IBM @server systems
- ▶ IBM TotalStorage Tape and Network Attached Storage
- ▶ Applications from leading security and surveillance vendors
- ▶ Implementation, consulting and network integration services from IBM Global Services

For more information on digital video security, refer to the following Web links:

<http://www.ibm.com/solutions/digitalmedia/doc/content/bin/Flexibleend-to-endsolution.pdf>
http://www.ibm.com/servers/storage/solutions/digitalmedia/solutions/video_security.html

26.2 Support

IBM Global Services and Finance, the world's largest information technology services provider, integrates a broad range of IBM capabilities services, hardware, software and research to help companies of all sizes realize the full value of information technology (IT).

IBM actively seeks to partner with industry-leading Independent Software Vendors (ISVs) who are key to the media creation space. You can find the current list of partners with applications certified on IntelliStation® workstations at the following Web link:

http://www-1.ibm.com/servers/storage/solutions/digitalmedia/pdf/certified_apps.pdf

26.3 More information

To find out more about the IBM Digital Media solutions, we invite you to talk to your IBM Sales Representative, or visit:

<http://www.ibm.com/industries/media>

<http://www.ibm.com/industries/digitalmedia>



Solutions for the life sciences

End-to-end infrastructure solutions from IBM provide the scalable tools and systems to help life sciences companies access, manage and develop content. The IBM Life Sciences organization offers a comprehensive set of innovative IT infrastructure products and services designed to create value for drug discovery, pharmaceutical and biotech companies, health care organizations and academic research institution.

A key component of a top-quality life sciences solution is reliable, disaster-tolerant storage. With scalable IBM storage systems, laboratories can increase storage capacity and reduce back-up time and bottlenecks. IBM solutions provide multi-platform accessibility and the flexibility for a range of applications that meet research and development computing needs.

IBM server and storage solutions coupled with IBM Life Sciences Consulting services and expertise are designed to maximize laboratory productivity and minimize cost. Key features, in terms of business value, include:

- ▶ Expediting time to market by increasing R&D productivity
- ▶ Highly available storage minimizes downtime
- ▶ Highly accessible data through networking
- ▶ Investment protection via control of data and network management costs
- ▶ Many data protection features
- ▶ High degree of scalability
- ▶ Complete line of open, scalable and reliable products

27.1 Life sciences solutions and services

Customers' storage concerns are seldom one-dimensional. IBM offers key solutions to address the specific needs of small, mid-sized, and large research organizations or pharmaceutical firms. The solutions are designed to provide a high-performance infrastructure that can easily scale to meet the growing demand for storage, regardless of the size of the organization.

Common to all solutions is a networked storage architecture to address the data intensive requirements and anticipate the rapidly changing environment of emerging research and discovery organizations. This networked storage architecture is designed to run heterogeneous application servers, provide high performance, reliability, scalability, availability, and ease of management. It also helps provide investment protection to our customers.

These solutions may also come with IBM DB2 database software, and IBM Tivoli management software, as well as other components included in essential SAN technologies, such as routers, switches, and host bus adapters.

Research and discovery are driving a data explosion resulting in the need for continuously available and highly functional storage systems. This IBM solution offers a means to optimize the use of storage through consolidation and resource optimization, while at the same time providing automated data protection.

Along with these solutions, IBM provides a wide range of consulting services that embrace and extend the value of these life sciences solutions. The services offered can help address current and future infrastructure needs, optimize key business applications and leverage existing storage, while helping to integrate new storage networking technologies to help reduce management costs. IBM Life Sciences Storage Services includes:

- ▶ Life Sciences Storage Health Assessment to understand the solution environment and to optimize performance of database, backup/restore and storage subsystems
- ▶ Life Sciences Storage Planning and Design to provide a blueprint for implementation and for managing security and scale
- ▶ Life Sciences Storage Implementation to provide implementation, verification and training for storage, storage management and software associated with the solution

Training is also available to help build skills needed for ongoing support. The components of the solution are modular to allow them to grow and adapt as business requirements evolve.

27.2 Storage for medical imaging

The effects of computers are widespread in all aspects of medicine, especially within radiology. Digital imaging has led to a revolution within the medical arena: the development of the Picture Archival and Communication System (PACS).

PACS encompasses a broad range of technologies that enable digital radiology. At its most fundamental level, PACS represents the integration of image acquisition devices, display workstations and storage systems all connected via a computer network. Current products allow completely digital acquisition, transfer, interpretation, storage, and transmission.

27.2.1 PACS overview

Current PACS products allow completely digital acquisition, transfer, interpretation, storage and transmission of medical images. Since digital medical images are typically very large and the volume of data produced by an imaging department depends on the type, quantity and distribution of studies performed, image volume can present a storage capacity challenge. In addition, the images must be safely stored for a long time and easy to access: hence the requirement for a reliable, flexible, high-capacity storage system to complement a hospital's PACS applications.

IBM TotalStorage and PACS

IBM TotalStorage products are highly scalable and can be configured to meet current and future medical imaging storage needs of health care providers. IBM storage offers disaster-tolerant systems that help avoid single points of failure and support high availability and protection for digital image files. IBM storage solutions are built on best practices to help health care institutions simplify storage management and they support open industry standards for easy integration with multivendor systems.

IBM TotalStorage quality products for digital imaging

IBM offers disk, tape, storage networking and storage management software products to help maximize PACS productivity and support lower total cost of ownership. Built on modular, open and integrated architectures and standards, some of the benefits include:

- ▶ Data sharing through storage networking
- ▶ Fast access for patient diagnoses
- ▶ Centralized storage management for easier management
- ▶ Fibre Channel connectivity to support increased network distances
- ▶ Highly scalable architecture to support capacity growth
- ▶ High-availability configurations for disaster tolerance
- ▶ LAN-less and server-less backups to provide data protection
- ▶ Heterogeneous server support for investment protection
- ▶ High capacity for long-term record retention
- ▶ Copy services to support data protection

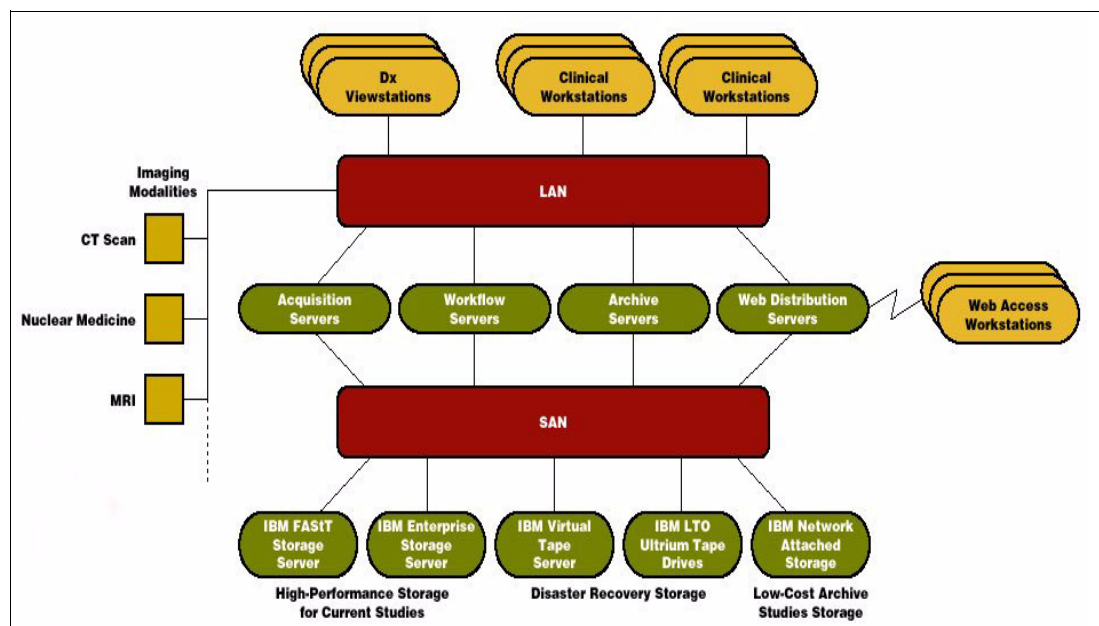


Figure 27-1 PACS storage solution

27.3 IBM Medical Archive Solution

Storage requirements to retain medical imaging data can become very substantial. Storage requirements can range from a few terabytes for small medical treatment facilities to hundreds of terabytes for large/multiple treatment facilities or imaging centers. Access requirements for medical image data can greatly depend on the age of a performed exam, which will generally affect the data flow and types of storage utilized to satisfy a particular long-term archive requirement.

