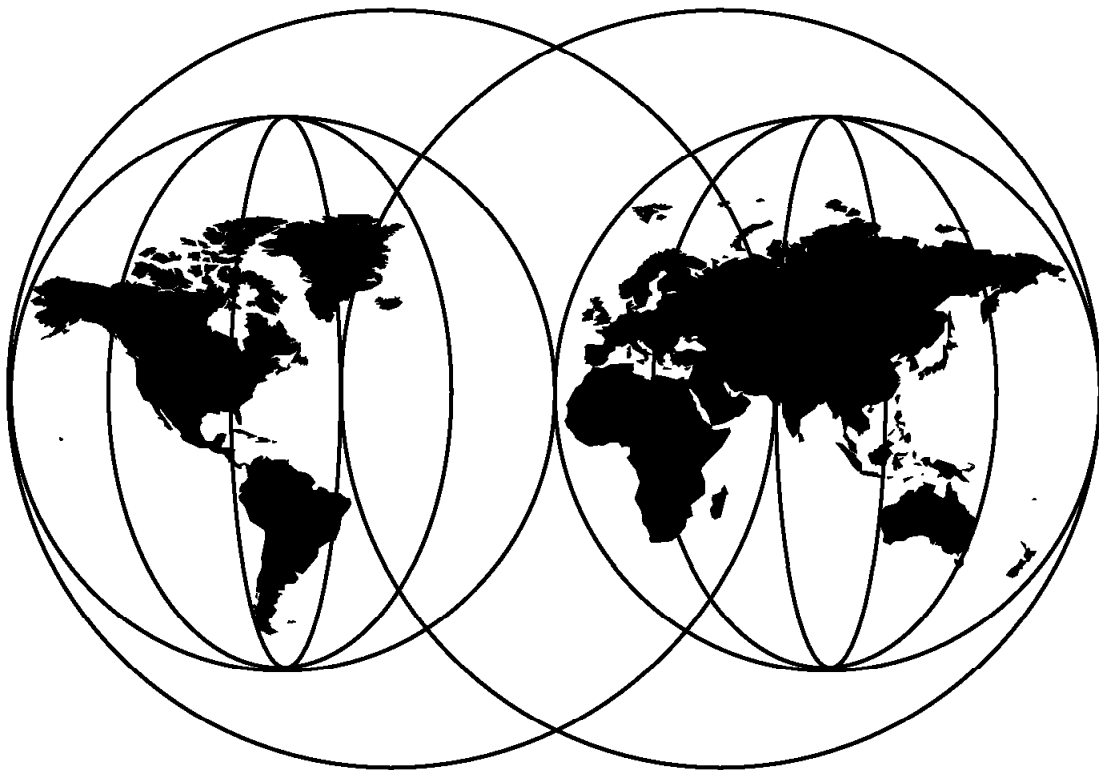




Data Sharing: Cross Platform Extension (XPE) Implementation Guide

Phil Norman Chris Owen Alison Pate Joerg Pohlmann



International Technical Support Organization

<http://www.redbooks.ibm.com>

This book was printed at 240 dpi (dots per inch). The final production redbook with the RED cover will be printed at 1200 dpi and will provide superior graphics resolution. Please see "How to Get ITSO Redbooks" at the back of this book for ordering instructions.

SG24-5256-00



International Technical Support Organization

SG24-5256-00

**Data Sharing: Cross Platform Extension (XPE)
Implementation Guide**

June 1998

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix C, "Special Notices" on page 109.

First Edition (June 1998)

This edition applies to the IBM 9399 Model 1 Cross Platform Extension

Note

This book is based on a pre-GA version of a product and may not apply when the product becomes generally available. We recommend that you consult the product documentation or follow-on versions of this redbook for more current information.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. QXXE Building 80-E2
650 Harry Road
San Jose, California 95120-6099

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© **Copyright International Business Machines Corporation 1998. All rights reserved.**

Note to U.S. Government Users — Documentation related to restricted rights — Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Preface	vii
The Team That Wrote This Redbook	vii
Comments Welcome	viii
Chapter 1. Notes on Terminology	1
Chapter 2. Introduction to Data Sharing	3
2.1 Levels of Data Sharing	3
2.1.1 Storage Sharing	4
2.1.2 Data Copy Sharing	4
2.1.3 "True" Data Sharing	5
2.2 Business Values of the IBM XPE Solution	6
Chapter 3. XPE Functional Description	7
3.1 Components	7
3.1.1 SCSI Adapters	8
3.1.2 ESCON Adapters	9
3.1.3 Ethernet Connection	10
3.2 Operation	10
3.2.1 Power-On Sequence	10
3.2.2 Power-Off Sequence	11
3.3 Physical Characteristics	11
3.4 Remote Service Facility	11
3.5 Installation	12
3.6 Supported Hardware Environments	12
Chapter 4. XPE Implementation Planning	15
4.1 Workload Considerations	15
4.2 Storage Considerations	15
4.3 Physical Hardware Planning Considerations	16
4.4 Installation Worksheets	16
4.5 Outline Plan of Activities	16
4.5.1 Installation	16
4.5.2 Backup and Recovery Processes	17
Chapter 5. XPE Configuration Steps	19
5.1 Configuration Worksheet	20
5.1.1 Server Information	22
5.1.2 XPE SCSI Definition	22
5.1.3 XPE ESCON Definition	23
5.1.4 MVS Device Information	25
5.2 ESCON Worksheet	25
5.3 Web Configuration Worksheet	27
Chapter 6. MVS Implementation Process	29
6.1 Selecting Volumes	29
6.2 Recommendations	29
6.3 Initializing XPE Volumes	30
6.3.1 ICKDSF	30
6.3.2 SnapShot	33
6.3.3 Other Utilities	34

6.4	Converting to SMS-Managed Volumes	34
6.5	Security Considerations for XPE Volumes and Data Sets	37
Chapter 7. RS/6000 Implementation Process		39
7.1	Initial Steps	39
7.2	Some Thoughts before Using The New SCSI Disks	40
7.3	Differences from Standard SCSI Disks	41
7.4	Output from a Sample Configuration	43
7.5	Completing the Configuration Worksheet	47
7.6	AIX Recommendations with XPE	47
Chapter 8. Windows NT Implementation Process		49
8.1	Before Connecting to the XPE	49
8.2	After Connecting to the XPE	50
8.3	Making the Disks Usable by Windows NT	52
8.3.1	Formatting the Disks and Assigning Drive Letters	54
8.3.2	SCSI and LUN Considerations	55
8.4	Completing the Configuration Worksheet	56
8.5	Windows NT Recommendations with XPE	56
Chapter 9. Sun Microsystems and Hewlett Packard Implementation Process		57
Chapter 10. The XPE Web Interface		59
10.1	IBM 9399 Cross Platform Extension - Current Configuration	59
10.1.1	Total Devices Defined	60
10.1.2	ESCON Status	60
10.1.3	SCSI Status	60
10.2	Options Bar	60
10.2.1	XPE DEVICES	61
10.2.2	Documentation	61
10.2.3	UTILITIES	61
10.3	Device Status	62
10.4	ESCON	63
10.5	SCSI	64
10.6	Configuration	65
Chapter 11. OS/390 Operational Considerations		67
11.1	Backing Up UNIX or Windows NT Data: General Considerations	67
11.1.1	Synchronization between OS/390 and UNIX or Windows NT	67
11.2	Using DFSMSdss to Manage UNIX or Windows NT Data	68
11.2.1	Concurrent Copy	68
11.2.2	Virtual Concurrent Copy	68
11.2.3	Disk to Disk Copy	68
11.3	Using SnapShot to Back Up UNIX or Windows NT Data	70
11.4	Recovery of Backed Up Disks	71
11.4.1	Snap-back	71
11.4.2	DFSMSdss Restore	71
11.4.3	Recovering an Individual UNIX or Windows NT File	71
11.5	SMS Considerations	71
11.5.1	Year 2000 Testing	72
11.5.2	Remote Copy	72
11.6	Operations	72
11.6.1	Service Information Messages	73
11.6.2	IXFP Reports	73
11.6.3	RMF Reports	74

Chapter 12. AIX Operational Considerations	75
12.1 Adding New SCSI Disks	75
12.2 Moving Disks to Other Servers	76
12.3 Data Backups	76
12.4 Creating a Clone Server	77
12.5 Recovering Data with ADSM	78
12.6 Recovering a Lost or Deleted File without ADSM	78
12.7 Year 2000 Testing	80
12.8 Recovering from an XPE Failure	81
Chapter 13. Windows NT Operational Considerations	83
13.1 Adding New SCSI Disks	83
13.2 Moving Disks to Other Servers	84
13.3 Data Backups	84
13.4 Creating a Clone Server	85
13.5 Recovering Data with ADSM	86
13.6 Recovering a Lost or Deleted File without ADSM	86
13.7 Year 2000 Testing	87
13.8 Recovering from an XPE Failure	88
Chapter 14. Performance Characteristics	89
14.1 Usage and Configuration Guidance	89
14.2 XPE Capacity	90
14.3 XPE Performance	90
Chapter 15. Increasing Systems Availability	93
15.1 RAMAC Array Media Failure	93
15.2 A SCSI Adapter or Cable Failure	93
15.3 A UNIX or Windows NT Server Failure	93
15.4 An XPE Failure	94
15.5 An ESCON Adapter or ESCON Link Failure	94
15.6 A RAMAC Array Storage Failure	94
15.7 Potential High Availability Configurations That We Have Not Tested	95
Appendix A. XPE Installation Worksheets	99
A.1 Configuration Worksheet	99
A.1.1 ESCON Worksheet	103
A.1.2 XPE Web Configuration Worksheet	103
Appendix B. The RAMAC Virtual Array and Snapshot	105
B.1 Overview of RVA	105
B.1.1 Log Structured File	105
B.1.2 Data Compression and Compaction	106
B.2 Overview of SnapShot	106
B.2.1 SnapShot Copy	106
B.2.2 Invoking SnapShot	107
B.2.3 Hardware and Software Requirements	107
Appendix C. Special Notices	109
Appendix D. Related Publications	111
D.1 International Technical Support Organization Publications	111
D.2 Redbooks on CD-ROMs	111
D.3 Other Publications	111

How to Get ITSO Redbooks	113
How IBM Employees Can Get ITSO Redbooks	113
How Customers Can Get ITSO Redbooks	114
IBM Redbook Order Form	115
Glossary	117
Index	119
ITSO Redbook Evaluation	121

Preface

The *Data Sharing: Cross Platform Extension Implementation Guide* is written to help storage administrators, system programmers, systems administrators, and technical operations managers and staff implement the IBM 9399 Cross Platform Extension Model 1 (XPE).

The XPE, which is based on the Seascape Architecture, is a storage sharing solution for the S/390 RAMAC family. It provides the connections and conversion to support the use of S/390 RAMAC disk capacity for RS/6000, HP, Sun, and Windows NT servers. With XPE, the data management functions of RAMAC Arrays and S/390 are available to non-S/390 servers.

This redbook provides detailed technical and operational guidance for implementing XPE in both UNIX and Windows NT server environments. It makes recommendations on the different ways you can implement XPE and provides examples throughout to introduce new XPE concepts or functions.

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.

Alison Pate is a project leader in the International Technical Support Organization, San Jose Center. She joined IBM in 1985 after completing an MSc in Information Technology. A large-systems storage specialist since 1990, Alison has eight years of practical experience in using and implementing IBM DASD solutions. She has acted as a consultant to some of the largest, leading-edge customers in the United Kingdom, providing technical support and guidance to their key business projects.

Phil Norman is a Consultant IT Specialist in the UK. He has more than 30 years of IBM experience and for the last 15 years has been UK and EMEA Northern Region specialist in S/390 storage. Phil has also been on assignment to IBM Poughkeepsie on IMS/XRF and has worked as a specialist SE with S/390 customers. Phil's areas of expertise include all areas of S/390 systems, and he has contributed to numerous S/390 and storage redbooks.

Chris Owen is a Senior Systems Specialist in the UK. He has more than 18 years of IBM experience, having been an S/370 assembler programmer, a Systems Engineer and for the last 6 years RS/6000 and AIX technical support with special emphasis on storage and high availability. Chris has written extensively on clustering.

Joerg Pohlmann is a Consulting (Certified) IT Specialist for Storage Systems in Vancouver, BC, Canada. Of his almost 30 years with IBM, he has more than 25 years of experience in storage systems. He has responsibility for technical sales support in Western Canada for IBM Storage Systems Division hardware and software products and for distributed, enterprisewide storage management.

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

Maggie Cutler, Editor, International Technical Support Organization, San Jose Center
Kjell Nystrom, International Technical Support Organization, San Jose Center

Doug Clem, IBM SSD, Tucson
Tom Connolly, IBM SSD, San Jose
Ken Day, IBM SSD, Tucson
Doug Dewey, IBM SSD, Tucson
Cindy Grossman, IBM SSD, Tucson
Norm Hanami, IBM SSD, Almaden
Dean Hanson, IBM SSD, Tucson
Paul Hurley, IBM SSD, San Jose
Allan McClure, IBM SSD, Tucson
Dave Weber, IBM SSD, Tucson

Thanks also to the following reviewers:

Scott Drummond, Data Sharing Competency Center, San Jose
Ed Quigley, Data Sharing Competency Center, Montana
Chris Saul, IBM SSD, San Jose
Tom Weaver, HACMP Development, Austin

Comments Welcome

Your comments are important to us!

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

- Fax the evaluation form found in "ITSO Redbook Evaluation" on page 121 to the fax number shown on the form.
- Use the electronic evaluation form found on the Redbooks Web sites:

For Internet users <http://www.redbooks.ibm.com/>

For IBM Intranet users <http://w3.itso.ibm.com/>

- Send us a note at the following address:

redbook@us.ibm.com

Chapter 1. Notes on Terminology

The IBM 9399 Model 1 Cross Platform Extension (XPE) is one of a new generation of products that closely integrate large, mainframe, MVS type installations with systems that have typically been small, entry, low-end, UNIX or NT departmental servers. To those devotees of their respective platforms, this is a rather generalistic statement and not in general true, but what is true is that we can no longer use a shared terminology.

If you read this book from cover to cover, terms such as channels, buses, SCSI IDs, LUNs, ESCON, servers, hosts, directors and switches, only some of which may be familiar to you now, should become much clearer.

When we discuss the devices on which the data (a word that everyone seems to agree on) is stored, the terminology of volumes, disks, hdisks, drives, and volume groups is common.

This book is not going to make an MVS systems programmer into a UNIX systems administrator or vice versa. The contributors to this book would not want to say that they are now specialists (or converts) in their colleagues' respective fields, but one thing that has come to light is that, although MVS and UNIX professionals use different terminology, it is possible to understand one another.

Once in a while you will see a word you may not prefer, but when you bring together the biases of professionals from different operating systems into one book, there is not always a correct answer.

Chapter 2. Introduction to Data Sharing

With today's increasingly sophisticated applications, such as business intelligence and e-business, information must be accessible to anyone who needs it. IBM's XPE data sharing solution addresses such challenges of data sharing in today's environment as open connectivity, flexibility, centralized storage management, and asset protection.

In this chapter we provide an overview of data sharing and discuss the three levels of data sharing. The term *data sharing* describes the access of common data for processing by multiple computer platforms or processors. Different platforms store and access data differently. For example, disk storage for OS/390, UNIX, and Windows NT have different methods for data encoding, file structures, disk subsystem attachment mechanisms, and other data management protocols. Given the complexity of these differences, it is important to determine whether data is shared among platforms that are similar or different. The term *homogeneous data sharing* indicates that all platforms are the same; *heterogeneous data sharing* indicates that at least two platforms are different.

Figure 1 shows the concept of data sharing.

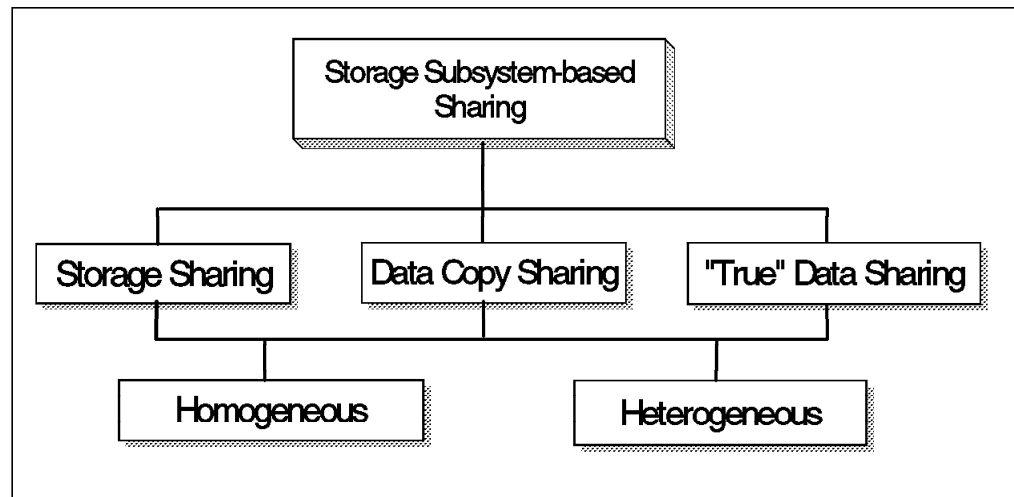


Figure 1. Data Sharing

2.1 Levels of Data Sharing

The three levels of data sharing involve varying degrees of complexity and data accessibility among homogeneous or heterogeneous computer platforms. The three levels of data sharing are:

- Storage sharing
- Data copy sharing
- "True" data sharing

2.1.1 Storage Sharing

With storage sharing (Figure 2), two or more homogeneous or heterogeneous platforms share a single storage subsystem whose capacity has been physically partitioned so that each attached platform can access only the units allocated to it. Multiple platforms can own the same partition, but this is only possible with homogeneous platforms such as in a S/390 Parallel Sysplex configuration.

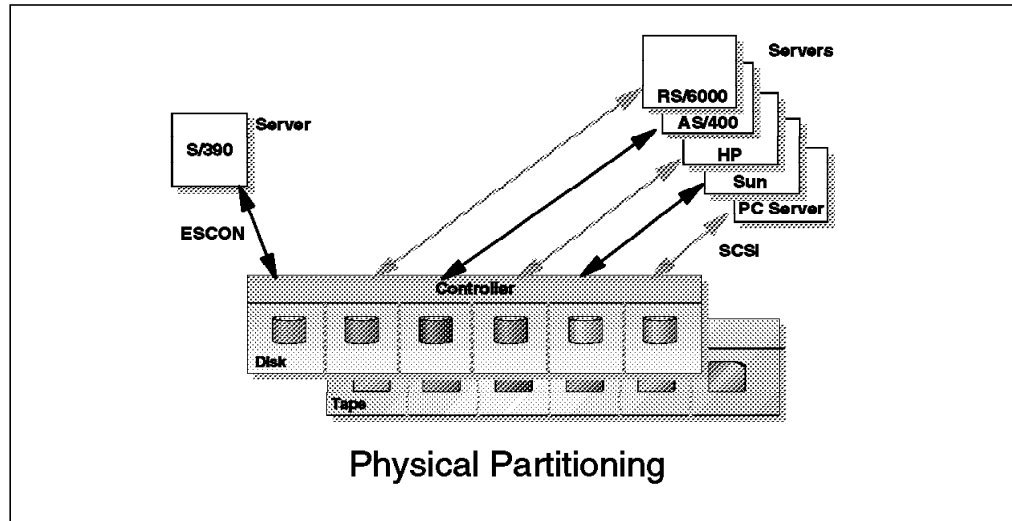


Figure 2. Storage Sharing

Different platform types use different attachment mechanisms for storage subsystems. For example, S/390 uses Enterprise Storage Connection (ESCON) channels, and UNIX systems or PC-based Windows NT servers use small computer storage interface (SCSI) attachments.

Many reference books on SCSI are available; one reference is Appendix A in *Data Sharing: Implementing InfoSpeed Copy Sharing Solutions Among Heterogeneous Systems*, SG24-4120.

To implement storage sharing, the storage subsystem must support each platform's attachment mechanism, or a functional component must be placed between the processor and the storage subsystem that translates the commands and protocols used by the attachment mechanism. IBM's XPE is a product that implements storage sharing.

2.1.2 Data Copy Sharing

One method of data sharing among different platforms is to send a copy of the data from one platform to the other. This is known as *data copy sharing* (see Figure 3 on page 5).

Data copy sharing allows different platforms to access data that was the same when the copy was made. The speed of the copy operation is important because the availability of the data directly depends on how quickly the data can be copied.

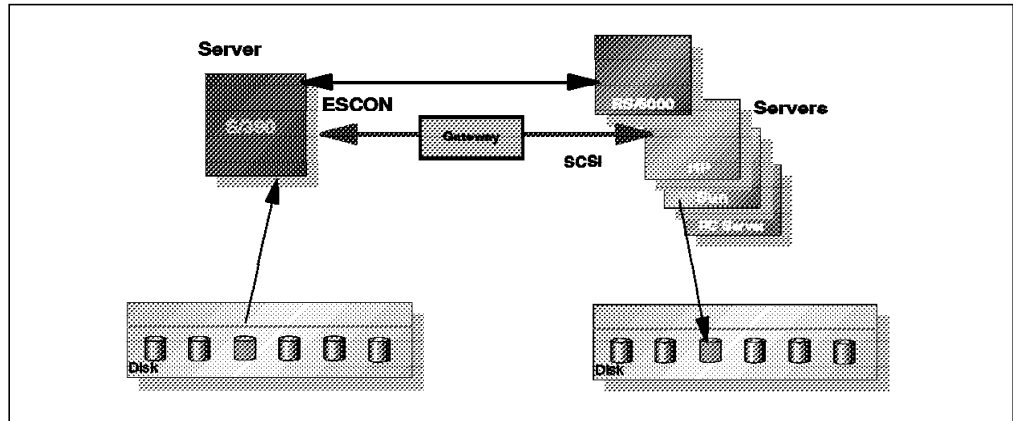


Figure 3. Data Copy Sharing

There are two approaches to data copy sharing between heterogeneous platforms: bulk file transfer and one-step data transfer. The bulk file transfer approach is typically a three-step serial process:

1. Extract the data onto a flat file
2. Transfer the flat file to another system
3. Load the flat file into a database

The one-step data transfer approach involves a direct application-to-application data transfer, which is significantly faster and cheaper than bulk file transfer. IBM's InfoSpeed products implement data copy sharing between OS/390 and UNIX or NT systems.

2.1.3 "True" Data Sharing

In true data sharing (Figure 4), only one copy of the data is accessed by multiple computer platforms, homogeneous or heterogeneous. Because there is only one copy of the data, there is never a need to replicate the data for use elsewhere. Each platform has read and write access to the single copy of the data. True data sharing reduces maintenance and storage costs and eliminates any data synchronization problems.

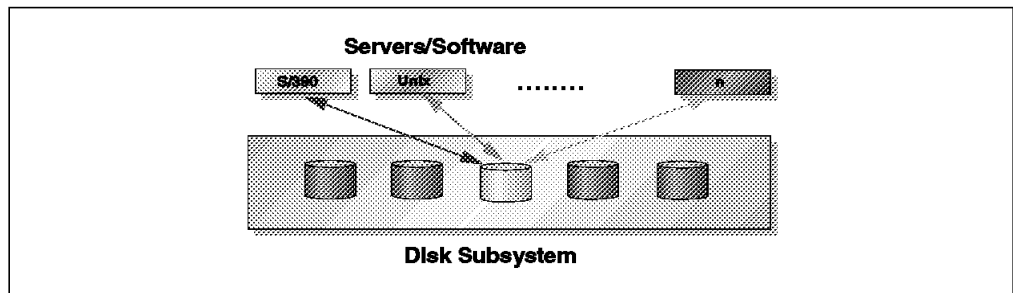


Figure 4. "True" Data Sharing

In today's computing environment, true data sharing exists only on homogeneous platforms. For example, in an S/390 Parallel Sysplex configuration, all processors in the Sysplex run on similar platforms and have read and write access to data on shared disk storage. Applications running on separate processors can simultaneously read the same copy of the data, but write access requires the data to be locked by a single application, to preserve

data integrity. The process of locking data is managed by hardware and software independent of the disk storage subsystem.

Another example of true data sharing is an RS/6000 cluster configuration with shared-disk architecture. Clustering technology allows a group of independent nodes to work together as a single system, and, in a shared-disk architecture, each node can be connected to a shared pool of disks and its own local disks. All of the nodes have concurrent read and write access to the data stored on the shared disks. As with the S/390 Parallel Sysplex, to preserve data integrity, write access requires the data to be locked by the node requesting the write. This locking process is also managed by software independent of the disk storage subsystem.

Figure 5 shows two examples of true data sharing in homogeneous environments.

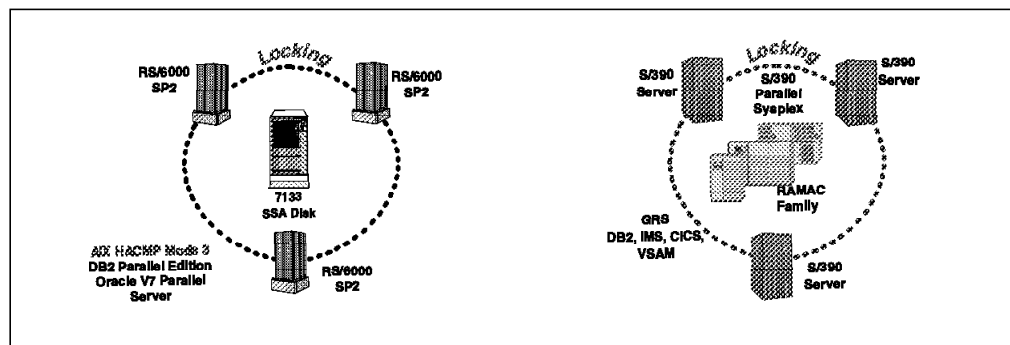


Figure 5. True Data Sharing in Homogeneous Environments

2.2 Business Values of the IBM XPE Solution

IBM XPE implements storage sharing. It provides high-speed ESCON communication to the RAMAC Array Storage and SCSI communications to UNIX and Windows NT Servers. The XPE software converts SCSI to ESCON (or vice versa).

Up to four UNIX or Windows NT servers can be connected to an XPE, and applications running on these servers have no need to know that their data is being read from or written to a RAMAC Array Storage.

Now that the application data resides on RAMAC Array Storage, it can be managed by sophisticated S/390 data management facilities.

In storage sharing environments, the key is to provide data as it is needed, a reliable backup environment, and support for those servers needing additional storage capacity as and when required. The IBM XPE fulfills all of these requirements.

Chapter 3. XPE Functional Description

The IBM 9399 Model 1 Cross Platform Extension allows users of IBM AIX, Sun Solaris, HP-UX, or Windows NT systems to use S/390 RAMAC Array Storage to meet their capacity and storage management needs.

Up to 64 logical volumes or disks are supported by the UNIX or Windows NT servers. Up to 180 GB of RAMAC Array Storage can be supported by one XPE. The servers can be located at distances of up to 26 km from the RAMAC storage. Up to four servers can be connected to one XPE. Logical volumes can be managed through traditional OS/390 storage management, including backup using the high performance SnapShot option of the RAMAC Virtual Array (RVA).

3.1 Components

The XPE is the size of a typical desktop workstation. It can be positioned either as a table top or floor standing unit. It has no display or keyboard; all your communication is through a Web interface. In addition, the IBM Service Representative can use a local CE Mobile Operation Service Terminal (MOST) or a Remote Service Facility (RSF) modem connection.

Figure 6 shows the rear view of the XPE box with the position of the various adapters and connections.

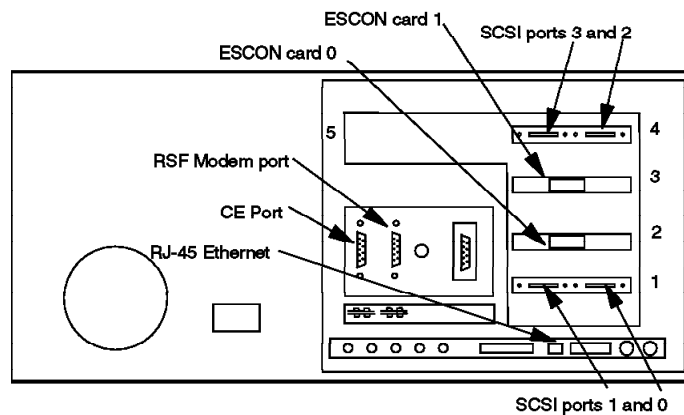


Figure 6. Rear View of XPE

- XPE slot 1
The first dual Ultra SCSI adapter. SCSI port 0 is on the right and SCSI port 1 on the left.
- XPE slot 2
An ESCON adapter configured in the XPE as ESCON card 0
- XPE slot 3
An ESCON adapter configured in the XPE as ESCON card 1
- XPE slot 4
The second dual Ultra SCSI adapter. SCSI port 2 is on the right and SCSI port 3 is on the left.

The XPE always comes with these four adapter cards preinstalled. Licenced Internal Code (LIC) is provided on CD-ROM and is installed by the IBM Service Representative.

- CE port (COM 1)

The CE port is solely used by the IBM Service Representative to attach a service terminal (MOST) and do the initial setup.

- RSF modem port (COM 2)

The RSF port is used to attach the RSF modem so that the IBM Support Center can perform remote diagnosis and apply LIC fixes if required.

- RJ-45 Ethernet port

The Ethernet port provides a network connection to the XPE internal Web server. This connection provides an external interface to the XPE through which configuration information and usage statistics can be viewed. If the XPE is to be installed in an environment with no Ethernet capability, then an appropriate bridge product into Ethernet, such as the IBM 8229 Bridge Model 2 (Token-Ring to Ethernet) is required. The Web interface also allows you to update the XPE configuration files, although shutdown and restart of the XPE are required to activate a new configuration. See Chapter 10, "The XPE Web Interface" on page 59 for details of the Web interface and functions.

- There are several additional ports shown that are not in use.

3.1.1 SCSI Adapters

SCSI overview for the MVS specialist

For the MVS specialist, a SCSI bus looks similar to an S/390 block multiplexer channel. Each device is chained to the next device, and each has its own unique ID on the bus (channel). Only one device on the bus transfers data at one time. The SCSI command set is similar to the S/390 channel commands. The host can initiate I/O to a device, disconnect while seeking, and then reconnect to transfer data.

The XPE contains two dual-port Ultra SCSI adapters. Note that the Ultra SCSI adapters are differential, so the SCSI adapters in the UNIX or Windows NT server must also be differential. Typically a Windows NT server will not have a differential SCSI adapter, so you will probably have to install a new adapter to connect to the XPE.

The cables used are very high density cable interconnect (VHDCI) cables and are supplied with the XPE. Two lengths are available, 10 m (feature code 7711) and 20 m (feature code 7712). Specify the length when ordering the XPE. Order one length for each SCSI interface to be used (maximum of four interfaces).

3.1.1.1 SCSI IDs and Logical Unit Numbers

SCSI ID and LUN overview for the MVS specialist

SCSI addressing takes the form *00-bb-cc-d,e* where

- bb is the host SCSI adapter slot number.
- cc is the connector identifier and is typically 00. On adapters with multiple connectors, it is used to identify the adapter connector to which a resource is attached.
- d is the SCSI ID, also known as the control unit address.
- e is the logical unit number.

