

HP OpenView Storage Data Protector Concepts Guide

Manual Edition: February 2006 (build label 249)



Manufacturing Part Number: B6960-90001

Release A.06.00

© Copyright Hewlett-Packard Development Company, L.P.2006.

Legal Notices

©Copyright 2004 Hewlett-Packard Development Company, L.P.

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

UNIX® is a registered trademark of The Open Group.

Microsoft®, Windows®, and Windows NT® are U.S. registered trademarks of Microsoft Corporation.

Oracle® is a registered U.S. trademark of Oracle Corporation, Redwood City, California.

Java™ is a U.S. trademark of Sun Microsystems, Inc.

ARM® is a registered trademark of ARM Limited.

1. About Backup and Data Protector

In This Chapter	2
About Data Protector	3
Introducing Backups and Restores	7
What Is a Backup?	7
What Is a Restore?	7
Backing Up a Network Environment	8
Direct Backup	9
Data Protector Architecture	10
Operations in the Cell.	12
Backup Sessions	12
Restore Sessions	14
Enterprise Environments	15
Splitting an Environment into Multiple Cells	16
Media Management	19
Backup Devices	20
User Interfaces.	22
Data Protector GUI.	23
Overview of Tasks to Set Up Data Protector	24

2. Planning Your Backup Strategy

In This Chapter	28
Backup Strategy Planning.	29
Defining the Requirements of a Backup Strategy	29
Factors Influencing Your Backup Strategy.	31
Preparing a Backup Strategy Plan	32
Planning Cells	35
One Cell or Multiple Cells?	35
Installing and Maintaining Client Systems	36
Creating Cells in the UNIX Environment.	37
Creating Cells in the Windows Environment	38
Creating Cells in a Mixed Environment	39
Geographically Remote Cells	40
Understanding and Planning Performance	41
The Infrastructure	41
Configuring Backups and Restores	44
Disk Performance	45
SAN Performance	46

Contents

Online Database Application Performance	46
Planning Security	48
Cells	48
Data Protector Users Accounts	49
Data Protector User Groups	49
Data Protector User Rights	50
Visibility of Backed Up Data	50
Data Encoding	50
Who Owns a Backup Session?	51
Clustering	52
Cluster Concepts	52
Cluster Support	55
Example Cluster Environments	56
Full and Incremental Backups	66
Full Backups	67
Incremental Backups	67
Considering Restore	71
Keeping Backed Up Data and Information About the Data	73
Data Protection	73
Catalog Protection	74
Logging Level	74
Browsing Files for Restore	74
Backing Up Data	77
Creating a Backup Specification	78
Selecting Backup Objects	78
Backup Sessions	80
Object Mirrors	80
Media Sets	80
Backup Types and Scheduled Backups	80
Scheduling, Backup Configurations, and Sessions	81
Scheduling Tips and Tricks	81
Automated or Unattended Operation	87
Considerations for Unattended Backups	87
Duplicating Backed Up Data	89
Copying Objects	90
Object Mirroring	96
Copying Media	98
Restoring Data	100

Restore Duration	100
Selection of the Media Set	101
Operators Are Allowed to Restore	102
End Users Are Allowed to Restore.	103
Disaster Recovery	104
Disaster Recovery Methods	105
Alternative Disaster Recovery Methods	107

3. Media Management and Devices

In This Chapter	110
Media Management	111
Media Life Cycle.	113
Media Pools	114
Free Pools	116
Media Pool Usage Examples	118
Implementing a Media Rotation Policy	121
Media Management Before Backups Begin	124
Initializing or Formatting Media	124
Labeling Data Protector Media	124
Location Field	125
Media Management During Backup Sessions	126
Selecting Media for Backups	126
Adding Data to Media During Backup Sessions	127
Writing Data to Several Media Sets During Backup	130
Calculating Media Condition	130
Media Management After Backup Sessions	131
Vaulting.	131
Restoring from Media in a Vault	133
Devices	134
Device Lists and Load Balancing	135
Device Streaming and Concurrency	136
Segment Size.	137
Block Size	139
Number of Disk Agent Buffers.	139
Device Locking and Lock Names	140
Standalone Devices	141
Small Magazine Devices	142
Large Libraries.	143

Contents

Handling of Media	143
Size of a Library	143
Sharing a Library with Other Applications	144
Enter / Eject Mail slots	144
Barcode Support	144
Cleaning Tape Support	145
Sharing a Library with Multiple Systems	145
Data Protector and Storage Area Networks	152
Storage Area Networks	152
Fibre Channel	153
Device Sharing in SAN	156
Indirect and Direct Library Access	159
Device Sharing in Clusters	161
4. Users and User Groups	
In This Chapter	164
Increased Security for Data Protector Users	165
Access to Backed Up Data	165
Users and User Groups	166
Using Predefined User Groups	167
Data Protector User Rights	167
5. The Data Protector Internal Database	
In This Chapter	170
About the IDB	171
The IDB on the Windows Cell Manager	172
The IDB on the UNIX Cell Manager	172
The IDB in the Manager-of-Managers Environment	172
IDB Architecture	173
Media Management Database (MMDB)	174
Catalog Database (CDB)	175
Detail Catalog Binary Files (DCBF)	176
Session Messages Binary Files (SMBF)	176
Serverless Integrations Binary Files (SIBF)	177
IDB Operation	178
Overview of IDB Management	180
IDB Growth and Performance	181
Key IDB Growth and Performance Factors	181

IDB Growth and Performance: Key Tunable Parameters	182
IDB Size Estimation	187

6. Service Management

In This Chapter	190
Overview	191
Data Protector and Service Management	192
Native Data Protector Functionality	194
Application Response Measurement Version 2.0 (ARM 2.0 API)	195
Integration with HP OpenView Operations	196
Integration with ManageX	196
SNMP Traps	197
The Monitor	197
Reporting and Notification	197
Event Logging and Notification	199
Windows Application Log	200
Java-Based Online Reporting	200
Data Protector Checking and Maintenance Mechanism	201
Central Management, Distributed Environment	201
Using the Data Provided by Data Protector	201
Service Management Integrations	203
Data Protector OVO-OVR Integration	203
Data Protector OVO-SIP	206
Data Protector SIP	206
Integrating Data Protector with HP OpenView Service Desk	207