IBM Medical Archive Solution (MAS) consists of the following components:

- ▶ IBM @server
- ▶ IBM TotalStorage
 - IBM online disk storage
 - IBM near-line tape storage
- ▶ IBM Tivoli Storage Manager Server
- ▶ IBM Tivoli Storage Manager for Space Management

The IBM MAS can utilize several different types of online storage in an effort to provide the most efficient storage to meet performance requirements, while providing the most cost-effective solution possible. In standard and typical installations SCSI, Fibre Channel (FC), and Serial ATA (SATA) disks are used for storage of medical image reference data. Using a tiered, hierarchical storage management (HSM) approach, the IBM MAS can move data to more cost-efficient storage based on data age, retention, and accessibility requirements. An example of a typical tiered storage could resemble the following data ages:

- ▶ Zero to 2 weeks old, Tier 1: Internal Ultra320 SCSI disks
- ▶ Two weeks to three years old, Tier 2: 2 GB FC disks
- ▶ Three years or older, Tier 3: SATA disks

As medical image reference data ages, it is generally more cost effective to migrate the primary copy of the data to slower performing spinning storage. The option also exists to insert tape, DVD, or other near-line and offline storage technology into the tiered storage approach. These media types can be used to generate backup and/or disaster recovery copies of image data, and to migrate primary copies of very old data from SATA storage to free up spinning disks for newer data.

In addition to the standard online storage types mentioned above, the IBM MAS can also utilize the new IBM Enterprise disk storage DS6000 series and DS8000 series for any level tier storage.

27.4 More information

For more information on Life Science solutions, refer to this Web page:

<http://www-1.ibm.com/industries/healthcare>

IBM Medical Archive Solution, REDP-9130, discusses the societal and technological factors for developing an image management solution. This Redpaper can be found at:

<http://w3.itso.ibm.com/abstracts/redp9130.html>



Part 7

Appendixes



Introduction to SMIS and CIM/WBEM

SMIS

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* (SMIS).

SMIS 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. SMIS 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, SMIS 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 languages 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 A-1 on page 461 shows a typical interactions between the components in the CIM/WBEM environment. 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 on 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.

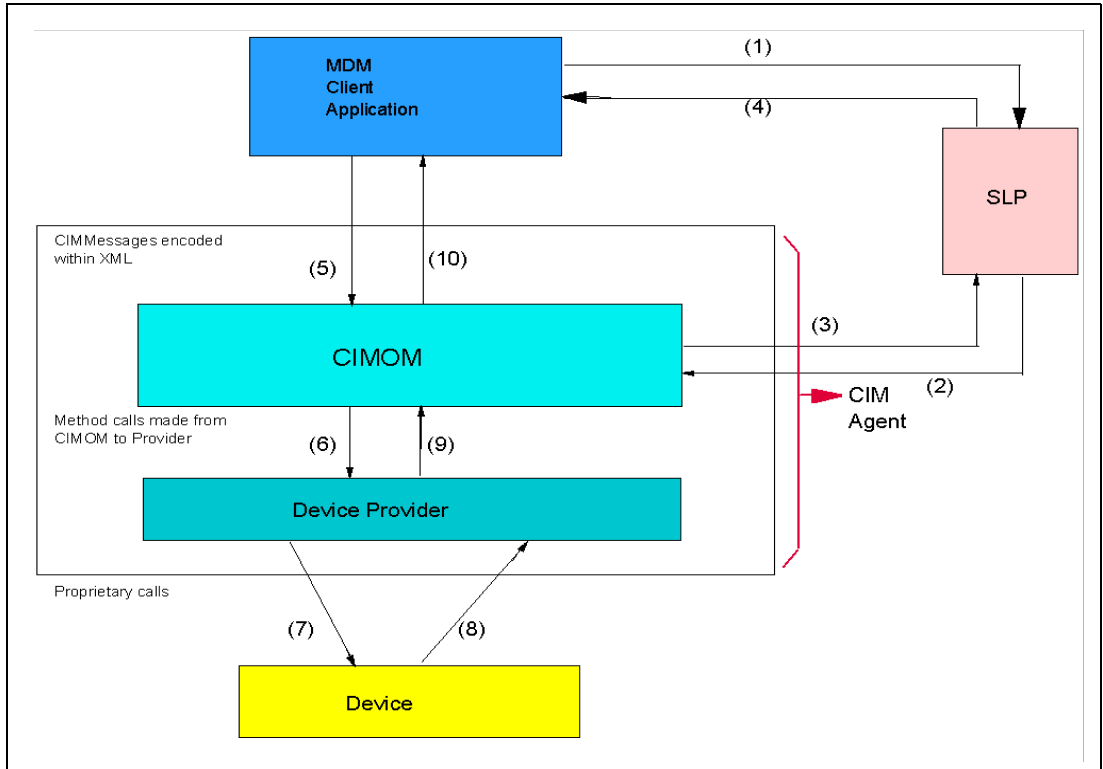


Figure A-1 Component Interaction in CIM Environment

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:

- ▶ 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 B-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.

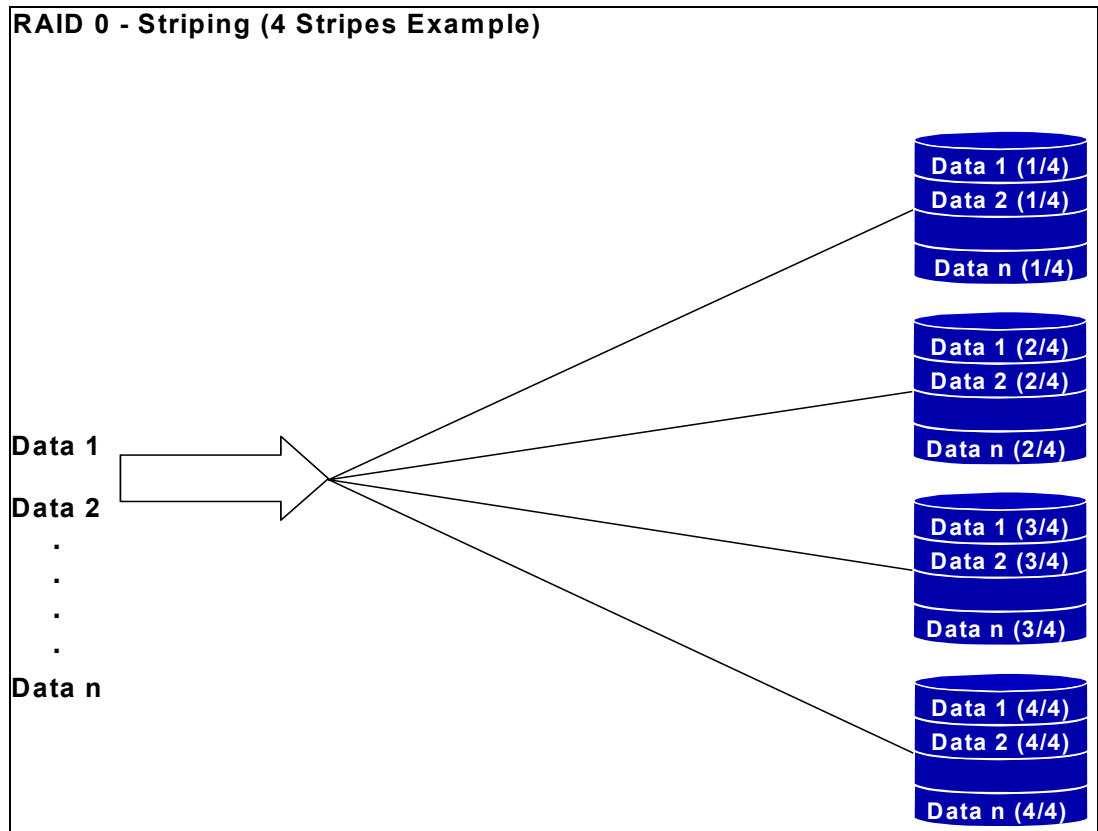


Figure B-1 RAID 0

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 B-2) employs data mirroring to achieve redundancy.