SCSI ID 7 is normally reserved for use by the server SCSI adapter. The 15 other IDs, 0-6 and 8-15, are available for use. The LUN is usually left as 0, but to provide more than 15 addresses the LUN can be changed to any number in the range 0-7. Hence SCSI ID 0 LUN 0, SCSI ID 0 LUN 1, SCSI ID 1 LUN 0, and SCSI ID 1 LUN 1 are four unique addresses.

Each SCSI interface on the XPE can attach to a different server, either UNIX or Windows NT. The SCSI bus cannot be shared with other servers or I/O devices. The supported and tested interfaces on each server type are listed in 3.6, “Supported Hardware Environments” on page 12.

The XPE implementation provides up to 64 devices among up to 4 servers.

3.1.2 ESCON Adapters

ESCON overview for the UNIX and Windows NT administrator

ESCON architecture can be thought of as the S/390 fiber-optic storage area network (SAN). ESCON channels on the S/390 processors connect to ESCON Directors, which are fiber-optic switches that connect processor to processor, processor to storage, and even storage to storage.

Two types of fiber are used with ESCON: multimode and singlemode. Multimode operates at short distances up to 3 km and uses 62.5/125 micron fiber, with light emitting diode (LED) optical drivers. Single-mode operates at distances up to 20 km using laser optical drivers. Single mode is used only between ESCON Directors (and is called Extended Distance Facility (XDF) in the S/390 world).

The XPE ESCON adapters require ESCON LED jumper cables with the IBM ESCON connector. The maximum distance that can be reached without going through an ESCON Director is 3 km. Typically connection without an ESCON Director would involve a jumper cable from the XPE to a fiber patch-panel, part of an installation’s ESCON fiber infrastructure, and another jumper cable to the RAMAC Array Storage. Short jumper cable lengths are available when the XPE and RAMAC Array Storage are located in close proximity (under 400 m). Contact your IBM Representative for details of how to order longer cables.

Longer distances of up to 26 km can be achieved by the use of ESCON Directors and ESCON XDF fiber—a single mode laser fiber that has a maximum distance of 20 km. Maximum distance would therefore be 3 km from XPE to an ESCON Director, 20 km of XDF fiber to another ESCON Director, and 3 km of multimode LED fiber to the RAMAC Array Storage. Figure 7 on page 10 shows an example of this configuration from an S/390 (top left) to the RAMAC Array Storage (bottom right).

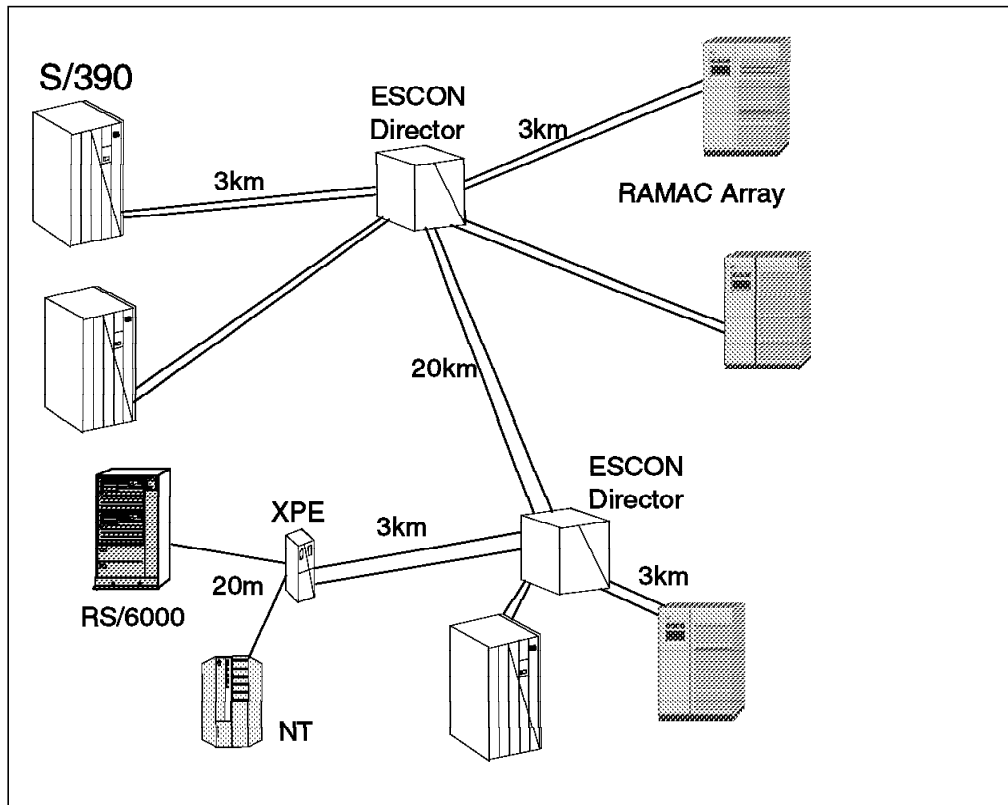


Figure 7. Sample ESCON Configuration

An ESCON Director is an ultra high-speed fiber-optic switch that enables multiple S/390 host systems to share control units or one S/390 host to share multiple control units on the same channel. An XPE acts like a S/390 host to the RAMAC Array Storage.

3.1.3 Ethernet Connection

Each XPE has an RJ-45 Ethernet connection, which provides a TCP/IP connection to the Web server function running in the XPE. See Chapter 10, "The XPE Web Interface" on page 59 for more details on the Web browser interface.

3.2 Operation

The XPE contains a Reduced Instruction Set Computer (RISC) processor and SCSI and ESCON adapters. It has 128 MB of memory that is used for the LIC and for buffers. It converts SCSI commands to S/390 channel commands and transfers data to and from the UNIX or Windows NT server and the RAMAC Array Storage.

3.2.1 Power-On Sequence

The XPE must be powered up and fully operational *before* the UNIX or Windows NT servers. When the XPE is initializing, it reads a configuration file that matches VOLSER devices with SCSI addresses. The XPE tries to connect through the ESCON paths to each device in turn in the RAMAC Array Storage.

Once the XPE has discovered all of the OS/390 volumes, the UNIX or Windows NT servers can be initialized. Only the successfully matched devices are passed

to the servers. Each unique MVS VOLSER is passed to only one server as a SCSI disk, so no two servers attached to the XPE can see the same disk.

3.2.2 Power-Off Sequence

If, for example, you want to move the XPE, you have to disconnect the power to it. To ensure a controlled move, you must shut down the UNIX or Windows NT servers *before* you shut down and power off the XPE. When you shut down the servers in a controlled manner, you ensure that any outstanding write data is written out to the RAMAC Array Storage. Once you have shut down *all* servers, you can power off the XPE.

The key message here is that there must be no activity between the servers and the RAMAC Array Storage when you power off the XPE, otherwise unpredictable results may be experienced.

3.3 Physical Characteristics

Table 1 and Table 2, respectively, list the dimensions and weight of the XPE.

	Front	Side	Height
mm	421	455	156
in.	16.8	18	6.1

kg	18
lb	39.6

The XPE is supplied with a single phase power cord and connector appropriate to each country.

3.4 Remote Service Facility

The XPE is connected through an RSF modem or switch to allow dial-in by the IBM Support Center so that it can maintain the XPE and diagnose and rectify faults.

The RSF feature codes (FCs) on the XPE are the same as those used by the 3494 Tape Library, and the phonenumber can be shared with the library and other XPEs.

The RSF features dial-in capability with password protection, which enables remote analysis of error logs, queries of system status, and download of LIC updates.

One of the following features is required on each XPE for remote support:

- FC 2710 Remote Support Facility

This feature provides a modem for use by the XPE RSF function. The modem is connected to COM 2 on the XPE. Order this feature if you do not have an FC 2710 on another unit nearby (for example, the 3494 Tape Library).

- FC 2711 Remote Support Switch

This feature is a switch that enables multiple devices to share the same phone line. Up to 14 devices can be connected to the same line through the switch. Order this feature if you already have FC 2710 on another unit but do not have a switch unit.

- FC 2712 Remote Support Attachment

This feature provides a cable to connect to a switch (FC 2711). Order this feature if you already have a modem (FC 2710) and a switch (FC 2711) on other units nearby in your installation.

3.5 Installation

Installation time for the XPE is 3 hours and is performed by an IBM Service Representative.

It is the customer's responsibility to ensure that the XPE can be located within cable distance (10 or 20 m) of the UNIX or Windows NT servers. Connection of the XPE SCSI cables to these servers is the customer's responsibility.

3.6 Supported Hardware Environments

The XPE supports the server hardware and software environments listed below.

Note: The adapters and products listed here are those that have been tested by IBM. We expect that other models of HP, Sun, and IBM RS/6000 servers will also work.

IBM RS/6000 and SP

- Adapters
 - FC 2412 (Enhanced SCSI-2 Differential Fast/Wide Adapter/A)
 - FC 6207 (PCI Differential Ultra SCSI Adapter)
- Servers
 - 7012 models G40, 397
 - 7013 models 590, 591, 595, 59H, J30, J40, and J50
 - 7017 model S70
 - 7024 models E20 and E30
 - 7025 models F30, F40, and F50
 - 7026 models H10 and H50
 - 9076 models 202, 203, 204, 205, 206, 207, 208, 2A2, 2A3, 2A4, 2A5, 2A7, 2A8, 302, 303, 304, 305, 306, 307, 308, 3A2, 3A3, 3A4, 3A5, 3A7, 3A8, 3B2, 3B3, 3B4, 3B5, 3B7, 3B8, 402, 403, 404, 405, 406, 407, and 408
- Operating systems
 - AIX 4.2.1 and 4.3

Hewlett Packard HP 9000-800

- Adapters
 - 28696A; HP-PB bus adapter (SCSI-2 Fast/Wide differential)
- Server
 - E series
- Operating system
 - HP-UX 10.01, 10.10, and 10.20

Sun Microsystems

- Adapters
 - X1062A SCSI-2 Fast/Wide Differential
 - X1065A Ultra SCSI
- Servers
 - SPARC 10
 - Ultra 2 mod 2200
- Operating system
 - Solaris 2.5.1 and 2.6

PC Servers

- Adapters (tested for PC servers running Windows NT)
 - Adaptec 2944 SCSI Fast/Wide differential
 - Symbios Logic SYM8751D SCSI Fast/Wide differential PCI
- Operating system
 - Windows NT Version 4.0 or later

A PTF for APAR PQ11919 is required to ICKDSF to support the initialization of S/390 disks for use by UNIX or Windows NT servers.

XPE supports the following RAMAC Array Storage:

IBM 3990 Model 6 Storage Control

IBM 9390 RAMAC Array Storage Control

IBM 9391/9392 RAMAC DASD (RAMAC, RAMAC 2, RAMAC 3)

IBM 3390-3 attached to 3990-6 only

IBM 9393 RAMAC Virtual Array (XPE devices must be configured as 3390-3s)

IBM 9396 RAMAC Scalable Array Storage (support available 9/98); XPE devices must be configured as 3390-3s.

Statement of Direction for support of IBM 9394/9395 RAMAC 2 Array Subsystem (with 9395 drawers formatted as 3390-3)

Note

An engineering change (EC) is required for the Service Information Message (SIM) support for each box. Please obtain the EC numbers from your IBM representative.

Throughout this book we use the term *RAMAC Array Storage* to represent all of the products listed above. Where there are differences, we point them out (for example, SnapShot is available only on the IBM 9393 RAMAC Virtual Array).

Chapter 4. XPE Implementation Planning

Detailed planning is required for an XPE implementation as, by the very nature of the product, you are bringing together technical professionals who until this point may never have worked together and whose terminology is completely different.

4.1 Workload Considerations

In many cases we would expect the UNIX or Windows NT capacity to be part of a larger OS/390 workload on the RAMAC Array Storage. Adding a UNIX or Windows NT workload to the RAMAC Array Storage is similar to attaching another S/390; all of the same capacity planning and performance considerations apply.

The characteristics of your workload will determine what performance you can expect from XPE-attached storage. For example, performance for *cache friendly* workloads will benefit significantly from the large cache on the RAMAC Array Storage and will give significantly better response times than *cache unfriendly* workloads.

In Chapter 14, "Performance Characteristics" on page 89 we discuss the performance characteristics of the XPE under various workload conditions and recommend maximum I/O rates. Review Chapter 14, "Performance Characteristics" on page 89 to determine the performance characteristics you can expect from your workload and how much capacity you should plan to attach through an XPE.

4.2 Storage Considerations

You need to decide how much disk capacity and therefore how many disks are to be provided to each server. Each 3390-3 volume is 2.83 GB, which appears as 2.5 GB of SCSI disk to the servers once formatted.

The server administrators need to plan on the disk configuration that will be implemented. For example, the UNIX server may require a large GB file system, which means a volume group containing many SCSI disks. Alternatively disk striping over many disks for performance with low capacity may be another option. The Windows NT server may require a large G: drive, which implies a drive consisting again of many SCSI disks.

The reconfiguration of the XPE is disruptive to the attached servers. Therefore if you have more than one server attached to an XPE, you may want to consider defining the maximum number of RAMAC Array Storage disks at the initial installation, to reduce the requirement to shut down all servers to add additional disks. Note: if you are using an RVA, defining your maximum configuration requires a device number for each volume but uses little RVA capacity as an empty XPE volume is compressed into a very small amount of RVA disk space. So for an RVA, defining a maximum configuration will cost little. This is not true, however, for other RAMAC Array Storage products where real capacity would need to be dedicated.

For larger server capacity configurations, you may want to attach multiple SCSI adapters from one server to the XPE. This will give you up to a maximum of 64 SCSI disks on one server (approximately 180 GB capacity).

4.3 Physical Hardware Planning Considerations

The XPE must be located within 10 or 20 m of the UNIX or Windows NT servers to which it will be attached. You need to order the correct cable length for each server.

The XPE is a desktop or floor standing unit, so you need to find a suitable location for it.

The XPE ESCON cables will either connect through a fiber patch-panel, through an ESCON Director, or directly to the RAMAC Array Storage. You need to ensure that your fiber infrastructure has the capacity for the additional links, and that your ESCON Director has the additional ports available. The XPE can share existing ports to the RAMAC devices through dynamic switching in the ESCON Director.

You should consider using an uninterruptible power supply (UPS) for the XPE to minimize any impact of power outages, particularly if you have multiple servers attached to one XPE.

4.4 Installation Worksheets

The initial configuration of the XPE is undertaken by an IBM Service Representative, who will require input from the customer. To help with the configuration and as a useful reference, we have designed new installation worksheets and include them in Appendix A, "XPE Installation Worksheets" on page 99. Some sections of the worksheets must be completed before the XPE is installed; others can be completed once the disks on the UNIX or Windows NT Servers are configured. Chapter 5, "XPE Configuration Steps" on page 19, Chapter 6, "MVS Implementation Process" on page 29, Chapter 7, "RS/6000 Implementation Process" on page 39, and Chapter 8, "Windows NT Implementation Process" on page 49 help in completing the new worksheets.

4.5 Outline Plan of Activities

Now that you are ready to start, let us outline here the activities, and people responsible for those activities, that will ensure a fully operational XPE environment. Customers may not use the titles we use for the person responsible for each activity but should know who owns the activity.

4.5.1 Installation

Table 3 lists the sequence of activities to perform to install an XPE.

Activity	Person Responsible
Determine capacity required for new UNIX or Windows NT server	Capacity planner
Select S/390 storage subsystem	Capacity planner or MVS storage administrator

<i>Table 3 (Page 2 of 2). Sample Installation Sequence</i>	
Activity	Person Responsible
Plan disk organization, volume groups, striping	UNIX or Windows NT system administrator
Plan ESCON connectivity to S/390 storage	MVS system administrator, MVS systems programmer
Select server name	Use installation standards
Decide on RAMAC device addresses	MVS storage administrator
Specify VOLSER and data set names	Use installation standards
Initialize volumes	MVS systems programmer
Install XPE near UNIX or Windows NT box	IBM Service Representative
Connect ESCON cables to S/390 storage	IBM Service Representative
Connect UNIX or Windows NT boxes to XPE	IBM Service Representative or customer technician
Fill in mandatory installation worksheets sections	MVS systems programmer, IBM Service Representative, UNIX or Windows NT systems administrator
Configure and activate XPE	IBM Service Representative
Configure SCSI disks on UNIX or Windows NT	UNIX or Windows NT systems administrator
Fill in final details of the installation worksheets	MVS systems programmer, IBM Service Representative, UNIX or Windows NT systems administrator

4.5.2 Backup and Recovery Processes

At the same time that you plan the initial installation process, it is good practice for you to develop the backup and recovery plan for the UNIX and Windows NT servers. See Table 4 and Table 5. The servers can continue to use their own backup and recovery processes for the XPE-configured disks. This redbook does not cover these processes but highlights some of the necessary considerations to implement OS/390 facilities.

<i>Table 4. Backup Activities for UNIX and Windows NT Servers</i>	
	Person Responsible
Define the backup requirements for each disk	UNIX or Windows NT system administrator
Implement the backup procedure	MVS systems programmer
Test the backup procedure	UNIX or Windows NT system administrator and MVS systems programmer

<i>Table 5. Recovery Activities for UNIX and Windows NT Servers</i>	
	Person Responsible
Request recovery process	UNIX or Windows NT system administrator
Run recovery procedure in OS/390	MVS systems programmer
Apply logs or journals on server	UNIX or Windows NT system administrator

Chapter 5. XPE Configuration Steps

Throughout this chapter, we use the configuration shown in Figure 8 to complete the sample installation worksheets. We hope working through the sample configuration will demonstrate many of the configuration issues that you may come across when installing your own XPE.

The configuration consists of two servers, an RS/6000 and a Windows NT server. After the XPE is configured, the RS/6000 sees three SCSI disks attached to its external SCSI bus (note that the RS/6000 also has two internal disks, *hdisk0* and *hdisk1*), starting at *hdisk2*.

The Windows NT server, which has two internal disks (*Disk 0* and *Disk 1*, not shown in Figure 8) will see two disks on its external XPE SCSI bus starting at *Disk 2*. *Disk 2* is formatted as the F: drive, and *Disk 3*, as the G: drive (SCSI IDs 0 and 3).

The SCSI IDs for the disks attached to the AIX and Windows NT servers are specified as part of the XPE configuration, and you have total freedom to use whatever values are appropriate.

We have taken an RVA as the S/390 storage. The XPE is attached to the RVA through two ESCON Directors, ESCON ports D1 and D5 on one Director and C7 and D4 on the other. Within the RVA we have up to four logical control units (0-3), which are referred to in the OS/390 I/O definition as control unit addresses (CUADDs) and thus are shown here as CUADD 0 through CUADD 3. The RVA disks we are using are part of a group of disks attached through CUADD 1.

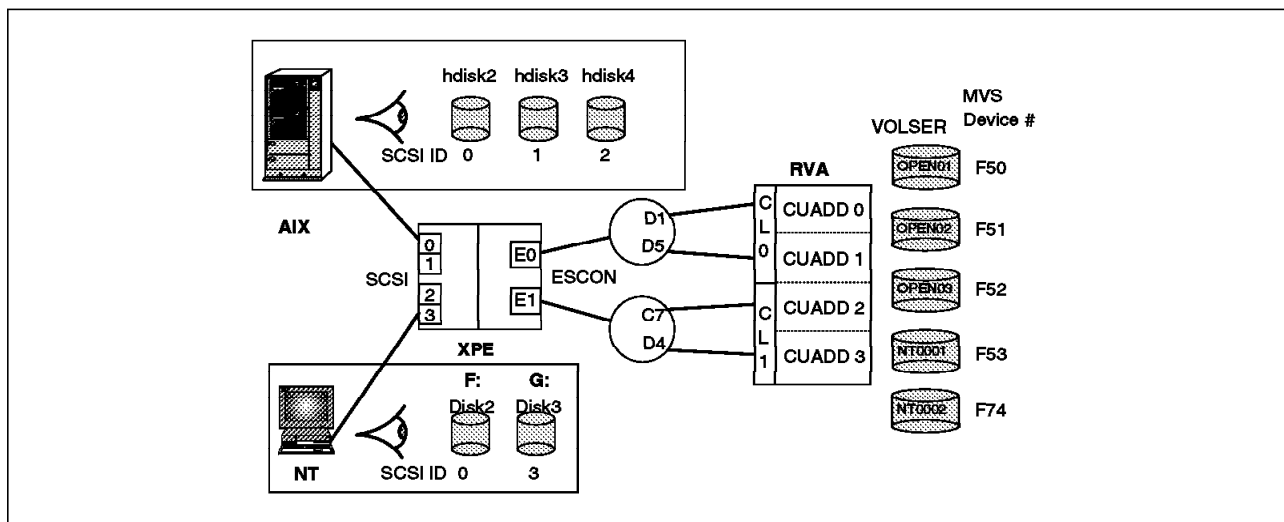


Figure 8. Sample Configuration Schematic Showing XPE Disks

The installation worksheets are useful for performing the initial setup of an XPE and as a long term reference document that will help you understand the relationships among the various devices and systems that use the XPE.

There are three installation worksheets:

Configuration worksheet

Information necessary to configure and manage UNIX and Windows NT disks with MVS volumes

ESCON worksheet	Information about the ESCON Director ports
Web configuration worksheet	Information relevant to configuring the XPE Web server

5.1 Configuration Worksheet

Table 6 on page 21 is a completed configuration worksheet for the configuration shown in Figure 8 on page 19. We discuss the worksheet in detail in the sections that follow.

Blank worksheets are found in Appendix A, “XPE Installation Worksheets” on page 99.

Table 6. XPE Configuration Worksheet

Server Information				XPE SCSI Definition				XPE ESCON Definition			MVS Definition
UNIX/NT Server Name	Logical Name	Physical Disk Name	Server Type	XPE Device Number	XPE SCSI Bus (0-3)	XPE SCSI ID (0-F)	XPE SCSI LUN (0-7)	RAMAC CUADD (0-3)	RAMAC ESCON Address	MVS VOLSER (6 Chars)	MVS Device Number
PAYSV1	PAYVG	hdisk2	RS	DEV000	0	0	0	1	10	OPEN01	F50
PAYSV1	PAYVG	hdisk3	RS	DEV001	0	1	0	1	11	OPEN02	F51
PAYSV1	DATCVG	hdisk4	RS	DEV002	0	2	0	1	12	OPEN03	F52
SQLS1	G:	Disk3	NT	DEV003	1	3	0	1	13	NT0001	F53
SQLS1	F:	Disk2	NT	DEV004	1	0	0	1	34	NT0002	F74
				DEV005							
				DEV006							
				DEV007							
				DEV008							
				DEV009							
				DEV00A							
				DEV00B							
				DEV00C							
				DEV00D							
				DEV00E							
				DEV00F							

Note: Keep this worksheet as a reference document after installation.

The worksheet is divided into four sections. The two center sections (XPE SCSI Definition and XPE ESCON Definition) are required to install the XPE. The other sections (Server Information and MVS Device Information) are optional sections that we recommend you complete. They will assist you in managing your XPE environment.

5.1.1 Server Information

The server information section consists of optional fields that help relate the server and its disks to an OS/390 device.

- **UNIX or NT Server Name** - This field identifies the UNIX or Windows NT server that is using devices through the XPE. This is a useful documentation aid when relating an MVS volume to a UNIX or Windows NT disk.
- **Logical Name** - When a volume is defined to UNIX, it is created as part of a volume group. Use this field to record the volume group name.

For a Windows NT system, this field can be used to indicate the drive letter.

- **Physical Disk Name** - For a UNIX system, disks are identified starting with *hdisk0*. There can be more than one internal disk, and the new external SCSI disks start at the next free number. The external SCSI disks are identified in ascending SCSI ID/LUN order and assigned successive *hdisk* numbers.

A Windows NT system is similar: the first disk, usually an internal disk, is *Disk0* (see the **Start**→**Administrator Tools**→**Disk Administrator**). Successive disks are identified in ascending SCSI/LUN order.

- **Server Type** - Use this field to keep track of the servers connected to the XPE. Suggested types are:
 - RS - RS/6000
 - NT - Windows NT
 - SN - Sun server
 - HP - HP server

5.1.2 XPE SCSI Definition

This section must be completed in order for the IBM Service Representative to configure the XPE.

- **XPE Device Number** - An XPE supports up to 64 devices. During installation, the IBM Service Representative will enter information for each device.

Note: When you are completing the worksheet all entries should be contiguous with no blank entries (a blank entry indicates the end of the list of devices).

- **XPE SCSI Bus** - The XPE supports four SCSI buses. These buses are identified as 0 through 3. (See Figure 6 on page 7 for the position of the four interfaces). An XPE SCSI bus is attached to only one server through a SCSI cable. The XPE supports from one to four servers, each of which can be any of the supported server types (RS/6000, HP, Sun, or Windows NT). If required for capacity reasons, more than one cable can be attached to the same server. Each XPE cable requires a separate SCSI interface at the server.
- **XPE SCSI ID** - This is the SCSI ID assigned to a physical disk being attached through the XPE. SCSI ID 7 is usually used by the UNIX or Windows NT server for its adapter ID on the bus, so 0-6 and 8-15(F) are available. This

leaves only 15 IDs. To add the sixteenth, a straightforward approach is to use one ID twice with a different LUN.

- **XPE SCSI LUN** - Within each SCSI ID, the XPE supports up to eight logical units (0-7). The selection of a LUN is arbitrary and is used only to uniquely identify several disks using the same SCSI ID (see comment above under XPE SCSI ID on identifying 16 devices).

Number of disks per SCSI bus

In fact the combination of SCSI ID and LUN allows more than 16 disks to be defined on one bus, although more than 16 is not recommended for performance reasons. Should you require high capacity without a high I/O load, up to 64 disks could be attached to one bus.

5.1.3 XPE ESCON Definition

This section is required for XPE installation. The MVS storage administrator or systems programmer decides which of the OS/390 volumes and addresses are to be used. Any devices on the selected control unit can be used as long as they emulate 3390-3 disks. The addresses need not be consecutive, but consideration should be given to grouping the volumes and naming the volumes so they can be easily identified as belonging to non-MVS systems.

Details on ESCON S/390 hardware connectivity can be obtained from configuration information in the OS/390 hardware configuration definition (HCD).

- **RAMAC CUADD** - The CUADD parameter is used in the S/390 I/O definition for ESCON control units that have multiple logical control units within a single physical control unit. An example of would be the RVA, which has four logical control units, each of 64 devices, or the RAMAC Scalable Array (RSA), which has four logical control units, each with 256 devices.

The RAMAC CUADD parameter indicates to which logical control unit on the RAMAC Array the RAMAC ESCON address (in the next field in the worksheet) applies.

For the 3990 Model 6 and 9390, the CUADD is always 0.

For the RVA and RSA, the CUADD can be determined from the output of the DEVSERV PATHS command on the MVS console or by referring to the Input Output Configuration Program (IOCP) or HCD definitions. Figure 9 on page 24 shows a sample DEVSERV output and Figure 10 on page 24 shows a sample IOCP fragment.

Using DEVSERV to determine CUADD

The MVS DEVSERV command for a specific device number gives details about VOLSER, active paths to the device, and status of caching options. It also give two fields that help identify the CUADD and ESCON addresses: CCA and DDC.

CCA indicates the channel address of the device. For the 3990 or 9390, the CCA indicates the ESCON address, and the CUADD will always be 0.

For the RVA and RSA, you can determine the CUADD from either the IOCP or HCD definitions, or by analyzing the CCA and DDC addresses in DEVSERV. If the CCA and DDC addresses are identical, the device is in the first CUADD (0), and the ESCON address is the CCA.

If the CCA and DDC addresses are different, the DDC indicates the ESCON address, and by analyzing the CCA you can determine the CUADD. To determine the CUADD you need to subtract the previous CUADDs in increments of 64 for the RVA and 256 for the RSA.

For example, for a CCA address in the 00-3F range use CUADD 0; in the 40-7F range use CUADD 1; in the 80-BF range use CUADD 2; and in the C0-FF range use CUADD 3.

```

DS P,F74
IEE459I 11.03.34 DEVSERV PATHS 152
UNIT DTYPE M CNT VOLSER  CHPID=PATH STATUS
      RTYPE  SSID CFW TC  DFW  PIN  DC-STATE CCA  DDC  ALT  CU-TYPE
0F74,33903 ,0,000,AMBJ74,56=+ 5A=+ 57=+ 5B=+
      33903C 014B Y YY. YY.  N  SIMPLEX  74  34      3990-3
***** SYMBOL DEFINITIONS *****
0 = ONLINE          + = PATH AVAILABLE
    
```

Figure 9. Sample DEVSERV Output

- **RAMAC ESCON Address** - The RAMAC ESCON address is the address that is used in combination with CUADD by the S/390 channel subsystem to address a specific device attached to a channel path. The RAMAC ESCON address is also shown as the UNITADD (rather than device number) in the IOCP definition. The number ranges from 00 to FF (256 devices). For any ESCON link, the UNITADD always starts at 00. Figure 10 shows the ICOP definition for device 0F74 indicating that it is attached to CUADD 1.

```

CNTLUNIT CUNMBR=003,          CNTLUNIT CUNMBR=004,
UNIT=3990,                   UNIT=3990,
UNITADD=((00,64)),           UNITADD=((00,64)),
PATH=(56,5A),               PATH=(57,5B),
CUADD=1                      CUADD=1
:
IODEVICE UNIT=3390,ADDRESS=(F40,64),CUNMBR=(003,004),
      UNITADD=00,STADET=Y,FEATURE=SHARED
    
```

Figure 10. Sample IOCP Fragment

Once the devices to be used by the RVA have been selected, the ESCON information can be obtained from the HCD or by using the DEVSERV command on the console.

- **MVS VOLSER** - This field identifies the MVS VOLSER of each of the disks being used by the XPE. At XPE startup, each ESCON CUADD address is interrogated, and the VOLSER on the device is checked against the VOLSER in the XPE initialization table. If the VOLSERs match, the volume is made available for use. If the VOLSERs do not match, the device will not be made available. Note that a VOLSER mismatch can cause unpredictable results if previously the Windows NT or UNIX server had been using the device. If the VOLSER is changed while the XPE is using the volume, XPE will detect the state change, and the volume will become unavailable, as highlighted on the XPE Web Status page.

5.1.4 MVS Device Information

This section provides additional information about MVS connections.

- **MVS Device Number** - For completeness and easy reference, the MVS device number should be filled in here. ICKDSF, the MVS utility that formats the *Opendisk* volumes, requires the disks to be offline to MVS when they are formatted and uses the MVS device number to identify them (See 6.3, “Initializing XPE Volumes” on page 30).

5.2 ESCON Worksheet

The ESCON worksheet defines the connections between the XPE and the RAMAC Array Storage. The connection can be either dedicated—the ESCON cable connected at one end to the XPE and the other end to an ESCON port on the RAMAC Array Storage—or through an ESCON Director. If connected through an ESCON Director, the XPE needs to know the destination port addresses to which the RAMAC Array Storage ESCON interface is connected. Each XPE ESCON link can specify up to four destination ports that provide paths to the RAMAC Array Storage.