7. How Data Protector Operates

In This Chapter	210
Data Protector Processes or Services	211
Backup Sessions	212
Scheduled and Interactive Backup Sessions	212
Backup Session Data Flow and Processes	212
Pre-Exec and Post-Exec Commands	215
Queuing of Backup Sessions	216
Mount Requests in Backup Sessions	216
Backing Up with Disk Discovery	217
Restore Sessions	218
Restore Session Data Flow and Processes	218

Contents

Queuing of Restore Sessions	219
Mount Requests in a Restore Session	220
Parallel Restores	220
Fast Multiple Single File Restore	221
Object Copy Sessions	223
Automated and Interactive Object Copy Sessions	223
Object Copy Session Data Flow and Processes	223
Queuing of Object Copy Sessions	225
Mount Requests in an Object Copy Session	225
Object Consolidation Sessions	226
Automated and Interactive Object Consolidation Sessions	226
Object Consolidation Session Data Flow and Processes	227
Queuing of Object Consolidation Sessions	228
Mount Requests in an Object Consolidation Session	228
Media Management Sessions	229
Media Management Session Data Flow	229

8. Integration with Database Applications

In This Chapter	232
Overview of Database Operation	233
Filesystem Backup of Databases and Applications	236
Online Backup of Databases and Applications	237

9. Direct Backup

In This Chapter	242
Overview	243
Direct Backup	244
How Direct Backup Works	245
Direct Backup Process Flow	248
Requirements and Support	250
Supported Configurations	251
Three Hosts: CM, Application, Resolve	251
Two Hosts: Cell Manager/Resolve Agent and Application	252
Basic Configuration: Single Host	252

10. Disk Backup

In This Chapter	254
Overview	255

Disk Backup Benefits	256
Data Protector Disk-Based Devices	258
11. Synthetic Backup	
In This Chapter	262
Overview	263
Synthetic Backup Benefits	264
How Data Protector Synthetic Backup Works	265
Restore and Synthetic Backup	267
12. Split Mirror Concepts	
In This Chapter	272
Overview	273
Supported Configurations	277
Local Mirror - Dual Host	277
Local Mirror - Single Host	278
Remote Mirror	278
Local/Remote Mirror Combination	280
Other Configurations	281
13. Snapshot Concepts	
In This Chapter	284
Overview	285
Storage Virtualization	285
Snapshot Concepts	286
Snapshot Backup Types	287
Instant Recovery	288
Replica Set and Replica Set Rotation	288
Types of Snapshots	289
Supported Configurations	291
Basic Configuration: Single Disk Array - Dual Host	291
Other Supported Configurations	293
Other Configurations	296
14. Microsoft Volume Shadow Copy Service	
In This Chapter	300
Overview	301
Data Protector Volume Shadow Copy Integration	306

Contents

VSS Filesystem Backup and Restore	308
---	-----

A. Backup Scenarios

In This Appendix	A-2
Considerations	A-2
Company XYZ.	A-5
Environment	A-5
Backup Strategy Requirements	A-8
Proposed Solution	A-9
Company ABC	A-22
Environment	A-22
Backup Strategy Requirements	A-24
Proposed Solution	A-26

B. Further Information

In This Appendix	B-2
Backup Generations	B-3
Examples of Automated Media Copying.	B-5
Example 1: Automated Media Copying of Filesystem Backups.	B-5
Example 2: Automated Media Copying of Oracle Database Backups	B-11
Internationalization	B-14
Localization	B-14
File Name Handling	B-14

Glossary

Index

Printing History

The manual printing date and part number indicate its current edition. The printing date will change when a new edition is printed. Minor changes may be made at reprint without changing the printing date. The manual part number will change when extensive changes are made.

Manual updates may be issued between editions to correct errors or document product changes. To ensure that you receive the updated or new editions, you should subscribe to the appropriate product support service. See your HP sales representative for details.

Table 1

Edition History

Part Number	Manual Edition	Product
B6960-90080	May 2003	Data Protector Release A.05.10
B6960-90105	October 2004	Data Protector Release A.05.50
B6960-90001	April 2006	Data Protector Release A.06.00

Conventions

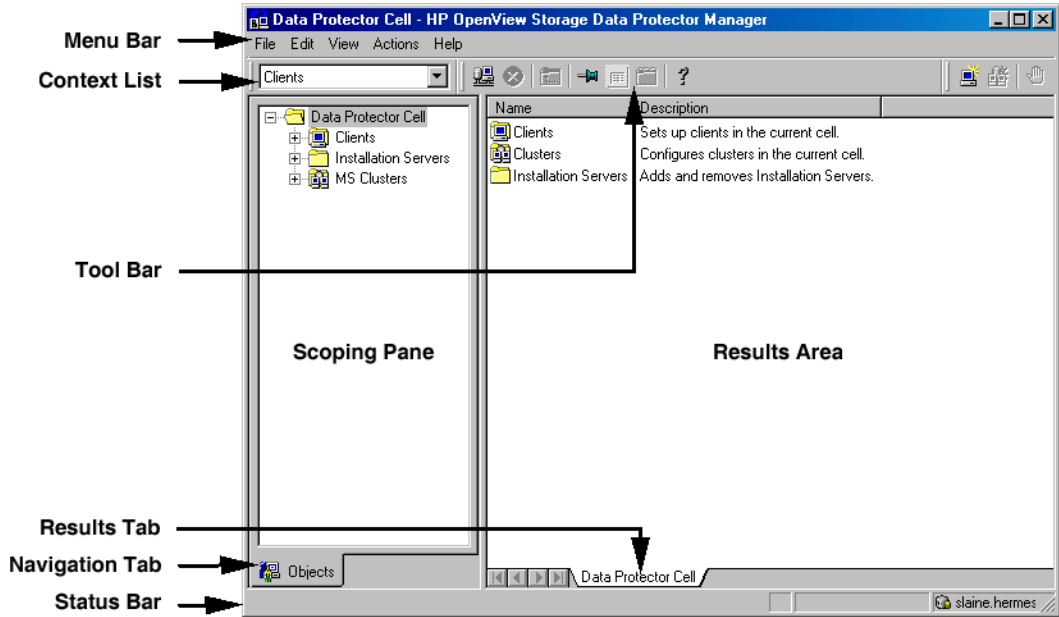
The following typographical conventions are used in this manual.