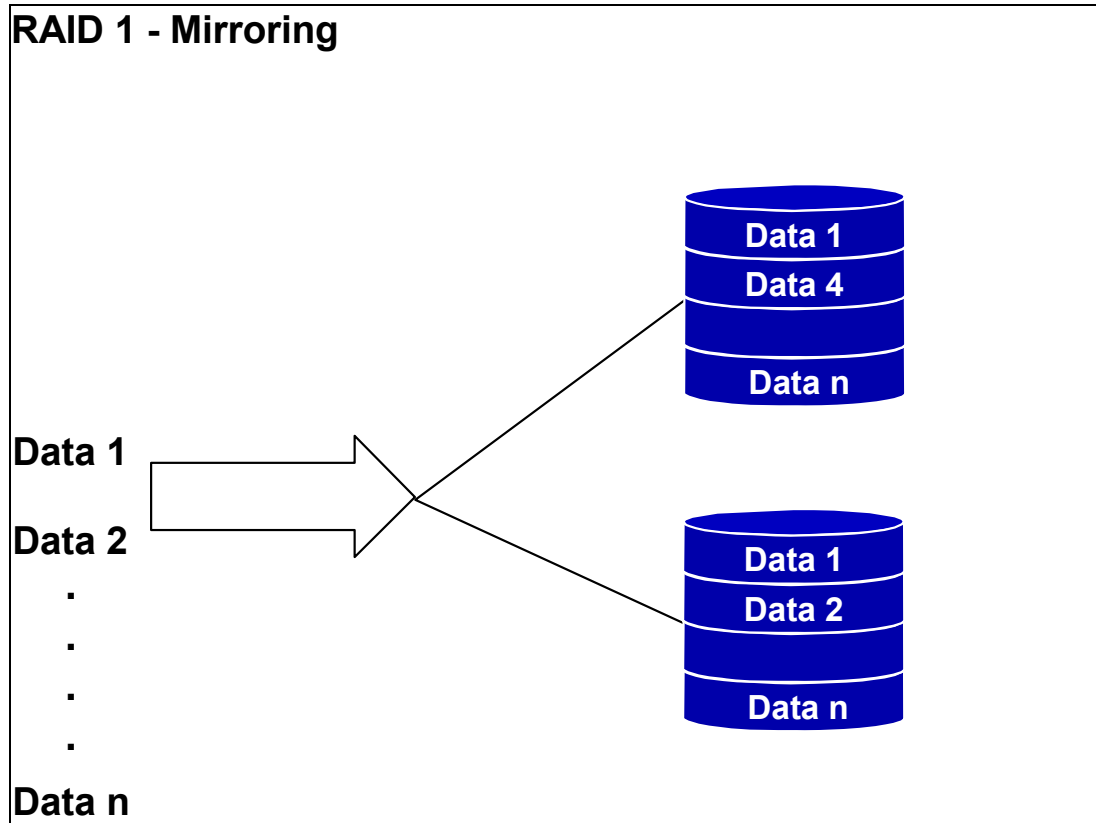


Figure B-2 RAID 1

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 B-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.

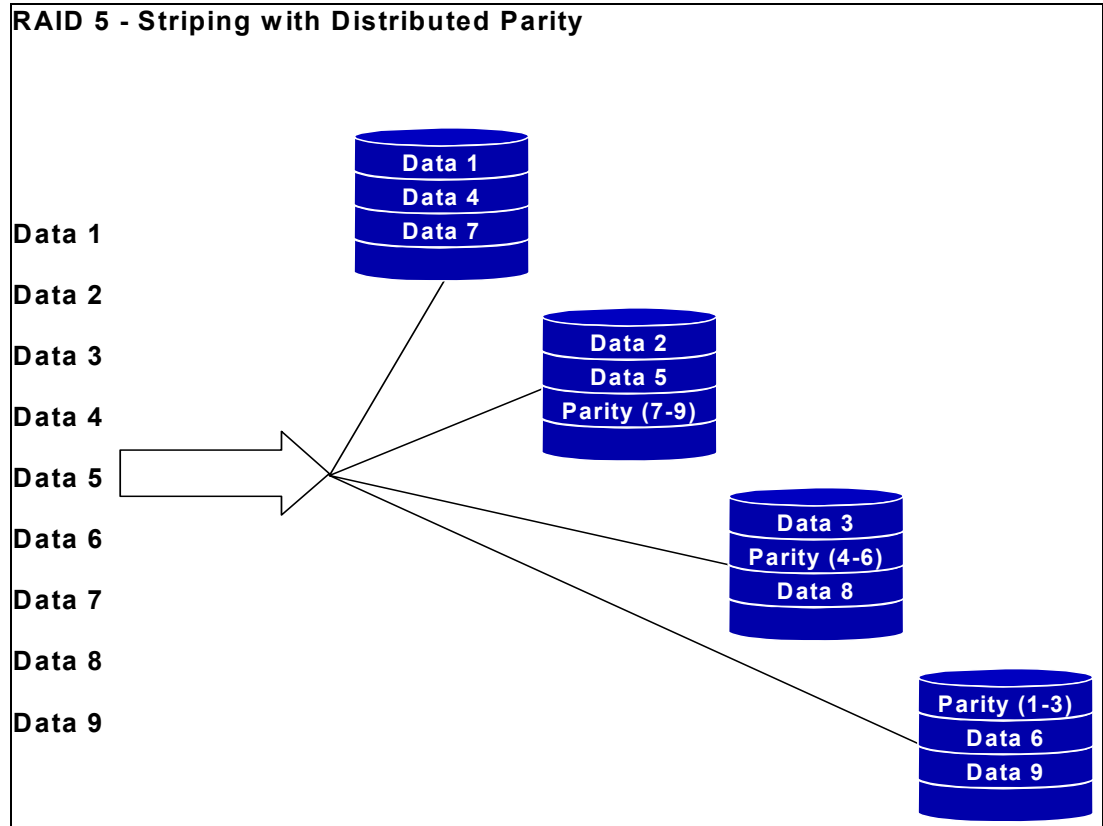


Figure B-3 RAID 5

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 B-4 is a combination of multiple, local RAID-0 arrays. It is possible RAID-00 set with two disks because it is possible to create a RAID-0 array using only one disk (and applying the second array level requires a minimum of two local arrays).

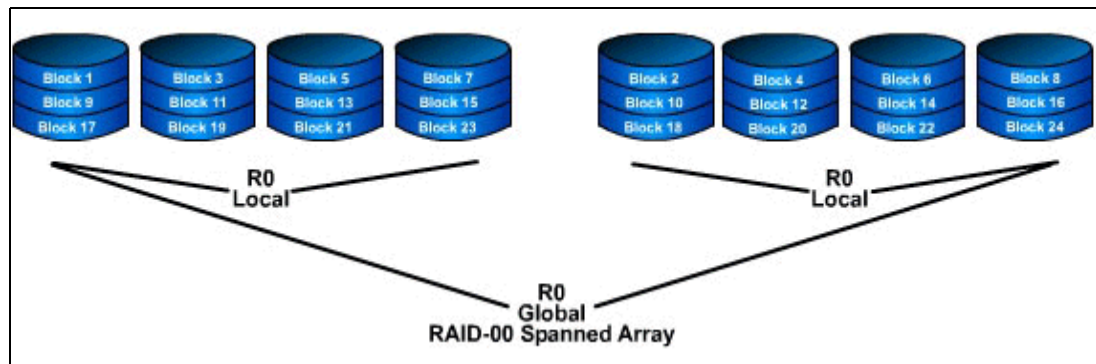


Figure B-4 RAID-00

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 B-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.

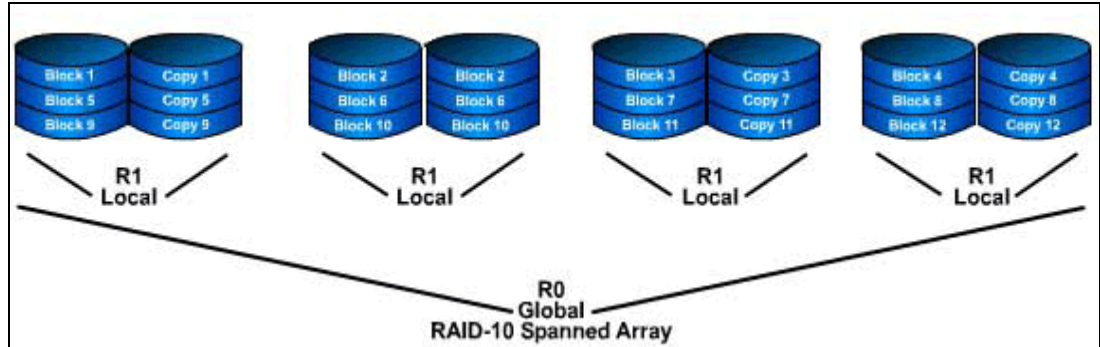


Figure B-5 RAID-10

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 B-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)