The ESCON Director switches data frames to the correct destination port by looking at the frame header containing the port number put there by the XPE for each I/O. The selection of which destination port is done by the ESCON card in the XPE. If one port is busy, the ESCON card chooses the next port (if you have defined more than one).

When an XPE ESCON connection passes through an ESCON Director, part of the XPE configuration data is to supply for each XPE ESCON card the port numbers within the ESCON Director to which the RAMAC Array Storage is attached. In the example in Figure 11, there are two XPE ESCON cards, each with one ESCON director port (C1 and D1).

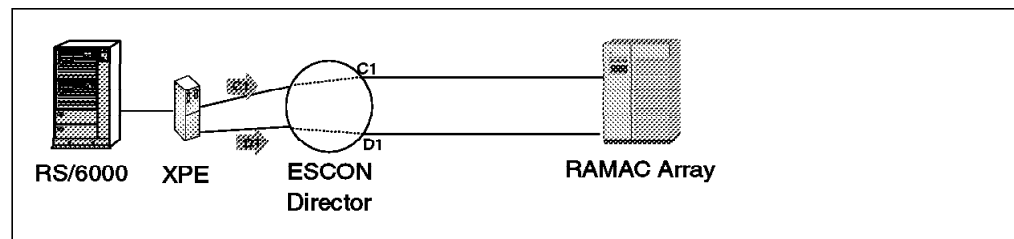


Figure 11. XPE ESCON Director Port Example: One Port per Card

If you have multiple links from the ESCON Director to the RAMAC storage, you can specify up to four ports for each XPE ESCON card. It is wise to specify more

than two ports if other systems are also using the RAMAC Array Storage. If one destination port in the ESCON Director is busy, the XPE will try the others. As an example, the RVA supports up to eight simultaneous I/O operations and so will usually have at least eight connections from one or more ESCON Directors to OS/390 systems. In the example in Figure 12 we have configured four ports for each card. If, for example, port D1 is busy because another system is using it, the XPE will try port D2 and so on until it finds a free port.

Recommendation

If you attach the XPE to a RAMAC Array Storage through an ESCON Director, we recommend that you have a minimum of two ESCON Director ports specified per XPE ESCON card. This will reduce the possibility of performance degradation in the event that one port is very busy with I/O from another system and locks out the XPE.

Define all ports on one card to one RAMAC Array Storage cluster, and all ports on the second card to the other RAMAC Array Storage cluster. This will give optimum XPE performance and at the same time protect against a cluster failure as each XPE card is connected to a different RAMAC Array Storage cluster.

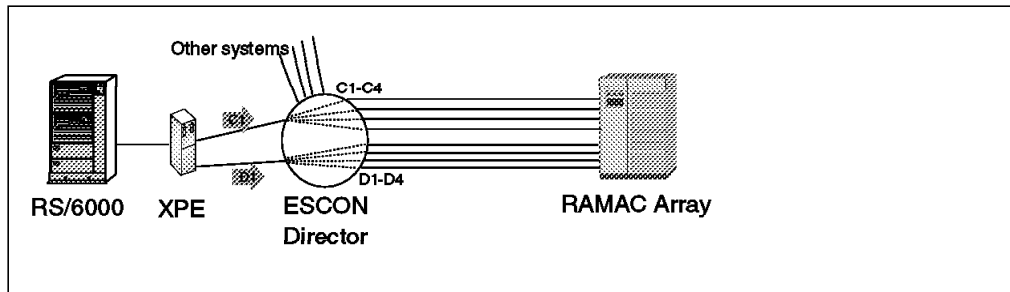


Figure 12. XPE ESCON Director Port Example: Four Ports per Card

In the configuration that we are using for our installation example (Figure 8 on page 19), two ports are defined to each of two directors. Table 7 shows a completed ESCON worksheet for this example.

Table 7. ESCON Worksheet

ESCON Card Number	Switched / Direct	ESCD Port Address	ESCD Port Address	ESCD Port Address	ESCD Port Address
ESCON Card 0	Switched	D1	D5		
ESCON Card 1	Switched	C7	D4		

Finally complete Table 12 on page 103 with your own XPE ESCON connections.

- **ESCON Card Number** - There are two ESCON cards in the XPE in slots 2 and 3. Slot 2 contains ESCON card 0 and slot 3 ESCON Card 1.

We recommend that you always connect both ESCON ports to the RAMAC Array Storage. As both paths may be used at the same time, connecting only one port may have a significant impact on throughput.

- **Switched/Direct**

If the ports are switched, the ESCON Director destination ports must be defined for each card. Up to four ports can be defined for each ESCON card. All ports must go to the same RAMAC Array Storage.

If the ESCON connection goes through an ESCON Director and the connections within the ESCON Director are dedicated, the links should be defined to the XPE as Direct.

- **ESCON Port Address** - Each ESCON card can connect directly to the &ramac, or the connection can be switched through an ESCON Director.

5.3 Web Configuration Worksheet

The Web Configuration Worksheet provides the IBM Service Representative with details needed to set up and operate the XPE Web Server.

The XPE runs a Web Server that provides a management interface. This interface is accessed through a standard Web browser such as Netscape Navigator.

The XPE has an Ethernet connection to either thick (10 Base T) or twisted pair Ethernet media. For connection to thick (10 Base 5) Ethernet media, the output connector is a thick 15-pin D-shell "slide-latch" type that is compatible with IEEE 802.3 section 7.6. For connection to twisted pair Ethernet media, an RJ-45 connector is provided. For connection to thin (10 Base 2) Ethernet media, an optional external transceiver is required.

Table 8 shows the XPE Web configuration worksheet with details of the setup used in the sample configuration.

Host name for this XPE	XPE3
IP address for this XPE	9.115.22.136
Network mask	255.255.254.0
IP address of your domain name server	9.115.0.250
Domain name	IBM.COM
IP address of your gateway	9.115.23.254

Once the IBM Service Representative has installed and configured the XPE and connected it to your Ethernet, you can point your Web browser to the IP address and use the XPE Web Interface. The universal resource locator (URL) for the example in Table 8 is <http://9.115.22.136> or XPE3.IBM.COM. For a complete discussion on the use of the XPE Web Interface, see Chapter 10, "The XPE Web Interface" on page 59.

Chapter 6. MVS Implementation Process

In this chapter we look at the XPE implementation from the perspective of the MVS system programmer or storage administrator. The MVS system programmer wants to know how to define and initialize the S/390 disks that are to be used by the UNIX or Windows NT servers. The storage administrator wants to know where to put the UNIX or Windows NT data and what it looks like. In Chapter 11, "OS/390 Operational Considerations" on page 67 we cover the backup and recovery by OS/390 of UNIX or Windows NT data stored on the RAMAC Array Storage.

6.1 Selecting Volumes

When you select volumes for use by a UNIX or Windows NT server, there are only a few restrictions as noted below. You may want to consider the existing load on the RAMAC Array Storage and the load implications of adding a number of open systems disks, just as you would adding any new workload to a storage subsystem.

Supported subsystems include:

- RAMAC Virtual Array (T42, T41, T82, T81, 001, 002)
- RAMAC 3 DASD Subsystem
- RAMAC Scalable Array
- 3990-6 or 9390 with RAMAC, RAMAC 2 Array DASD, or 3390-3 disks

Certain criteria must be met when you select the RAMAC disks and subsystem to use for an open system:

1. The disk must be defined as a 3390-3 volume.
2. The two ESCON links from an XPE must go to the same RAMAC Array Storage (usually one link to each cluster if connected directly).

The 3990, 9390, and RVA subsystems support peer-to-peer remote copy (PPRC), (RVA PPRC will be available 4Q98). If required, you can mirror the open systems disks by using PPRC at distances of up to 26 km.

Extended remote copy (XRC) is supported by 3990 and 9390 subsystems and enables remote copy of both MVS disks and XPE-configured disks at unlimited distances.

6.2 Recommendations

- We recommend that you identify the XPE-configured disks as non-OS/390 through their VOLSER, for example, OPEN01, or NT0001, or SUN001, to enable operator commands to clearly identify volumes that are not being used by OS/390. As an example, resource measurement facility (RMF) will show no activity on the XPE volumes even though they may be in active use by the UNIX or Windows NT servers. The ability to identify XPE volumes by their names may eliminate some potential confusion.
- By placing the disks in their own storage group you can apply different management policies and control access from the managing OS/390 systems.

6.3 Initializing XPE Volumes

Once the devices to be used by the UNIX or Windows NT server have been selected, they must be initialized for open systems use. You can initialize all the devices through ICKDSF on an OS/390 system, or you can initialize a single device, and use SnapShot to replicate the initialized device. You can also use other utilities to replicate preinitialized devices.

6.3.1 ICKDSF

ICKDSF has a new operand, OPENDISK, that indicates that the disk must be formatted in a new way, for both OS/390 use and open systems use. You need a PTF to APAR PQ11919 installed on ICKDSF to provide the support for the OPENDISK operand and the allocation and formatting of the new XPE data sets.

The target volumes must be offline, and the operator must reply “U” to the message on the console.

The JCL shown in Figure 13 gives an example of a job used to initialize a 3390-3 volume.

```
//AMBI380 JOB SLT,MSGCLASS=H,PERFORM=7
//*JOBPARM SYSAFF=*
//*
//*
//STRNG89B EXEC PGM=ICKDSF,REGION=6000K
//SYSPRINT DD SYSOUT=*
//SYSIN DD *
        INIT UNIT(F50) NOVERIFY PURGE -
            VOLID(AMBJ50) VTOC(OPENDISK)
/*
```

Figure 13. Sample OPENDISK Initialization Job

Figure 14 on page 31 shows the output from an ICKDSF job to initialize an XPE volume.


```

                                J E S 2  J O B  L O G  --  S Y S T E M  3 0 9 0  --  N O D E  T P C 9 7 2

10.09.29 JOB00035 IRR010I USERID IBMUSER IS ASSIGNED TO THIS JOB.
10.09.30 JOB00035 ICH70001I IBMUSER LAST ACCESS AT 10:05:08 ON THURSDAY, APRIL 30, 1998
10.09.30 JOB00035 $HASP373 IBMUSER1 STARTED - INIT 1 - CLASS A - SYS 3090
10.09.30 JOB00035 IEF403I IBMUSER1 - STARTED - TIME=10.09.30
10.09.31 JOB00035 *09 ICK003D REPLY U TO ALTER VOLUME OF50 CONTENTS, ELSE T
10.09.37 JOB00035 R 9,U
10.25.24 JOB00035 ICK061I OF50 VTOC INDEX CREATION SUCCESSFUL: VOLUME IS IN INDEX FORMAT
10.25.25 JOB00035 IBMUSER1 STEP1 ICKDSF 0000
10.25.25 JOB00035 IEF404I IBMUSER1 - ENDED - TIME=10.25.25
10.25.25 JOB00035 $HASP395 IBMUSER1 ENDED

.
1 //IBMUSER1 JOB CLASS=A,MSGLEVEL=(1,1),MSGCLASS=H,REGION=4096K
2 //STEP1 EXEC PGM=AMBJDSF
//*****
//* ICKDSF JOB TO FORMAT DASD FOR USE WITH XPE
//*****
3 //SYSPRINT DD SYSOUT=*
4 //SYSIN DD *

.
.

ICKDSF - MVS/XA/ESA R16 PQ13687 +AMBJ 04/03/98 TIME: 10:09:31 04/30/98 PAGE 1

INIT UNIT(F50) VOLID(AMBJ50) NVFY VTOC(OPENDISK) 00090003
ICK00700I DEVICE INFORMATION FOR OF50 IS CURRENTLY AS FOLLOWS:
    PHYSICAL DEVICE = 3390
    STORAGE CONTROLLER = 3990
    STORAGE CONTROL DESCRIPTOR = EC
    DEVICE DESCRIPTOR = 0A
    ADDITIONAL DEVICE INFORMATION = 71002160
ICK31061I OF50 VTOC INDEX CREATION SUCCESSFUL: VOLUME IS IN INDEX FORMAT
ICK01314I VTOC IS LOCATED AT CCHH=X'0000 0001' AND IS 1 TRACKS.
ICK00001I FUNCTION COMPLETED, HIGHEST CONDITION CODE WAS 0
    10:25:24 04/30/98

ICK00002I ICKDSF PROCESSING COMPLETE. MAXIMUM CONDITION CODE WAS 0

```

Figure 14. Output from Job to Initialize XPE Volume

The job initializes the volume and formats the OPENDISK data set with binary zeros. On an RVA, the job takes about 15 min, on a 3390-3 or RAMAC 3 it takes about 90 min.

Figure 15 is the JCL for the job used to format the disks used in the sample installation.

```

//IBMUSER1 JOB CLASS=A,MSGLEVEL=(1,1),MSGCLASS=H,REGION=4096K
//STEP1 EXEC PGM=ICKDSF
//*****
//* ICKDSF JOB TO FORMAT DASD FOR USE WITH XPE
//*****
//SYSPRINT DD SYSOUT=*
//SYSIN DD *
    INIT UNIT(F74) VOLID(AMBJ74) NVFY VTOC(OPENDISK)
    INIT UNIT(F50) VOLID(AMBJ50) NVFY VTOC(OPENDISK)
    INIT UNIT(F51) VOLID(AMBJ51) NVFY VTOC(OPENDISK)
    INIT UNIT(F52) VOLID(AMBJ52) NVFY VTOC(OPENDISK)
    INIT UNIT(F53) VOLID(AMBJ53) NVFY VTOC(OPENDISK)

```

Figure 15. JCL Used to Initialize Sample XPE Volumes

The ICKDSF job formats the volumes and creates a volume label, a one-track volume table of contents (VTOC), a VTOC index, and a small data set on cylinders 0 and 1. Ten tracks are left free on cylinder 0 for a VSAM volume data set (VVDS). The VVDS is created by CONVERTV when the volume is converted to SMS managed. Listing the VTOC shows no free tracks on the volume (see Figure 16).

The main data set, SYS1.OVDTA.Vvvvvvv.P00 (where vvvvvv is the VOLSER) occupies from cylinder 2 to the end of the volume on cylinder 3338. After formatting, the last block pointer in the VTOC is updated to indicate that the data set is full, stopping utilities from attempting to free unused tracks. Formatted capacity is approximately 2.5 GB (which is the value seen by both the UNIX and Windows NT servers).

Data set SYS1.OVDS.Vvvvvvv is reserved for future use.

The names assigned to the two data sets are not options within ICKDSF, but they can be changed after the volume has been initialized. You may want to change the names if the use of the prefix SYS1. will cause you problems with your installation standards or operation. Any name can be used, but we recommend that you retain the VOLSER information in the data set name.

You could also include the server name as part of the data set name to help you identify which server is using which volumes. For example, on a Windows NT server with an RVA disk with VOLSER of OPEN01, the open disk data set could be SYSOPEN.NT001.OVDTA.VOPEN01.P00, and the reserved data set could be SYSOPEN.NT001.OVDS.VOPEN01.

Or, for a multiple disk RS/6000 database server called RS53 running a SAP database on three disks, SAPDB1, SAPDB2, SAPDB3:

```
SYSXPE.RS53.OVDTA.VSAPDB1.P00
SYSXPE.RS53.OVDTA.VSAPDB2 P00
SYSXPE.RS53.OVDTA.VSAPDB3.P00
```

Figure 16 and Figure 17 on page 33 are IEHLIST output of an OPENDISK showing the data set names and their characteristics for a volume that has been converted to SMS managed.

```

                                SYSTEMS SUPPORT UTILITIES---IEHLIST                                PAGE    1
DATE: 1998.119  TIME: 11.12.09
                                CONTENTS OF VTOC ON VOL AJS381 <THIS IS AN SMS MANAGED VOLUME>
THERE IS A 1 LEVEL VTOC INDEX
DATA SETS ARE LISTED IN ALPHANUMERIC ORDER
-----DATA SET NAME-----   CREATED  DATE.EXP  FILE TYPE  SMS.IND  EXT  DS.VOL1  VOL.SEQ  SECURITY
SYS1.OVDS.VAJS381             1998.117  1999.365  SEQUENTIAL S        001  AJS381  00001  NONE
SYS1.OVDTA.VAJS381.P00       1998.117  1999.365  SEQUENTIAL S        001  AJS381  00001  NONE
SYS1.VTOCIX.AJS381           1998.117   00.000  SEQUENTIAL SU       001  AJS381  00001  NONE
SYS1.VVDS.VAJS381            1998.118   00.000  VSAM       S        001  AJS381  00001  PWD
THERE ARE    0 EMPTY CYLINDERS PLUS    0 EMPTY TRACKS ON THIS VOLUME
THERE ARE   44 BLANK DSCBS IN THE VTOC ON THIS VOLUME
THERE ARE   14 UNALLOCATED VIRS IN THE INDEX

```

Figure 16. IEHLIST Output Showing Data Sets Created by ICKDSF

```

SYSTEMS SUPPORT UTILITIES---IEHLIST
PAGE 2
DATE: 1998.119 TIME: 11.12.09
CONTENTS OF VTOC ON VOL AJS381 <THIS IS AN SMS MANAGED VOLUME>
THERE IS A 1 LEVEL VTOC INDEX
DATA SETS ARE LISTED IN ALPHANUMERIC ORDER
FORMAT 4 DSCB NO AVAIL/MAX DSCB /MAX DIRECT NO AVAIL NEXT ALT FORMAT 6 LAST FMT 1 VTOC EXTENT THIS DSCB
VI DSCBS PER TRK BLK PER TRK ALT TRK TRK(C-H) (C-H-R) DSCB(C-H-R)/LOW(C-H) HIGH(C-H) (C-H-R)
81 45 50 45 0 3339 0 0 1 50 0 1 0 1 0 1 1
-----DATA SET NAME----- SER NO SEQNO DATE.CRE DATE.EXP DATE.REF EXT DSORG RECFM OPTCD BLKSIZE
SYS1.OVDS.VAJS381 AJS381 1 1998.117 1999.365 00.000 1 PS FB S 00 27648
SMS.IND LRECL KEYLEN INITIAL ALLOC 2ND ALLOC EXTEND LAST BLK(T-R-L) DIR.REM F2 OR F3(C-H-R) DSCB(C-H-R)
S 512 TRKS CONTIG 0 16 2 0 0 1 3
EXTENTS NO LOW(C-H) HIGH(C-H)
0 0 13 1 14
----ON THE ABOVE DATA SET, THERE ARE 0 EMPTY TRACK(S).

-----DATA SET NAME----- SER NO SEQNO DATE.CRE DATE.EXP DATE.REF EXT DSORG RECFM OPTCD BLKSIZE
SYS1.OVDTA.VAJS381.P00 AJS381 1 1998.117 1999.365 1998.118 1 PS FB S 00 27648
SMS.IND LRECL KEYLEN INITIAL ALLOC 2ND ALLOC EXTEND LAST BLK(T-R-L) DIR.REM F2 OR F3(C-H-R) DSCB(C-H-R)
S 512 CYLS CONTIG 0 50054 2 0 0 1 4
EXTENTS NO LOW(C-H) HIGH(C-H)
0 2 0 3338 14
----ON THE ABOVE DATA SET, THERE ARE 0 EMPTY CYLINDER(S) PLUS 0 EMPTY TRACK(S).

-----DATA SET NAME----- SER NO SEQNO DATE.CRE DATE.EXP DATE.REF EXT DSORG RECFM OPTCD BLKSIZE
SYS1.VTOCIX.AJS381 AJS381 1 1998.117 00.000 00.000 1 PS F 00 2048
SMS.IND LRECL KEYLEN INITIAL ALLOC 2ND ALLOC EXTEND LAST BLK(T-R-L) DIR.REM F2 OR F3(C-H-R) DSCB(C-H-R)
SU 2048 TRKS CONTIG 0 0 21 0 0 1 5
EXTENTS NO LOW(C-H) HIGH(C-H)
0 0 2 0 2
----ON THE ABOVE DATA SET, THERE ARE 0 EMPTY TRACK(S).

-----DATA SET NAME----- SER NO SEQNO DATE.CRE DATE.EXP DATE.REF EXT DSORG RECFM OPTCD BLKSIZE
SYS1.VVDS.VAJS381 AJS381 1 1998.118 00.000 00.000 1 VS * 80 4096
SMS.IND LRECL KEYLEN INITIAL ALLOC 2ND ALLOC EXTEND LAST BLK(T-R-L) DIR.REM F2 OR F3(C-H-R) DSCB(C-H-R)
S 0 TRKS 0 0 0 58786 0 1 6
EXTENTS NO LOW(C-H) HIGH(C-H)
0 0 3 0 12
----UNABLE TO CALCULATE EMPTY SPACE.
SYSTEMS SUPPORT UTILITIES---IEHLIST
PAGE 3
VPSM A = NUMBER OF TRKS IN ADDITION TO FULL CYLS IN THE EXTENT
TRK FULL TRK FULL TRK FULL TRK FULL TRK FULL
ADDR CYLS A ADDR CYLS A ADDR CYLS A ADDR CYLS A ADDR CYLS A
THERE ARE 0 EMPTY CYLINDERS PLUS 0 EMPTY TRACKS ON THIS VOLUME
THERE ARE 44 BLANK DSCBS IN THE VTOC ON THIS VOLUME
THERE ARE 14 UNALLOCATED VIRS IN THE INDEX

```

Figure 17. IEHLIST Formatted Output of Data Sets Created by ICKDSF

6.3.2 SnapShot

You might want to use SnapShot to rapidly create a large number of XPE volumes. Formatting one volume takes about 15 min, and SnapShot just tens of seconds for each subsequent volume. So you could initialize all 64 volumes in 30 min. The process is as follows:

1. Use ICKDSF to initialize the first of your open disk volumes as OPENDISK devices.
2. Format additional volumes without the OPENDISK keyword.
3. Use SnapShot and COPYVOLID(NO) to rapidly create further volume copies.
4. Rename the data sets on each volume to reflect the VOLSER and naming convention.

6.3.3 Other Utilities

You can also use DFSMSdss, PPRC, or XRC to initialize XPE volumes.

Attention: If a disk has been used on one server and you plan to move the volume to a different server of a different platform, then you should reinitialize the disk before use to remove any old data and file system information specific to that platform.

XPE devices on an RVA will benefit from good compression ratios when they are first initialized. Once the open server updates the data, net capacity load (NCL) use will increase. There is no deleted data space release (DDSR) function for files that are deleted within the XPE data set. If you remove an XPE device from use, initializing it will also remove all the open server data and will reduce the NCL.

6.4 Converting to SMS-Managed Volumes

We recommend that you put the XPE volumes in their own storage group, to isolate them from other OS/390 volumes. You do not want to use traditional storage management tools such as DFSMSHsm to manage XPE volumes, as the volumes need to be managed in their entirety, rather than by data set.

Follow these steps to put XPE volumes in their own storage group:

1. Define the new storage group, for example, XPEGRP.

Defining a separate storage group ensures that DFSMSHsm migration will not occur on it. The last reference date for the XPE data sets is not updated. Therefore the data sets will appear to MVS as aging, whereas the XPE is continually using the volume.

You can use AUTODUMP for the storage group, but you must assess the practicality of using it. Usually, you will want to coordinate the MVS dump of a volume with the UNIX or Windows NT system. In real-life situations, you will most likely use an automated operations system on the UNIX or Windows NT system that in turn will initiate a procedure on MVS to dump the XPE volumes. You can use ABARS to back up the XPE data sets, but the same synchronization question arises.

2. Set up a storage group for the target volumes if you plan to use SnapShot to prevent allocations on those volumes. See 11.3, "Using SnapShot to Back Up UNIX or Windows NT Data" on page 70 for more detail on using SnapShot with XPE volumes.
3. Add the XPE VOLSERS to the XPEGRP storage group.
4. Add an automatic class selection (ACS) fragment to ensure that CONVERTV will be allowed by SMS (see Figure 18 on page 35). Place the fragment at the very beginning of the logic in your storage class ACS routine so that all data sets on the volume being converted receive a storage class called *standard*. The rationale for this approach is one of "brute force" assignment of a storage class in order to have the volume converted successfully. If you want to change storage classes for specific data sets later, you can use the IDCAMS ALTER command to achieve the desired storage class assignment for the data set or data sets. The key purpose of this fragment in this book is to remind you to assign a storage class for CONVERTV. Many installations do not assign storage classes to SYS1 data sets; the XPE data set names

start with SYS1, and this fragment ensures a successful conversion to SMS management.

You may also want to disable all allocations to the XPE storage group once the conversion is complete. Disabling all allocations would prevent allocations to the XPE data sets. If you are using data set dumps, set the storage group to disable new allocations only.

```
/* ITSO TEST (XPE) */  
IF (&ACSENVIR = ' CONVERT') THEN  
  DO  
    SET &STORCLAS = ' STANDARD'  
  EXIT  
END
```

Figure 18. Sample ACS Fragment for CONVERTV

5. Once the XPE volumes have been initialized, run CONVERTV to convert the volumes to SMS managed. CONVERTV automatically adds a VVDS. Once the VVDS has been added, there is no unallocated space on the volume. Figure 19 on page 36 is the output from a job we used to convert three XPE volumes to SMS managed.

```

11.13.26 JOB00044 IRR010I USERID IBMUSER IS ASSIGNED TO THIS JOB.
11.13.27 JOB00044 ICH7000I IBMUSER LAST ACCESS AT 11:04:39 ON THURSDAY, APRIL 30, 1998
11.13.27 JOB00044 $HASP373 IBMUSER7 STARTED - INIT 1 - CLASS A - SYS 3090
11.13.27 JOB00044 IEF403I IBMUSER7 - STARTED - TIME=11.13.27
11.13.31 JOB00044 IBMUSER7 STRNG89B ADDRSSU 0000
11.13.31 JOB00044 IEF404I IBMUSER7 - ENDED - TIME=11.13.31
11.13.31 JOB00044 $HASP395 IBMUSER7 ENDED

      .
      .
1 //IBMUSER7 JOB CLASS=A,MSGLEVEL=(1,1),MSGCLASS=H,REGION=4096K          JOB00044
2 //STRNG89B EXEC PGM=ADDRSSU,REGION=0M
3 //SYSPRINT DD SYSOUT=*
4 //SYSIN DD *

      .
      .
PAGE 0001      5695-DF175 DFSMSDSS VIR3.0 DATA SET SERVICES      1998.120 11:13
  CONVERTV SMS DYNAM((AMBJ50,3390)) CATALOG
ADR101I (R/I)-RI01 (01), TASKID 001 HAS BEEN ASSIGNED TO COMMAND 'CONVERTV'
  CONVERTV SMS DYNAM((AMBJ51,3390)) CATALOG
ADR101I (R/I)-RI01 (01), TASKID 002 HAS BEEN ASSIGNED TO COMMAND 'CONVERTV'
  CONVERTV SMS DYNAM((AMBJ52,3390)) CATALOG
ADR101I (R/I)-RI01 (01), TASKID 003 HAS BEEN ASSIGNED TO COMMAND 'CONVERTV'
ADR109I (R/I)-RI01 (01), 1998.120 11:13:28 INITIAL SCAN OF USER CONTROL STATEMENTS COMPLETED.
ADR016I (001)-PRIME(01), RACF LOGGING OPTION IN EFFECT FOR THIS TASK
ADR006I (001)-SETUP(01), 1998.120 11:13:28 EXECUTION BEGINS
ADR860I (001)-KVSMS(01), PROCESSING BEGINS ON VOLUME AMBJ50
ADR873I (001)-KVSMS(01), VOLUME AMBJ50 IN STORAGE GROUP $XPEVOL5 IS ELIGIBLE FOR CONVERSION TO SMS MANAGEMENT
ADR877I (001)-KVSMS(01), THE FOLLOWING DATA SETS ON VOLUME AMBJ50 WERE SUCCESSFULLY PROCESSED
      SYS1.OVDS.VAMBJ50          CATALOG: SYS1.MVSRES.MASTCAT
      STORCLAS: STANDARD          MGMTCLAS: MCLASS1
ADR879I (001)-KVSMS(01), THE FOLLOWING DATA SETS ON VOLUME AMBJ50 WERE ALREADY IN SMS STATUS
      SYS1.OVDTA.VAMBJ50.P00     CATALOG: SYS1.MVSRES.MASTCAT
      STORCLAS: STANDARD          MGMTCLAS: MCLASS1
ADR885I (001)-KVSMS(01), VOLUME AMBJ50 HAS BEEN SUCCESSFULLY CONVERTED TO SMS MANAGEMENT

PAGE 0002      5695-DF175 DFSMSDSS VIR3.0 DATA SET SERVICES      1998.120 11:13
-ADR892I (001)-KVRPT(01), THE STATUS OF EACH VOLUME IS AS FOLLOWS
      VOLUME          FINAL STATUS          REASON FOR FAILURE
      -----          -----          -----
      AMBJ50 - CONVERTED          SMS
ADR013I (001)-CLTSK(01), 1998.120 11:13:29 TASK COMPLETED WITH RETURN CODE 0000
ADR016I (002)-PRIME(01), RACF LOGGING OPTION IN EFFECT FOR THIS TASK
ADR006I (002)-SETUP(01), 1998.120 11:13:29 EXECUTION BEGINS
ADR860I (002)-KVSMS(01), PROCESSING BEGINS ON VOLUME AMBJ51
ADR873I (002)-KVSMS(01), VOLUME AMBJ51 IN STORAGE GROUP $XPEVOL5 IS ELIGIBLE FOR CONVERSION TO SMS MANAGEMENT
ADR877I (002)-KVSMS(01), THE FOLLOWING DATA SETS ON VOLUME AMBJ51 WERE SUCCESSFULLY PROCESSED
      SYS1.OVDS.VAMBJ51          CATALOG: SYS1.MVSRES.MASTCAT
      STORCLAS: STANDARD          MGMTCLAS: MCLASS1
ADR879I (002)-KVSMS(01), THE FOLLOWING DATA SETS ON VOLUME AMBJ51 WERE ALREADY IN SMS STATUS
      SYS1.OVDTA.VAMBJ51.P00     CATALOG: SYS1.MVSRES.MASTCAT
      STORCLAS: STANDARD          MGMTCLAS: MCLASS1
ADR885I (002)-KVSMS(01), VOLUME AMBJ51 HAS BEEN SUCCESSFULLY CONVERTED TO SMS MANAGEMENT

PAGE 0003      5695-DF175 DFSMSDSS VIR3.0 DATA SET SERVICES      1998.120 11:13
ADR892I (002)-KVRPT(01), THE STATUS OF EACH VOLUME IS AS FOLLOWS
      VOLUME          FINAL STATUS          REASON FOR FAILURE
      -----          -----          -----
      AMBJ51 - CONVERTED          SMS
ADR013I (002)-CLTSK(01), 1998.120 11:13:30 TASK COMPLETED WITH RETURN CODE 0000
ADR016I (003)-PRIME(01), RACF LOGGING OPTION IN EFFECT FOR THIS TASK
ADR006I (003)-SETUP(01), 1998.120 11:13:30 EXECUTION BEGINS
ADR860I (003)-KVSMS(01), PROCESSING BEGINS ON VOLUME AMBJ52
ADR873I (003)-KVSMS(01), VOLUME AMBJ52 IN STORAGE GROUP $XPEVOL5 IS ELIGIBLE FOR CONVERSION TO SMS MANAGEMENT
ADR877I (003)-KVSMS(01), THE FOLLOWING DATA SETS ON VOLUME AMBJ52 WERE SUCCESSFULLY PROCESSED
      SYS1.OVDS.VAMBJ52          CATALOG: SYS1.MVSRES.MASTCAT
      STORCLAS: STANDARD          MGMTCLAS: MCLASS1
ADR879I (003)-KVSMS(01), THE FOLLOWING DATA SETS ON VOLUME AMBJ52 WERE ALREADY IN SMS STATUS
      SYS1.OVDTA.VAMBJ52.P00     CATALOG: SYS1.MVSRES.MASTCAT
      STORCLAS: STANDARD          MGMTCLAS: MCLASS1
ADR885I (003)-KVSMS(01), VOLUME AMBJ52 HAS BEEN SUCCESSFULLY CONVERTED TO SMS MANAGEMENT

PAGE 0004      5695-DF175 DFSMSDSS VIR3.0 DATA SET SERVICES      1998.120 11:13
ADR892I (003)-KVRPT(01), THE STATUS OF EACH VOLUME IS AS FOLLOWS
      VOLUME          FINAL STATUS          REASON FOR FAILURE
      -----          -----          -----
      AMBJ52 - CONVERTED          SMS
ADR013I (003)-CLTSK(01), 1998.120 11:13:31 TASK COMPLETED WITH RETURN CODE 0000
ADR012I (SCH)-DSSU (01), 1998.120 11:13:31 DFSMSDSS PROCESSING COMPLETE. HIGHEST RETURN CODE IS 0000

```

Figure 19. Output from CONVERTV for XPE Volumes

6.5 Security Considerations for XPE Volumes and Data Sets

The data stored on XPE disks most likely represents an asset that, like any other asset in an enterprise, must be protected from accidental or intentional destruction or disclosure.