Table 2

Convention	Meaning	Example
<i>Italic</i>	Book or manual titles, and manual page names	Refer to the <i>HP OpenView Storage Data Protector Integration Guide</i> for more information.
	Provides emphasis	You <i>must</i> follow these steps.
	Specifies a variable that you must supply when entering a command	At the prompt type: rlogin <i>your_name</i> where you supply your login name.
Bold	New terms	The Data Protector Cell Manager is the main ...
Computer	Text and items on the computer screen	The system replies: Press Enter
	Command names	Use the grep command ...
	File and directory names	/usr/bin/X11
	Process names	Check to see if Data Protector Inet is running.
	Window/dialog box names	In the Backup Options dialog box...
	Text that you must enter	At the prompt, type: ls -l
Keycap	Keyboard keys	Press Return .

Data Protector provides a cross-platform (Windows and UNIX) graphical user interface. Refer to the online Help for information about the Data Protector graphical user interface.

Figure 1 Data Protector Graphical User Interface



Contact Information

General Information

General information about Data Protector can be found at

<http://www.hp.com/go/dataprotector>

Technical Support

Technical support information can be found at the HP Electronic Support Centers at

<http://support.openview.hp.com/support.jsp>

<http://www.hp.com/support>

Information about the latest Data Protector patches can be found at

http://support.openview.hp.com/patches/patch_index.jsp

For information on the Data Protector required patches, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

HP does not support third-party hardware and software. Contact the respective vendor for support.

Documentation Feedback

Your comments on the documentation help us to understand and meet your needs. You can provide feedback at

http://ovweb.external.hp.com/lpe/doc_serv/

Training Information

For information on currently available HP OpenView training, see the HP OpenView World Wide Web site at

<http://www.openview.hp.com/training/>

Follow the links to obtain information about scheduled classes, training at customer sites, and class registration.

Data Protector Documentation

Data Protector documentation comes in the form of manuals and online Help.

Manuals

Data Protector manuals are available in printed format and in PDF format. Install the PDF files during the Data Protector setup procedure by selecting the User Interface component on Windows or the OB2-DOCS component on UNIX. Once installed, the manuals reside in the `<Data_Protector_home>\docs` directory on Windows and in the `/opt/omni/doc/C/` directory on UNIX. You can also find the manuals in PDF format at http://ovweb.external.hp.com/lpe/doc_serv/

HP OpenView Storage Data Protector Concepts Guide

This manual describes Data Protector concepts and provides background information on how Data Protector works. It is intended to be used with the task-oriented online Help.

HP OpenView Storage Data Protector Installation and Licensing Guide

This manual describes how to install the Data Protector software, taking into account the operating system and architecture of your environment. This manual also gives details on how to upgrade Data Protector, as well as how to obtain the proper licenses for your environment.

HP OpenView Storage Data Protector Troubleshooting Guide

This manual describes how to troubleshoot problems you may encounter when using Data Protector.

HP OpenView Storage Data Protector Disaster Recovery Guide

This manual describes how to plan, prepare for, test and perform a disaster recovery.

HP OpenView Storage Data Protector Integration Guide

This manual describes how to configure and use Data Protector to back up and restore various databases and applications. It is intended for backup administrators or operators. There are four versions of this manual:

- *HP OpenView Storage Data Protector Integration Guide for Microsoft Applications: SQL Server, Exchange Server, and Volume Shadow Copy Service*

This manual describes the integrations of Data Protector with the following Microsoft applications: Microsoft Exchange Server 2000/2003, Microsoft SQL Server 7/2000/2005, and Volume Shadow Copy Service.

- *HP OpenView Storage Data Protector Integration Guide for Oracle and SAP*

This manual describes the integrations of Data Protector with Oracle, SAP R3, and SAP DB.

- *HP OpenView Storage Data Protector Integration Guide for IBM Applications: Informix, DB2, and Lotus Notes / Domino*

This manual describes the integrations of Data Protector with the following IBM applications: Informix Server, IBM DB2, and Lotus Notes/Domino Server.

- *HP OpenView Storage Data Protector Integration Guide for Sybase, Network Node Manager, and Network Data Management Protocol*

This manual describes the integrations of Data Protector with Sybase, Network Node Manager, and Network Data Management Protocol.

HP OpenView Storage Data Protector Integration Guide for HP OpenView

This manual describes how to install, configure, and use the integration of Data Protector with HP OpenView Service Information Portal, and HP OpenView Reporter. It is intended for backup administrators. It discusses how to use the OpenView applications for Data Protector service management.

HP OpenView Storage Data Protector Integration Guide for HP OpenView Operations for UNIX

This manual describes how to monitor and manage the health and performance of the Data Protector environment with HP OpenView Operations (OVO), HP OpenView Service Navigator, and HP OpenView Performance (OVP) on UNIX.

HP OpenView Storage Data Protector Integration Guide for HP OpenView Operations for Windows

This manual describes how to monitor and manage the health and performance of the Data Protector environment with HP OpenView Operations (OVO), HP OpenView Service Navigator, and HP OpenView Performance (OVP) on Windows.

There are two versions of the manual:

- for OVO 7.1x, 7.2x
- for OVO 7.5

HP OpenView Storage Data Protector Zero Downtime Backup Concepts Guide

This manual describes Data Protector zero downtime backup and instant recovery concepts and provides background information on how Data Protector works in a zero downtime backup environment. It is intended to be used with the task-oriented *HP OpenView Storage Data Protector Zero Downtime Backup Administrator's Guide* and the *HP OpenView Storage Data Protector Zero Downtime Backup Integration Guide*.

HP OpenView Storage Data Protector Zero Downtime Backup Administrator's Guide

This manual describes how to configure and use the integration of Data Protector with HP StorageWorks Virtual Array, HP StorageWorks Enterprise Virtual Array, EMC Symmetrix Remote Data Facility and TimeFinder, and HP StorageWorks Disk Array XP. It is intended for backup administrators or operators. It covers the zero downtime backup, instant recovery, and the restore of filesystems and disk images.

HP OpenView Storage Data Protector Zero Downtime Backup Integration Guide

This manual describes how to configure and use Data Protector to perform zero downtime backup, instant recovery, and standard restore of Oracle, SAP R/3, Microsoft Exchange Server 2000/2003, and Microsoft

SQL Server 2000 databases. The manual also describes how to configure and use Data Protector to perform backup and restore using the Microsoft Volume Shadow Copy Service.

HP OpenView Storage Data Protector MPE/iX System User Guide

This manual describes how to configure MPE/iX clients and how to back up and restore MPE/iX data.