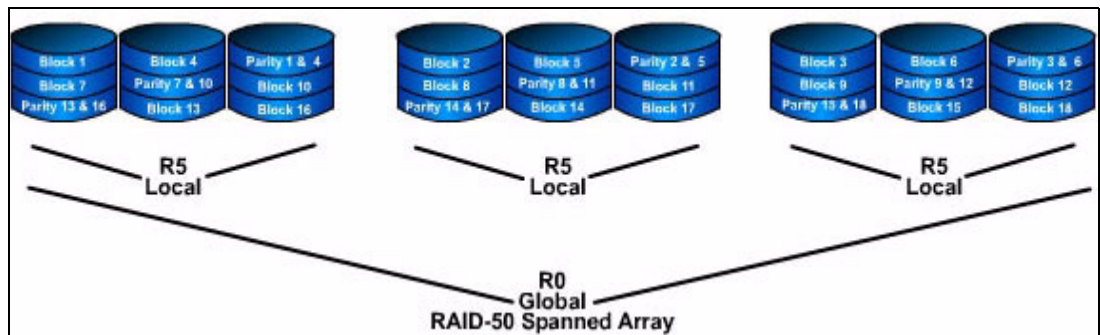


Figure B-6 RAID-50

For more information on RAID Levels see the following site:

<http://www.acnc.com>

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	LVDD	Logical Volume Device Driver
IDE	Integrated Drive Electronics	LVM	Logical Volume Manager
IEEE	Institute of Electrical and Electronic Engineers	MAN	Metropolitan Area Network
IETF	Internet Engineering Task Force	MBR	Master Boot Record
IFCP	Internet Fibre Channel Protocol	MB/s	Megabytes per second
IGMP	Internet Group Management Protocol	MCA	Micro Channel® Architecture
IGS	IBM Global Services	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	Multi-protocol 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	Journalized 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
		ODBC	Open Database Connectivity
		ODM	Object Data Manager

OLTP	OnLine Transaction Processing	RMC	Resource Monitoring and Control
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
		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
VT	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

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.

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.

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.

MDM See Multiple Device Manager.

MDS See Metadata Server Engine.

Metadata Server Engine Within SAN File System, an MDS is the physical hardware on which a Metadata server and an Administrative server run.

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.

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 A hardware based remote copy option that provides a synchronous volume copy across storage subsystems for disaster recovery, device migration, and workload migration.

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 Peer-to-Peer remote copy.

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 FASTT product line.

Remote volume mirroring A hardware based remote copy option that provides a synchronous volume copy with the FASTT 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.

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 481. Note that some of the documents referenced here may be available in softcopy only.

- ▶ *IBM TotalStorage Enterprise Storage Server Model 800*, SG24-6424
- ▶ *IBM TotalStorage Enterprise Storage Server Implementing ESS Copy Services with IBM @server zSeries*, SG24-5680
- ▶ *FICON Native Implementation and Reference Guide*, SG24-6266
- ▶ *IBM TotalStorage Enterprise Storage Server PPRC Extended Distance*, SG24-6568
- ▶ *The IBM TotalStorage DS6000 Series: Concepts and Architecture*, SG24-6471
- ▶ *The IBM TotalStorage DS8000 Series: Concepts and Architecture*, SG24-6452
- ▶ *iSeries in Storage Area Networks: Implementing Fibre Channel Disk and Tape with iSeries*, SG24-6220
- ▶ *IBM TotalStorage DS4000 Series and Storage Manager 9.10*, SG24-7010
- ▶ *IBM TotalStorage FAStT Best Practices*, 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
- ▶ *The IBM TotalStorage Tape Selection and Differentiation Guide*, SG24-6946
- ▶ *IBM TotalStorage Tape Libraries Guide for Open Systems*, SG24-5946
- ▶ *Implementing IBM Tape in Open 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
- ▶ *Tivoli Storage Manager Version 5.1 Technical Guide*, SG24-6554

- ▶ *IBM Tivoli Storage Resource Manager: A Practical Introduction*, SG24-6886
- ▶ *DFSMSHsm ABARS and Mainstar Solutions*, SG24-5089
- ▶ *IBM TotalStorage Solutions for Disaster Recovery*, SG24-6547
- ▶ *Understanding the IBM TotalStorage Open Software Family*, SG24-7098
- ▶ *IBM TotalStorage SAN Volume Controller*, SG24-6423
- ▶ *IBM TotalStorage: Introducing the SAN File System*, SG24-7057
- ▶ *IBM TotalStorage Productivity Center: Getting Started*, SG24-6490
- ▶ *Implementing the IBM TotalStorage SAN Volume Controller Storage Software on the Cisco MDS 9000*, SG24-7059
- ▶ *IBM TotalStorage: Implementing an Open IBM SAN*, SG24-6116
- ▶ *IBM Tivoli Storage Area Network Manager: A Practical Introduction*, SG24-6848
- ▶ *IBM Tivoli Storage Manager Implementation Guide*, SG24-5416
- ▶ *IBM Enterprise Storage Server*, SG24-5465
- ▶ *Understanding the IBM TotalStorage DR550*, SG24-7091

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM Tivoli Storage Manager V5.3 for Windows Administrator's Guide*, GC32-0782
- ▶ *IBM Tivoli Storage Manager V5.3 for Windows Installation Guide*, GC32-1602
- ▶ *IBM TotalStorage Expert Installation Guide*, GC26-7436
- ▶ *IBM Tivoli Storage Manager for Windows Backup-Archive Client Installation and User's Guide*, GC32-0788

Online resources

These Web sites and URLs are also relevant as further information sources:

- ▶ IBM TotalStorage Web Site
<http://www.storage.ibm.com/>
- ▶ IBM Redbooks
<http://www.redbooks.ibm.com/>
- ▶ IBM Tivoli Storage Manager
<http://www.ibm.com/software/tivoli/>
- ▶ IBM TotalStorage Productivity Center
<http://www-1.ibm.com/servers/storage/software/center/index.html>
- ▶ IBM Storage networking products
<http://www.storage.ibm.com/snetwork/index.html>
- ▶ IBM TotalStorage proven
<http://www-1.ibm.com/servers/storage/proven/Tivoli>
- ▶ LTO Technology Web site
<http://www.lto-technology.com/>

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

2027-256 201
2027-C40 200
2027-R16 199
2045-N16 204
2062-D04 202
2062-D07 203
2109-A16 197
2109-M14 196
3494 139
 Base Frame 140
 Enterprise Tape Library 139
 High Availability Frame 142
 S10 tape storage 142
 Tape Drive Expansion Frame 141
 Virtual Tape Server 143
3580 model L33 113
3582 117
3584 123
3588 124
3588 Model F3A 109
3590 133
 cartridge 139
 H1A 135
3592 128, 405
 cartridge 139
3953 132
3953 Tape Frame and Library manager 132
7133 98
7204 External Disk Drive 152
7205 152
7206 153
7207 154

A

Accelis 108
Access Control List 25
accessor 127
activation plan for DRP 426
adaptive subfile backup 330
Administration Center 324, 326
 wizards 326
Advanced Library Management System (ALMS) 127
Advanced Metal Evaporated (AME) 155
Advanced Performance Monitoring 183
Advanced Policy Management 146
Advanced Provisioning 250, 258
Advanced Security 189
Advanced Security (AS) 183
Advanced Web Tools 182
Advanced Zoning 183
agent 259
Agent Manager 280
AIX 349

alert 267–268
Alert Manager 54
allowable outage times 419
Alternate Pathing Support 135
alternative sites 422
Andrew file system (AFS) 235
ANSI 163
any-to-any 167
API 318, 331
appliance 177
Application Programming Interface (API) 444
arbitrated loop 62, 165
archive 317
archive data 402
Archive management 7
archive retention
 chronological 406–407
 event based 407
Array Across Loops (AAL) 75, 82–83
Array Support Library (ASL) 46
asset management 389, 392
asset reporting 264, 266
asymmetry compensation 110
Asynchronous Cascading PPRC 92
Asynchronous data transfer 164
Asynchronous Mirroring 51
Asynchronous PPRC 90
Asynchronous PPRC-XD 89
Asynchronous Transfer Mode (ATM) 171, 184
ATM. See Asynchronous Transfer Mode
Automated Cartridge Facility (ACF) 135
automated management 2
Automatic Failover 135
Automatic file system extension 257
availability reporting 263, 266
avoidance system 419

B

backhitch 110
backup and restore of data 317
backup reporting 267
backup topology 335
bandwidth 165–166
barcode reader 119–120, 125
barcode scanner 118
bare machine recovery 316
bare metal recovery 331, 344
base logical drive 47
Basic Sequential Access Method (BSAM) 369
battery 31, 78
battery backup unit (BBU) 68
BBU 68
BCP 413
 business process 413

- BIA 419, 421
 - allowable outage times 419
 - disruptions impacts 419
 - recovery priorities 419
 - risk management 417
- block aggregation 226
- BM 7206 Model VX2 154
- Boot from SAN 25
- boss node 223
- Brocade (b-type) 182
- BRP 413
- buffer credits 170
- Business Continuity 1–2, 207, 209, 411
 - Plan 413
- Business Continuity Plan 413
- Business Impact Analysis (BIA) 357
- Business Recovery
 - Plan 413
- Business Recovery Plan see BRP
- Business Resumption 412

C

- cache hit. 390
- cache holding time 391
- Calibrated Vektored Cooling 68
- calibration sensor 126
- call home 123
- capacity management 390, 393
- Capacity on Demand (CoD) 127
- capacity reporting 266
- Cartridge Memory 108
- cascade 166–167
- CDB 172
- chargeback 268
- CIFS 162, 177–178
- CIM 272
 - agent 253, 272, 299
 - object manager 272
- CIMOM 460
- CIMS 268
- Cisco CWDM 179
- Cisco MDS 9000 283
- Cisco MDS 9506 Multilayer Director 202
- Cisco MDS 9509 Multilayer Director 203
- CKD volume 65
- CLI 85
- client options file 354
- Cloudscape 278, 282
- cluster 75
 - cache 225
 - node 66
 - resources 353
- clustering 207–208, 349, 353
- clutch 110
- CMIP 459
- CNT (n-type) 182
- CNT UltraNet Multi-service Director 205
- cold sites 422
- Command Line Interface. See CLI.
- Command Line Interface. See CLI.