The XPE volume and the XPE data sets must be protected with the OS/390 Security Server or a functionally equivalent product. As for any other disk data set or volume, the DASDVOL and DATASET resource classes are available and should be used to protect the UNIX or Windows NT data. Both resource classes are used by facilities such as DFSMSdss and SnapShot when an attempt is made to access data. In the XPE context, authorize appropriate operations user IDs or groups to back up the XPE data.

Chapter 7. RS/6000 Implementation Process

In this chapter we describe what you as an AIX systems administrator need to do to set up an RS/6000 to use the disks attached through the XPE. The examples we use are based on an XPE that has been configured to provide four SCSI disks to an RS/6000 which has two internal SCSI disks.

7.1 Initial Steps

Once the devices have been initialized on the RAMAC Array Storage and configured to the XPE, you can start using the new disks on the RS/6000.

AIX scripts are available on the XPE LIC CD. These scripts will:

- Change the queue depth for XPE-configured disks to 8 from 3
- Change the `rw_timeout` value on the XPE-configured disks to 600 sec from 30 sec

This change allows for the rare cases when the RAMAC Array Storage may be undertaking some error recovery that will take precedence over any outstanding I/O operation.

- Add definitions for a 9399 disk device type and associated device attributes.

The install script can only be installed and run on AIX 4.1 or higher. You must have a minimum of 1 MB disk space available.

To run the install script, attach the XPE to the RS/6000 server and power on the server and all other attached devices, excluding the XPE. From the superuser ID follow these steps to install the script from the XPE CD-ROM:

1. Insert the CD-ROM into your CD-ROM drive and mount the drive as a file system, using this command:

```
/usr/sbin/mount -prv cdrfs /dev/cd0 /cdrom
```

where `/cdrom` is your CD-ROM mount directory.

2. Run the `rs6k9399_setup` script to attach the XPE, using this command:

```
/cdrom/rs6k9399_setup /cdrom
```

where `/cdrom` is your CD-ROM mount directory. The `rs6k9399_setup` script will un-tar all related files to the `/usr/lpp/9399` directory.

3. Invoke the `/usr/lpp/9399/9399inst` script to start the installation. Use this command:

```
/usr/lpp/9399/9399inst
```

4. Once you have installed the script, shut down the RS/6000 and the XPE, connect the SCSI cable between the RS/6000 and XPE, and reboot the XPE.
5. Reboot the RS/6000 only when the XPE is fully operational, to ensure that the XPE has enabled its SCSI interface to the RS/6000. There are two ways of telling when the XPE is operational. The most certain way is to use the Web Interface on a networked machine, as explained in detail in Chapter 10, "The XPE Web Interface" on page 59. An alternative, but less reliable approach, is to issue ping commands to the Internet Protocol (IP) address of the XPE defined in the XPE Web configuration worksheet.

After the XPE has been configured and the RS/6000 is booted, new disks will become visible to AIX. Figure 20 on page 40 shows the results of issuing the `lsdev -Cc` and `lspv` commands on an RS/6000 connected to an XPE.

```
lsdev -Cc adapter | grep 02
scsi0 Available 00-02 Wide SCSI I/O Controller Adapter

lsdev -Cc disk

hdisk0 Available 00-01-00-0,0 2.2 GB SCSI Disk Drive
hdisk1 Available 00-01-00-1,0 400 MB SCSI Disk Drive
hdisk2 Available 00-02-01-0,0 IBM 9399001
hdisk3 Available 00-02-01-1,0 IBM 9399001
hdisk4 Available 00-02-01-2,0 IBM 9399001
hdisk5 Available 00-02-01-3,0 IBM 9399001

lspv

hdisk0 0001247451612c6e rootvg
hdisk1 0000149621346efb None
hdisk2 none None
hdisk3 none None
hdisk4 none None
hdisk5 none None
```

Figure 20. `lsdev` and `lspv` Command Output

In Figure 20, 9399001 indicates that the disks are configured by and connected through the XPE. The SCSI adapter connected to the XPE is in slot 2, and its SCSI address, although not visible here, is the standard ID of 7. The new disks have SCSI IDs of 0, 1, 2, and 3. The LUN is 0 for all of the disks. If 16 disks were being defined for the RS/6000, the LUN portion would have to be set for at least one disk, because only 15 SCSI IDs would be available (0 to 15, minus 7). The LUN can be a value ranging from 0 to 7.

Unlike the two internal disks (`hdisk0` and `hdisk1`), the new SCSI disks do not have a physical volume identifier (PVID) until a volume group is created on them.

7.2 Some Thoughts before Using The New SCSI Disks

Several options are open to you when planning the layout of the disks:

- One volume group assigned to each XPE disk

This simple option allows a one-to-one relationship between an AIX volume group and an MVS volume and simplifies backup and restore procedures.

- Multiple logical volumes on each XPE disk

The backup and recovery procedures from MVS work on a single MVS volume, so all of the AIX logical volumes on the one MVS volume will be backed up or restored at the same time. Depending on the application, this option may be an acceptable solution.

- Volume group across several XPE disks

If an application needs a large logical volume, you can group the physical disks together to make very large logical volumes. When backing up a large logical volume spanning many disks, you must back up all disks at the same point in time. If you are using SnapShot, reserve sufficient spare RVA volumes to snap all disks in the volume set.

- Striping data across multiple disks

The RAMAC Array Storage provides RAID 5 protection for data availability, and AIX can configure striping at a per logical volume level.

Striping improves performance of sequential reads and writes by reading or writing the stripes in parallel. The RVA also has a large disk cache and detects sequential operations, staging the data from disk into its cache. Because the XPE has two concurrent paths to the RAMAC Array Storage, striping can improve throughput.

One advantage the RAMAC Array Storage does bring is that the RVA has spare disks so that in the event of a disk failure, the failing disk data is rebuilt automatically on a spare disk.

- Mirroring

Mirroring (or RAID 1) provides fault tolerance to a disk failure. Choosing mirroring is largely redundant because of the RVA RAID 6 design (which handles two simultaneous disk failures) and hot spare.

In Chapter 15, “Increasing Systems Availability” on page 93 we discuss these issues of disk and component failure and explain how to address them.

The standard AIX logical volume manager (LVM) commands are used to create volume groups, logical volumes, and file systems. All of the AIX LVM restrictions, such as maximum number of physical volumes in a volume group, still apply.

LVM overview for the MVS specialist

In AIX physical disks are known as hdisks and are first placed in a volume group (VG) with a name that you specify. There is a maximum of 32 physical disks in a single volume group. You then create logical volumes (LVs) in the VG and specify such options as LV name, striping, mirroring, and placement of LV on the disks. If you have an application that uses file system I/O, then in addition you will create a unique file system name that mounts on top of a specific LV. For data integrity purposes, when using file system I/O, you also need to create a log logical volume (loglv) in the same VG as the LV.

7.3 Differences from Standard SCSI Disks

There are two notable differences between XPE-connected disks and non-XPE-connected disks when logical volumes are created or changed, typically through the System Management Interface Tool (SMIT). Figure 21 on page 42 shows the SMIT Add a Logical Volume panel.

Add a Logical Volume	
Type or select values in entry fields. Press Enter AFTER making all desired changes.	
[TOP]	[Entry Fields]
Logical volume NAME	<input type="text"/>
* VOLUME GROUP name	xpelvg
* Number of LOGICAL PARTITIONS	<input type="text"/>
PHYSICAL VOLUME names	<input type="text"/>
Logical volume TYPE	<input type="text"/>
POSITION on physical volume	outer_middle
RANGE of physical volumes	minimum
MAXIMUM NUMBER of PHYSICAL VOLUMES to use for allocation	<input type="text"/>
Number of COPIES of each logical partition	1
Mirror Write Consistency?	yes
Allocate each logical partition copy on a SEPARATE physical volume?	yes
RELOCATE the logical volume during reorganization?	yes
Logical volume LABEL	<input type="text"/>
MAXIMUM NUMBER of LOGICAL PARTITIONS	[128]
Enable BAD BLOCK relocation?	yes
SCHEDULING POLICY for writing logical partition copies	parallel
Enable WRITE VERIFY?	no
File containing ALLOCATION MAP	<input type="text"/>
Stripe Size?	[Not Striped]

Figure 21. SMIT Add a Logical Volume Panel

Three parameters on the Add a Logical Volume panel require careful consideration:

- POSITION on physical volume

States where on the disk the logical volume data is to reside. The options are inner edge, inner middle, center, outer middle, and outer edge. This performance-based option is to minimize the disk arm movement. RAMAC Array Storage controllers have large cache to aid high performance in an S/390 environment which masks the performance impact of physical disk arm movement. Leave the default option unchanged as, except for the 3390-3, the RAMAC Array Storage has its own technique for data placement. See B.1, "Overview of RVA" on page 105 for more details on the RVA.

The one exception to this recommendation is when you are using the real 3390 Model 3 disks, where you can specify the position of the data on disk.

- Number of COPIES of each logical partition

If any enhanced data availability such as RAID 5 or RAID 1 (dual copy) is implemented on the RAMAC Array Storage, there is really no need to specify a number other than 1.

- Enable BAD BLOCK relocation?

If enabled, this parameter redirects read-write requests from a disk block that can no longer retain data to one that can. The process is transparent to any application. However, the RAMAC Array Storage controllers have sophisticated error checking and correction built in, and, because they will be managing all I/O requests, set this parameter to NO.

7.4 Output from a Sample Configuration

Figure 22 shows the output for the `lspv` command issued against `hdisk2` and `hdisk3` and demonstrates how the XPE disks appear similar to “normal” SCSI disks when you use the following commands:

- volume group named `xpe1vg` using `hdisk2` and `hdisk3`.
Note: because the volume group size is more than 4 GB, the physical partition size must be 8 MB.
- logical volume `xpebiglv` in `xpe1vg`
- logical volume `xpe1lv` in `xpe1vg`
- log logical volume `xpebiglvlog` in `xpe1vg`
- volume group named `xpe2vg` using `hdisk4`.
- logical volume `xpe2lv` in `xpe2vg`
- log logical volume `xpe2lvlog` in `xpe2vg`
- filesystem `/xpebigfs` using `xpebiglv`

```
lspv hdisk2
PHYSICAL VOLUME:   hdisk2                VOLUME GROUP:   xpe1vg
PV IDENTIFIER:    0001247424110cc1        VG IDENTIFIER   0001247424111633
PV STATE:         active
STALE PARTITIONS: 0
PP SIZE:          8 megabyte(s)          ALLOCATABLE:    yes
TOTAL PPs:        317 (2536 megabytes)   LOGICAL VOLUMES: 1
FREE PPs:         0 (0 megabytes)        VG DESCRIPTORS: 2
USED PPs:         317 (2536 megabytes)
FREE DISTRIBUTION: 00..00..00..00..00
USED DISTRIBUTION: 64..63..63..63..64

lspv hdisk3
PHYSICAL VOLUME:   hdisk3                VOLUME GROUP:   xpe1vg
PV IDENTIFIER:    0001247424112f08        VG IDENTIFIER   0001247424111633
PV STATE:         active
STALE PARTITIONS: 0
PP SIZE:          8 megabyte(s)          ALLOCATABLE:    yes
TOTAL PPs:        317 (2536 megabytes)   LOGICAL VOLUMES: 3
FREE PPs:         193 (1544 megabytes)    VG DESCRIPTORS: 1
USED PPs:         124 (992 megabytes)
FREE DISTRIBUTION: 00..44..23..62..64
USED DISTRIBUTION: 64..19..40..01..00
```

Figure 22. Sample Output of `lspv` Command on `hdisk2` and `hdisk3`

Figure 23 on page 44 shows the sample output of the `lspv` command issued against `hdisk4`.

```

lspv hdisk4

PHYSICAL VOLUME:   hdisk4                VOLUME GROUP:   xpe2vg
PV IDENTIFIER:    000124742411d81c         VG IDENTIFIER   000124742411de8b
PV STATE:         active
STALE PARTITIONS: 0
PP SIZE:          4 megabyte(s)           ALLOCATABLE:    yes
TOTAL PPs:        634 (2536 megabytes)    LOGICAL VOLUMES: 2
FREE PPs:         583 (2332 megabytes)    VG DESCRIPTORS: 2
USED PPs:         51 (204 megabytes)
FREE DISTRIBUTION: 127..77..126..126..127
USED DISTRIBUTION: 00..50..00..01..00

```

Figure 23. Sample Output of Lspv Command on hdisk4

Looking at Figure 23, you see that hdisk4 is in the xpe2vg volume group, a volume group with only one disk. The disk capacity is less than 4 GB, so the default partition size of 4 MB is satisfactory.

Figure 24 shows the output of issuing the lspv -p command against hdisk2 and hdisk3.

```

lspv -p hdisk2

hdisk2:
PP RANGE  STATE  REGION      LV NAME          TYPE      MOUNT POINT
 1-64     used   outer edge   xpebiglv         jfs       /xpebigfs
 65-127   used   outer middle xpebiglv         jfs       /xpebigfs
128-190   used   center       xpebiglv         jfs       /xpebigfs
191-253   used   inner middle xpebiglv         jfs       /xpebigfs
254-317   used   inner edge   xpebiglv         jfs       /xpebigfs

lspv -p hdisk3

hdisk3:
PP RANGE  STATE  REGION      LV NAME          TYPE      MOUNT POINT
 1-64     used   outer edge   xpebiglv         jfs       /xpebigfs
 65-83    used   outer middle xpebiglv         jfs       /xpebigfs
 84-127   free   outer middle
128-167   used   center       xpe1lv           jfs       N/A
168-190   free   center
191-191   used   inner middle xpebiglvlog      jfslog    N/A
192-253   free   inner middle
254-317   free   inner edge

```

Figure 24. Sample Output of Lspv -p Command on hdisk2 and hdisk3

As you can see in Figure 24, xpebiglv is a large logical volume of 400 physical partitions spanning two disks, whereas xpe1lv is a smaller logical volume contained within one disk, hdisk3.

Figure 25 on page 45 shows the output of issuing the lspv -p command against hdisk4.

```

lspv -p hdisk4

hdisk4:
PP RANGE  STATE  REGION          LV ID          TYPE          MOUNT POINT
  1-127   free   outer edge
128-177   used   outer middle   xpe2lv         jfs           /xpe2fs
178-254   free   outer middle
255-380   free   center
381-381   used   inner middle   xpe2lvlog      jfslog        N/A
382-507   free   inner middle
508-634   free   inner edge

```

Figure 25. Sample Output of `lspv -p` Command on `hdisk4`

In Figure 25, `xpe2lv` is a logical volume on `hdisk4`.

In all cases LVM mirroring is not implemented. There is unlikely to be a disk error as seen by AIX, as the RAMAC Array Storage has implemented RAID to provide data availability. Providing for greater data availability is discussed in Chapter 15, “Increasing Systems Availability” on page 93.

Figure 26 shows the output from a command to show the logical volumes existing on specific `hdisks` and the `hdisks` that contain a specific logical volume.

```

lspv -l hdisk2

hdisk2:
LV NAME          LPs  PPs  DISTRIBUTION          MOUNT POINT
xpebiglv         317  317  64..63..63..63..64  /xpebigfs

lspv -l hdisk3

hdisk3:
LV NAME          LPs  PPs  DISTRIBUTION          MOUNT POINT
xpebiglv         83   83   64..19..00..00..00  /xpebigfs
xpe1lv           40   40   00..00..40..00..00  N/A
xpebiglvlog      1    1    00..00..00..01..00  N/A

lspv -l hdisk4

hdisk4:
LV NAME          LPs  PPs  DISTRIBUTION          MOUNT POINT
xpe2lv           50   50   00..50..00..00..00  /xpe2fs
xpe2lvlog        1    1    00..00..00..01..00  N/A

lslv -l xpe1lv

xpe1lv:N/A
PV              COPIES          IN BAND          DISTRIBUTION
hdisk3          040:000:000    100%             000:000:040:000:000

lslv -l xpe2lv

xpe2lv:/xpe2fs
PV              COPIES          IN BAND          DISTRIBUTION
hdisk4          050:000:000    100%             000:050:000:000:000

```

Figure 26. Sample Output of `lspv -l` and `lslv -l` Commands

In Figure 26, the distribution of the data in the LVs can be seen on hdisk2, hdisk3, and hdisk4. Even though AIX is issuing read and write requests to a specific region of the disk, placement is meaningless with the RAMAC Array Storage except for the real 3390 Model 3 disks. Figure 27 on page 46 shows the output of issuing the `lsvg` command with various options against `xpe1vg`.

```

lsvg xpe1vg

VOLUME GROUP:  xpe1vg                VG IDENTIFIER:  0001247424111633
VG STATE:      active                 PP SIZE:        8 megabyte(s)
VG PERMISSION: read/write            TOTAL PPs:      634 (5072 megabytes)
MAX LVs:       256                    FREE PPs:       193 (1544 megabytes)
LVs:           3                       USED PPs:       441 (3528 megabytes)
OPEN LVs:      2                       QUORUM:         2
TOTAL PVs:     2                       VG DESCRIPTORS: 3
STALE PVs:     0                       STALE PPs       0
ACTIVE PVs:    2                       AUTO ON:        yes

lsvg -p xpe1vg

xpe1vg:
PV_NAME      PV STATE  TOTAL PPs  FREE PPs  FREE DISTRIBUTION
hdisk2       active    317        0          00..00..00..00..00
hdisk3       active    317        193       00..44..23..62..64

lsvg -l xpe1vg

xpe1vg:
LV NAME      TYPE      LPs   PPs   PVs  LV STATE  MOUNT POINT
xpebiglv    jfs      400   400   2    open/syncd  /xpebigfs
xpe1lv      jfs      40    40    1    closed/syncd  N/A
xpebiglvlog jfslog   1     1     1    open/syncd   N/A

```

Figure 27. Sample Output of Various `lsvg` Commands

As you can see in Figure 20 on page 40 through Figure 27, when you use the standard AIX LVM commands of `lspv`, `lslv`, and `lsvg`, there is nothing to inform you that the SCSI disks are connected through the XPE.

Even in the `/etc/filesystems` file, there is no indication of the data being stored through XPE. This level of detail is shown in the `/xpebigfs` and `/xpe2fs` file system definitions in Figure 28 on page 47.


```

tail -16 /etc/filesystems

/xpebigfs:
    dev      = /dev/xpebiglv
    vfs      = jfs
    log      = /dev/xpebiglvlog
    mount    = false
    options  = rw
    account  = false

/xpe2fs:
    dev      = /dev/xpe2lv
    vfs      = jfs
    log      = /dev/xpe2lvlog
    mount    = false
    options  = rw
    account  = false

```

Figure 28. Sample Output: Display File System Definitions

Once the required volume groups, logical volumes, and file systems are created, they can be treated as per normal SCSI or SSA disks.

In Chapter 12, “AIX Operational Considerations” on page 75 we discuss the facilities in an S/390 environment for managing RS/6000 data.

7.5 Completing the Configuration Worksheet

Before moving on, the configuration worksheet in Appendix A, “XPE Installation Worksheets” on page 99 needs to be completed with:

- UNIX/NT Server Name** The hostname of the RS/6000
- Logical Names** A unique name for AIX support administration purposes, typically the volume group name
- Physical Disk Name** The hdisk number

For AIX, multiple hdisks could be contained within any volume group.

With these fields completed, you can correlate any AIX hdisk with its corresponding MVS VOLSER and device number.

7.6 AIX Recommendations with XPE

The most straightforward action is to keep a simple relationship with one disk configured to be one volume group.

If a one-to-one relationship is not possible, it should be noted that, with multiple disks per drive, either for capacity or disk striping, if you have to recover data from one of the disks, you must recover all of the XPE disks containing the LV.

Where possible, give the VGs, LVs, and file systems names indicating the data is actually behind the XPE.

It is good discipline to update the XPE configuration worksheet and keep it up to date after any changes or additions, either in hardcopy or on the XPE Web Interface.

Chapter 8. Windows NT Implementation Process

In this chapter we describe what the you as the Windows NT administrator need to do to set up a Windows NT system to use disks attached through the XPE.

8.1 Before Connecting to the XPE

Before you connect the Windows NT server to the XPE, we recommend that you boot the server to check the existing internal disk numbers. It is difficult with Windows NT to match a SCSI ID or LUN with a drive or drive letter.

After booting the stand-alone Windows NT Server, select the **Start**→**Programs**→**Administrative Tools**→**Disk Administrator** to look at the existing disk configuration (see Figure 29 and Figure 30).

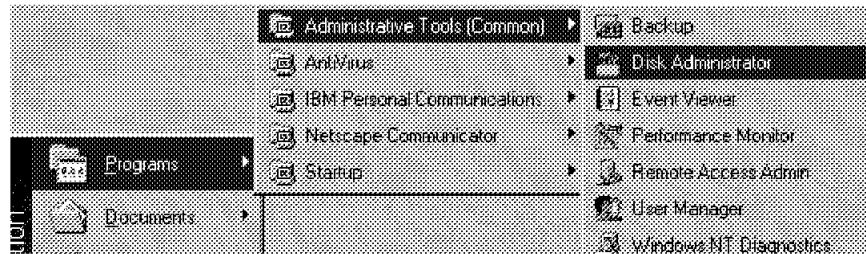


Figure 29. Selecting the Windows NT Disk Administrator Panel

Identify the disks that are currently installed and record the results. The very first disk in the list should be *Disk 0*, typically on the C: Drive. If there are other drives in the Windows NT server, they are listed in the configuration.

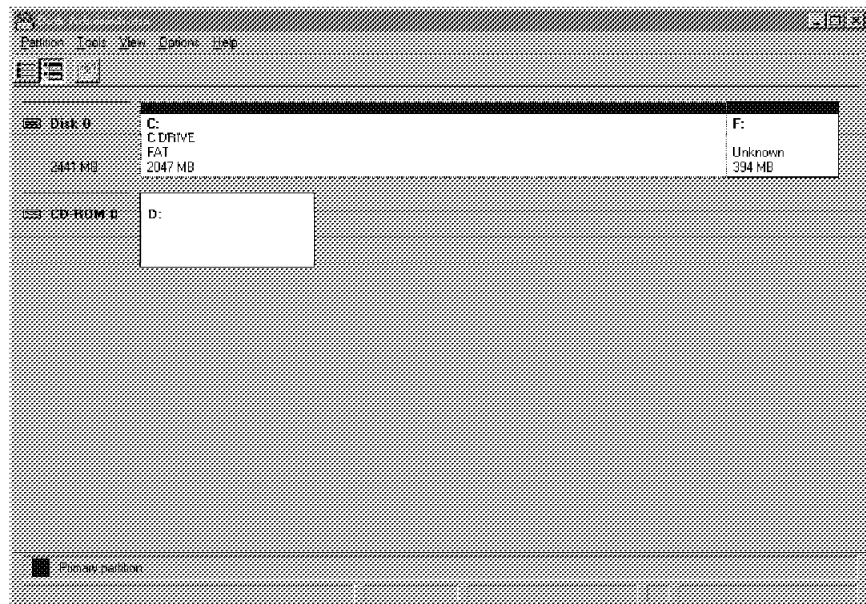


Figure 30. Windows NT Disk Administrator Panel

Now shut down the Windows NT server and connect the SCSI cable between the server and the XPE.

8.2 After Connecting to the XPE

After you have updated the XPE configuration and booted the XPE, you can boot the Windows NT server. This sequence ensures that the XPE has enabled its SCSI interface to the Windows NT server before Windows NT boots.

During the Windows NT boot process, any new disks will be discovered, and Windows NT will issue a message indicating that it has found additional disks. Details of the new disks can be found by looking at the SCSI information on the SCSI Adapters icon on the Control Panel (see Figure 31).

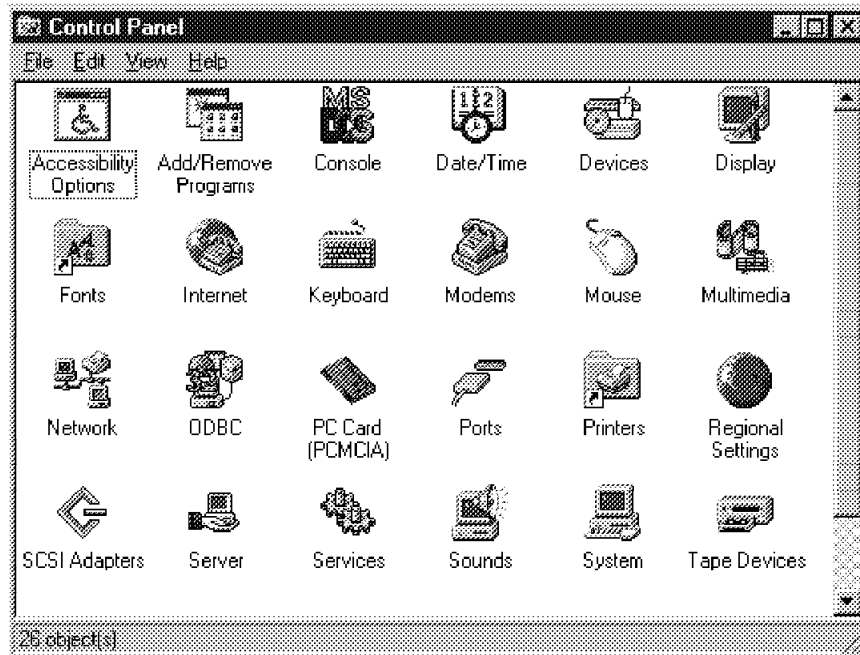


Figure 31. SCSI Adapters Icon on Windows NT Control Panel

Double-click on the SCSI Adapters icon to see a list of installed adapters and devices (see Figure 32 on page 51).

The IBM 9399001 description in the SCSI Adapters panel with all the disks exploded into view indicates that these disks are configured by and connected through the XPE.

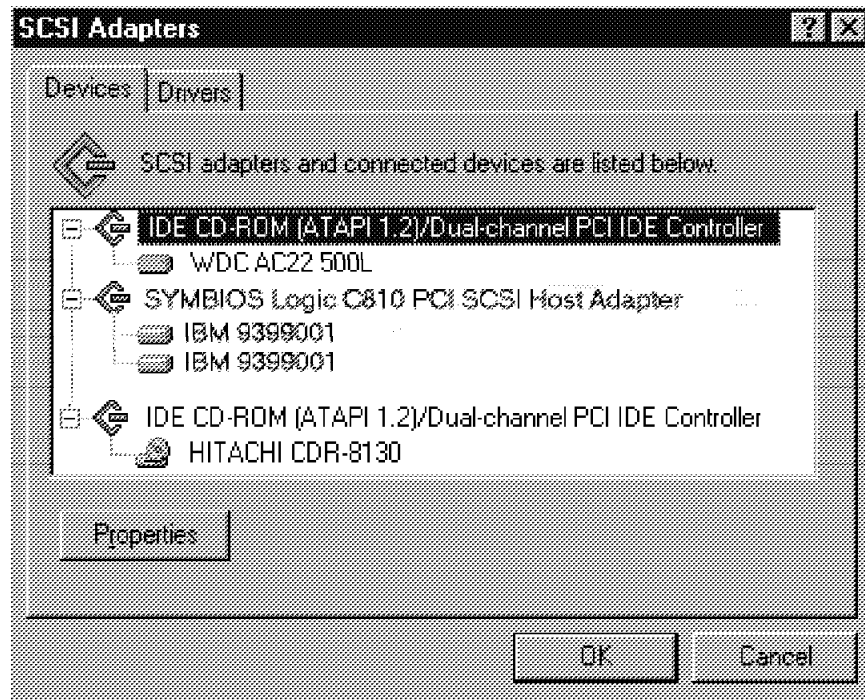


Figure 32. SCSI Adapters Panel Showing XPE SCSI Devices

The disks attached to the XPE will show immediately under the SCSI Differential adapter definition that is connected to the XPE. Expand the list by clicking on the “+” symbol, and you will see a list of XPE devices. These can be identified by the name *IBM 9399001*. Select each in turn and click on the **Properties** button to get to the Properties window. Then click on the **Settings** tab to get to the Settings page (see Figure 33 on page 52).