HP OpenView Storage Data Protector Media Operations User's Guide

This manual provides tracking and management of offline storage media. It is intended for network administrators responsible for maintaining and backing up systems. It describes the tasks of installing and configuring the application, performing daily media operations and producing reports.

HP OpenView Storage Data Protector Product Announcements, Software Notes, and References

This manual gives a description of new features of HP OpenView Storage Data Protector A.06.00. It also provides information on supported configurations (devices, platforms and online database integrations, SAN, and ZDB), required patches, and limitations, as well as known problems and workarounds. An updated version of the supported configurations is available at http://www.openview.hp.com/products/datapro/spec_0001.html

There are also four other *Product Announcements, Software Notes and References*, which serve a similar purpose for the following:

- OVO UNIX integration
- OVO 7.1x/7.2x Windows integration
- OVO 7.5 Windows integration
- Media Operations

Online Help

Data Protector provides context-sensitive (F1) Help and Help Topics for Windows and UNIX platforms.

Documentation Map

Abbreviations

Abbreviations in the documentation map that follows are explained below. The manual titles are all preceded by the words “HP OpenView Storage Data Protector”

Abbreviation	Manual
CLI	Command Line Interface Reference Guide
Concepts	Concepts Guide
DR	Disaster Recovery Guide
GS	Getting Started Guide
Help	Online Help
IG-IBM	Integration Guide—IBM Applications
IG-MS	Integration Guide—Microsoft Applications
IG-O/S	Integration Guide—Oracle & SAP
IG-OV	Integration Guide—HP OpenView Service Information Portal/OpenView Reporter
IG-OVOU	Integration Guide—HP OpenView Operations, UNIX
IG-OVOW	Integration Guide—HP OpenView Operations 7.1x, 7.2x, Windows
IG-OVOW	Integration Guide—HP OpenView Operations 7.5, Windows
IG-Var	Integration Guide—Sybase, Network Node Manager & NDMP
Install	Installation and Licensing Guide
MO GS	Media Operations Getting Started Guide
MO RN	Media Operations Product Announcements, Software Notes, and References
MO UG	Media Operations User Guide
MPE/iX	MPE/iX System User Guide
PA	Product Announcements, Software Notes, and References

Abbreviation	Manual
Trouble	Troubleshooting Guide
ZDB Admin	ZDB Administrator's Guide
ZDB Concpt	ZDB Concepts Guide
ZDB IG	ZDB Integration Guide

Map

The following table shows where to find information of different kinds. Shaded squares are a good place to look first.

	Help	GS	Concepts				Integration Guides								ZDB			MO			MPE/iX	CLI		
			Install	Trouble	DR	PA	MS	O/S	IBM	Var	OV	OVOU	OVOW	Concept	Admin	IG	GS	User	PA					
Backup	X	X	X					X	X	X	X					X	X	X					X	
CLI																								X
Concepts/Techniques	X		X					X	X	X	X	X	X	X	X	X	X	X					X	
Disaster Recovery	X		X			X																		
Installation/Upgrade	X	X		X			X					X	X	X					X	X			X	
Instant Recovery	X		X												X	X	X							
Licensing	X			X			X													X				
Limitations	X				X		X	X	X	X	X			X			X					X		
New features	X						X															X		
Planning strategy	X		X									X												
Procedures/Tasks	X			X	X	X		X	X	X	X	X	X	X		X	X			X				
Recommendations			X				X								X							X		
Requirements				X			X	X	X	X	X			X				X	X	X				
Restore	X	X	X					X	X	X	X				X	X							X	
Support matrices							X																	
Supported configurations															X									
Troubleshooting	X			X	X			X	X	X	X	X				X	X							

Integrations

Look in these manuals for details of the following integrations:

Integration	Guide
HP OpenView Operations (OVO)	IG-OVOU, IG-OVOW
HP OpenView Reporter (OVR)	IG-OV
HP OpenView Reporter Light	IG-OVOW
HP OpenView Service Information Portal (OVSIP)	IG-OV
HP StorageWorks Disk Array XP	all ZDB
HP StorageWorks Enterprise Virtual Array (EVA)	all ZDB
HP StorageWorks Virtual Array (VA)	all ZDB
IBM DB2 UDB	IG-IBM
Informix	IG-IBM
Lotus Notes/Domino	IG-IBM
Media Operations	MO User
MPE/iX System	MPE/iX
Microsoft Exchange Servers	IG-MS, ZDB IG
Microsoft Exchange Single Mailbox	IG-MS
Microsoft SQL Servers	IG-MS, ZDB IG
Microsoft Volume Shadow Copy Service (VSS)	IG-MS, ZDB IG
NDMP Server	IG-Var
Network Node Manager (NNM)	IG-Var
Oracle	IG-O/S
Oracle ZDB	ZDB IG
SAP DB	IG-O/S
SAP R/3	IG-O/S, ZDB IG
Sybase	IG-Var
Symmetrix (EMC)	all ZDB

In This Book

The *HP OpenView Storage Data Protector Concepts Guide* describes Data Protector concepts. Read this manual to fully understand the fundamentals and the model of Data Protector.

Audience

The manual is intended for users interested in understanding the concepts of Data Protector operation and for people who plan company backup strategies. Depending on the required level of detail, you can also use this manual together with the Data Protector online Help.

Organization

The manual is organized as follows:

Chapter 1	“About Backup and Data Protector” on page 1.
Chapter 2	“Planning Your Backup Strategy” on page 27.
Chapter 3	“Media Management and Devices” on page 109.
Chapter 4	“Users and User Groups” on page 163.
Chapter 5	“The Data Protector Internal Database” on page 169.
Chapter 6	“Service Management” on page 189.
Chapter 7	“How Data Protector Operates” on page 209.
Chapter 8	“Integration with Database Applications” on page 231.
Chapter 9	“Direct Backup” on page 241.
Chapter 10	“Disk Backup” on page 253.
Chapter 11	“Split Mirror Concepts” on page 271.
Chapter 12	“Snapshot Concepts” on page 283.
Chapter 13	“Microsoft Volume Shadow Copy Service” on page 299.
Appendix A	“Backup Scenarios” on page A-1.
Appendix B	“Further Information” on page B-1.
Glossary	Definition of terms used in this manual.

1 About Backup and Data Protector

In This Chapter

This chapter provides an overview of backup and restore concepts. It introduces Data Protector architecture, media management, user interfaces, backup devices, and other features. The chapter concludes with an overview of Data Protector configuration and other tasks needed to set up Data Protector.