- Common Base 298
- Common Information Model (CIM) 272, 459
- Common Information Model/Object Messaging (CIM/OM) 258
- Common Internet file system (CIFS) 235
- Compass architecture 225
- Compliance framework 403
- compression technology 109
- consistency 28
- Consistency Group 51, 87
- console 278
- consolidation 2, 59, 93
- contingency plan 418
- Contingency Planning 412
- Continuity of Operations Plan see COOP
- Continuous Operations 3
- Control Path Failover 121
- Control Procedures 420
- control statements 384
- Control Unit Port (CUP) 197
- Controller service mode 46
- cooling 78
- COOP 413
- Copy Services 65, 86
- copy storage pool 346, 354
- copy-on-write 24, 48
- cost of downtime 355
- Cost of Service (CoS) 174
- Cristie 331
- critical records 420
- CSV output 265
- CWDM 179

D

- DASD 366
- data
 - backup and restore 317
 - security 356
- Data agent 253
- Data Backup with no Hot-Site 429
- data buffer 109
- data migration 89, 229
- data path 63
- Data Path Failover 121
- data protection 208
- data protection phase 420
- data protection solution 208
- data retention 7, 318, 341
 - event based 408
- data set 366
- data striping 465
- database
 - asset reporting 264
 - trend reporting 264
- dataspace sorting 384
- DB2 318, 337
- DBCS Ordering 382
- deletion
 - hold 409
 - release 409

- delta logging 28
- DES 331
- Device Manager 249, 298
- DFSMS 234, 372
 - Advanced Copy Services 371
 - overview 372
- DFSMSdfp 364, 367, 370
 - data class 367
 - management class 368
 - storage class 367
- DFSMSdss 371
 - overview 371
 - Stand-Alone Services 372
- DFSMSHsm 372
 - Availability Management 373
 - Space Management 373
 - tape management 374
- DFSMSrmm 375
 - environment 379
 - location management 377
 - overview 376
 - policies 379
 - validation 377
 - volume management 376
- DFSORT 382
 - ICETOOL 383
- DFSORT/MVS
 - DFSORT symbols 383
 - ICEGENER 383
 - OUTFIL 384
- DFSORT/VSE
 - control statements 384
 - dataspace sorting 384
 - getvis 385
 - getvis sorting 385
- digital content creation 450
- digital media 448
- Digital Media Center 448
- Digital Video Surveillance and Security 450
- Disaster Recovery 3, 90, 342, 412, 442
 - activation of team 421
- Disaster Recovery Manager. see DRM
- Disaster Recovery Plan file 342
- Disaster Recovery planning
 - change management 415
 - critical needs 414
 - data collection 414
 - data protection 415
 - notification and activation procedures 424
 - project initiation 414
 - recovery plans 415
 - risk analysis 415
 - training and testing 415
- discovery 272
 - full 274
 - in-band 273
 - iSCSI 274, 277
 - topology 274
- disk
 - allocated space 262
 - instant copy 337
 - remote replication 356
- DISK CONTRL 69
- DISK EXP 69
- disk system
 - LUN 288
- disruption impacts 419
- Distributed Management Task Force (DMTF) 272, 459
- DLT 123, 152
- DMTF 272
- DR550 401
- DRM
 - PREPARE 344, 347
 - TSM clients 344
- DRP 414, 422
 - activation 424
 - activation plan 426
 - assumptions 416
 - data protection 420
 - end user recovery 421
 - maintenance and changes 424
 - network recovery 421
 - notification 424
 - project initiation 416
 - project phases 416
 - recovery plans 421
 - recovery strategy 421
 - systems recovery 421
 - team selection 416
 - testing and training 423
- DS CLI 61
- DS Command-Line Interface (CLI) 61, 85
- DS Management Console 84
- DS Open application programming interface (API) 85
- DS Storage Manager 84
- DS300 18
- DS400 18, 22
- DS4000 Storage Manager 45
- DS4100 29–31, 404
- DS4300 29
- DS4300 Turbo 29
- DS4400 29
- DS4500 29
- DS6000 58
- DS6000 expansion enclosure 68
- DS6800 66
- DS8000 59, 72
- DS8100 Model 921 79
- DS8300 Model 922 80
- DS8300 Model 9A2 73, 80
- DSUTIL 366
- dual accessor 127
- dummy carrier 67
- DWDM 88, 356
- Dynamic Capacity Addition 35
- Dynamic Capacity Expansion (DCE) 47
- Dynamic Load Balancing 135
- Dynamic LUN 65
- Dynamic Multipathing Driver (DMP) 46
- dynamic speed matching 110

Dynamic Volume Expansion (DVE) 35, 37, 47

E

EDFI 275, 293
EHPCT 139
electronic tape vaulting 355
electronic vaulting 348, 430
Element Manager 190–191
enclosure 78
Enclosure Services Module 24
end user recovery 421
endpoint devices 273
Enhanced Remote Mirroring (ERM) 50
Enterprise Fabric Connectivity Manager (EFCM) 183, 191, 199, 201
enterprise Remote Copy Management Facility (ERCMF) 432, 442
Enterprise Repository 259
Enterprise Storage Server (ESS) 58–59, 442
Enterprise Storage Server Expert 387
Enterprise Tape Library 139
Enterprise Tape Library Expert 387
Enterprise Tape products
 overview 132
error detection 275
Error Detection and Fault Isolation 293
ESCON 76, 134–135, 356
ESCON Director 167
ESM 24, 44
ESS 81, 334, 337
ESS Expert 86
ESS Master Console 84
ESS Model 750 83
ESS Model 800 82
ESS Specialist 83
Ethernet switch 77
EXP Plus 103
EXP Plus 320 103
EXP100 40
EXP400 24
EXP700 43
 upgrading 44
EXP710 29, 44
expansion enclosure 68
expansion frame 78
expansion unit
 EXP400 24
EXTENDED COPY 337
Extended Fabric feature 183
Extended Remote Copy XRC 90
External Disk Drive 152

F

fabric 164–167
Fabric agent 253
Fabric Manager 183, 188, 196, 280
Fabric Services Architecture 182
Fabric Shortest Path First (FSPF) 192, 203
Fabric Watch 183, 196

failover 223, 349, 352
fanout 165
FAST 92
 expansion units 43
 service alert 31, 35, 53
FAStT 200 30, 32, 34, 36, 39
 Specification 32, 34, 39
FAStT 700 37
 overview 37
 specification 38
FAStT Management Suite Java 54
FAStT MSJ 54
FAStT600 29
FAStT600 Turbo 29
FAStT700 29
FAStT900 29
fast-write 67
fault detection 294
fault isolation 275
fault management 275
fault-tolerant 61
FC-AL 62, 66, 75, 93
FC-FC Routing Service 198
FCIP 171, 173, 175, 179
FCIP Tunneling Service 198
FC-MI 272
FCP 67
feature key 49
Fiber Channel over IP 173
Fibre Alliance 272
Fibre Channel 163–166
Fibre Channel Arbitrated Loop 66
Fibre Channel network 273
fibre extenders 356
FICON 61, 67, 76, 134–135, 196, 356
FICON CUP Activation 197
FICON In-Band Management Server 193
FICON Management Server 199
file copy migration 245
file placement 240
file system 235
filemark 111
files
 most at risk 265
 statistics 264
fileset 239
filesystem capacity 262
fixed-length records (FLR) 385
FlashCopy 24, 37, 48, 86, 93, 223, 228, 241, 244, 318, 334, 337
 Data Set 87
 Incremental 87
 Multiple Relationship 87
FlashCopy repository 48
FlashCopy service 444
FlexPar 201–202
FlexPort 183
free space 262
full duplex 164
full synchronization 52