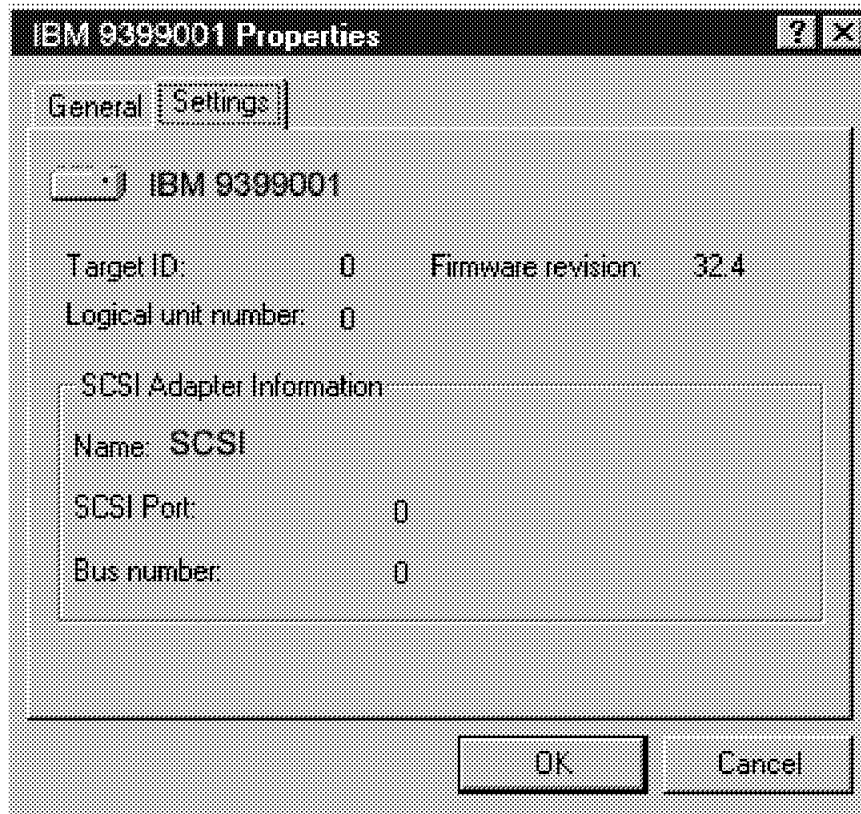


Figure 33. Windows NT SCSI ID and LUN Information

On the Settings page for each SCSI disk, the four fields of interest are:

- Target ID:
- Logical unit number:
- SCSI Port:
- Bus number:

The number of disks and the SCSI ID and LUNs should match the number of disks and the SCSI IDs and LUNs recorded in the XPE configuration worksheet.

If there are any differences, compare the Windows NT SCSI ID and LUN with the active configuration on the XPE Web Interface and with the completed XPE configuration worksheet. Missing disks may be caused by duplicate SCSI IDs or LUNs or by the OS/390 disk VOLSER not matching the VOLSER in the active XPE configuration.

8.3 Making the Disks Usable by Windows NT

Once you have determined that the disks are visible to Windows NT, you can go ahead and format the disks and assign drive letters.

Figure 34 on page 53 shows the new disks before they have been formatted after Windows NT has detected them during boot (Disk 0 is the internal disk).

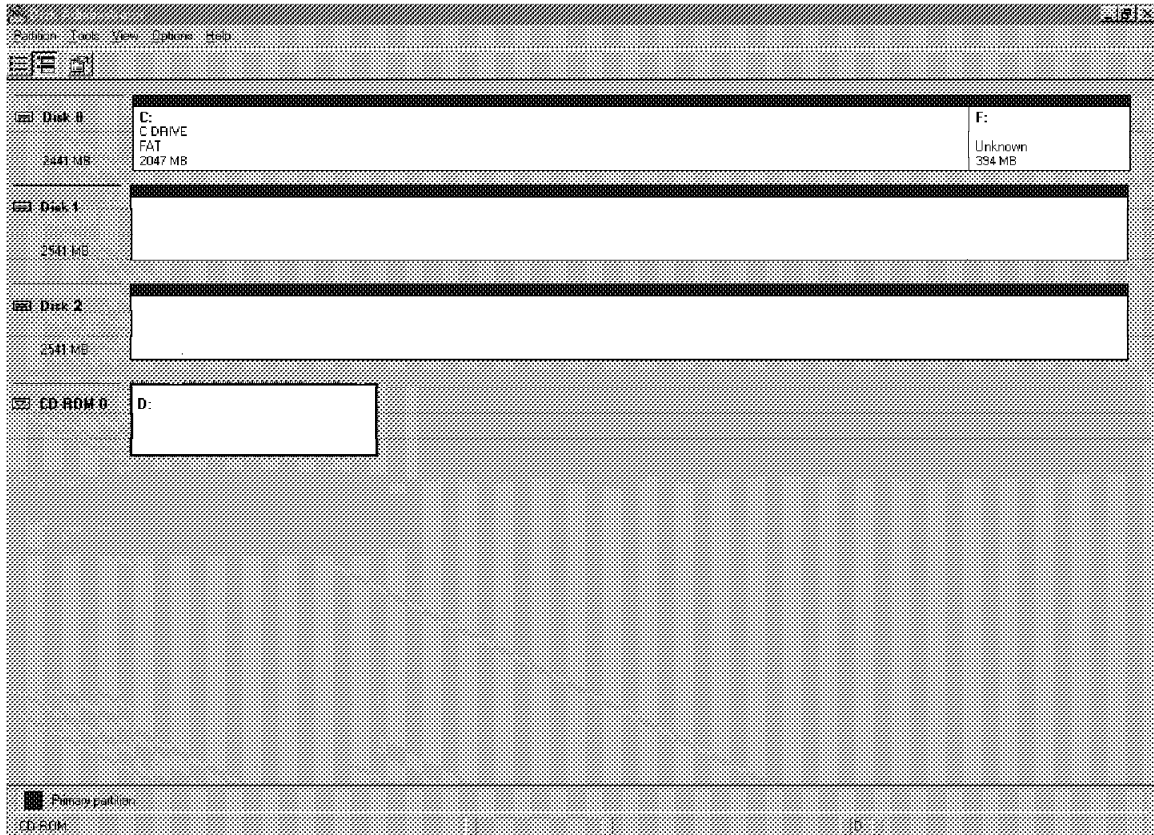


Figure 34. Windows NT Disk Administrator with Added XPE Disks

Before formatting the disks have a capacity of 2541 MB.

The order of the disks in the Disk Administrator panel is LUN within SCSI ID. As an example, if there are three disks configured in the XPE with these addresses:

XPE Device Number	SCSI ID	LUN Number	MVS VOLSER
DEV000	1	0	OPEN01
DEV001	2	0	OPEN05
DEV002	1	1	OPEN09

the order in which these disks display in Windows NT is:

Disk Number	SCSI ID	LUN Number
DISK 1	1	0
DISK 2	1	1
DISK 3	2	0

The devices do not *have* to be created in the SCSI-LUN order on the XPE configuration worksheets. However, we recommend that you enter them on the XPE configuration worksheet in SCSI-LUN sequence to simplify systems administration. So, careful analysis is required to match the SCSI ID with the correct disk number in Windows NT.

8.3.1 Formatting the Disks and Assigning Drive Letters

Several options are available to you when planning the layout of the disks. These are selected from the Partition Create panel (see Figure 35).



Figure 35. Windows NT Partition Create Panel

- One NT logical disk (disk letter) assigned to each XPE disk
This simple option allows a one-to-one relationship between the Windows NT volume and the OS/390 volume.
- Multiple logical volumes on each XPE disk
The backup and recovery procedures from OS/390 work on a single OS/390 volume, so all of the Windows NT logical volumes on the one OS/390 volume will be backed up or restored at the same time. Depending on the application, this may be an acceptable solution.
- Volume group across several XPE disks
If an NT application needs a large logical volume, you can group the physical disks together to make very large logical volumes. When backing up a large logical volume spanning many disks, you must back up all disks at the same point in time. If you are using SnapShot, reserve sufficient spare RVA volumes to snap all disks in the volume set.
- Striping data across multiple disks with or without parity
Striping improves performance of sequential reads and writes by reading or writing the stripes in parallel. The RVA also has a large disk cache and detects sequential operations, staging the data from disk into its cache.
Striping with parity (a form of RAID 5) is largely redundant when the data resides on a RAMAC Array Storage product—all of which support RAID 5 operations. The RVA has additional advantages over a simple Windows NT parity striping setup: The RVA has spare disks, so that in the event of a disk failure, the failing disk data is rebuilt automatically onto a spare disk.
- Mirroring

Mirroring (or RAID 1) provides fault tolerance to a disk failure. This again is redundant because of the RVA RAID 6 design, which can handle two simultaneous disk failures and hot spare.

Because the disks are preformatted by MVS when they are initialized, it is not necessary to do a full Windows NT disk format. You can use the quick format option (see Figure 36).

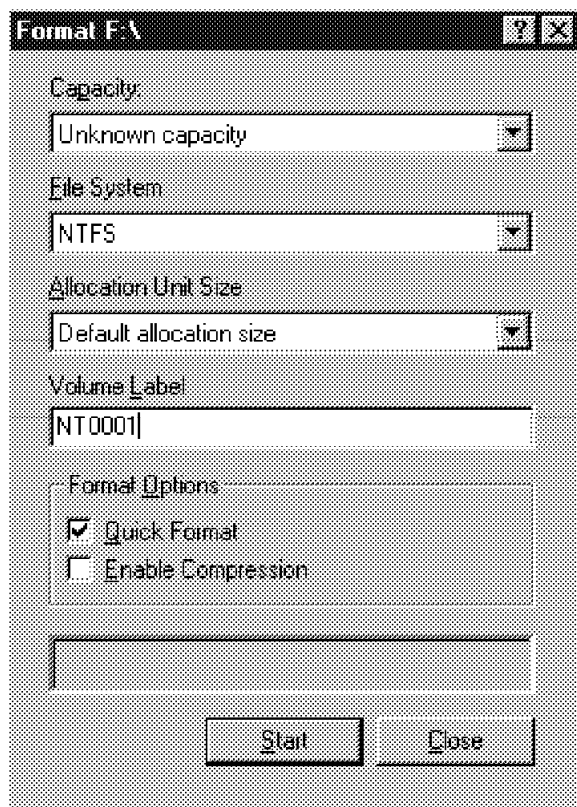


Figure 36. Windows NT Disk Format Panel

To assign drive letters use the **Tools** drop down panel. Once you have assigned drive letters, remember to complete the XPE configuration worksheet, which you can use to answer this question: My J: disk has been corrupted, can you restore it? The configuration worksheet will tell you which MVS volume contains the J: disk.

A word of warning: As drive letters can easily be changed, make sure that you update your XPE configuration worksheet after any changes.

8.3.2 SCSI and LUN Considerations

SCSI supports 16 IDs, 0 to 15. The XPE supports more than 16 disks per SCSI bus. However, the SCSI adapter in the Windows NT server will take a SCSI address, leaving 15 SCSI addresses. To support 16 disks, the LUN portion must be set for at least one SCSI ID. The LUN can be a value ranging from 0 to 7.

Once the required disks and drive numbers are created, they can be treated as per other Windows NT disks.

In Chapter 13, "Windows NT Operational Considerations" on page 83, we discuss the facilities in the S/390 environment for managing Windows NT data.

8.4 Completing the Configuration Worksheet

Before moving on, the configuration worksheet in Table 9 on page 100 needs to be completed with:

UNIX/NT Server Name	The hostname of the Windows NT server
Logical Name(s)	A unique name for NT administration purposes, typically the disk letter, for example, J:
Physical Disk Name	The disk number, for example, Disk 2

For Windows NT there can be multiple disks within any drive letter and multiple drive letters within one disk.

With these fields completed, you can correlate any Windows NT drive with its corresponding MVS VOLSER and device number.

8.5 Windows NT Recommendations with XPE

The Windows NT recommendations are similar to those for an RS/6000 environment, where RS/6000 volume groups could be replaced with a Windows NT drive letter.

The most straightforward action is to keep a simple relationship with one physical disk configured to be one logical volume (drive letter).

Multiple disks, configured as one logical volume (drive letter) for either capacity or disk striping, complicate disk recovery as *all* the XPE disks must be recovered to recover from any failure such as lost files or logical disk error.

It is good discipline to keep the configuration worksheet up-to-date, either in hardcopy or on the XPE Web Interface. If nothing else, it will act as the interpreter of the language of an NT systems administrator and that of an MVS storage administrator.

Chapter 9. Sun Microsystems and Hewlett Packard Implementation Process

Although we did not have access to Sun Microsystems or Hewlett Packard servers during the project, we believe they should work in a manner similar to that of the RS/6000.

The XPE-connected disks will be "seen" on a specific SCSI bus with XPE-configured SCSI IDs. For example, on HP-UX systems, the System Administration Manager (SAM) task-oriented interface (similar to AIX's SMIT) provides access to the HP LVM software that configures the disks for productive use. On Sun Microsystems servers this configuration work could be achieved with Solstice DiskSuite and Solaris. The applications need not be changed to use the "new" disks and therefore run unmodified. Use the same management practices as for an AIX environment (see Chapter 12, "AIX Operational Considerations" on page 75).

The RW_timeout for the SCSI connection with the XPE must be set to 600 sec. This setting allows for a tolerant connection for the rare cases where the RAMAC Array Storage may be undertaking some error recovery that will take precedence over any outstanding I/O operation.

Chapter 10. The XPE Web Interface

Note

The Web pages were under development while we were writing this redbook. The panels shown here are only examples of the kind of information that will be provided. Actual layout and content may change in the final version.

The Web interface to the XPE provides you with information about the status of the XPE, configuration details, and performance data.

The XPE Web interface is defined at XPE installation time by an IBM Service Representative with customer-supplied data (see Table 8 on page 27).

10.1 IBM 9399 Cross Platform Extension - Current Configuration

Figure 37 shows the first page you see when you access the XPE.

Color is used to indicate the status of the XPE. Green indicates that all is OK, and red indicates that there is a problem.

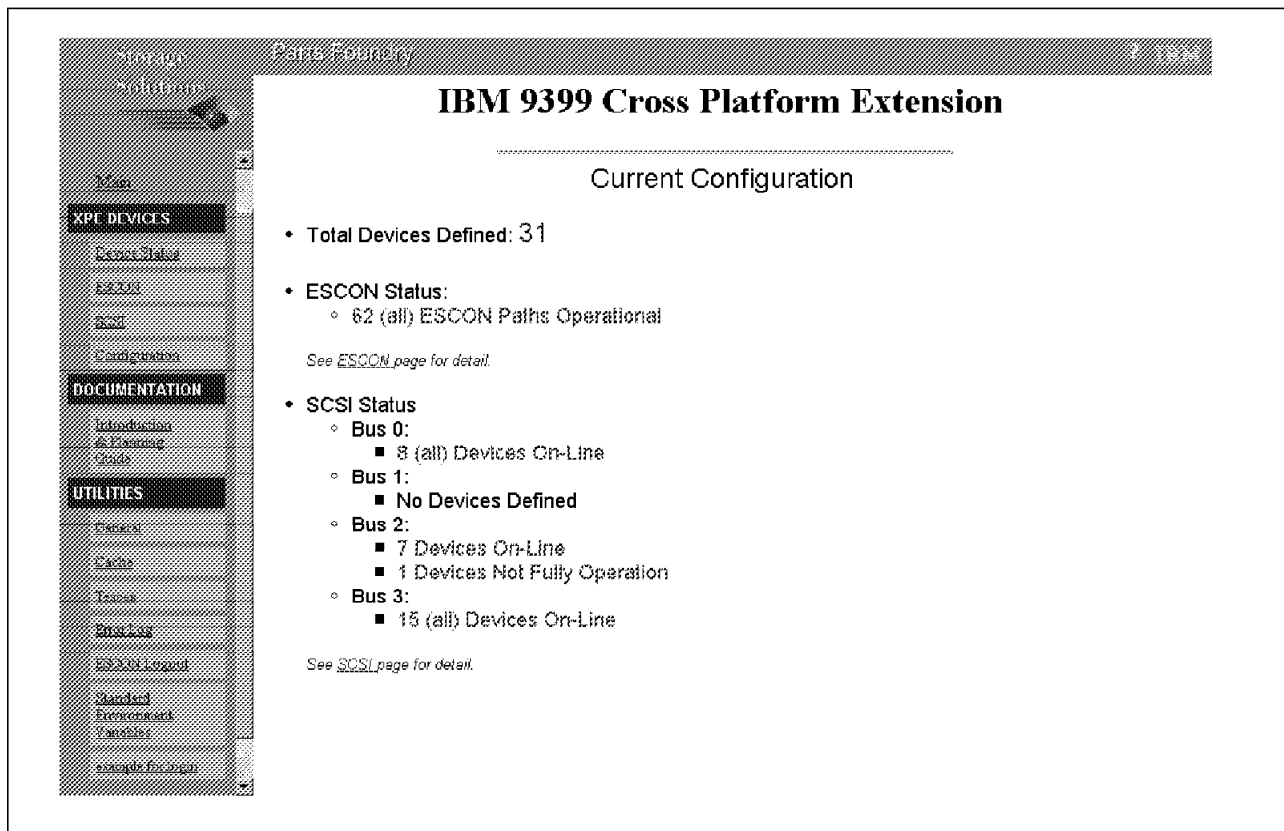


Figure 37. XPE Main Web Page: Current Configuration

10.1.1 Total Devices Defined

The total number of devices defined to this XPE is 31. A maximum of 64 devices can be defined.

10.1.2 ESCON Status

ESCON Status shows the status of the paths from the XPE to the RAMAC Array Storage. The Total Devices Defined is 31, and the number of paths is 62. The paths are *logical paths* (the same as logical paths from an S/390 host to RAMAC Array Storage). Each device can be accessed through both ESCON cards, giving a total of 62 paths. If only one path were operational, you would see (in this example) only 31 paths.

Clicking on **See ESCON Page for detail** takes you to Figure 41 on page 64.

10.1.3 SCSI Status

SCSI Status shows the status of the four available SCSI buses:

Bus 0 has all devices connected and online. The color is green.

Bus 1 is not being used.

Bus 2 has eight devices connected, but one of them has an error. This is shown in red. The problem could be that the match failed when the XPE tried to match its VOLSER with the VOLSER on the RAMAC Array Storage.

Bus 3 has 15 devices defined and all are available.

For more detail, click on **See SCSI page for detail** (see Figure 42 on page 65).

10.2 Options Bar

Figure 38 on page 61 shows the current options bar on the XPE Main Web Page.

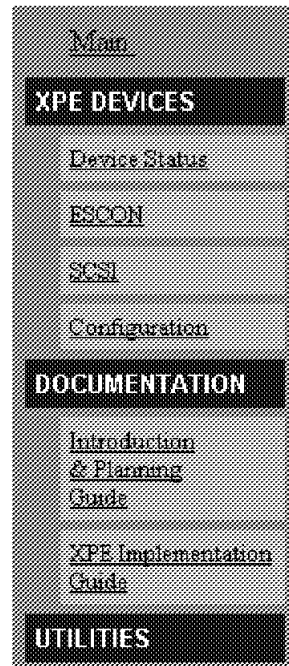


Figure 38. Web Options Bar

Clicking on **Main** always takes you back to the XPE Main Web Page (Figure 37 on page 59).

10.2.1 XPE DEVICES

Device Status takes you to a list of all installed devices in the order in which they were entered on the configuration worksheet (see 10.3, “Device Status” on page 62).

ESCON takes you to an ESCON status panel.

SCSI takes you to a SCSI status panel.

Configuration takes you to a set of panels where you can view and change the configuration.

10.2.2 Documentation

Introduction and Planning Guide is an online version of the *Installation, Planning, and User’s Guide*, GA32-0383.

XPE Implementation Guide is an online version of this book, *Data Sharing: Cross Platform Extension (XPE) Implementation Guide*, SG24-5256.

10.2.3 UTILITIES

The utilities section was still under development at the time of writing this book. The utilities are primarily for an IBM Service Representative to use for diagnostics and traces.

10.3 Device Status

The Device Status panel (Figure 39) lists the installed devices in sequential order from DEV000 to the last installed. The devices are listed in the same order in which they were entered on the configuration worksheet.

Bs:ID:Ln is the SCSI bus, SCSI ID, and LUN to which this device will respond.

The *VOLSER* is the OS/390 *VOLSER* for the RAMAC Array Storage volume.

The last six columns are counters showing I/O activity on each of the devices and the last command used.

IBM 9399 Cross Platform Extension

Device Status

DEVICE	Bs:ID:Ln	VOLSER	Blocks Written	Blocks Read	Total Cmds	Last Cmd	Read Bytes	Write Bytes
DEV00	2:0:0	JACK00	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV01	2:1:0	JACK01	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV02	2:2:0	JACK02	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV03	2:3:0	JACK03	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV04	2:4:0	JACK04	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV05	2:5:0	JACK05	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV06	2:6:0	JACK06	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV07	2:8:0	JACK07	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV08	0:9:0	JACK08	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV09	0:10:0	JACK09	0.00 B	27.65 KB	0	0	0.00 B	0.00 B
DEV0A	0:11:0	JACK0A	0.00 B	27.65 KB	0	0	0.00 B	0.00 B

Figure 39. XPE Device Status Page

Clicking on **DEV01** gives more detail on the device (see Figure 40 on page 63).

Device01 has been defined as following:

<u>BUS</u>	<u>ID</u>	<u>LUN</u>	<u>HOST</u>	<u>SECTOR</u>	<u>UA</u>	<u>VOL</u>	<u>BLCKSRD</u>	<u>BLCKSWR</u>
0	0	1	0	512	21	JACK21	8943	8807

<u>Card</u>	<u>Port</u>	<u>Address</u>	<u>Defined</u>	<u>Failed</u>	<u>Fenced</u>	<u>Verified</u>	<u>RC</u>	<u>Volser</u>	<u>Blocks Read</u>	<u>Blocks Write</u>
0	0	64	1	0	0	1	0	JACK21	4621	4254
0	1	6c	1	0	0	0	0	*****	0	0
0	2	0	0	0	0	0	0	*****	0	0
0	3	0	0	0	0	0	0	*****	0	0
1	0	64	1	0	0	1	0	JACK21	4322	4553
1	1	6c	1	0	0	0	0	*****	0	0
1	2	0	0	0	0	0	0	*****	0	0
1	3	0	0	0	0	0	0	*****	0	0

Figure 40. Device Detail Page

10.4 ESCON

The ESCON Status panel (Figure 41 on page 64) shows the ESCON status for each path on each ESCON card.

IBM 9399 Cross Platform Extension

ESCON STATUS

	CARD 0	CARD 1
	64	6C
Paths Verified	31	31
Paths Pending	**	**
Paths Fenced	**	**
Paths Failed	**	**
Total Paths	31	31

Figure 41. XPE ESCON Status Page

In Figure 41, 64 is the ESCON Director port number for Card 0, and 6C is the port number for Card 1. If four ports were defined (the maximum) on each ESCON card, you would see eight columns, four for each card. The XPE verifies that it can access each of the RAMAC Array Storage devices down each path. If all paths verify successfully, Total Paths should be the same as Paths Verified.

The other rows indicate paths that are not operational. This status would be indicated on the *XPE Main Web Page*.

10.5 SCSI

The SCSI Status panel (Figure 42 on page 65) shows the status of the four SCSI buses, which is in effect the status as presented by the XPE to the four SCSI servers.

IBM 9399 Cross Platform Extension

SCSI STATUS

	Bus 0	Bus 1	Bus 2	Bus 3
On Line	8	**	7	15
Long Busy	**	**	**	**
Pending	**	**	**	**
Intervention Req	**	**	**	**
Off Line	**	**	1	**
Failed	**	**	**	**
Total	8	0	8	15

Figure 42. XPE SCSI Status Page

The On Line row indicates the number of devices available on each bus. In this example Bus 1 has no devices attached.

Bus 2 has 8 devices attached but one of them has an error and is Offline. The cause of this error is not clear. It may be that the validation of the VOLSER failed.

10.6 Configuration

The configuration pages are still under development and may be added to a revision of this book.

Chapter 11. OS/390 Operational Considerations

In this chapter we discuss the options to manage UNIX or Windows NT data residing on S/390 storage. Many options are available, including SnapShot with the RVA, remote copy with the 3990 and 9390 (and RVA 4Q98), and concurrent copy on the 3990 and 9390. Software options include DMSMSdss or any compatible OS/390 dump and restore package. We discuss the benefits of each of these options and explain how to implement them.

11.1 Backing Up UNIX or Windows NT Data: General Considerations

From an OS/390 point of view, an XPE disk is no different from any other OS/390 disk. It has a volume label, VTOC, and a VTOC index and can be SMS managed. The data sets on the volumes occupy the whole volume, so there is no free space (see Figure 15 on page 31 for a list of the data sets on an XPE volume).

The volumes are accessible from OS/390 (they can be online or offline), so you can use OS/390 utilities to dump or restore or move the data. Because the volumes are always available to the OS/390 systems, you need to exercise caution when accessing or changing the data. The volumes may still be in use on the UNIX or Windows NT system and not be visible to OS/390.

The recommendations in Chapter 6, "MVS Implementation Process" on page 29 about naming volumes, storage groups, and data sets should be followed as they will help operations and OS/390 storage management identify disks being used by the XPE. Even from RMF, there is typically no I/O to these disks, so to an observer they may appear unused.

Such questions as "Which OS/390 volumes contains my AIX volume group?" or "Which OS/390 volume holds my NT H: disk?" will arise when the UNIX or Windows NT administrator wants OS/390 to back up or restore a UNIX or Windows NT disk, volume group, or drive letter. Keeping the XPE installation worksheets up to date will make it easy to answer those questions. The UNIX system administrator knows what the volume group name is, and the MVS storage administrator knows what the VOLSER and device number are, and the configuration worksheet allows the connection to be made.

No issues apply to sharing the XPE volumes between OS/390 systems. Dumping a volume can be done from any host that has access to the volumes. Restoring a disk will primarily affect the UNIX or Windows NT system, not the OS/390 systems, as generally there is no activity to those volumes from OS/390. Similarly, no issues apply to a Parallel Sysplex except perhaps the use of *SYS1*. as the high-level qualifier for the XPE data sets. However, the XPE data set names can be changed to conform to your installation standards.

11.1.1 Synchronization between OS/390 and UNIX or Windows NT

In many cases an OS/390 system and a UNIX or Windows NT system do not communicate, so you may need a manual process to coordinate the running of the OS/390 volume dumps with application availability on the servers. Such products as Tivoli Global Enterprise Manager allow automation across heterogeneous server environments and may enable you to perform server backups and recovery automatically, but we did not investigate them during the project.

Backing up the server data involves closing down the application (to ensure that all buffers are flushed to the disks) and then running a job on OS/390 to perform the backup. Once the backup is complete, the server system can restart the application. (Although it is possible to take an OS/390 backup of the UNIX or Windows NT disks while in use, this backup option may not produce a recoverable environment. However, if your application can recover from a “fuzzy copy,” there is nothing to prohibit its use).

11.2 Using DFSMSdss to Manage UNIX or Windows NT Data

A simple way of backing up an XPE volume or volumes is to use DFSMSdss to perform a full volume dump to tape, exactly the same process as dumping any other MVS volume.

A full dump to tape is appropriate if you are not running a 24x7 operation, you have plenty of time to take the backups, or you do not have concurrent copy or SnapShot.

A quicker option is to use either concurrent copy or virtual concurrent copy to take the backup to tape. This option also minimizes outage time.

11.2.1 Concurrent Copy

Concurrent copy is a function of DFSMSdftp, DFSMSdss, and 3990-6 or 9390. It allows you to take a dump of a volume in parallel with accessing and updating the data. Concurrent copy does this by intercepting changes in the 3990 or 9390 and keeping the original data in cache while writing the new data to disk. As far as the UNIX or Windows NT server is concerned, the dump is complete as soon as the DFSMSdss has logically completed the copy, which typically takes less than a minute. The advantage is that the outage time for the UNIX or Windows NT server is dramatically reduced and dumping the data occurs in parallel with access. See *Implementing Concurrent Copy*, GG24-3990, for more information.

11.2.2 Virtual Concurrent Copy

Virtual concurrent copy is similar to concurrent but uses the RVA storage. Virtual concurrent copy uses the SnapShot capability of the RVA to take a copy of the volumes and DFSMSdss copies the resulting copy to tape. The outage time while virtual concurrent copy makes the copy is typically less than a minute. See *Implementing DFSMSdss SnapShot and Virtual Concurrent Copy*, SG24-5268, for more information.

11.2.3 Disk to Disk Copy

You can use DFSMSdss COPY to copy the XPE disk to another OS/390 disk. If the volume resides on an RVA, DFSMSdss automatically invokes SnapShot as shown in Figure 43 on page 69.

```

                J E S 2  J O B  L O G  --  S Y S T E M  3 0 9 0  --  N O D E  T P C 9 7 2

10.40.09 JOB00039 IRR010I  USERID  IBMUSER  IS  ASSIGNED  TO  THIS  JOB.
10.40.10 JOB00039 ICH70001I  IBMUSER  LAST  ACCESS  AT  10:38:55  ON  THURSDAY,  APRIL  30,  1998
10.40.10 JOB00039 $HASP373  IBMUSER4  STARTED  -  INIT    1  -  CLASS  A  -  SYS  3090
10.40.10 JOB00039 IEF403I  IBMUSER4  -  STARTED  -  TIME=10.40.10
10.40.13 JOB00039 *11  ADR369D  AUTHORIZE  FOR  WRITE  ACCESS  A  VTOCIX  DATA  SET  ON  AMBJ51,....
10.40.24 JOB00039 R 11,U
10.40.30 JOB00039 IEC604I  VTOC  CONVERT  ROUTINE  ENTERED  ON  OF51,AMBJ51,DOS
10.40.32 JOB00039 ICK502I  BUILDIX  FUNCTION  STARTED
10.40.32 JOB00039 ICK503I  OF51  REQUEST  RECEIVED  TO  CONVERT  VTOC  TO  IXFORMAT
10.40.32 JOB00039 ICK504I  OF51  VTOC  FORMAT  IS  CURRENTLY  OSFORMAT,  REQUEST  ACCEPTED
10.40.32 JOB00039 ICK513I  OF51  BUILDIX  PROCESSING  COMPLETED:  VTOC  IS  NOW  IN  IXFORMAT
10.40.33 JOB00039          IBMUSER4  STEP001  ADRDSSU    0000
10.40.33 JOB00039 IEF404I  IBMUSER4  -  ENDED  -  TIME=10.40.33
10.40.33 JOB00039 $HASP395  IBMUSER4  ENDED

      ..
      1 //IBMUSER4  JOB  CLASS=A,MSGLEVEL=(1,1),MSGCLASS=H,REGION=4096K          JOB00039
      2 //STEP001  EXEC  PGM=ADRDSSU                                           00040000
      3 //SYSPRINT DD  SYSOUT=*                                               00050000
      4 //SYSIN   DD  *                                                       00060000

      ..
IEF373I STEP/STEP001 /START 1998120.1040
IEF374I STEP/STEP001 /STOP  1998120.1040 CPU    OMIN 00.81SEC SRB    OMIN 00.14SEC ...
IEF375I JOB/IBMUSER4/START 1998120.1040
IEF376I JOB/IBMUSER4/STOP  1998120.1040 CPU    OMIN 00.81SEC SRB    OMIN 00.14SEC

PAGE 0001    5695-DF175  DFSMSDSS V1R3.0  DATA SET SERVICES    1998.120 10:40
COPY FULL -                00070000
  INDYNAM ( -                00080000
    (AMBJ50) -                00090005
  ) -                00100000
  OUTDYNAM ( -                00110000
    (AMBJ51) -                00120005
  ) -                00130000
  PURGE                00140005
ADR101I (R/I)-RI01 (01), TASKID 001 HAS BEEN ASSIGNED TO COMMAND 'COPY '
ADR109I (R/I)-RI01 (01), 1998.120 10:40:12 INITIAL SCAN OF USER CONTROL STATEMENTS COMPLETED.
ADR016I (001)-PRIME(01), RACF LOGGING OPTION IN EFFECT FOR THIS TASK
ADRO06I (001)-SETUP(01), 1998.120 10:40:12 EXECUTION BEGINS
ADR241I (001)-DDTFP(01), TARGET VTOC BEGINNING AT 000000:0001 AND ENDING AT 000000:0001 IS OVERLAID
ADR806I (001)-TOMI (02), VOLUME COPIED USING SNAPSHOT FUNCTION.
ADR013I (001)-CLTSK(01), 1998.120 10:40:33 TASK COMPLETED WITH RETURN CODE 0000
ADRO12I (SCH)-DSSU (01), 1998.120 10:40:33 DFSMSDSS PROCESSING COMPLETE. HIGHEST RETURN CODE IS 0000

```

Figure 43. SNAP Using DFSMSdss COPY

If taking a copy of an XPE volume, consider having the target volume of the copy outside the range of volumes owned by the XPE. This approach will ensure that copies of data are not mixed with the originals, and that UNIX and Windows NT do not find the additional copies if they reboot.