It is organized as follows:

“About Data Protector” on page 3

“Introducing Backups and Restores” on page 7

“Data Protector Architecture” on page 10

“Enterprise Environments” on page 15

“Media Management” on page 19

“Backup Devices” on page 20

“User Interfaces” on page 22

“Overview of Tasks to Set Up Data Protector” on page 24

About Data Protector

HP OpenView Storage Data Protector is a backup solution that provides reliable data protection and high accessibility for your fast growing business data. Data Protector offers comprehensive backup and restore functionality specifically tailored for enterprise-wide and distributed environments. The following list describes major Data Protector features:

- **Scalable and Highly Flexible Architecture**

Data Protector can be used in environments ranging from a single system to thousands of systems on several sites. Due to the network component concept of Data Protector, elements of the backup infrastructure can be placed in the topology according to user requirements. The numerous backup options and alternatives to setting up a backup infrastructure allow the implementation of virtually any configuration you want. Data Protector also enables the use of advanced backup concepts, such as synthetic backup and disk staging.

- **Easy Central Administration**

Through its easy-to-use graphical user interface (GUI), Data Protector allows you to administer your complete backup environment from a single system. To ease operation, the GUI can be installed on various systems to allow multiple administrators to access Data Protector via their locally installed consoles. Even multiple backup environments can be managed from a single system. The Data Protector command-line interface allows you to manage Data Protector using scripts.

- **High Performance Backup**

Data Protector enables you to perform backup to several hundred backup devices simultaneously. It supports high-end devices in very large libraries. Various backup possibilities, such as local backup, network backup, online backup, disk image backup, synthetic backup, backup with object mirroring, and built-in support for parallel data streams allow you to tune your backups to best fit your requirements.

- **Supporting Mixed Environments**

As Data Protector supports heterogeneous environments, most features are common to the UNIX and Windows platforms. The UNIX and Windows Cell Managers can control all supported client platforms (UNIX, Windows, and Novell NetWare). The Data Protector user interface can access the entire Data Protector functionality on all supported platforms.

- **Easy Installation for Mixed Environments**

The Installation Server concept simplifies the installation and upgrade procedures. To remotely install UNIX clients, you need an Installation Server for UNIX. To remotely install Windows clients, you need an Installation Server for Windows. The remote installation can be performed from any client with an installed Data Protector GUI. For supported platforms for the Installation Server, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

- **High Availability Support**

Data Protector enables you to meet the needs for continued business operations around the clock. In today's globally distributed business environment, company-wide information resources and customer service applications must always be available. Data Protector enables you to meet high availability needs by:

- Integrating with clusters to ensure fail-safe operation with the ability to back up virtual nodes. For a list of supported clusters, see *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.
- Enabling the Data Protector Cell Manager itself to run on a cluster.
- Supporting all popular online database Application Programming Interfaces.
- Integrating with advanced high availability solutions like EMC Symmetrix, HP StorageWorks Disk Array XP, HP StorageWorks Virtual Array, or HP StorageWorks Enterprise Virtual Array.
- Providing various disaster recovery methods for Windows and UNIX platforms.

— Offering methods of duplicating backed up data during and after the backup to improve fault tolerance of backups or for redundancy purposes.

- **Easy Restore**

Data Protector includes an internal database that keeps track of data such as which files from which system are kept on a particular medium. In order to restore any part of a system, simply browse the files and directories. This provides fast and convenient access to the data to be restored.

- **Automated or Unattended Operation**

With the internal database, Data Protector keeps information about each Data Protector medium and the data on it. Data Protector provides sophisticated media management functionality. For example, it keeps track of how long a particular backup needs to remain available for restoring, and which media can be (re)used for backups.

The support of very large libraries complements this, allowing for unattended operation over several days or weeks (automated media rotation).

Additionally, when new disks are connected to systems, Data Protector can automatically detect (or discover) the disks and back them up. This eliminates the need to adjust backup configurations manually.

- **Service Management**

Data Protector is the first backup and restore management solution to support service management. The integration with Application Response Management (ARM) and Data Source Integration (DSI) enables powerful support of Service Level Management (SLM) and Service Level Agreements (SLA) concepts by providing relevant data to management and planning systems.

The DSI integration provides a set of scripts and configuration files from which users are able to see how to add their own queries using Data Protector reporting capabilities.

- **Monitoring, Reporting and Notification**

Superior web reporting and notification capabilities allow you to easily view the backup status, monitor active backup operations, and customize reports. Reports can be generated using the Data Protector GUI, or using the `omnirpt` command on systems running UNIX or Windows, as well as using Java-based online generated web reports.

You can schedule reports to be issued at a specific time or to be attached to a predefined set of events, such as the end of a backup session or a mount request.

- **Integration with Online Database Applications**

Data Protector provides online backup of Microsoft Exchange Server, Microsoft SQL Server, Oracle, Informix Server, SAP R/3, Lotus Notes/Domino Server, IBM DB2 UDB, and Sybase database objects. For a list of supported versions for a particular operating system, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

- **Integration with Other Products**

Additionally, Data Protector integrates with EMC Symmetrix, Microsoft Cluster Server, MC/ServiceGuard and other products.

For detailed documentation describing the features of Data Protector, including integrations, as well as the latest platform and integration support information, consult the HP OpenView Storage Data Protector home page at

http://www.openview.hp.com/products/datapro/spec_0001.html.

Introducing Backups and Restores

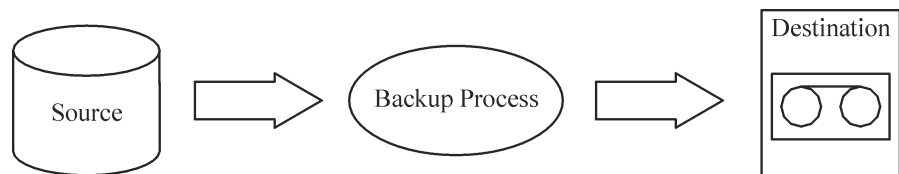
This section explains basic backup and restore concepts.

What Is a Backup?

A backup is a process that creates a copy of data on backup media. This copy is stored and kept for future use in case the original is destroyed or corrupted.

A high-level presentation of a backup is shown in Figure 1-1.

Figure 1-1 Backup Process



In most cases, the **source** is data on a disk, such as files, directories, databases, and applications. If the backup is expected to be used for disaster recovery, it needs to be consistent.