G

- gateway 178
- gauge 303
- GDOC 438
- GDPS/PPRC 433
 - management for Open Systems LUNs 435
 - Multi-Platform Resiliency for zSeries 435
 - Near Continuous Availability of data with HyperSwap 434
 - overview 433
- GDPS/PPRC HyperSwap Manage
 - overview 435
- GDPS/XRC 433
 - overview 436
 - topology 436
- getvis 385
- Global Copy 28, 51, 88–89
- Global Mirror 52, 88, 90
 - Utility 443
- Global Mirroring 28, 52
- global spare 67
- GMU tool 443
- GPFS (General Parallel File System) 449
- grain 224
- group
 - collocation 322

H

- HACMP 349, 405
- HACMP/XD
 - hardware requirements 440
 - service and ordering 441
 - software requirements 440
 - solution components 440
 - solution description 439
- Hardware Management Console 74, 77
- hardware zoning 275
- HBA 273
- head actuator 109
- head cleaner 111
- health monitoring 299
- heartbeat 349
- Hierarchical File System 370
- Hierarchical Space Management (HSM) 316, 319
- Hierarchical Storage Management 7
- Hierarchical Storage Manager 372
- High Availability 3, 349
- High Voltage Differential (HVD) 116
- high-availability server clustering 208
- hop 174
- hot sites 423
- hot spare 31, 33
- HotCAT 186
- hotsite 355
- hotswappable 191
- HPCT 139
- HTML output 265
- hub 165
- Hypervisor 73

I

- IBD 53
- IBM 7206 Model 110 153
- IBM 7206 Model 220 153
- IBM 7206 Model 336 154
- IBM 7207 Model 122 154
- IBM 7207 Model 330 154
- IBM 7208 Model 345 155
- IBM 7210 External DVD-RAM Drive 155
- IBM 7212 Storage Enclosure 156
- IBM 7212-312 VXA-2 autoloader 157
- IBM 7332 Model 110 157
- IBM 7332 Model 220 158
- IBM lossless compression algorithm 111
- IBM SAN Data Gateway 336
- IBM Tivoli Provisioning Manager 250
- IBM Tivoli Service Level Advisor (ITSLA) 281
- IBM Tivoli Storage Manager 116, 444
 - API 405
 - Client API 405
- IBM Tivoli Storage Manager for Data Retention 406
- IBM TotalStorage
 - Business Continuity solution 432
 - DS4000 444
 - DS4000 Integrated Backup 445
 - DS4000 Storage Server 445
 - Enterprise Storage Server 442
 - SAN Volume Controller 443
- IBM TotalStorage Expert 86, 388
- ICEGENER 383
- IDCAMS 370
- IEBCOMPR 370
- IEBCOP 370
- IEBDG 370
- IEBEDIT 370
- IEBGENER 370
- IEBPTPCH 370
- IEBUPDTE 370
- IEHINITT 370
- IEHLIST 370
- IETF (Internet Engineering Task Force) 276
- iFCP 174, 176, 179
- ILM 1–2
- image backup 339
- in-band 273
- Incident Response Plan see IRP
- Information Lifecycle Management 1–2, 402
- Information on demand 2
- information on demand 1
- Infrastructure Simplification 1–2, 207, 209
- infrastructure simplification 184
- instant copy 241, 337
- Integrated Backup for Databases (IDB) 53
- Integrated Service Console 320
- Integrated Solutions Console 324
- intelligent disk subsystem 318
- Interactive Storage Management Facility (ISMF) 365
- intermix 46
- internet Storage Name Service (iSNS) 276–277
- Interoperability Matrix 96

- inter-switch link 167
- invoices 268
- IP network management 284
- IRP 413
- ISCSI 171
- iSCSI 18, 21, 170, 172, 175, 179, 276, 285, 357
- ISL Trunking 184
- ISM 366
- ISMF 365
- iSNS 276
- ISV 331
- ITSLA 281

J

- J cartridge 139
- JBOD 44
- Job Scheduler 259

K

- Kerberos 331

L

- lambda 178
- LAN-free backup 330, 335
- launchpad 248
- LDAP 237
- library partitions, 119
- LIC 93–94
- license
 - REGISTER LICENSE 316
- licensed internal code (LIC) 393
- Licensed Machine Code 94
- life cycle 402
- Life sciences 454
- line module 201
- linear density 109
- Linear Tape-Open (LTO) 107
- Linux 31
- Load Balancing 121
- load-balancing 223
- logical drive
 - primary 48
- logical partition 73, 76
- logical subsystem 65
- logical topology display 287
- logical unit (LU) 224
- logical volume 289, 340
- Logical Volume Storage Agent 340
- loop 164–165
- Lotus Domino 318
- Lotus Notes 318
- Low Voltage (LVD) 118
- Low Voltage Differential (LVD) 119
- LPAR 73, 76
- LTO 108
 - overview 112
- LTO Ultrium 3 109
- LTO-CM 111

- LTO-DC 111
- LUN 224, 288
- LUN expansion 25

M

- Magstar 133
 - 3590 135
- MAN 356
- Managed disk 224
- managed disk group (MDG) 224
- Management Information Base (MIB) 273
- management station 23
- Master Console 84, 242
- McData (m-type) 182
- MDisk 224
- MDS 9120 193
- MDS 9140 193
- MDS 9216 194
- MDS 9506 202
- MDS 9509 203
- Medial Archive Solution 456
- medical image 455
- metadata 236
- metadata controller 234
- Metadata server 236, 241
- metal particle 109
- Metro Mirror 24, 88, 209, 228, 442
 - Tier 6 443
- Metro Mirror (Synchronous PPRC) 51
- Metro Mirroring 28, 51
- MIB 273, 277, 291
- Microsoft Cluster Server. see MSCS
- Microsoft Exchange 318
- Microsoft SQL Server 318
 - 2000 database 445
 - database 445
- Microsoft Volume Shadow Copy Service 444
- Microsoft VSS 444
- Midmarket tape products
 - Overview 108
- migration level 373
- mining 402
- mirrored sites 423
- mirroring 464
- mobile sites 423
- monitoring agent 23
- most at risk files 265
- MPP 46
- MS SQL
 - Server 445
- MSCS 352–353
 - and TSM 352
- Multipath 121
- Multipath architecture 117
- multipath architecture 119
- Multi-Path Proxy 46
- multipathing 60
- Multiple Allegiance 60
- Multiple Device Manager 249

- Multiple Device Manager (MDM) 298
- Multiple Device Manager Replication Manager 308
- Multiprotocol Services Module 195
- multi-tiered storage 62
- mutual suspicion algorithm 331

N

- namespace 234
- NAS 258, 339
 - directory level backup 339
 - TSM 338
- NAS Gateway 178
- NAS pooling 177
- NaviQuest 366
- NDMP 316
- NetView 284
 - explorer display 284
 - historical reporting 291
 - icon display 284
 - real-time reporting 292
 - reporting 291
- Network attached storage (NAS) 176
- network attached storage (NAS) 338, 357
- network bandwidth 348
- Network file system (NFS) 235
- network recovery 421
- Network Recovery Objective (NRO) 417
- NFS 162, 177–178
- NOLIMIT 407
- non-cascaded 167
- notification
 - personnel 425
- NTFS 337
- NVS 59, 67, 391
- nvsniffer 277

O

- Object Access Method (OAM) 369
- Occupant Emergency Plan see OEP
- OEP 413
- off-site backups 89
- off-site data storage 355
- off-site storage 342, 354, 357
- on demand 1–2, 72, 80
- on demand storage environment 5
- open software 6
- Open Trunking 192
- optical media 323
- Oracle 318, 337
- outband discovery 273
- OUTFIL 384
- out-of-band 273
- out-of-band virtualization 236

P

- Parallel Access Volumes (PAV) 60
- PAV 60
- Pay-as-you-grow 182

- Peer-to-Peer VTS 146
- performance analysis 300
- performance management 389–390
- Performance Manager 249, 298
- personnel 425
- PFA 64, 293
- pin retention 109
- ping 263, 266
- PL/I, 382
- POC 424
- Point-in-Time 48–49
- Points of Contact see POC
- point-to-point 164–165, 167
 - virtual 165
- Policy-based management 257
- polling interval 292
- power 78
- power management 110
- power subsystem 68
- power supply 68
- POWER5 72
- PowerPC architecture 66
- PPRC 356
- predictive failure analysis 75, 275, 293
- Predictive Failure Analysis (PFA) 64
- Preferred Path Option 192
- prevention
 - physical 420
 - procedural 420
- primary 48
- primary storage pool 346, 354
- probe 264–268
- probes 260
- procedures
 - recovery 427
- processor 66, 72
- processor complex 74
- profile 264
- progressive backup 330
- progressive incremental backups 340
- properties dialog 275
- provisioning 3, 235, 247, 250
- PtP VTS 146