When using DFSMSdss to copy (with or without the Snapshot capability), use the COPYVOLID(NO) option. This option leaves the target volume with its original VOLSER, and the volume can remain online to OS/390. You can dump the volume to tape or leave it online to enable a fast restore with a Snap-back.

The data set names on the target volume will be the same as the source, but this will only be a concern if you are running an entirely SMS-managed environment, and even then, as no OS/390 job needs to access the data sets except to dump the data, the data set names need not be changed.

You could consider naming the VOLSERs of the target volumes with a name that reflects their usage, for example, SNAP001 or OSNAP01, or include the server name plus a backup indicator such as NT001B.

11.3 Using SnapShot to Back Up UNIX or Windows NT Data

As an alternative to DFSMSdss you can use SnapShot with the SIBBATCH program (see Figure 44). In this case you perform a volume snap but do not copy the VOLSER. Thus the volume is available to the OS/390 system to take a dump to tape.

```
//IBMUSER2 JOB  MSGCLASS=H,CLASS=A,MSGLEVEL=(1,1),REGION=4M
//STEP1  EXEC  PGM=SIBBATCH
//SYSTEM DD  SYSOUT=*
//SYSPRINT DD  SYSOUT=*
//SYSIN   DD   *

    SNAP VOLUME ( SOURCE(UNIT(F52)) -
                  TARGET(UNIT(F75)) -
                  DATAMOVERNAME(NONE) COPYVOLID(NO) REPL(YES))

/*
```

Figure 44. JCL Using SIBBATCH to Snap

Figure 45 shows the result of a job to snap volume F50 to F52. Note that after the snap both volumes remain online. F52 retains its own VOLSER (using the COPYVOLID(NO) parameter) and can now be backed up as normal.

```
      J E S 2  J O B  L O G  --  S Y S T E M  3 0 9 0  --           N O D E  T P C 9 7 2

10.35.05 JOB00036 IRR010I  USERID  IBMUSER  IS ASSIGNED TO THIS JOB.
10.35.05 JOB00036 ICH70001I  IBMUSER  LAST ACCESS AT 10:09:30 ON THURSDAY, APRIL 30, 1998
10.35.06 JOB00036 $HASP373  IBMUSER2  STARTED - INIT    1 - CLASS A - SYS 3090
10.35.06 JOB00036 IEF403I  IBMUSER2  - STARTED - TIME=10.35.06
10.35.15 JOB00036 VARY F52,ONLINE XSA/SNAPSHOT INTERNAL VARY
10.35.15 JOB00036      IBMUSER2  STEP1      SIBBATCH    0000
10.35.15 JOB00036 IEF404I  IBMUSER2  - ENDED - TIME=10.35.15
10.35.15 JOB00036 $HASP395  IBMUSER2  ENDED

      ..
      1 //IBMUSER2 JOB  MSGCLASS=H,CLASS=A,MSGLEVEL=(1,1),REGION=4M                JOB00036
      2 //STEP1  EXEC  PGM=SIBBATCH                                           00040000
      3 //SYSTEM DD  SYSOUT=*                                                 00050000
      4 //SYSPRINT DD  SYSOUT=*                                              00060000
      5 //SYSIN   DD   *                                                       00070000

      ....
IEF373I STEP/STEP1  /START 1998120.1035
IEF374I STEP/STEP1  /STOP  1998120.1035 CPU    OMIN 01.09SEC SRB    OMIN 00.12SEC ..
IEF375I JOB/IBMUSER2/START 1998120.1035
IEF376I JOB/IBMUSER2/STOP  1998120.1035 CPU    OMIN 01.09SEC SRB    OMIN 00.12SEC

SIB0715I IXFP 2.1.1 (SIBBATCH) invoked on Thu Apr 30, 1998 10:35:09.
SIB0700I 10:35:09
SIB0700I 10:35:09  SNAP VOLUME ( SOURCE(UNIT(F50)) -
SIB0700I 10:35:09      TARGET(UNIT(F52)) -
SIB0700I 10:35:09      DMNM(NONE) COPYVOLID(NO) REPL(YES))
SIB4751I No VVDS data set found on volume AMBJ52 (0F52).
SIB4630I A VARY ONLINE request has been issued for device F52.
SIB4617I 10:35:15 SnapShot completed, rc=0.
```

Figure 45. Output from SNAP Job

11.4 Recovery of Backed Up Disks

Having taken backups of your UNIX or Windows NT data on your OS/390 system and stored the backups in your tape library, you will get a request to restore the backed-up data from your OS/390 system.

11.4.1 Snap-back

If you used SnapShot or DFSMSdss COPY to copy the UNIX or Windows NT data disk to another OS/390 volume and the copy is still available online, you can use SnapShot in reverse to Snap-back the copy over the top of the original volume.

Before Snapping-back the volume, you have to shut down or deactivate the volumes on the UNIX or Windows NT system. Follow the procedures outlined in Chapter 12, "AIX Operational Considerations" on page 75 or Chapter 13, "Windows NT Operational Considerations" on page 83.

Once the UNIX or Windows NT system is ready, the Snap-back is simple and is the same as the original Snap, but in reverse. (In Figure 44 on page 70, SOURCE would now be F75, and TARGET would be F52.)

11.4.2 DFSMSdss Restore

If the UNIX or Windows NT data had been backed up to tape, either through a SnapShot or simple DFSMSdss dump, the data can be restored through a normal DFSMSdss restore job pointing to the UNIX or Windows NT volume as the target for the restore.

11.4.3 Recovering an Individual UNIX or Windows NT File

You cannot recover an individual UNIX or Windows NT file directly from an DFSMSdss dump tape or even from a Snapped copy of a volume because OS/390 treats the whole of the UNIX or Windows NT disk as a single data set and does not know about the individual files on the disk.

What you can do is restore the volume to a spare disk on the UNIX or Windows NT server and recover the file from the volume. (This restore is similar to what used to be done on MVS before you could restore a data set from a full volume dump.)

There are some limitations to consider before recovering a file on Windows NT and UNIX. These limitations are documented in 13.6, "Recovering a Lost or Deleted File without ADSM" on page 86 for Windows NT and in 12.6, "Recovering a Lost or Deleted File without ADSM" on page 78 for UNIX.

11.5 SMS Considerations

If you run with all your data managed by SMS, you will want to add the XPE volumes too. Recommendations on how you manage the XPE volumes with DFSMS/MVS are covered in 6.4, "Converting to SMS-Managed Volumes" on page 34.

Operationally, you may want to keep the XPE data isolated from your traditional MVS data to prevent an MVS job or user from accessing or modifying the UNIX or Windows NT data.

You also need to stop allocation (allocation should not be possible as the XPE volumes have no free space), deletion, or management of the data by DFSMSHsm or other storage management utilities.

The only valid management function is backup and restore of the volume.

11.5.1 Year 2000 Testing

The RVA with SnapShot is being used extensively for Year 2000 testing on OS/390. The same benefits can be achieved in a UNIX or Windows NT environment for Year 2000 testing using the following procedure:

1. Create your Year 2000 test system and put all your test files on XPE managed volumes.
2. Take a SnapShot of all your test data volumes to a set of OS/390 owned volumes. *This SnapShot is your master copy of test data .*
3. Run your Year 2000 programs against the test data.
4. Check your results.
5. Fix problems.
6. Restore the test data from the master copy, using Snap-back.
7. Retest.

For further details see 12.7, "Year 2000 Testing" on page 80 for AIX and 13.7, "Year 2000 Testing" on page 87 for Windows NT systems.

11.5.2 Remote Copy

If the XPE volumes are part of either a 3990/9390 or RVA storage array, you can use PPRC or XRC (XRC is 3990/9390 only) to maintain a disaster backup of your UNIX or Windows NT data.

PPRC is a synchronous copy solution that keeps the two copies of data, at two locations, identical. The RAMAC Array Storage is connected by ESCON links at up to 26 km. One copy of the data is on the *primary*, the other on the *secondary* storage. Should the primary site fail, you can switch to the backup site and restart the applications, using the secondary copy of data. This applies to both OS/390 data and any UNIX or Windows NT data on the RAMAC Array Storage.

XRC is an asynchronous remote copy solution that uses 3990/9390 as the primary storage and any RAMAC Array Storage as the secondary. Because XRC is an asynchronous solution, the distance between the primary storage and the secondary storage is virtually unlimited. The OS/390 Datamover function is used to manage the copying of the data from primary to secondary storage and at the same time ensure data integrity.

For more information about remote copy, see *DFSMS/MVS Version 1 Remote Copy Administrator's Guide and Reference*, SC35-0169.

11.6 Operations

From an operator view, the XPE disks look like any other OS/390 3390-3 disks. All commands operate as normal (see the example of the DEVSERV command on Figure 9 on page 24). The volumes can be online or offline to OS/390. It is usually recommended that the volumes be online so that SIMs can be offloaded to an S/390 host.

From an RVA or 3990/9390 perspective, all of the volumes should be cached with DASD Fast Write active. The 3990/9390 uses QuickWrite (the 3990/9390 function that gives write cache hits for regular format data) as all the blocks are fixed length.

The ESCON links from the XPE to the S/390 storage can be either direct or through an ESCON Director. When through an ESCON Director, they can be switched or dedicated. The ESCON links are not seen by any S/390 host, so care should be taken when reconfiguring ESCON Directors. ESCON link errors are reported by the RAMAC Array Storage or by the ESCON Director if used.

11.6.1 Service Information Messages

SIMs are produced by the RAMAC Array Storage controllers after an error. They indicate the type of error that has occurred, the severity of the error condition, and the recovery action needed or taken by the storage control. For RAMAC Array Storage, a SIM is sent to all attached processors. The SIM is then displayed on the operator console and logged in the OS/390 error log (LOGREC). Because the XPE cannot handle a SIM (it has no console to report the error), the LIC in the 3990-6, 9390, RVA, and RSA has been changed to handle SIMs differently.

As neither UNIX nor Windows NT understands or handles SIMs, and the XPE has no method of forwarding a SIM, a method has been introduced to ensure that SIMs go only to hosts that can handle them.

Although very unlikely because of the RAID design of the RAMAC Array Storage, should a RAMAC Array Storage device error occur, the UNIX or Windows NT server gets an I/O error condition, and the application and/or server may fail. Any recovery from a device error should be performed through ICKDSF from MVS, not with any UNIX or Windows NT functions.

11.6.2 IXFP Reports

IBM Extended Facilities Product (IXFP) provides reports on device activity by attached host that enable you to see the I/O rate and response time that your RVA delivers to the XPE.

The *IBM Extended Facilities Product: Subsystem Reporting Manual*, SC28-7184, shows how to produce reports on RVA performance and capacity. Some IXFP reports useful for looking at XPE device performance are the Device Performance Reports, the Channel Interface Performance Report, the Cache Effectiveness Report, and the Space Utilization Summary Report.

The Device Performance Reports (interval and overall) give details by individual device number (XPE volume). So for an XPE volume, you can see the I/O rate, KB/sec, I/O service time (total, disconnect time, connect time), and %device utilization. You can also see the distribution of I/O activity across the XPE volumes and the performance of the XPE volumes.

The Channel Interface Performance report shows the utilization of the ESCON channel interfaces to the RVA. Depending on whether you have a dedicated ESCON connection or not, you can see the overall utilization of the two XPE ESCON channels. High channel utilization may affect the performance of any system using the channels.

The Cache Effectiveness Report shows which percentage of the XPE I/Os to the RVA are handled by the RVA cache. The higher the percentage, the better response time is achieved. Cache details for each XPE volume include overall read and write activity, read to write ratio, read cache hit %, and write hit %.

The Space Utilization Summary reports show the amount of data stored on the XPE disks and the compression ratio for that data. High compression ratios (in excess of 4:1) indicate that very efficient use is being made of the RVA disk capacity.

11.6.3 RMF Reports

RMF does not show any activity for XPE volumes as there is no host activity to these volumes (apart from an MVS initiated backup or restore). Cache RMF Reporter (CRR) and RMF Cache Reports show cache statistics and I/O rates for the XPE volumes, similar to those described for IXFP.

Chapter 12. AIX Operational Considerations

In this chapter we discuss some operational considerations for an AIX environment.

In normal operation, the AIX server should be started after the XPE to ensure that the SCSI disks are available for use.

The main operational processes to consider are:

- Adding new SCSI disks
- Moving disks to other servers
- Data backups
- Creating a clone server
- Recovering data with ADSM
- Recovering a lost or deleted file without ADSM
- Year 2000 testing
- Recovering from an XPE failure

In an AIX environment, exercise caution when manipulating disks either within a server or between servers. Each SCSI disk has a unique PVID, and, with the possibilities available with SnapShot to back up and restore data, serious problems can arise when two different hdisks have the same PVID.

12.1 Adding New SCSI Disks

Follow these steps to create a new SCSI disk:

1. Create a new disk in the RAMAC Array Storage, using ICKDSF.
2. Add the new disk (with a unique SCSI ID and LUN) to the XPE configuration nondisruptively.
3. Shut down and restart the XPE to activate the new configuration.

All servers are affected. Therefore a quiet period of operation must be planned for this short interrupt.

4. Run `cfgmgr` or reboot the RS/6000 to create the new hdisk in AIX.

You do not have to reboot the other servers if you can guarantee there is no communication or data transfer with the XPE during the XPE shutdown and restart process. Otherwise the safer approach is to shut down all servers before shutting down the XPE and then restart them after the XPE is fully operational.

5. Use standard LVM commands to configure the disk into a new or existing volume group.

Follow these steps to delete an existing disk:

1. Stop all use of the disk with `exportvg`, or from a volume group with `rm\vcopy` and `reducevg`.
2. Remove the disk from the XPE configuration.
3. If the new XPE configuration has to be activated, there must be a quiet period in which to shut down and restart the XPE.
4. Delete the disk from the RAMAC Array Storage configuration.

12.2 Moving Disks to Other Servers

The steps to move a disk to other servers differ according to whether the disk is being moved:

1. Within the same XPE or to another XPE
2. To a like technology server such as RS/6000 to RS/6000

This example concentrates on moves within the same XPE and between RS/6000 servers:

1. Stop all use of the disk in AIX (with `exportvg` or `reducevg` commands).
2. Delete the disk from AIX, for example, `rmdev -l hdiskx -d`.
3. Reconfigure the XPE and move the disk definition to a new server.
4. If the disk is being moved to a different technology server, reinitialize it in MVS, using `ICKDSF`. When an RS/6000, for example, discovers a disk, it looks for a PVID, and a disk that was in use by Windows NT will not be in the correct format.
5. If the disk is being moved to the same technology server, rename the XPE OS/390 data set to something meaningful to match the new server, for manageability purposes.
6. Shut down and restart the XPE to load the new configuration table.
7. Run `cfgmgr` or reboot the RS/6000 to create the new `hdisk` in AIX, which can then be manipulated through the standard LVM commands.

The other servers do not have to be rebooted if you can guarantee that there is no communication or data transfer with the XPE during the XPE shutdown and restart process.

12.3 Data Backups

Many data backup options are available, including continuing with the operational backup procedures in use on the RS/6000 before moving to the XPE environment.

However, MVS provides a centralized, sophisticated, and automated backup and recovery environment. One major benefit of the MVS SnapShot facility is that it creates a backup in a matter of seconds, thereby greatly reducing the backup window.

Figure 46 on page 77 shows an example of using SnapShot copy.

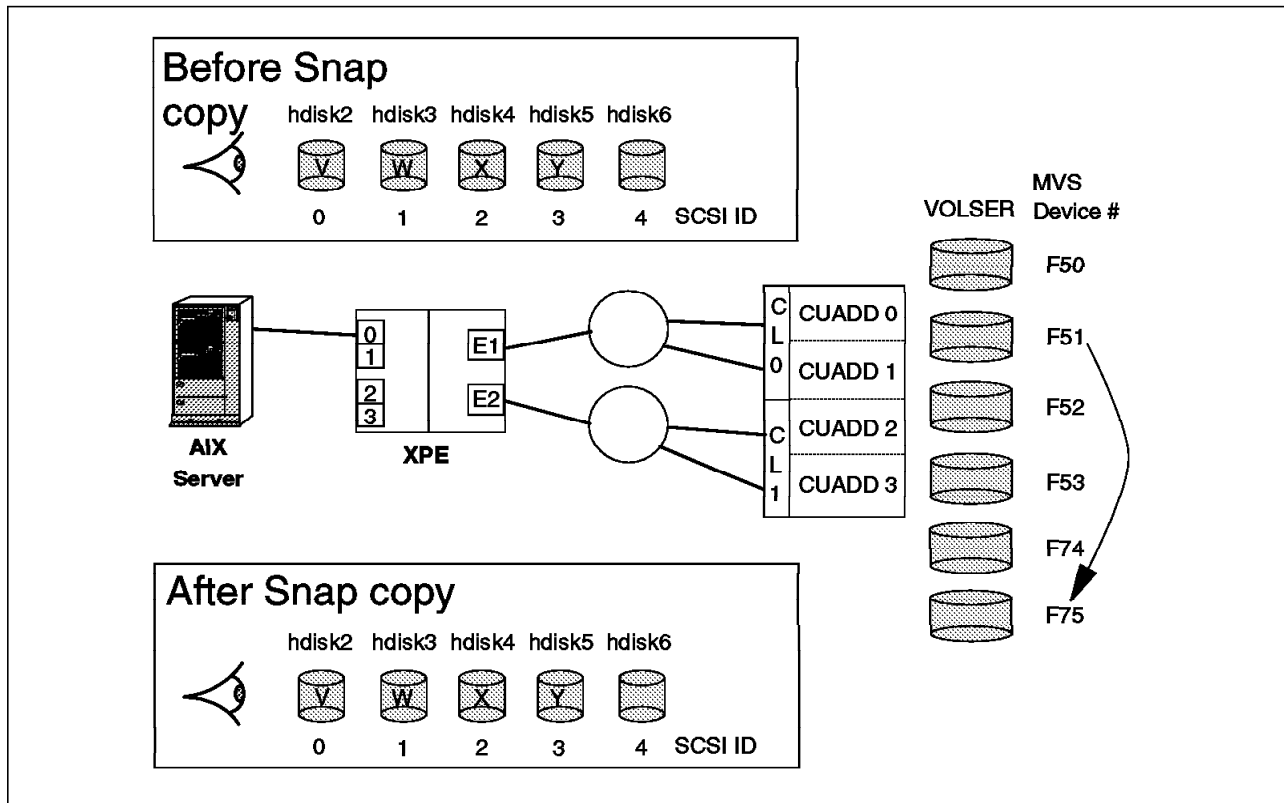


Figure 46. SnapShot Copy: AIX

Before running SnapShot, shut down the RS/6000. Then run SnapShot to copy the data to another MVS volume. In Figure 46 MVS device number F51 is an MVS volume also defined to the XPE, and the RS/6000 sees it as hdisk3, which contains data noted as W. MVS device number F75 is an MVS volume not defined to the XPE and therefore is not a known disk in AIX. There are no changes made in the RS/6000 after the SnapShot is made.

F75 can then be backed up to tape before the next SnapShot copy is made. For a detailed discussion of SnapShot, see Appendix B, "The RAMAC Virtual Array and Snapshot" on page 105.

12.4 Creating a Clone Server

If an exact copy of a server is required in another part of the customer installation, cloning, as it is known, can easily be achieved in an XPE environment.

Follow these steps to create a cloned RS/6000 system:

1. Use SnapShot from OS/390 to make copies of the XPE volumes.
2. Reconfigure the XPE connected to the new server to point to the new SnapShot disks. Note the new SCSI ID definitions.
3. Install the server. The cloned server system may be a new install or built using a mksysb from the original RS/6000.
4. Boot the new server.
5. Cross reference the new SCSI ID and LUN with valid hdisk numbers.

6. Run `importvg` for each volume group.

PVIDs need to be unique on an RS/6000

The PVIDs of the disks on the cloned server are identical to the PVIDs of the original disks. If you intend to reuse the cloned disks on the original servers, you must reinitialize the disk in MVS, using ICKDSF to delete the duplicate PVIDs.

12.5 Recovering Data with ADSM

From a procedural point of view, an ADSM environment with XPE disks and an ADSM environment with non-XPE disks are the same. ADSM operates on a file basis. Files can be recovered from incremental backup. Entire file systems can be recovered from the root down the hierarchy.

One option is that with `hdisk`-level backups through MVS, using SnapShot, ADSM file system recovery can be speeded up. Use the `-ifnewer` specification on the `dsmc restore` command. Depending on the frequency of SnapShot backup copies that you take of the physical volumes, the `-ifnewer` forward recovery can potentially reduce the overall recovery time of the file systems.

12.6 Recovering a Lost or Deleted File without ADSM

Using XPE and the RVA, you can recover lost files or bad logical volumes. The process documented below assumes that there is a 1:1 ratio of one `hdisk` per volume group and you are restoring missing files from a recent SnapShot copy. The process of restoring from a SnapShot is also known as Snap-back.

The recovery activities should be executed either overnight or during a quiet period of the day as the volume group is unavailable for productive use during the recovery process. However, the XPE and the other three servers connected to the XPE can continue to operate.

If there are multiple disks in the volume group, you must Snap-back all the relevant disks in the volume group, using SnapShot from OS/390.

Follow these steps to recover a lost or deleted file without ADSM:

1. Stop use of the production disks on the RS/6000 in a controlled manner. Either shut down the RS/6000 or unmount the file system and `varyoffvg` the volume group.
2. Make a Snapshot of the production disks.
3. Recover the last SnapShot of the production disks onto the *same* `hdisks`. For example, this recovery SnapShot could be from the previous night's SnapShot.
4. Reboot the server or run the `varyonvg` and `mount` commands to activate the data.
5. Copy the recovered files to a temporary area, for example, the `/tmp` directory.
6. Stop use of the temporary production data on the RS/6000. Either shut down the RS/6000 or unmount the file system and `varyoffvg` the volume group.

7. Recover the recent SnapShot (step 2) of the production disks.
8. Reboot the server or run the varyonvg and mount commands to activate the production data.
9. Move the files in /tmp to the production file system.

Figure 47 shows how to use Snap-back together with two SnapShot copies as part of a process to recover lost files.

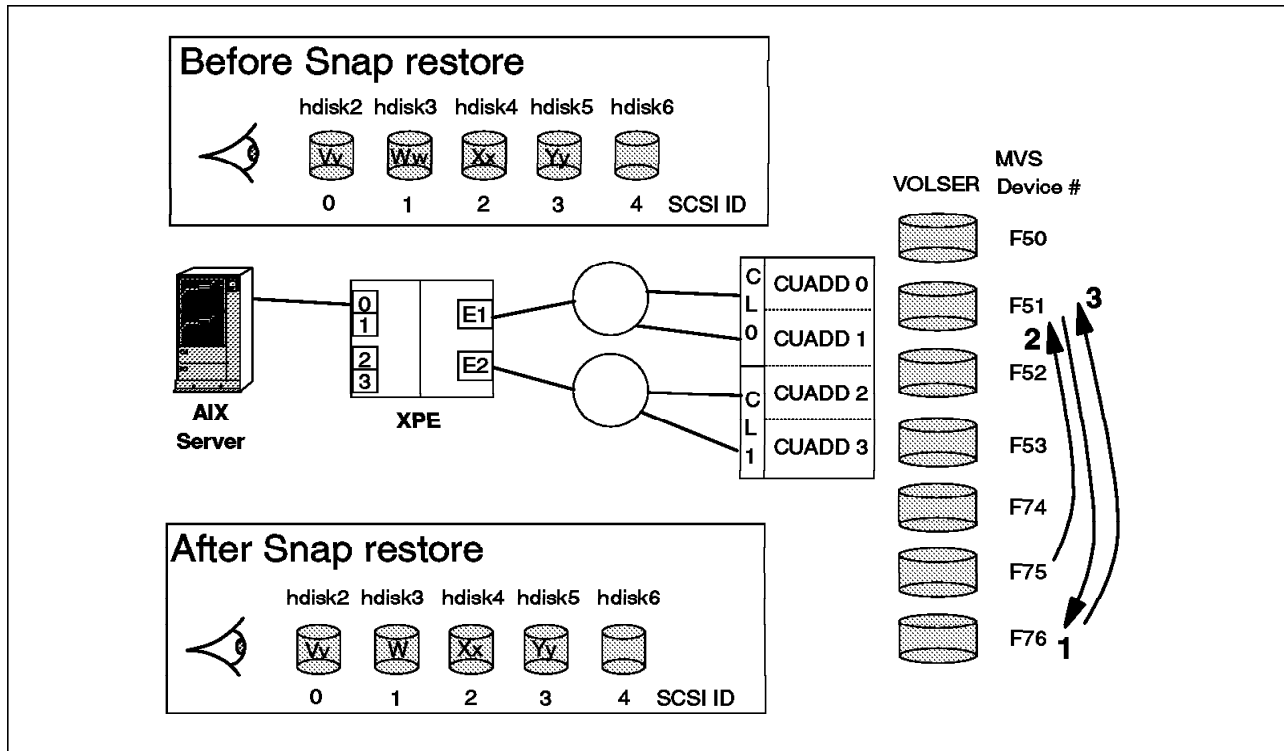


Figure 47. Using Snap-back to Recover Lost Files:AIX

MVS Volume F75 has a recent SnapShot copy of MVS Volume F51, which is the RS/6000 hdisk3.

Before you restore the lost files, you have to SnapShot the production data to MVS Volume F76 and Snap-back the recent SnapShot (on MVS Volume F75) over hdisk3. Extra care needs to be taken after the Snap-back to ensure that the restored files are saved in a temporary staging area. Then you have to Snap-back the production data from MVS Volume F76, and copy the restored files from the temporary staging area into the production data, thereby fully restoring the live production data of hdisk3.

Figure 48 on page 80 shows the command sequence to restore files that were wrongly deleted from /xpefs in xpevg on hdisk3.

```

umount /xpefs
varyoffvg xpevg
  (undertake SnapShot of production data to MVS Volume F76)
  (undertake Snap-back of last SnapShot from MVS Volume F75)
varyonvg xpevg
mount /xpefs
cp /xpefs/testfile /tmp/testfile      (copy file(s) to temporary area)
umount /xpefs
varyoffvg xpevg
  (undertake Snap-back of production data from MVS Volume F76)
varyonvg xpevg
mount /xpefs
cp /tmp/testfile /xpefs/testfile      (restores file(s))

```

Figure 48. File Restore Process

12.7 Year 2000 Testing

The XPE environment is ideal for Year 2000 testing. You could have your normal production environment separate from XPE and just use XPE to keep copies of the production data and to do Year 2000 testing. Alternatively you could use an XPE environment as your production system and use the SnapShot type facilities to switch between Year 2000 testing and normal productive work.

The key benefit here is that you can save the data that is about to be tested and easily restore that data back to its original state for normal operation, thereby allowing for short test slots on the production machine rather than dedicating longer periods of time on possibly additional machines.

Follow these steps to undertake a Year 2000 test:

1. Use SnapShot from OS/390 to copy the production data.
2. Note the hdisk number in each production volume group for a later importvg step.
3. Change the RS/6000 time to a future year 2000 relevant date.
4. Undertake the tests.
5. Disconnect the RS/6000 from the XPE.
6. Once the testing is finished, rebuild the RS/6000, possibly reinstalling AIX.
7. Snap-back the original production data.
8. Reconnect the RS/6000 to the XPE.
9. Reboot the RS/6000.
10. After a new install, the RS/6000 will not know about the original production data. Use the importvg command to restore each volume group in turn and therefore the original data. As a parameter for each importvg command use the hdisk number saved in step 2.

For the latest information about Year 2000, check the IBM Web site:

<http://www.software.ibm.com/year2000/>

12.8 Recovering from an XPE Failure

If an XPE should fail, any servers with outstanding I/O activity will see SCSI bus and SCSI disks errors. Shut down all servers immediately until the XPE is recovered, fixed, or replaced.

The whole area of system availability requires serious consideration especially in a production environment. See Chapter 15, "Increasing Systems Availability" on page 93.

Chapter 13. Windows NT Operational Considerations

In this chapter we describe the operation of a Windows NT server with attached XPE disks.

In normal operation, the Windows NT server should be started after the XPE to ensure that the SCSI disks are available for use.

The main operational processes to consider are:

- Adding new SCSI disks
- Moving disks to other servers
- Data backups
- Creating a clone server
- Recovering data with ADSM
- Recovering a lost or deleted file without ADSM
- Year 2000 testing
- Recovering from an XPE Failure

13.1 Adding New SCSI Disks

Follow these steps to create a new SCSI disk:

1. Initialize a new disk in the RAMAC Array Storage, using ICKDSF and specifying a new VOLSER.
2. Add the new disk (with a unique SCSI ID, LUN, and VOLSER) to the XPE configuration nondisruptively.
3. Activate the new configuration, by shutting down the XPE and restarting it.
All servers connected to the XPE are affected. Therefore a quiet period of operation must be planned for this short interrupt.
4. Reboot the server. The server will indicate that the configuration has changed. Go to the Disk Administrator panel, and you will see the new disk, for example *Disk 3*, with 2541 MB capacity.

The other servers may need to be rebooted if there is any activity during the XPE shutdown and restart process.

5. Use standard Windows NT commands to configure this disk into a new or existing drive with the appropriate file system, either file allocation table (FAT) or NT file system (NTFS).

Follow these steps to delete an existing disk:

1. Stop all use of the disk.
If the disk to be deleted is one of many disks acting as one drive, you have to back up the data and restore it later, using a Windows NT utility.
2. Remove the disk from the XPE configuration by updating the XPE configuration through the Web Interface.
3. Activate the new XPE configuration. A suitable quiet time must be planned for the shutdown and restart of the XPE.
4. Release the RAMAC Array Storage disk for other use.

13.2 Moving Disks to Other Servers

The steps to move a disk to other servers differ according to whether the disk is being moved:

1. Within the same XPE or to another XPE
2. To a like technology server, such as Netfinity to Netfinity

This example concentrates on moves within the same XPE and between Windows NT servers:

1. Stop all use of the disk by the Windows NT server.
2. Delete the drive from Windows NT.
3. Reconfigure the XPE.

Select a free SCSI ID and LUN on the new server, update the SCSI bus ID field to the bus number of the bus attached to the new server, and change the VOLSER (if there is some relationship between the OS/390 VOLSER and server name).