Software that actually copies data to the destination is a backup application. The **destination** is a backup device, such as a tape drive, with media to which a copy of the data is written.

What Is a Restore?

A restore is a process that recreates the original data from a backup copy. This process consists of the preparation and actual restore of data, and some post-restore actions that make that data ready for use.

Figure 1-2 **Restore Process**

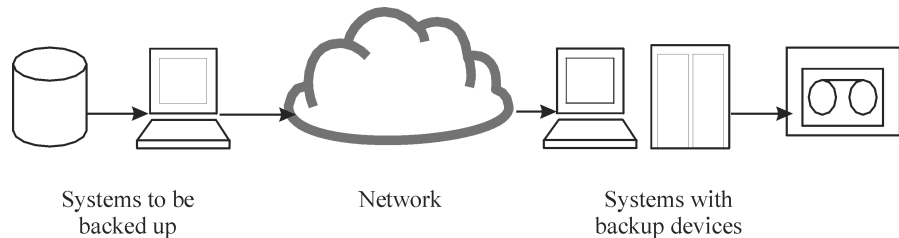


The **source** is a backup copy. A restore application is software that actually writes data to a destination. The **destination** is usually a disk to which the original data is written.

Backing Up a Network Environment

During backups in a network environment, data is transferred over the network from systems to be backed up to media on systems with backup devices, where the data is stored.

Figure 1-3 **Network Backup**



To accomplish backup of a network environment you need an application that allows you to:

- Attach backup devices to any system in the network

This enables local backups of systems with large volumes of data and network backups in order to reduce backup device costs.

- Route backup data flow to any network path
- Route backup data away from the LAN and onto a SAN when data volume or network traffic makes LAN transfer inefficient
- Manage backup activities from any system

- Integrate into the IT management framework
- Support many different types of systems to be backed up

Direct Backup

A **direct** backup is one in which you can send data directly from disk to tape in the SAN without involving a dedicated backup server for the data movement.

The use of a filesystem-independent ability to resolve data is fully integrated with the industry-standard XCOPY functionality that is embedded in supported disk arrays and bridges, removing the need for a separate data mover appliance.

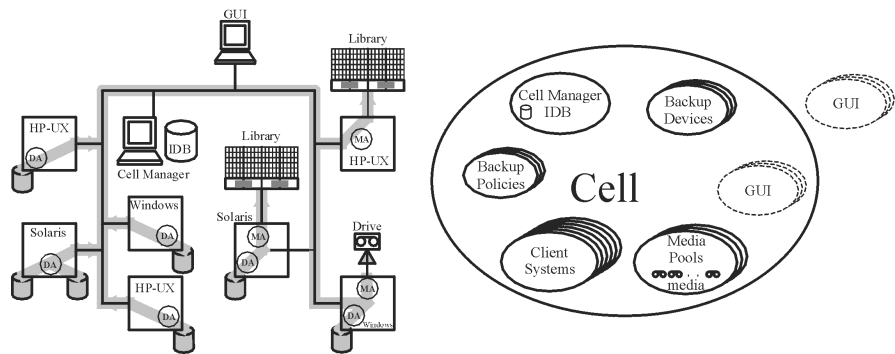
Data Protector Architecture

The Data Protector **cell**, shown in Figure 1-4, is a network environment that has a **Cell Manager**, **client systems**, and **devices**. The Cell Manager is the central control point where Data Protector software is installed. After installing Data Protector software, you can add systems to be backed up. These systems become Data Protector client systems that are part of the cell. When Data Protector backs up files, it saves them to media in backup devices.

The **Data Protector internal database (IDB)** keeps track of the files you back up so that you can browse and easily recover the entire system or single files.

Data Protector facilitates backup and restore jobs. You can do an immediate (or interactive) backup using the Data Protector user interface. You can also schedule your backups to run unattended.

Figure 1-4 The Data Protector Cell (Physical View and Logical View)



NOTE

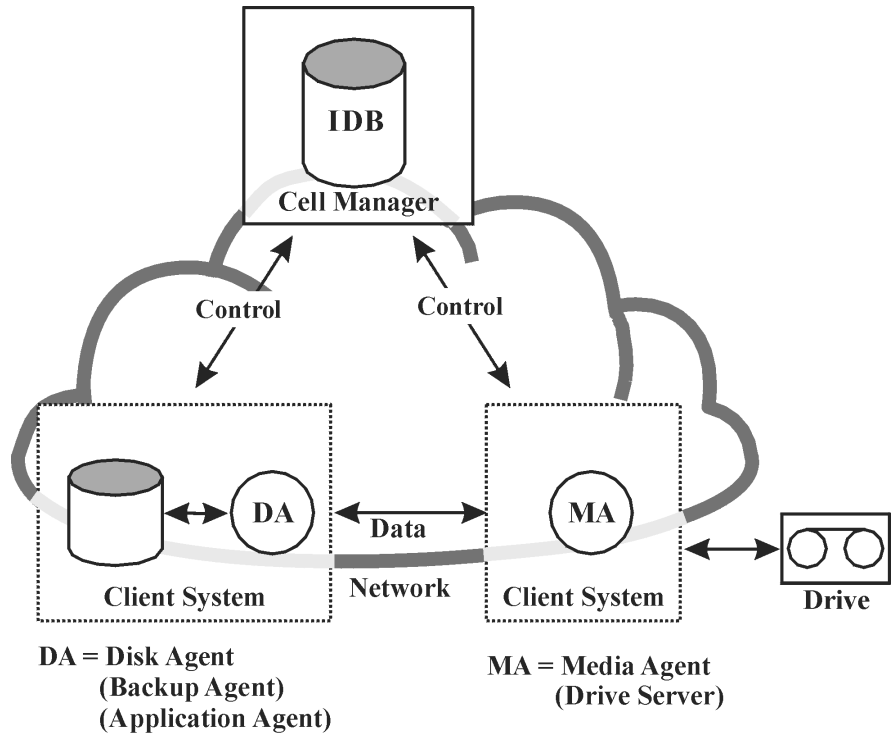
The GUI and the Cell Manager systems can run on UNIX and Windows operating systems; they do not have to run the same operating system. For a list of supported operating systems for a particular Data Protector component, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