Q

- Quality of Service (QoS) 174
- Queued Sequential Access Method (QSAM) 369

R

- R/3 337
- Rack Mount Option 116
- RAID 61
 - distributed parity 466
 - independent access 466
 - levels
 - RAID 1 465
 - mirroring 465
- RAID 0 464
- RAID 1 464

- RAID 10
 - implementation 468
- RAID 5 466
- RAID level 37
- Rapid Data Recovery 442
- RCMF/PPRC 433
- RCMF/XRC 433
- RDAC 46
- rebranding 28–29
- recovery
 - end user 421
 - network 421
 - plans 421
 - priorities 419
 - procedures 427
 - strategy 421
 - systems 421
- Recovery management 7
- Recovery Point Objective (RPO) 417
- Recovery Time Objective (RTO) 417
- Redbooks Web site 481
 - Contact us xxiii
- reference data 402
- REGISTER LICENSE 316
- regulation 402
- regulatory compliance 341
- remote disk replication 355
- Remote Management Unit (RMU) 122
- remote mirror 87
- Remote Mirror and Copy 88
- Remote Switch 184
- Remote Volume Mirroring 37
- Removable Media Manager 375
- rendering 402
- replication 90
- Replication Manager 249, 298
- reporting
 - assets 266
 - availability 266
 - by userID 267
 - capacity 266
 - usage 266
 - usage violation 267
 - Web site 267
- repository 258
- repository logical drive 48
- Resource Management Facility (RMF) 388
- resource pooling 177
- Resume 28
- Resume Mirror 52
- retention managed data 402
- retention period 317, 402
- RETINIT 406–407
- RETMIN 406–408
- RETVR 406–408
- RIO 72
- RIO-G 76
- risk management 417
- RNID 273
- robotic cartridge 125

- robotic system 118, 120
- Routing Service 198
- rule set 276

S

- SAN 356
 - connections 286
 - Error Predictor 283
 - fault detection 294
 - physical topology 285
 - topology 285
 - transmission distance 356
 - zones 286
- SAN Data Gateway 336
- SAN Director M14 196
- SAN File System 217, 219, 234, 258
 - architecture 236
 - cache 244
 - client 444
 - fileset 239
 - FlashCopy 241
 - FlashCopy solution 444
 - global namespace 236, 238, 243
 - GUI 237
 - instant copy 241
 - locking 244
 - Master Console 242
 - MDS 236
 - Metadata server 236
 - policy 240, 243
 - quota management 244
 - storage pool 239
- SAN management
 - vendor applications 295
- SAN solutions 206
- SAN Switch
 - 08 188
 - H08 185
- SAN switch
 - features 182
- SAN Switch F32 188
- SAN Switch H08 184
- SAN Switch L10 184
- SAN Switch SAN12M-1 184, 186
- SAN Switch SAN24M-1 190
- SAN switches
 - b-type 181
 - Cisco 181
 - m-type 181
 - n-type 181
- SAN switches and gateways 184
- SAN topologies 164
 - extended distance fabrics 170
 - loops and hubs 165
 - point-to-point 164
 - switched fabric without cascading 167
- SAN Volume Controller 217, 219, 221–222
 - architecture 225
 - capacity utilization 230
 - compatibility 230

- Configuration node 223
- extent 224
- GUI 223
- licensing 228
- master console 224
- scalability 222
- storage utilization 222
- usage 228
- vDisk 224
- SAN Volume Controller (SVC) 442
- SAN12M-1 207
- SAN140M 198
- SAN16B-R 197
- SAN16M-R 199
- SAN256M 201
- SAN256N 204, 209
- SANC40M 200
- SANtegrity 191
- SANvergence Manager 200
- SAP R/3 318
- SARC 64
- SAS System 382
- SATA 28, 30, 40, 46, 404
- SATA Interface Card 41
- SBOD 28–29, 44, 69
- scalable performance 166
- scan 260, 264–265
 - TPC for Data 268
- scheduler 258
- scratch pool 374
- SCSI 98
- SCSI protocol 273
- SCSI reserve 350
- SCSI-3 337
- SDD 60
- SEC 404
- self-discovering 165
- Sequential Prefetching in Adaptive Replacement Cache (SARC) 64
- Serial Storage Architecture 82, 98
- ServeRAID Manager 22
- server-free backup 336
- Service Alert 53
- service alert 31, 35, 53
- Service Level Protocol (SLP) 298, 308
- Service Location Protocol 460
- service processor 75
- Services Module 195
- servo 110
- Simple Network Management Protocol (SNMP) 279
- sites
 - alternative 422
 - cold 422
 - hot 423
 - mirrored 423
 - mobile 423
 - warm 422
- SLP 298, 460
- SmartClean 155
- SMI-S 272, 310
- SMIS 459
- snapshot 283
- SNIA 218–219, 298
- SNIA Storage Model 218
- SNMP 263, 273, 275
 - events 275
 - trap 268, 291
- SNMP trap 294
- software zoning 275
- solutions 399
- spanning 44
- spare 67
- spare drive 64
- speed matching 110
- Spill backup tape 374
- split-mirror backup 334
- SQL Server
 - 2000 445
- SSA 62, 82, 98–99
- SSA Optical Extender 100
- Standby Capacity on Demand (Standby CoD) 80
- Statistical Analysis and Reporting System (SARS) 111, 134
- storage
 - asset discovery 264
 - capacity 266
 - growth 266
 - reporting 265
 - usage trends 264
- Storage Area Network Manager 271
- storage consolidation 207–208
- storage infrastructure 6
- Storage LPAR 73
- Storage Management Initiative Specification (SMIS) 272
- Storage Management Interface Specification (SMI-S) 298
- Storage Manager 35
- Storage Networking Industry Association (SNIA) 272
- Storage orchestration 7
- storage partition 37
- storage policies 257
- storage pool 323
- storage solutions 399
- storage subsystem 445
- storage utilization 272
- storage virtualization 6
- Streaming Lossless Data Compression (SLDC) 137
- Subsystem Device Driver (SDD) 223, 227
- Subsystem Device Driver (SSD) 60
- Supervisor Module 203
- Surface Control Guiding Mechanism 110
- Suspend 28
- Suspend Mirror 52
- switch 31, 33, 165–167
 - API 286
 - management applications 295
- switched (SAN topology) 164
- switched point-to-point configuration 167
- switches 273
 - nameserver 294

Symantec 331
Synchronous data transfer 164
synchronous mirroring 51
Synchronous PPRC 88
systems recovery 421