4. If the disk is being moved to a different technology server, reinitialize it in OS/390, using ICKDSF.
5. If the disk is being moved to the same technology server, rename the data set to something meaningful to match the new server, for manageability purposes.
6. Shut down and restart the XPE to load the new configuration.
7. Reboot the server to create the new disk in Windows NT. The new disk can then be manipulated with the standard Windows NT Disk Administrator commands.

The other servers may not have to be rebooted if there is no activity during the XPE shutdown and restart process.

13.3 Data Backups

Many data backup options are available, including continuing to use the operational backup procedures in use before connecting to the XPE.

However, MVS provides a centralized, sophisticated, and automated backup and recovery environment. One major benefit of the MVS SnapShot facility is that it creates a backup in a matter of seconds, thereby greatly reducing the backup window.

Figure 49 on page 85 shows an example of using SnapShot copy.

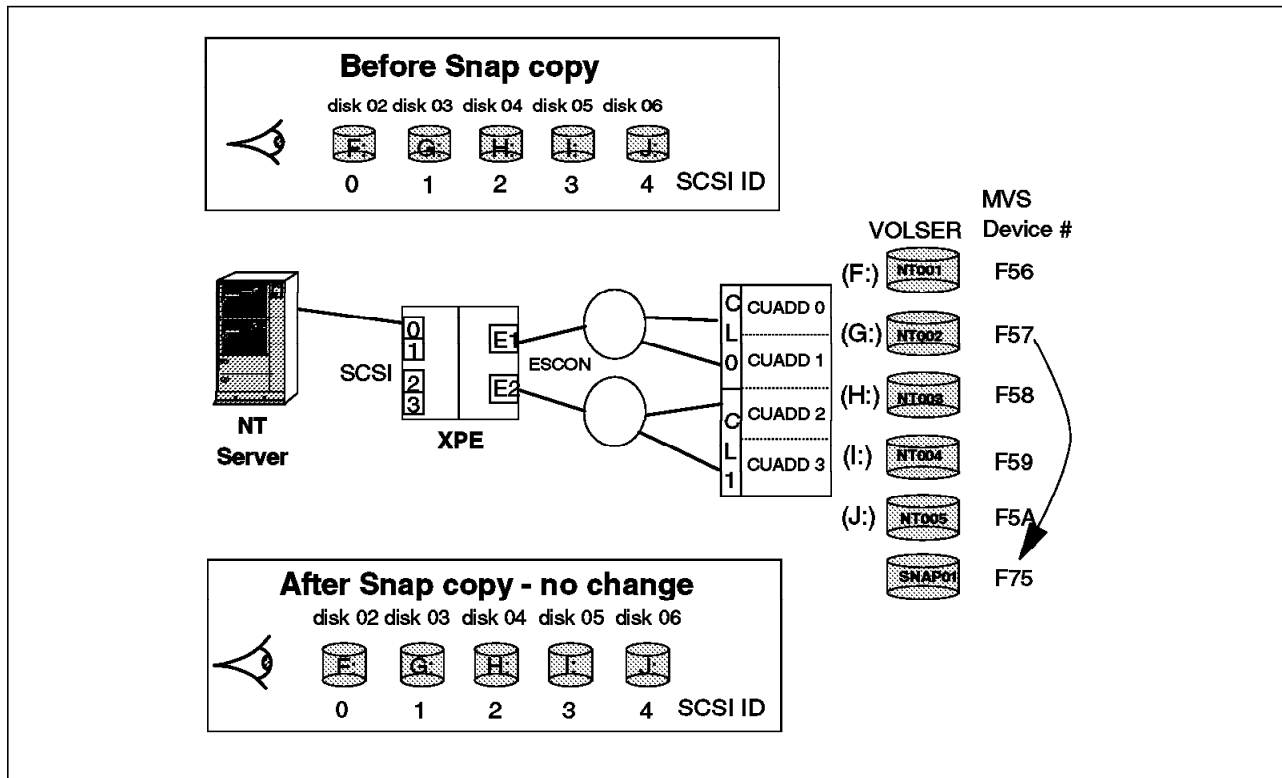


Figure 49. SnapShot Copy: Windows NT

Before running SnapShot, shut down the Windows NT Server. Then take a SnapShot, which copies the data to another OS/390 volume. OS/390 device number F57 is an OS/390 volume also defined to the XPE, and the Windows NT server sees it as *Disk 3*. OS/390 device number F75 is an OS/390 volume not defined to the XPE and is therefore not a known disk in Windows NT. There are no changes made in the Windows NT server after the SnapShot is made.

F75 can then be backed up to tape before the next SnapShot copy is made. For a more detailed discussion of SnapShot, see Appendix B, "The RAMAC Virtual Array and Snapshot" on page 105.

13.4 Creating a Clone Server

If an exact copy of a server is required in another part of the customer installation, cloning can be easily achieved in an XPE environment. Follow these steps to clone a Windows NT system:

1. Use SnapShot from O/390 to make copies of the SCSI disks.
2. Reconfigure the XPE connected to the new server to point to the new SnapShot disks. Note the new SCSI ID definitions.
3. Either install a new Windows NT server or build a copy from the original server.
4. Boot the new server.
5. Cross reference the new SCSI ID and LUN with valid disk numbers.
6. Windows NT automatically discovers the new disks and updates the NT Registry.

13.5 Recovering Data with ADSM

From a procedural point of view, an ADSM environment with XPE-configured disks, and an ADSM environment with non-XPE disks are the same. ADSM operates on a file basis. Files can be recovered from incremental backup. Entire Windows NT drives can be recovered from the root down the hierarchy.

One option is that with Windows NT disk-level backups through MVS, using SnapShot, ADSM NT drive-level recovery can be speeded up using the `-ifnewer` specification on the `dsmc restore` command. Depending on the frequency you take SnapShot copies of the XPE volumes, the `-ifnewer` forward recovery can potentially reduce the overall recovery time of the Windows NT drives.

13.6 Recovering a Lost or Deleted File without ADSM

Using XPE and the RAMAC Array Storage, you can recover lost files. The process documented below assumes that there is a 1:1 ratio of one *disk* per drive letter.

If there are multiple disks in a drive letter, all relevant disks in that drive must be snapped by OS/390, and all the new SCSI IDs and LUNs for the new disks need to be added to the XPE configuration ready for the Windows NT server to discover upon reboot.

Follow these steps to recover a lost or deleted file without ADSM:

1. Snap-back the snapped backup copy to a new disk. As the target disk is not being used, the Snap-back can be performed nondisruptively.
2. Add the new disk (with a unique SCSI ID and LUN) to the XPE configuration nondisruptively.
3. To activate the new configuration, shut down and restart the XPE.

All servers are affected, so a quiet period of operation must be planned for this short interrupt.

4. Reboot the server to discover the new disk.

The other servers do not have to be rebooted if there is no communication or data transfer with the XPE during the XPE shutdown and restart process.

5. Copy the files between the new drive and the drive with the missing files.

Alternatively, using Figure 50 on page 87 as an example, if you already have a spare unused disk (J:) on your Windows NT server, you can use SnapShot to restore the backup copy from the backup disk and copy the files you need back to the G: drive.

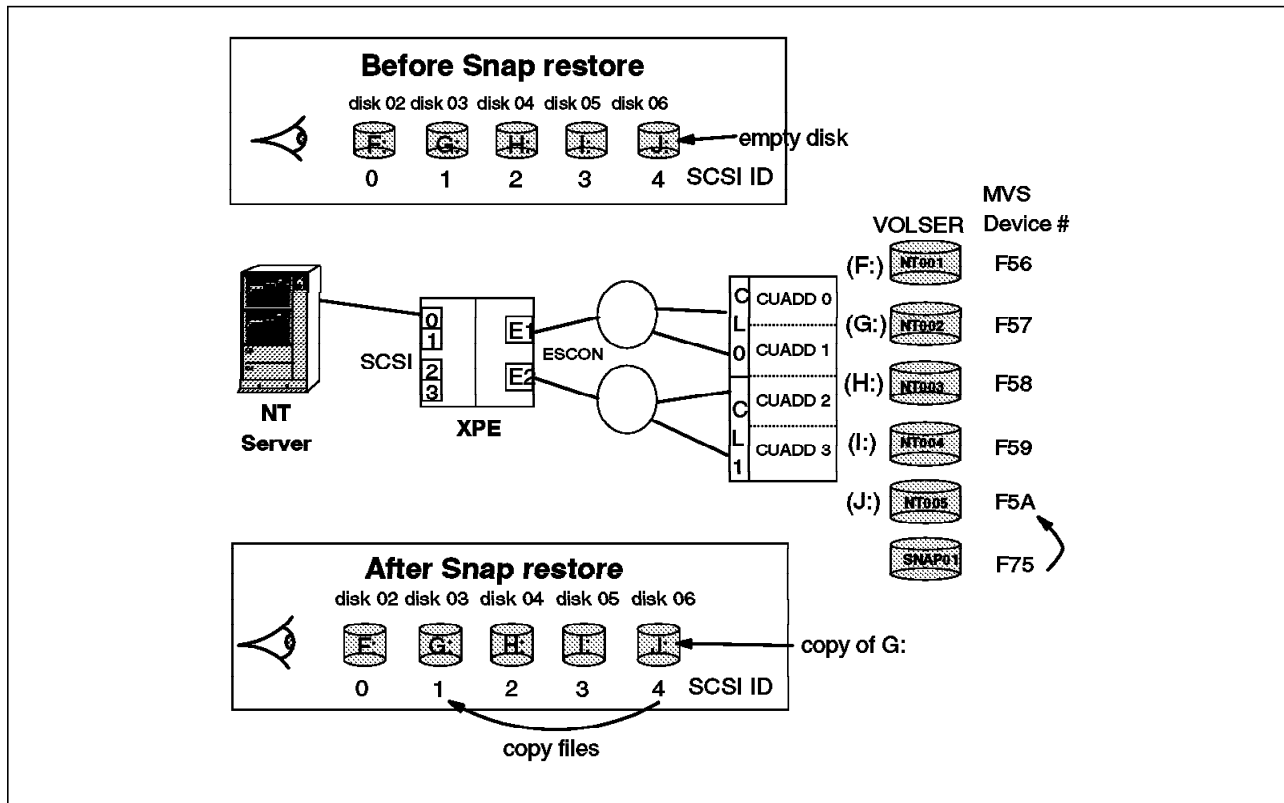


Figure 50. Using Snap-back to Recover Lost Files: Windows NT

In Figure 50, *disk 6* is available on the Windows NT server as empty drive J: connected to the XPE. MVS volume F75 has a recent SnapShot copy of MVS device F57, which is the Windows NT server *disk 3* (G: drive). You Snap-back MVS device F57 to F5A (the NT J: drive). The Windows NT system can access the data on J: and copy the required data back to the G: drive.

After restoring the files, drive J: should not be used any further, as the data on it is a copy of the G: drive when the last SnapShot copy was made, and many changes could have occurred since that time. You can erase all data on the J: drive, reformat it, or remove the drive from the XPE.

13.7 Year 2000 Testing

The key benefit here is the ability to protect the data that is being tested.

Follow these steps to undertake a Year 2000 test:

1. Use SnapShot in OS/390 to save the live data.
2. Change the Windows NT server date time to a future Year 2000 relevant date.
3. Undertake the tests.
4. Disconnect the Windows NT server from the XPE.
5. Rebuild the Windows NT server, possibly reinstalling Windows NT.
6. Reconfigure the XPE to include the SnapShot disks.
7. Reconnect the Windows NT server to the XPE.

8. Reboot the Windows NT server.
9. Cross reference the SCSI IDs and LUNs with their correct *disk* number. All the drives prior to the Year 2000 test will be rediscovered by Windows NT.

13.8 Recovering from an XPE Failure

If an XPE should fail, any servers with outstanding I/O activity will see SCSI bus and SCSI disks errors. Shut down all servers immediately until the XPE is recovered, fixed, or replaced.

This area of system availability requires serious consideration, especially in a production environment. See Chapter 15, "Increasing Systems Availability" on page 93.

Chapter 14. Performance Characteristics

XPE is designed to offer investment protection for customers using IBM RAMAC Array Family storage. As such, it enables customers to invest in RAMAC Array Family storage today with the confidence that, regardless of how their needs for additional capacity change in the future, they are making a sound investment. XPE is one of IBM's cross-platform data sharing solutions. With it, users can take advantage of S/390 data management functions such as SnapShot to create nearly instantaneous data backup or data replicates useful as the basis for facilitating Year 2000 compliance testing.

When customers acquire new storage systems for their most demanding business applications, they have traditionally "cascaded" older storage systems for use with less demanding, "second-tier" applications. XPE enhances this process by making it possible for customers to "cascade" RAMAC Array Family storage as an alternative storage medium for open application data.

S/390 server and open systems customers frequently choose to separate their most critical and performance-sensitive applications from each other by using dedicated storage and, in some cases, dedicated servers. By operating in this way, customers reduce contention and ease storage management. IBM's strategy for XPE and the 7133 matches this manner of customer operation.

In general, where customers require new storage for open systems—especially for production applications—IBMs advice remains unchanged, that is, the 7133 offers outstanding performance for open systems applications as a consequence of its advanced SSA interface and high-speed IBM disk drives. The Versatile Storage Server (VSS) offers high performance in combination with storage sharing among multiple UNIX, Windows NT, and IBM AS/400 servers, as well as data sharing capability between heterogeneous platforms.

XPE is likely to be most valuable when combined with the installation of a new, or use of an existing, RVA as a means of investment protection for the RVA. In most cases where customers require both open systems and S/390 storage for applications today, a combination of the RVA and 7133 will be the best choice.

14.1 Usage and Configuration Guidance

XPE was not designed to be a replacement for high-performance open systems disk. It does, however, serve as a vehicle for extending additional options relevant to maximizing the effectiveness of space used on the open systems platform.

Generally 90 GB should be regarded as the maximum practical capacity that a single XPE can support, although low activity workloads may be able to use up to 180 GB. XPE may be a good choice in support of server consolidation, where extended ESCON channel distances may be used instead of a LAN connection. For example, without XPE, a system might have used a LAN to access data stored on disk attached to a remote server. With XPE, the same system can be directly attached to an XPE, which is directly attached to S/390 storage. The XPE solution should offer competitive performance with the LAN solution.

Care should be taken when selecting data for storage using XPE. Although staff operating with S/390 systems will be familiar with such activity, many open systems staff may not be familiar with the process of selecting data to match the characteristics of a storage system. The tasks associated with this process may be simplified particularly in environments where test and production applications are clearly defined.

For most customers, ADSM file-level backups provide greater flexibility and are easier to manage than S/390 server backup tools. You should consider ADSM as IBM's premier backup solution both for "normal" open systems storage and for XPE. Nevertheless, SnapShot and/or DFSMSdss provide very high-speed backup capabilities, which may be able to offset some of the performance characteristics of XPE.

14.2 XPE Capacity

In terms of capacity, up to 64 devices (or disks) are supported, connecting to up to four servers using four SCSI connections. Each SCSI adapter in the XPE has two connections. Each of these 64 disks in an unformatted state, as far as the server is concerned, is approximately 2.8 GB. This capacity decreases once the disks have been formatted to about 2.5 GB.

Normally this maximum capacity is thought of as connecting up to 16 different disks to up to four different servers. However, the only restriction in the XPE configurations are a maximum of 64 disks and a maximum of four SCSI adapters, the latter being a hardware restriction in the XPE.

Technically many disk configurations are possible, among them the following:

1. Up to 64 disks configured in the XPE to connect to just one server through one SCSI connection
2. Up to 64 disks configured in the XPE to connect to just one server using all four SCSI connections
3. Up to 64 devices configured in the XPE to connect to four servers, each server being supplied with 16 disks

14.3 XPE Performance

Configurations for maximum performance are rarely the same as for maximum capacity. For example, four independent servers each with 16 disks can put a greater load on the XPE than one server with all 64 disks.

To understand the potential performance of the XPE it is important to understand a number of issues and the exact customer environment in which it is being implemented:

- The software in the XPE converts SCSI commands to and from ESCON commands. These commands usually include data in the transmission.
- The XPE uses caching techniques to reduce the effect of the SCSI-to-ESCON conversion.
- In the UNIX server, `queue_depth` parameters are set to allow multiple outstanding requests.

- The mainframe storage device has cache, for example, up to 4 GB in the RVA, RSA, and 3990/9390 storage subsystems.

We recommend that you plan to have 3-5 MB of cache available for every gigabyte of XPE storage to optimize performance.

- The Windows NT and UNIX servers will assume they have sole access to their disks but the RAMAC Array Storage may also be servicing I/O requests from S/390 mainframes.
- On sequential reads, there is a high chance of a buffer hit in the XPE or the cache in the RAMAC Array Storage already having the data.
- On random reads, there is a lesser chance of a buffer hit in the XPE or the cache in the RAMAC Array Storage already having the data.
- On sequential writes, there is a high chance of the buffering in the XPE or the cache in the RAMAC Array Storage making less writes necessary to the actual disk.
- On random writes, there is a high chance that buffering up of writes cannot be made to the actual disk.

Both ESCON connections should always be active for availability but also to enhance performance. Striping the data across multiple hdisks in AIX or multiple disks in Windows NT will also enhance the performance and utilize both ESCON connections.

Chapter 15. Increasing Systems Availability

In this chapter we discuss increasing the availability of an XPE system that is providing data to up to four servers.

We recommend that you plan for the failure of any of the following components, to increase systems availability:

- RAMAC Array media
- A SCSI adapter or cable
- A UNIX or Windows NT Server
- An XPE
- An ESCON adapter or ESCON link
- A RAMAC Array Storage

You can then decide what you are prepared to invest to avoid such a problem or, in the very least, recover as quickly as possible.

15.1 RAMAC Array Media Failure

A RAMAC Array media failure should not really be visible, as the RAMAC Array Storage provides RAID capability, and any disk error would be hidden from the UNIX or Windows NT servers.

The only possibility is a logical disk error, which occurs when a component between the user and the actual RAMAC Array Storage disks has failed. Sections 15.2, "A SCSI Adapter or Cable Failure" through 15.6, "A RAMAC Array Storage Failure" on page 94 discuss protecting your installation against component failure.

15.2 A SCSI Adapter or Cable Failure

To avoid an adapter or cable error, you can mirror between different disks on two different SCSI buses that are connected to the XPE. This is the same technique you would use if you had an array of SCSI disks that you wanted to protect from failure.

When you use the mirroring technique, remember that a fully configured XPE can support only two servers. If there is a SCSI adapter or cable error, the server will still have a path into half its UNIX disks or Windows NT drives.

15.3 A UNIX or Windows NT Server Failure

If a server fails, there will be an interruption in service for those users connected to it.

You can physically unplug the failed server from the XPE and connect a new server, boot the new server, manually reestablish the links to the SCSI disks (for example, in AIX, `importvg` the volume groups), and then start productive use.

Ideally, products like High Availability Cluster Multi-Processing for AIX (HACMP) or the Microsoft Cluster Server feature (codename Wolfpack) in NT Server Enterprise Edition 4.0 would be very useful here, but the difficulty today is that to move the SCSI disks to a new server means you have to shut down the XPE, reconfigure the disk definitions in the XPE, and reboot the XPE, all of which is a manual process.

You cannot attach devices on a single bus to more than one server which means that typical HACMP configurations of sharing the devices between two systems, are not supported.

15.4 An XPE Failure

If the XPE fails, any servers with outstanding I/O activity will see SCSI bus and SCSI disks errors. All the servers should be shut down immediately until the XPE is recovered, fixed, or replaced.

After the first configuration of an XPE, the IBM Service Representative can provide a diskette that contains the XPE configuration. If the XPE should fail, the diskette can be used in a replacement XPE to effectively clone the failed XPE.

Although we did not do any specific testing, a situation worth considering is where you connect a server that has two SCSI adapters to two different XPEs. These XPEs are connected to the same RAMAC Array Storage.

Each XPE is configured to contain half of all the disks, for example, the 16 disks configured through one XPE are a different set of 16 disks configured through the second XPE.

You could configure the server to use mirroring and any data written to one disk is written to a mirrored disk connected through the other XPE. If the XPE fails, the server will still be able to access the mirrored copy. The mirrored copy could be a mirrored logical volume in UNIX or a mirrored drive in Windows NT.

Detailed planning and systems management are required to implement this scenario; otherwise data could be corrupted or lost completely.

15.5 An ESCON Adapter or ESCON Link Failure

A failure in the ESCON path from the XPE to the RAMAC Array Storage should not be a problem, because there are two ESCON adapters in the XPE and thus a second route to complete the I/O. The XPE has built-in error recovery to mask any ESCON failure from the servers.

15.6 A RAMAC Array Storage Failure

The RAMAC Array Storage subsystems are designed to S/390 standards of availability; they are fault tolerant and have dual power feeds and RAID array design. The exception in the devices supported by XPE is the 3390 Model 3, which is a non-RAID disk array. The 3390 Model 3 can be mirrored within a single 3990 with the dual copy (RAID 1) feature.

A failure of a disk in the RAID RAMAC Array Storage can be recovered without loss of data or access. There are obviously areas that RAID cannot protect

against; for example, floods, earthquakes, and total power outages. In those circumstances, you can use one of the remote copy solutions that enable you to mirror the volumes on the RAMAC Array Storage at long distance. PPRC manages synchronous mirroring at distances of up to 26 km and is designed to offer no data loss in the event of a failure at the primary site. XRC can manage an asynchronous copy at virtually unlimited distances, with minimal data loss, but with data integrity.

15.7 Potential High Availability Configurations That We Have Not Tested

The examples discussed have not been put through any formal test and are included as discussion items only for potential users as they gain experience in using the XPE.

Figure 51 is an example of an HACMP idle or rotating standby configuration (with one server inactive). The two RS/6000 servers could be more than 50 km apart, providing a basic disaster recovery possibility, but, unlike standard HACMP over two sites (using SSA optical extender technology), the disks are in one place.

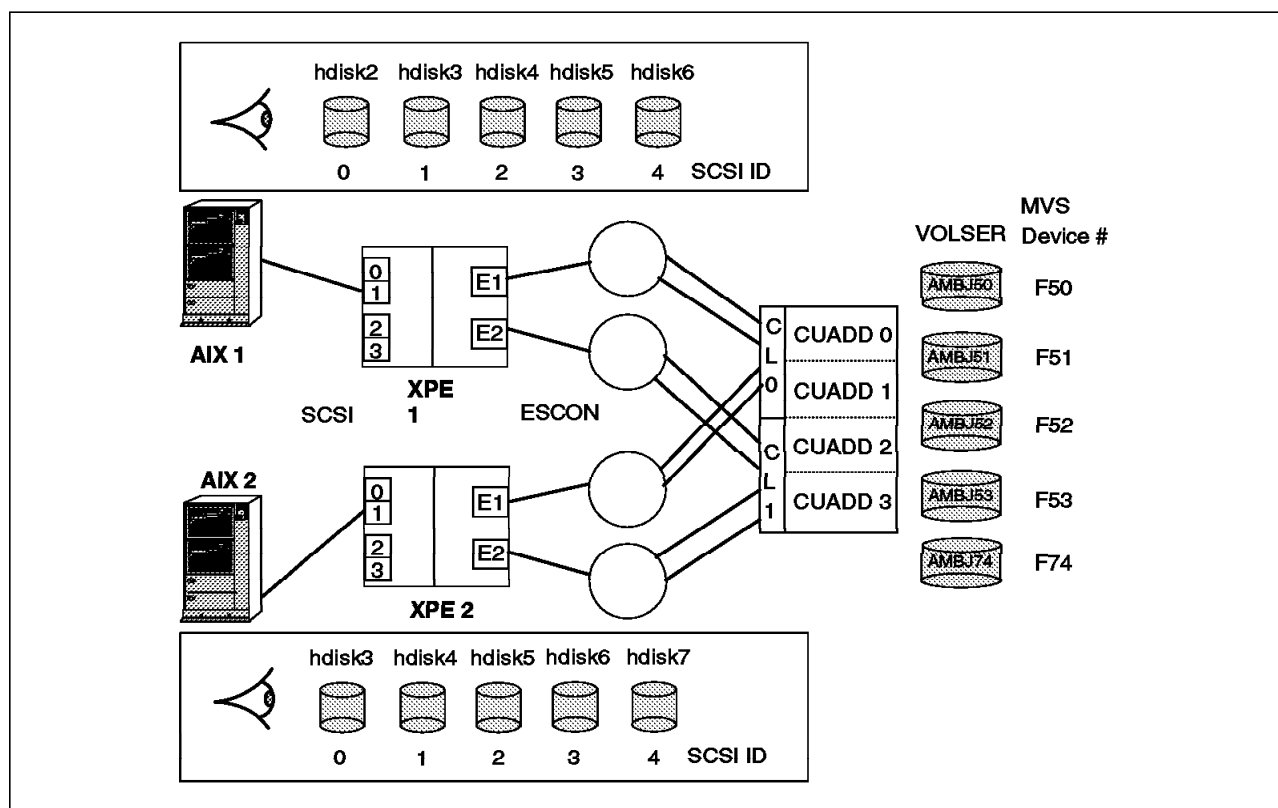


Figure 51. Sample Potential HACMP Configuration

This solution requires some customization work in HACMP; you have to use the HACMP post-event script facility to vary on all the disks on the system doing the takeover.

However, the danger of this type of configuration is that both RS/6000 servers can vary on their "shared" volume groups at the same time and write to the disks independently of each other, which in standard high-availability configurations, the HACMP software will prevent.

The XPE is also a single point of failure, but the situation is manageable as each RS/6000 can be customized to detect an XPE failure (detected as loss of LVM quora or hdisk errors) in an AIX error notification method.

More complex configurations could be considered when:

1. There is one site with two RS/6000s, two XPEs, and one RAMAC Array Storage. Each RS/6000 server has two SCSI adapters, each of which is connected to a different XPE. By cross-connecting the RS/6000s to the two XPEs, a mutual takeover/dual active configuration becomes a possibility (see Figure 52).

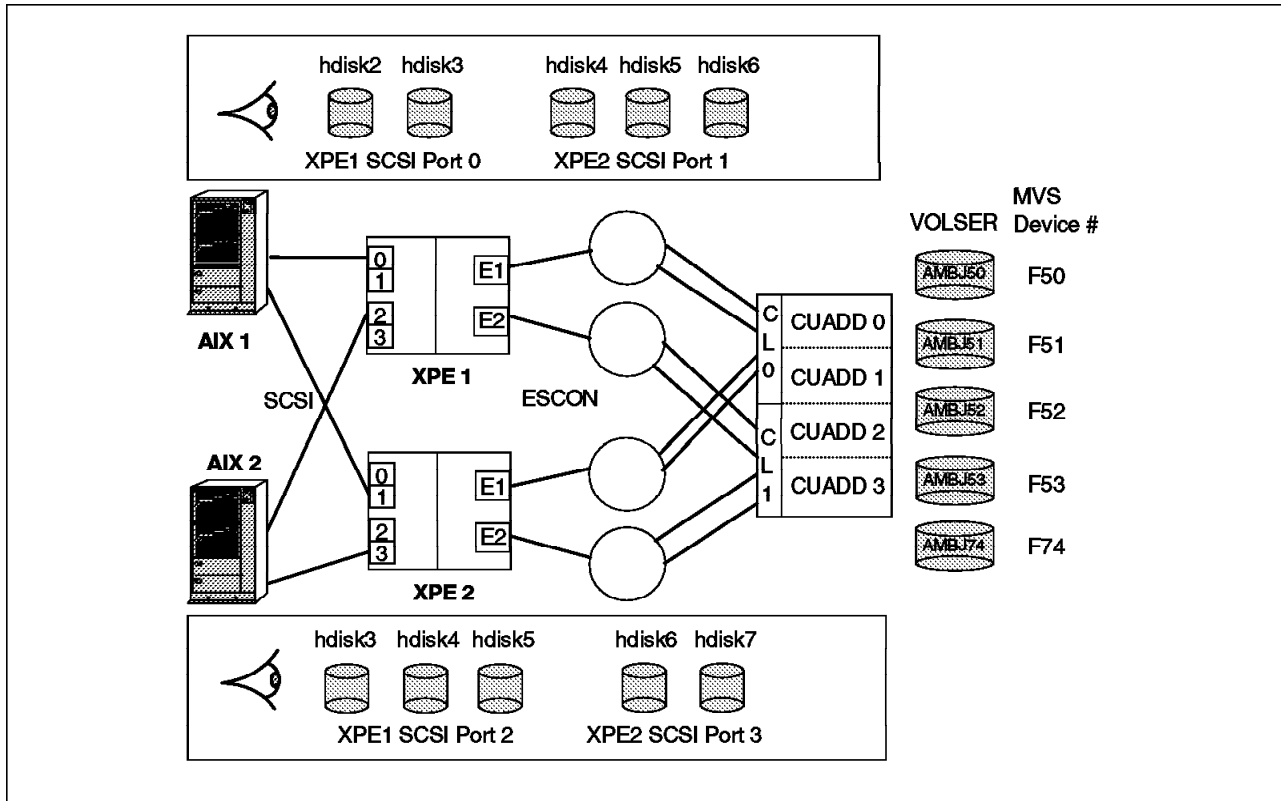


Figure 52. A Complex High-Availability Configuration

In the example in Figure 52, each RS/6000 is configured to see all the disks, in this case, five disks. Each XPE is configured to pass two disks to one RS/6000 and three disks to the second RS/6000. There is no relevance to which SCSI ports on the XPE are used except for the purposes of highlighting which XPE and SCSI port provide which disk to each RS/6000. The XPEs are configured independently, so careful planning is required when defining the disk configuration to ensure that when the two RS/6000s are booted they both see all five disks. Again you have to customize HACMP to use post-event scripts to vary on the disks to the system doing the takeover.

2. There are two sites and each site has one RS/6000 server, two XPEs, and one RAMAC Array Storage, so this could imply an S/390 at each site (see Figure 53 on page 97). In this instance, if:
 - The RS/6000 server is connected to both XPEs
 - Each XPE is connected to a different RAMAC Array Storage (one local and one remote)

- An HACMP idle standby configuration is implemented
 - Appropriate heartbeat paths are configured between the two AIX systems
- any data written to one RAMAC Array Storage would also be written to the second RAMAC Array Storage.