- Cell Manager** The Cell Manager is the main system in the cell. The Cell Manager:
- Manages the cell from a central point
 - Contains the IDB
- The IDB contains information about backup details such as, backup durations, media IDs, and sessionIDs
- Runs core Data Protector software
 - Runs Session Managers that start and stop backup and restore sessions and write session information to the IDB
- Systems to Be Backed Up** Client systems you want to back up must have the Data Protector Disk Agent (DA), also called **Backup Agent**, installed. To back up online database integrations, install the **Application Agent**. In the rest of the manual, the term Disk Agent will be used for both agents. The Disk Agent reads or writes data from a disk on the system and sends or receives data from a Media Agent. The Disk Agent is also installed on the Cell Manager, thus allowing you to back up data on the Cell Manager, the Data Protector configuration, and the IDB.
- Systems with Backup Devices** Client systems with connected backup devices must have a Data Protector **Media Agent** (MA) installed. Such client systems are also called **Drive Servers**. A backup device can be connected to any system and not only to the Cell Manager. A Media Agent reads or writes data from or to media in the device and sends or receives data from the Disk Agent.
- Systems with a User Interface** You can manage Data Protector from any system on the network on which the Data Protector graphical user interface (GUI) is installed. Therefore, you can have the Cell Manager system in a computer room while managing Data Protector from your desktop system.
- Installation Server** The **Installation Server** holds a repository of the Data Protector software packages for a specific architecture. The Cell Manager is by default also an Installation Server. At least two Installation Servers are needed for mixed environments: one for UNIX systems and one for Windows systems.

Operations in the Cell

The Data Protector Cell Manager controls backup and restore sessions, which perform all the required actions for a backup or restore, respectively, as shown in Figure 1-5.

Figure 1-5 Backup or Restore Operation



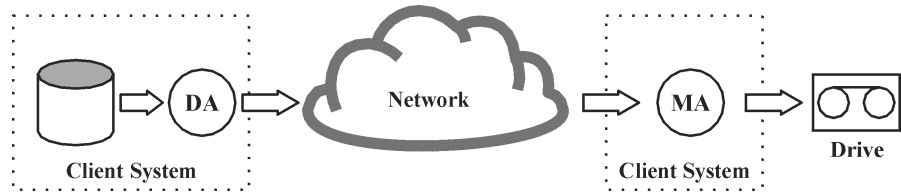
Backup Sessions

What Is a Backup Session?

A backup session, shown in Figure 1-6, is a process that creates a copy of data on storage media. It is started either interactively by an operator using the Data Protector user interface, or unattended using the Data Protector Scheduler.

How Does It Work? The Backup Session Manager process starts Media Agent(s) and Disk Agent(s), controls the session, and stores generated messages to the IDB. Data is read by the Disk Agent and sent to a Media Agent, which saves it to media.

Figure 1-6 Backup Session



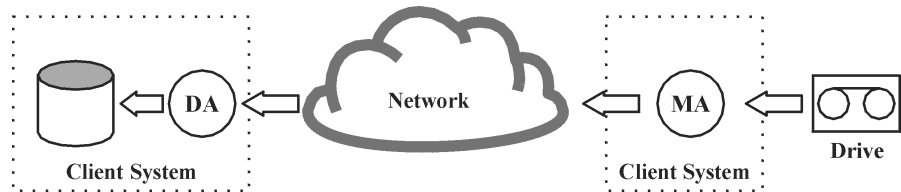
A typical backup session is more complex than the one shown in Figure 1-6. A number of Disk Agents read data from multiple disks in parallel and send data to one or more Media Agents. For more information on complex backup sessions, refer to Chapter 7, “How Data Protector Operates,” on page 209.

Restore Sessions

What Is a Restore Session? A restore session, shown in Figure 1-7, is a process that restores data from previous backups to a disk. The restore session is interactively started by an operator using the Data Protector user interface.

How Does It Work? After you have selected the files to be restored from a previous backup, you invoke the actual restore. The Restore Session Manager process starts the needed Media Agent(s) and Disk Agent(s), controls the session, and stores messages in the IDB. Data is read by a Media Agent and sent to the Disk Agent, which writes it to disks.

Figure 1-7 Restore Session



A restore session may be more complex than the one shown in Figure 1-7. Refer to Chapter 7, “How Data Protector Operates,” on page 209 for more information on restore sessions.

Enterprise Environments

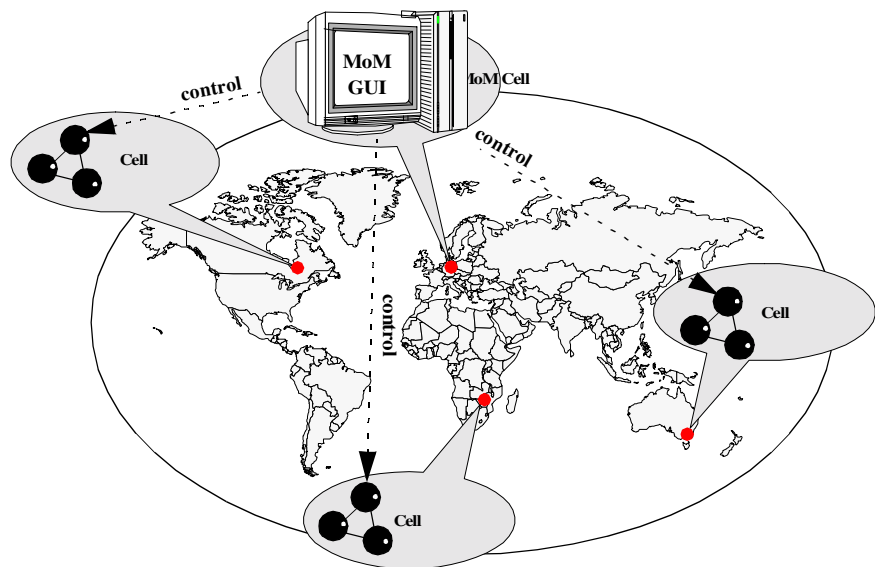
What Is an Enterprise Environment?

A typical enterprise network environment, shown in Figure 1-8, consists of a number of systems from different vendors with different operating systems. The systems may be located in different geographical areas and time zones. All the systems are connected with LAN or WAN networks operating at various communication speeds.

When to Use an Enterprise Environment

This solution can be used when several geographically separated sites require common **backup policies** to be used. It can also be used when all departments at the same site want to share the same set of backup devices.

Figure 1-8 Large Data Protector Enterprise Environment



Configuring and managing backups of such a heterogeneous environment is challenging. Data Protector functionality has been designed to highly simplify this task. For information about the Manager of Managers (MoM), refer to “MoM” on page 16.