T

tape

3580 114
3581 115
3582 117
3583 118
3590 135
3592 135
barcode reader 116

Tape Controller 137

Tape Library 139

tape vaulting 354

TAPECOPY 374

target server 346

TCP/IP 347

TDP 318, 334

TEDW 281

testing and training 423

three-site configuration 92

Tier 0 429

Tier 1 429

Tier 2 429

Tier 3 430

Tier 4 430, 442, 444

Tier 5 430

Tier 6 430, 442

Tier 7 431

Tiers

blending 431

recovery 428

TimeFinder 318, 334

Tivoli Common Agent Services 277

Tivoli Data Protection. *see* TDP

Tivoli Disaster Recovery Manager 342

Tivoli Enterprise Console 276

Tivoli Enterprise Data Warehouse (TEDW) 281

Tivoli NetView 275

Tivoli Ready certification 182

Tivoli SAN Manager

application launch 295

LUN display 288

navigation tree 290

physical topology 285

Tivoli Storage Area Network Manager 271

Tivoli Storage Manager (TSM) 444

Tivoli Storage Manager Disaster Recovery Manager 319

Tivoli Storage Manager for Application Servers 318

Tivoli Storage Manager for Data Retention 341

Tivoli Storage Manager for Databases 318

Tivoli Storage Manager for Enterprise Resource Planning (ERP) 318

Tivoli Storage Manager for Hardware 318

Tivoli Storage Manager for Mail 318

Tivoli Storage Manager for Space Management 319

Tivoli Storage Manager for Storage Area Networks 319

Tivoli Storage Manager. *See* TSM

Tivoli Storage Resource Manager 262

alert 268

chargeback 268

CIMS 268

Cycle Panels 262

invoices 268

overview 258

security levels 261

Tivoli Storage Resource Manager for Chargeback 268

Tivoli Storage Resource Manager for Databases

asset reporting 264

trend reporting 264

Tivoli Storage Resource Manager. *see* TSRM

topology 274

topology display 284

topology map 273

TotalStorage Expert 306

TotalStorage Productivity Center 86, 219, 248

architecture 251

Common Base 308

launchpad 248

TotalStorage Productivity Center for Data 248, 256

TotalStorage Productivity Center for Disk 248

TotalStorage Productivity Center for Disk (TPC for Disk) 298

TotalStorage Productivity Center for Fabric 248

TotalStorage Productivity Center for Replication 248, 308

TotalStorage Rapid Data Recovery 442

TotalStorage Data Manager (TDM) 444

TPC for Data

agent 259

agent statistics 262

Agents 259

alert 267

alert log 268

alerts 263

allocated disk space 262

architecture 258

asset reporting 264, 266

availability reporting 263

backup reporting

reporting

backup 267

capacity reporting 266

chargeback 268

CSV output 265

dashboard 262

discovery 263

Enterprise Repository 259

Enterprise-wide Summary 262

features 262

file statistics 264

filesystem capacity 262

Job Scheduler 259

most at risk files 265

OS User Groups 268

ping 263, 266

- probe 260, 264–268
- profile 264
- reporting 265
- scan 260, 264–265
- scheduled jobs 263
- script 267
- summary reporting 266
- trend reporting 264
- Triggered Action 267
- usage reporting 266
- user space summary 263
- username reporting 267
- Web reporting 267
- TPC for Disk 298
 - database information function 304
 - database purge 304
 - gauges 303
 - sampling frequency 303
 - threshold settings 301
 - workload profile 305
- TPC for Fabric
 - agent 280
 - API 282
 - Command-line interface 278
 - components 279
 - console 278, 296
 - database 273
 - Device Centric View 287
 - fault detection 294
 - filesystem display 288
 - historical reporting 291
 - Host Centric View 287–288
 - in-band 285
 - iSCSI 285
 - launch 283, 295
 - logical topology 287
 - logical volume display 289
 - navigation display 290
 - object properties 290
 - out-of-band 285
 - real-time reporting 292
 - reporting 291
 - summary display 290
 - topology map 273, 285, 294
 - WebSphere 280
 - zone control 281
 - zone display 286
- TPC for Replication 308
 - copyset 313
 - FlashCopy 309
 - Metro Mirror 309
 - replication group 310
 - replication session 311
 - replication task 310
 - sequence 314
 - session monitoring 312
 - storage group 312
 - storage pool 312
- tracks 110
- transaction integrity 430
- trend reporting 264
- Triggered Action 267
- TSM
 - adaptive subfile backup 330, 333
 - adaptive subfile differencing 321
 - Administration Center 324
 - and HACMP 349
 - and MSCS 352–353
 - API 318, 331
 - archive 317, 324, 334, 347
 - archive copy group 347
 - backup 324
 - backup and restore 317
 - backup methods 332
 - backup set 335
 - backup topology 335
 - backup/archive client 323, 329
 - backup/restore operations 332
 - Bare Machine Recovery 319
 - checkpoint restart 321
 - client backup methods 332
 - client configuration information 342, 344
 - client data 346
 - client interface 329
 - client node 328
 - client option set 346
 - client options file 351
 - client Web browser interface 330
 - client with MSCS 353
 - clustering applications 349
 - CLUSTERNODE 351, 354
 - collocation 322, 357, 359
 - command routing 345–346
 - command-line interface 346
 - COMMRESTARTDURATION 350
 - COMMRESTARTINTERVAL 350
 - configuration manager 346
 - copy storage pool 323, 347, 357–358
 - CPU utilization 332
 - cyclical redundancy check 321
 - data movement 323
 - data movements 321
 - Data Retention 318
 - database 323, 346, 357
 - database backup 346–348, 355
 - database mirroring 355
 - device class 347
 - Disaster Recovery 319
 - Disaster Recovery plan file 347–348
 - distributed configuration 345
 - DRM 344, 348, 354
 - electronic tape vaulting 355
 - electronic vaulting 348
 - encryption 321
 - enterprise administration 345
 - enterprise console 346
 - event receiver 346
 - expiration 354, 357
 - Extended Edition 342
 - HACMP 351

- high availability clustering 349
- HSM 319, 351
- image backup 333, 339
- instant archive 335
- LAN-based clients 335
- LAN-free backup 335–336
- Licensing 316
- logical volume backup 340
- Logical Volume Storage Agent 340
 - managed server 346
 - management class 340, 347
 - metadata 323
 - migration 357
 - monitoring 345
- NAS 338
- NAS backup 338
- off-site data copy 342, 354, 357
- Original Block File 340
- overview 316
- parallel backup 330
- point-in-time recovery 340
- policy 324, 346
- policy domain 328, 347
- primary storage pool 347, 357
- profile subscription 346
- progressive backup 321, 330, 332
- progressive incremental backups 340
- reclamation 357–359
- REGISTER LICENSE 316
- remote server 347
- scheduled operations 351, 353
- Scheduler Service 351, 354
- SCSI reserve 353
- security 331
- security concepts 331
- selective backup 333
- server 323
 - server administration 324
 - server consolidation 348
 - server group 346
 - server maintenance 329
 - server recovery 342
 - server-free backup 336–337
 - server-to-server communication 345
- snapshot image backup 340
- source server 346–348
- split-mirror backup 334, 337
- startserver script 350
- stopserver script 350
- Storage Agent 336
- storage device 327
- storage pool backup 346, 354
- storage pool hierarchy 323
- storage pools 357
- tape reclamation 322
- tape vaulting 354
- target server 347–348
- vaulting 354
- Version 5 330
- versions 332

- virtual volumes 345–348
- volume history 354
- volume tracking 354
- Web browser interface 346
- TSM commands
 - DEFINE DEVCLASS 347
 - EXPORT 346
 - IMPORT 346
 - QUERY RPFIL 348
- TSRM
 - architecture 258
 - security 261
- TTPC for Data
 - free space 262
 - HTML output 265
- Tunneling Service 198
- Turbo
 - FAST600 29
- Turbo II 82

U

- Ultra320 102
- UltraBac 331
- Ultrium 107–108, 123
 - compatibility 112
- Ultrium 1 112, 122
- Ultrium 2 112
- Ultrium 3 109
- UMD 205
 - logical domains 206
- UMD physical director 206
- usage reporting 266
- user space summary 263

V

- validate lanfree 336
- variable-length records (VLR) 385
- vDisk 224–225
- VERITAS 438
- Virtual Device Interface (VDI) 445
- Virtual disk (vDisk) 224
- Virtual Disk Service (VDS) 46
- Virtual SAN (VSAN) 193, 203
- Virtual Storage Access Method (VSAM) 369
- Virtual Tape Controller (VTC) 146
- Virtual Tape Server (VTS) 135
- virtualization 2, 65, 73, 214
- virtualized SAN (VSAN) 282
- vital record retention, archive and retrieval 317
- volume drain 244
- volume management 226
- Volume Performance Advisor 305
- volume serial (VOLSER) 127
- Volume Shadow Copy Service (VSS) 46, 444
- VolumeCopy 48
- VSAM 367
- VSAN 282
- VTS 143
 - models 146

Peer-to-Peer 146

W

WAN 170

warm sites 422

Wavelength Division Multiplexing (WDM) 178

WBEM 272

Web Based Enterprise Management (WBEM) xxi, 272, 459

Web reporting 267

WebSphere Application Server 318

WEBTOOLS 188, 196

Windows

 event log 263

Windows 2000

 Logical Volume Storage Agent 340

 snapshot image backup 340

WORM 341, 405

X

XML 296

Z

z/OS Global Mirror 90

z/OS Workload Manager (WLM) 60

zero data loss 430

zone 274

zone control 278

zone member 274

zone set 274

zones 286

zoning

 hardware 275

 software 275

zSeries File System(zFS) 370



Redbooks

IBM TotalStorage Solutions Handbook

(1.0" spine)

0.875" x 1.498"

460 <-> 788 pages



IBM TotalStorage 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 on the most current IBM TotalStorage® products, showing how IBM delivers the right mix of products for nearly every aspect of business continuance, information lifecycle management, and infrastructure simplification.

- ▶ Part 1 describes disk products, from entry-level offerings such as the DS400, to mid-range with the DS4000 family and DS6000, to high end with the DS8000.
- ▶ Part 2 is an overview of tape products, with the recently announced LTO3 support, and enterprise system like the 3494 library and Virtual Tape Server.
- ▶ Part 3 describes the different storage networking infrastructure, protocols, and SAN switches.
- ▶ Part 4 discusses IBM TotalStorage software portfolio for open systems and includes Storage Virtualization products such as the SAN Volume Controller and SAN File System, the TotalStorage Productivity Center and Tivoli Storage Manager.
- ▶ Part 5 describes z/OS storage management software.
- ▶ Part 6 presents some of the storage solutions and services offered by IBM, like Disaster Recovery and Business Continuity, the DR550 life sciences, and digital media.

This redbook is intended as a reference for basic and comprehensive information on the IBM TotalStorage products portfolio.

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**