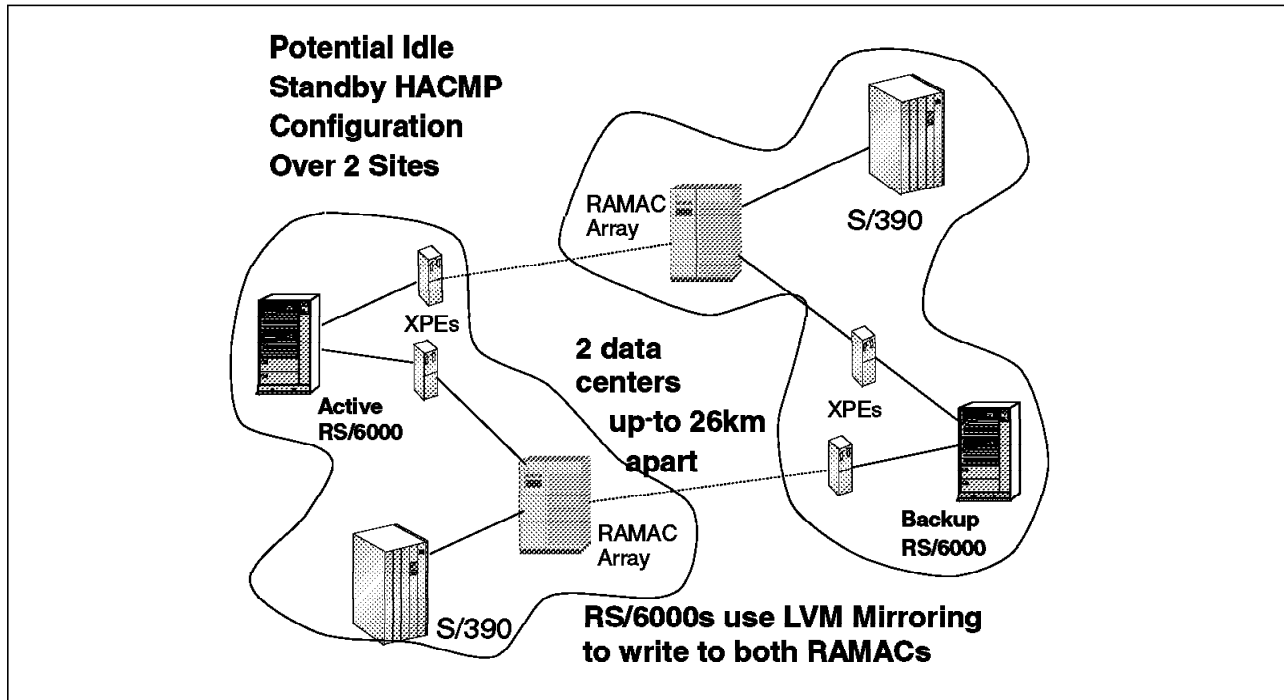


Figure 53. A More Complex High-Availability Configuration

This is a fairly complex configuration to manage and typically requires skilled resources at each site, but it could provide an automated environment with the fast backup capability of S/390 facilities like SnapShot.

(HACMP for AIX is not currently announced as supporting XPE.)

Similarly, adding the Microsoft High Availability Clustering software into this environment and running in an idle standby configuration (with one server inactive) could provide an automated takeover facility. Microsoft has made no statement as yet as to whether the Microsoft Cluster Server feature in NT Server Enterprise Edition 4.0 supports XPE.

Appendix A. XPE Installation Worksheets

This appendix contains three worksheets to be completed for a satisfactory XPE installation and as operational reference material:

- Configuration worksheet (Table 9 on page 100, Table 10 on page 101, and Table 11 on page 102)
- ESCON worksheet (Table 12 on page 103)
- Web configuration worksheet (Table 13 on page 103)

The XPE SCSI Definition and XPE ESCON Definition sections of the configuration worksheet, ESCON worksheet, and Web configuration worksheet must be completed for the IBM Service Representative to undertake the installation.

You can complete the other sections of the worksheet as you make decisions about the layout and organization of the XPE disks.

A.1 Configuration Worksheet

Table 9. XPE Configuration Worksheet: Devices DEV000-DEV015

Server Information				XPE SCSI Definition				XPE ESCON Definition			MVS Definition
UNIX/NT Server Name	Logical Name	Physical Disk Name	Server Type	XPE Device Number	XPE SCSI Bus	XPE SCSI ID	XPE SCSI LUN	RAMAC CUADD	RAMAC ESCON Address	MVS VOLSER	MVS Device Number
				DEV000							
				DEV001							
				DEV002							
				DEV003							
				DEV004							
				DEV005							
				DEV006							
				DEV007							
				DEV008							
				DEV009							
				DEV00A							
				DEV00B							
				DEV00C							
				DEV00D							
				DEV00E							
				DEV00F							
				DEV010							
				DEV011							
				DEV012							
				DEV013							
				DEV014							
				DEV015							

Note: We recommended that you keep this worksheet as a readily accessible reference document after installation.

Table 10. XPE Configuration Worksheet: Devices DEV016-DEV02B

Server Information				XPE SCSI Definition				XPE ESCON Definition			MVS Definition
UNIX/NT Server Name	Logical Name	Physical Disk Name	Server Type	XPE Device Number	XPE SCSI Bus	XPE SCSI ID	XPE SCSI LUN	RAMAC CUADD	RAMAC ESCON Address	MVS VOLSER	MVS Device Number
				DEV016							
				DEV017							
				DEV018							
				DEV019							
				DEV01A							
				DEV01B							
				DEV01C							
				DEV01D							
				DEV01E							
				DEV01F							
				DEV020							
				DEV021							
				DEV022							
				DEV023							
				DEV024							
				DEV025							
				DEV026							
				DEV027							
				DEV028							
				DEV029							
				DEV02A							
				DEV02B							

Note: We recommended that you keep this worksheet as a readily accessible reference document after installation.

Table 11. XPE Configuration Worksheet: Devices DEV02C-DEV03F

Server Information				XPE SCSI Definition				XPE ESCON Definition			MVS Definition
UNIX/NT Server Name	Logical Name	Physical Disk Name	Server Type	XPE Device Number	XPE SCSI Bus	XPE SCSI ID	XPE SCSI LUN	RAMAC CUADD	RAMAC ESCON Address	MVS VOLSER	MVS Device Number
				DEV02C							
				DEV02D							
				DEV02E							
				DEV02F							
				DEV030							
				DEV031							
				DEV032							
				DEV033							
				DEV034							
				DEV035							
				DEV036							
				DEV037							
				DEV038							
				DEV039							
				DEV03A							
				DEV03B							
				DEV03C							
				DEV03D							
				DEV03E							
				DEV03F							

Note: We recommended that you keep this worksheet as a readily accessible reference document after installation.

A.1.1 ESCON Worksheet

<i>Table 12. ESCON Worksheet</i>					
ESCON Card Number	Switched / Direct(1)	ESCD Port Address	ESCD Port Address	ESCD Port Address	ESCD Port Address
ESCON Card 0					
ESCON Card 1					
Note: (1) If Direct, an ESCD port address is not required.					

A.1.2 XPE Web Configuration Worksheet

<i>Table 13. XPE Web Configuration Worksheet</i>	
Host name for this XPE	
IP address for this XPE	
Network mask	
IP address of your domain name server	
Domain name	
IP Address of your gateway	

Appendix B. The RAMAC Virtual Array and Snapshot

This appendix gives the reader who is unfamiliar with the RAMAC Virtual Array (RVA) an overview of the architecture and the unique RVA SnapShot function that is frequently referenced in the main part of the book. For more information about the RVA and Snapshot, read *RAMAC Virtual Array*, SG24-4951, and *Implementing SnapShot*, SG24-2241.

B.1 Overview of RVA

Traditional storage subsystems, such as the 3880, 3990, and RAMAC Array DASD family, use the count key data (CKD) architecture. The CKD architecture defines how and where on the disk device the data is stored. Any updates to the data are written directly to the same position on the physical disk where the updated data was originally. This is referred to as *update in place*.

IBM's RVA provides a high-availability, high-performance storage solution, thanks to its revolutionary virtual disk architecture. To the host, the RVA appears as up to four traditional 3990 subsystems, with up to 256 3380 and/or 3390 volumes. These devices do not physically exist in the subsystem and are referred to as functional devices. In reality, the subsystem contains RAID-protected arrays of fixed block architecture (FBA) disk devices.

The RVA translates I/O requests from CKD format to FBA format.

B.1.1 Log Structured File

The RAID-protected FBA disk arrays that make up the RVA's physical disk space are sequentially filled with data. New and updated data is placed at the end, as it is on a sequential or log file. We call this a *log structured file*.

Updates leave areas of data that are no longer needed in the RVA log file. A microcode recycle process called *freespace collection* ensures that these areas are removed so that there is always enough freespace for writing. You can control the freespace by observing the net capacity load (NCL) of the RVA and using the IBM Extended Facilities Product (IXFP) program. See the RVA redbook *IBM RAMAC Virtual Array*, SG24-4951, for more information. The RVA's physical disk space typically should not be filled above 75% NCL. Above that level, the freespace collection process runs with higher priority, and performance degradation may result.

The following tables are used to map the tracks of functional devices to the FBA blocks related to those tracks:

- Functional device table

The functional device table (FDT) holds the information about the functional volumes that have been defined to the RVA.

- Functional track directory

The functional track directory (FTD) is the collective name for two tables that together map each functional track to an area in the RVA's physical storage:

- Functional track table

The functional track table (FTT) contains the host-related, that is, the functional-device-related track pointers of the FTD.

- Track number table

The track number table (TNT) contains the back-end data pointers of the FTD. A reference counter is also part of this table.

Although the FTD consists of two tables, we discuss it as one entity. In fact, each functional track has an entry in the FTD. If a functional track contains data, its associated FTD entry has a pointer to the FBA block where the data starts. The data of a functional track may fit on one or more FBA blocks.

If the data of a functional track is updated, the RVA puts the new data in a new place in the log structured file, and the FTD entry points to the new data location. There is no update in place.

B.1.2 Data Compression and Compaction

Because update in place does not occur in the log structured file, data can be compressed. In fact, all data, as soon as it enters an RVA, is subject to compression. In addition, the gaps between the records, as they appear on the CKD devices, are removed before the data is written to disk. This is called *compaction*. Experience shows that RVA customers can achieve an overall compression and compaction ratio of 3.6:1.

B.2 Overview of SnapShot

The RVA's virtual disk architecture enables you to produce almost instantaneous copies of data sets, volumes, and VM minidisks. We call this function SnapShot. Snapshot produces copies without data movement. We call making a copy with SnapShot a *snap*, the result of a SnapShot is also called a *snap*.

Conventional methods of copying data on DASD consist of making a physical copy of the data on either DASD or tape. Host processors, channels, tape, and DASD controllers are involved in these conventional copy processes. Copying may take a lot of time, depending on available system resources.

With SnapShot, there is no movement of the actual data. Snapping can take just seconds rather than minutes or hours, because data is not moved and the host processor and the channels are not involved in the data transfer. Furthermore, additional physical disk space is not required to accommodate the snap until either the source or the target is updated. As far as the operating system is concerned, the snap is a real copy of the data; as far as the RVA hardware is concerned, it is a virtual copy of the data.

B.2.1 SnapShot Copy

A functional device is represented by a certain number of pointers in the FTD. Every used track has a pointer in the FDT to its back-end data.

The RVA's virtual disk architecture enables you to make a copy of a data set, a VM minidisk, or volume, by copying its FTD pointers. As you can imagine, this is a very fast process. No data movement takes place, and no additional back-end physical space is used. Both FTD pointers, the original and the copy, point to the same physical data location.

Only when either the original or the copied track is updated is its associated FTD pointer changed to point to the new data location. The other FTD pointer remains unchanged. Additional space is needed in this case. As long as there is a pointer to a data block in the physical disk storage, the block cannot become free for the freespace collection process. A reference counter in the TNT prevents the block from becoming free.

On the RVA side, the SnapShot function is quite simple; on the MVS side, some additional operations are necessary. If a data set is snapped, a new data set name is required and given in the snap command. After the RVA performs the snap, MVS must update the catalog and VTOC, and the VVDS if the snap is performed on a VSAM data set.

Although the creation of the copy is a logical manipulation of pointers, the data exists on disk so in this sense is not a logical copy. As far as the operating system is concerned, the two copies are completely independent of each other.

When you make a duplicate with SnapShot, both the source and the target must be within the same RVA.

B.2.2 Invoking SnapShot

SnapShot is a combination of software and RVA microcode functions. It is invoked from the host and executed in the RVA. It can snap entire DASD volumes, VM minidisks, or data sets.

You invoke SnapShot commands with IAFP from:

- A batch job
- A REXX EXEC or CLIST
- The TSO environment

SnapShot commands can be set up to be invoked by end users as well as by database administrators and system programmers. See the *SnapShot Installation for MVS/ESA* manual for more information.

B.2.3 Hardware and Software Requirements

SnapShot is provided through a combination of hardware and software. You need to have both the correct hardware feature at the correct level of Licensed Internal Code (LIC) and the software program for your environment.

B.2.3.1 Hardware

To use SnapShot you need an RVA with the SnapShot enablement feature #6001 installed. This is a chargeable feature.

You also need to ensure that you are at the correct level of LIC. SnapShot Release 1.2 requires a minimum LIC level of 4.03.26 or later.

B.2.3.2 Software

To use SnapShot you need to have the correct SnapShot program product for your environment:

- In an OS/390 or MVS/ESA environment, SnapShot is provided by IBM RAMAC SnapShot for MVS/ESA, 5648-A12, which has the following prerequisites:

- IBM Extended Facilities Product Version 2 (5648-A17), which in turn requires:
 - OS/390 Version 1 Release 1 (5645-001) or later
 - or
 - MVS/SP 4.2.0 (5695-047/5695-048) with either MVS/DFP 3.1.0 (5665-XA3) or MVS/DFP V3.1.0 (5665-XA3), or later
- In an VM/ESA environment, SnapShot is provided by IBM RAMAC SnapShot for VM/ESA (5654-A03), which has the following prerequisites:
 - IBM Extended Facilities Product V2 R1 (5648-A17) plus PTF L170471, which in turn requires:
 - VM/ESA V1 R2.2 (5684-112)
 - ISPF V3.2 (5684-043)

Before installing SnapShot, check with your local IBM support center to get the latest information about the product and its prerequisites.

Appendix C. Special Notices

This publication is intended to help storage administrators, systems programmers, and systems administrators to implement the IBM 9399 Cross Platform Extension. The information in this publication is not intended as the specification of any programming interfaces that are provided by the IBM 9399 Cross Platform Extension. See the PUBLICATIONS section of the IBM Programming Announcement for the IBM 9399 Cross Platform Extension for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, 500 Columbus Avenue, Thornwood, NY 10594 USA.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The information about non-IBM ("vendor") products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other

operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

The following document contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples contain 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.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.

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

AIX®	DFSMS/MVS®
DFSMSdfp	DFSMSdss
DFSMShsm	ESCON XDF
ESCON®	IBM®
OS/390	RAMAC
RMF	RS/6000
S/390®	

The following terms are trademarks of other companies:

Microsoft, Windows, Windows NT, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

SnapShot is a trademark and property of Storage Technology Corporation for a duplication product

IXFP Storage Technology Corporation

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

Appendix D. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

D.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see "How to Get ITSO Redbooks" on page 113.

- *Implementing DFSMSdss SnapShot and Virtual Concurrent Copy*, SG24-5268
- *Implementing Concurrent Copy*, GG24-3990
- *Implementing SnapShot*, SG24-2241
- *RAMAC Virtual Array*, SG24-4951
- *Data Sharing: Implementing InfoSpeed Copy Sharing Solutions Among Heterogeneous Systems*, SG24-4120 (Available August 1998)

D.2 Redbooks on CD-ROMs

Redbooks are also available on CD-ROMs. **Order a subscription** and receive updates 2-4 times a year at significant savings.

CD-ROM Title	Subscription Number	Collection Kit Number
System/390 Redbooks Collection	SBOF-7201	SK2T-2177
Networking and Systems Management Redbooks Collection	SBOF-7370	SK2T-6022
Transaction Processing and Data Management Redbook	SBOF-7240	SK2T-8038
Lotus Redbooks Collection	SBOF-6899	SK2T-8039
Tivoli Redbooks Collection	SBOF-6898	SK2T-8044
AS/400 Redbooks Collection	SBOF-7270	SK2T-2849
RS/6000 Redbooks Collection (HTML, BkMgr)	SBOF-7230	SK2T-8040
RS/6000 Redbooks Collection (PostScript)	SBOF-7205	SK2T-8041
RS/6000 Redbooks Collection (PDF Format)	SBOF-8700	SK2T-8043
Application Development Redbooks Collection	SBOF-7290	SK2T-8037

D.3 Other Publications

These publications are also relevant as further information sources:

- *Remote Copy Administrator's Guide and Reference*, SC35-0169
- *IBM 9399 Cross Platform Extension Installation, Planning, and User's Guide*, GA32-0383

How to Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, CD-ROMs, workshops, and residencies. A form for ordering books and CD-ROMs is also provided.

This information was current at the time of publication, but is continually subject to change. The latest information may be found at <http://www.redbooks.ibm.com/>.

How IBM Employees Can Get ITSO Redbooks

Employees may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Redbooks Web Site on the World Wide Web**

<http://w3.itso.ibm.com/>

- **PUBORDER** — to order hardcopies in the United States

- **Tools Disks**

To get LIST3820s of redbooks, type one of the following commands:

```
TOOLCAT REDPRINT
TOOLS SENDTO EHONE4 TOOLS2 REDPRINT GET SG24xxxx PACKAGE
TOOLS SENDTO CANVM2 TOOLS REDPRINT GET SG24xxxx PACKAGE (Canadian users only)
```

To get BookManager BOOKs of redbooks, type the following command:

```
TOOLCAT REDBOOKS
```

To get lists of redbooks, type the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET ITSOCAT TXT
```

To register for information on workshops, residencies, and redbooks, type the following command:

```
TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ITSOREGI 1998
```

- **REDBOOKS Category on INEWS**

- **Online** — send orders to: USIB6FPL at IBMMAIL or DKIBMBSH at IBMMAIL

Redpieces

For information so current it is still in the process of being written, look at "Redpieces" on the Redbooks Web Site (<http://www.redbooks.ibm.com/redpieces.html>). Redpieces are redbooks in progress; not all redbooks become redpieces, and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

How Customers Can Get ITSO Redbooks

Customers may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Online Orders** — send orders to:

In United States:
In Canada:
Outside North America:

IBMAIL
usib6fpl at ibmail
caibmbkz at ibmail
dkibmbsh at ibmail

Internet
usib6fpl@ibmail.com
lmannix@vnet.ibm.com
bookshop@dk.ibm.com

- **Telephone Orders**

United States (toll free)
Canada (toll free)

1-800-879-2755
1-800-IBM-4YOU

Outside North America
(+45) 4810-1320 - Danish
(+45) 4810-1420 - Dutch
(+45) 4810-1540 - English
(+45) 4810-1670 - Finnish
(+45) 4810-1220 - French

(long distance charges apply)
(+45) 4810-1020 - German
(+45) 4810-1620 - Italian
(+45) 4810-1270 - Norwegian
(+45) 4810-1120 - Spanish
(+45) 4810-1170 - Swedish

- **Mail Orders** — send orders to:

IBM Publications
Publications Customer Support
P.O. Box 29570
Raleigh, NC 27626-0570
USA

IBM Publications
144-4th Avenue, S.W.
Calgary, Alberta T2P 3N5
Canada

IBM Direct Services
Sortemosevej 21
DK-3450 Allerød
Denmark

- **Fax** — send orders to:

United States (toll free)
Canada
Outside North America

1-800-445-9269
1-403-267-4455
(+45) 48 14 2207 (long distance charge)

- **1-800-IBM-4FAX (United States) or (+1)001-408-256-5422 (Outside USA)** — ask for:

Index # 4421 Abstracts of new redbooks
Index # 4422 IBM redbooks
Index # 4420 Redbooks for last six months

- **On the World Wide Web**

Redbooks Web Site
IBM Direct Publications Catalog

<http://www.redbooks.ibm.com/>
<http://www.elink.ibm.com/pbl/pbl>

Redpieces

For information so current it is still in the process of being written, look at "Redpieces" on the Redbooks Web Site (<http://www.redbooks.ibm.com/redpieces.html>). Redpieces are redbooks in progress; not all redbooks become redpieces, and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

IBM Redbook Order Form

Please send me the following:

Title	Order Number	Quantity

First name Last name

Company

Address

City Postal code Country

Telephone number Telefax number VAT number

- Invoice to customer number _____
- Credit card number _____

Credit card expiration date Card issued to Signature

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

Glossary

address. (1) To refer to a device or an item of data by its address. (2) The location in the storage of a computer where data is stored. (3) In data communication, the unique code assigned to each device or workstation connected to a network. (4) The identifier of a location, source, or destination.

AIX. Advanced Interactive Executive, the name of the operating system for the RS/6000.

CKD. Count key data is a means of recording data to disk traditionally used by large systems such as MVS.

DASD. Direct access storage device, the term an MVS systems programmer uses to refer to disks in a disk subsystem.

DASD subsystem. A storage control and its attached direct access storage devices.

destage. The asynchronous write of new or updated data from cache or nonvolatile storage to DASD.

differential SCSI. A hardware configuration for connecting SCSI devices. It uses a pair of lines for each signal transfer (as opposed to single-ended SCSI, which references each SCSI signal to a common ground).

ESCON. Enterprise Systems Connection Architecture, a standard introduced by IBM for connecting to I/O devices quickly over extended distances.

ESCON channel. A channel having an Enterprise Systems Connection channel-to-control-unit I/O interface that uses optical cables as a transmission medium.

ESCON Director (ESCD). A switch for ESCON optical channels that provides connectivity capability and control for attaching any two links to each other.

HCD. Hardware configuration definition. The HCD is the OS/390 process for configuring an OS/390 system, in particular, the I/O definition and ESCON path specifications.

LIC. Licensed Internal Code. Software provided for use on specific IBM machines and licensed to customers under the terms of IBM's Customer Agreement. Microcode can be LIC and licensed as such.

logical partition (LP). One to three physical partitions (copies). The number of logical partitions within a logical volume is variable.

logical volume (in AIX). A set of logical partitions, each of which is stored on one or more physical

partitions from one or more of the physical volumes of a given volume group. A logical volume has a device name (of the form /dev/hdn) and contains a single file system.

logical volume (in MVS). The storage space identified by a common label for example, VOLSER. When requesting access to the logical volume, the host system translates the label to an address that is meaningful to the storage subsystem, for example, device number.

logical volume manager (LVM). Manages disk space at a logical level. It controls fixed-disk resources by mapping data between logical and physical storage and by allowing data to span multiple disks and be discontinuous, replicated and dynamically expanded.

peripheral component interconnect (PCI). A local bus specification that allows connection of peripherals directly to computer memory.

physical partition (PP). The smallest unit of disk-space allocation for a logical volume. The physical partition is contiguous space on a physical volume. A fixed-size portion of a physical volume. One or more physical partitions constitute the underlying physical storage medium for a logical partition.

physical volume (PV). The portion of a single unit of storage accessible to a single read/write mechanism; for example, a drum, a disk pack, or part of a disk storage module.

Redundant Array of Independent Disks - Level 5 (RAID 5). A method of striping data and parity across several disk drives associated in a logical array of disk drives. RAID 5 is characterized by striping increments that support independent accesses to the disk drives. This architecture allows multiple host I/Os to be serviced concurrently, thereby improving array throughput. In a RAID 5 array, parity is distributed across all disks in the array so that the required write operation accesses to parity are not constrained by a single parity disk drive.

RVA. RAMAC Virtual Array, a S/390 storage device using a virtual disk architecture. The RVA utilizes a log structured array architecture to store up to 720 GB of data in a compact unit. Features enabled by the virtual disk architecture and log structured array include SnapShot, compression, and automatic tuning.

SCSI. A standard hardware interface that enables a variety of peripheral devices to communicate with one another. Initially associated with connecting I/O devices to small processors, SCSI is supported on S/390.

SMIT. System Management Interface Tool available with AIX.

Snap. The making of a copy with SnapShot or the result of a SnapShot.

SnapShot. A function of the RVA's virtual disk architecture that provides the ability to almost instantaneously create copies of data sets and volumes, without data movement.

striping. A technique that distributes data in bit, byte, multibyte, record, or block increments sequentially across multiple disk drives.

Ultra SCSI. A standard for SCSI data transfers. It allows a transfer rate of 20 MB/sec over an 8-bit SCSI bus and up to 40 MB/sec over a 16-bit SCSI bus.

VHDCI. Very high density cable interconnect, a cable usually used when connecting to a dual channel SCSI controller.

XPE. The short-form name for the IBM 9399 Cross Platform Extension.

Index

Numerics

9399001 device type 40

A

ABARS 34
ACS routine 34
ADSM 78, 86
AIX
 9399001 device 40
 adding disks 75
 backup 76
 bad block 42
 cloning 77
 configuration worksheet 47
 disk layout 40
 file set 39
 file systems 46
 logical volume 44
 lsdev -cc command 40
 lslv command 45
 lspv command 40, 44
 lsvg command 46
 LVM 41
 mirroring 41
 queue depth 39
 recommendations 47
 rw_timeout 39
 scripts 39
 SMIT 41
 striping 41
 volume group 40
availability
 adapter 93
 ESCON 94
 media 93
 planning 93
 PPRC 95
 server 93
 XRC 95

B

backup
 AIX 76
 automation 67
 MVS 67
 planning 17
 process 67
 SnapShot 68, 70
 Windows NT 84

C

cache 90
capacity planning 15
CE port 8
compression
 XPE volume 34
concurrent copy 68
configuration
 Web interface 61
configuration worksheet
 AIX 47
 description 22, 99
 ESCON definition 23
 field definitions 20
 MVS device information 25
 SCSI definition 22
 server information 22
CONVERTV 34

D

data set
 format 32
 names 32
data sharing
 concepts 3
 levels 3
 storage sharing 4
DEVSERV 24
DFSMSHsm 34
disk
 capacity 90
disk capacity 15
 Windows NT 53
documentation
 Web interface 61

E

ESCON
 adapters 9
 addressing 24
 CUADD 23
 DEVSERV 23
 director 9
 distances 9
 fiber 9
 logical control unit 23
 overview 9
 port address 25
 recommendations 25
 status 63
ESCON definition
 CUADD 23

ESCON Director 9, 25
ESCON worksheet 25, 103
Ethernet
 connection 10, 27
 port 8
extended remote copy
 See XRC

H

HACMP 94
hardware
 planning 16
 supported environments 12
 supported storage subsystems 13
Hewlett Packard
 adapters 12
 implementation 57
 operating system 12
 rw_timeout 57
 SAM 57
 server 12
High Availability Cluster Multi-Processing for AIX
 See HACMP

I

IBM Extended Facilities Product
 See IXFP
ICKDSF
 data set 32
 OPENDISK 30
 output 31
 ptf 13
 sample 30
implementation
 Hewlett Packard 57
 planning 15
 RS/6000 39
 Sun Microsystems 57
 Windows NT 49
installation plan 16
installation worksheets
 configuration worksheet 20
 description 16
 ESCON worksheet 25
 sample 19
 Web configuration worksheet 27
IXFP 73

L

logical disk
 Windows NT 54
logical volume manager
 See LVM
LUN
 overview 9
 Windows NT 52

LVM 41

M

maximum configuration 15
Microsoft Cluster Server feature 94
mirroring
 Windows NT 55
MVS
 ICKDSF 30
 initialization 29, 30
 operations 67
 performance reports 73
 volume selection 29

N

Netscape Navigator 27

O

OPENDISK
 compression 34
 format 29
 IEHLIST output 32
 reinitialization 34
 SnapShot 33
 utilities 34
 VOLSER 29
operations
 ESCON 72
OS/390 Security Server 37

P

Parallel Sysplex 67
PC servers
 Windows NT 13
peer-to-peer remote copy
 See PPRC
performance
 caching 90
 capacity planning 15
 ESCON 91
 recommendations 89
 reports, MVS 73
physical volume ID
 See PVID
planning
 backup 17
 recovery 17
power-off sequence 11
power-on sequence 10
power-on sequence, AIX 75
power-on sequence, Windows NT 83
PPRC 72
PVID 75

Q

quick format 55

R

RACF 37

RAMAC Array Storage

device support 13

SIM support 13

supported products 13

RAMAC Virtual Array

See RVA

recovery

file 71

file, AIX 78

file, Windows NT 86

planning 17

process 71

SnapShot 71

remote copy 72

RS/6000

adapters 12

file set 39

implementation 39

operating system 12

queue depth 39

rw_timeout 39

SCSI addressing 40

RSF

description 8

feature codes 11

modem 11

RVA

architecture 105

capacity 15

compression 106

space management 105

S

SAM 57

SCSI

adapters 8

addressing 9, 23

cables 8

configuration worksheet 22

ID 9

LUN 9

overview 8

status 60

SCSI adapters

Windows NT 50

server

availability 93

server consolidation 89

service information message

See SIM

SIM 73

SMIT 41

SMS

ACS routine 34

CONVERTV 34

recommendations 34, 71

storage class 34

storage group 34

XPE volume 34

SnapShot

AIX 77

backup 70

description 105

DFSMSdss 68

prerequisites 107

recovery 71

SMS 34

Windows NT 84, 86

Year 2000 72

storage class 34

storage group 34

striping

AIX 41

Windows NT 54

Sun Microsystems

adapters 12

implementation 57

operating system 13

rw_timeout 57

servers 13

Solstice DiskSuite 57

System Administration Manager

See SAM

System Management Interface Tool

See SMIT

T

terminology 1

Tivoli 67

U

universal resource locator

See URL

UNIX

backup 68

backup plan 17

URL 27

V

virtual concurrent copy 68

VOLSER 69

W

Web configuration worksheet

definitions 27

Web configuration worksheet (*continued*)

sample 103

Web interface

configuration 61, 65

definition 59

device status 61

documentation 61

ESCON status 63

main page 59

options bar 60

SCSI status 60

URL 27

utilities 61

Windows NT

adapters 13

add disks 83

assigning drive letters 54

backup 68, 84

backup plan 17

cloning 85

control panel 50

disk capacity 53

disk configuration 49

disk number 49

file allocation 83

format 52

implementation 49

logical disk 54

LUN 55

LUN information 52

mirroring 55

PC servers 13

quick format 55

recommendations 56

SCSI adapters 50

SCSI addressing 55

SCSI ID 52

SnapShot 84

striping 54

version 13

volume group 54

Wolfpack 94

worksheets

configuration 47

configuration worksheet 22, 99

ESCON worksheet 25, 103

XPE Web configuration 27, 103

X

XPE

adapters 7

availability 93

cable distance 12

caching 90

CE port 8

components 7

dimensions 11

Ethernet port 8

XPE (*continued*)

initialization 10, 25

installation time 12

installation worksheets 19

maximum capacity 89

maximum configuration 15

memory 10

MVS implementation 29

performance 15, 89

physical planning 16

power cord 11

processor 10

reconfiguration 15

RSF 8

status 59

Web interface 59

weight 11

XPE volume

compression 34

data set 32

format 32

ICKDSF 31

IEHLIST output 32

reinitialization 34

security 37

sharing 67

SMS 34

SnapShot 33

utilities 34

XRC 72

Y

Year 2000

SnapShot 72

testing, AIX 80

testing, Windows NT 87

ITSO Redbook Evaluation

Data Sharing: Cross Platform Extension (XPE) Implementation Guide
SG24-5256-00

Your feedback is very important to help us maintain the quality of ITSO redbooks. **Please complete this questionnaire and return it using one of the following methods:**

- Use the online evaluation form found at <http://www.redbooks.ibm.com>
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Which of the following best describes you?

Customer **Business Partner** **Solution Developer** **IBM employee**
 None of the above

Please rate your overall satisfaction with this book using the scale:
(1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor)

Overall Satisfaction _____

Please answer the following questions:

Was this redbook published in time for your needs? Yes____ No____

If no, please explain:

What other redbooks would you like to see published?

Comments/Suggestions: **(THANK YOU FOR YOUR FEEDBACK!)**