Splitting an Environment into Multiple Cells

You may decide to split large environments into multiple cells for a number of reasons:

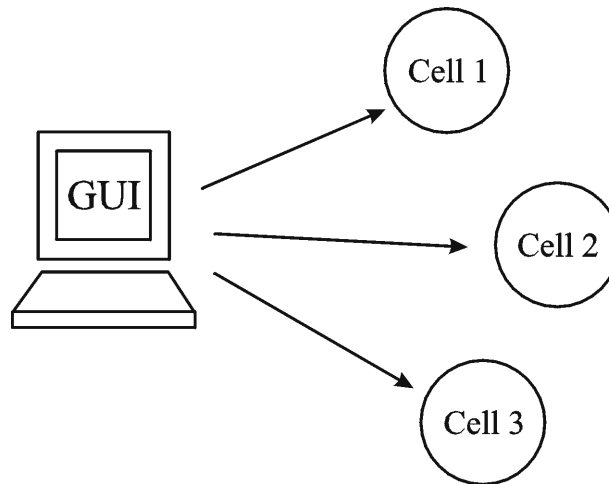
Why Split Large Environments into Multiple Cells?

- Geographical grouping of systems.
- Logical grouping of systems, for example, departments.
- Slow network connection between some systems.
- Performance considerations.
- Separate administrative control.

Refer to Chapter 2, “Planning Your Backup Strategy,” on page 27 for a list of considerations in planning your environment.

Data Protector allows you for managing multiple cells from a single point.

Figure 1-9 Single-Point Management of Multiple Cells

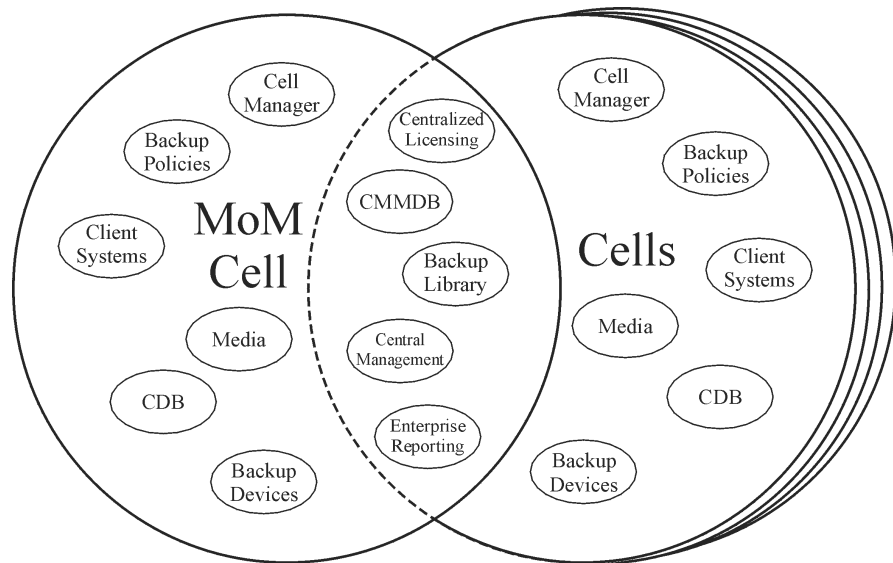


MoM

Data Protector provides the Manager-of-Managers to manage large environments with multiple cells. The MoM allows you to group multiple cells into a larger unit, called a MoM environment that can be managed from a single point, as shown in Figure 1-9. The MoM enables virtually unlimited growth of your backup environment. New cells can be added or existing ones split.

A MoM environment does not require a reliable network connection from Data Protector cells to the central MoM cell, because only controls are sent over the long distance connections, and backups are performed locally within each Data Protector cell. However, this is based on the assumption that each cell has its own Media Management Database.

Figure 1-10 **Manager-of-Managers Environment**



Manager-of-Managers provides the following features:

- **Centralized licensing repository**

This enables simplified license management. This is optional but useful for very large environments.

- **Centralized Media Management Database (CMMDB)**

The CMMDB allows you to share devices and media across several cells in a MoM environment. This makes devices of one cell (using the CMMDB) accessible to other cells that use the CMMDB. The CMMDB, if used, must reside in the MoM cell. In this case, a reliable network connection is required between the MoM cell and the other Data Protector cells. Note that it is optional to centralize the Media Management Database.

Enterprise Environments

- **Sharing libraries**

With the CMMDB, you can share high-end devices between cells in the multi-cell environment. One cell can control the robotics, serving several devices that are connected to systems in different cells. Even the Disk Agent to Media Agent data path can go across cell boundaries.

- **Enterprise reporting**

The Data Protector Manager-of-Managers can generate reports on a single-cell basis as well as for the entire enterprise environment.

Media Management

Data Protector provides you with powerful media management, which lets you easily and efficiently manage large numbers of media in your environment in the following ways:

Media Management Functionality

- Grouping media into logical groups, called **media pools**, which allows you to think about large sets of media without having to worry about each medium individually.
- Data Protector keeps track of all media and the status of each medium, data protection expiration time, availability of media for backup, and a catalog of what has been backed up to each medium.
- Fully automated operation. If Data Protector controls enough media in the library devices, the media management functionality lets you run the backup sessions without operator intervention.
- Automated media rotation policies that allow media selection for backups to be performed automatically.
- Recognition and support of barcodes on large library devices and silo devices with barcode support.
- Recognition, tracking, viewing, and handling of media used by Data Protector in large library devices and silo devices.
- The possibility of having information about the media in a central place and the sharing of this information among several Data Protector cells.
- Interactive or automated creation of additional copies of the data on the media.
- Support for media vaulting.

What Is a Media Pool?

Data Protector uses media pools to manage large numbers of media. A media pool is a logical collection of media of the *same* physical type with common usage policies (properties). Usage is based on the data on the media. The structure and quantity of the pools and which pool contains what type of data on its media, depends entirely on your preferences.

When a device is configured, a default media pool is specified. This media pool is used if no other media pool is defined in the backup specification.

Backup Devices

Data Protector defines and models each device as a physical device with its own usage properties, such as the default pool.

This device concept is used because it allows you to easily and flexibly configure devices and use them in conjunction with backup specifications. The definition of the devices is stored in the Data Protector Media Management Database.

Figure 1-11 How Backup Specifications, Devices, and Media Pools Are Related

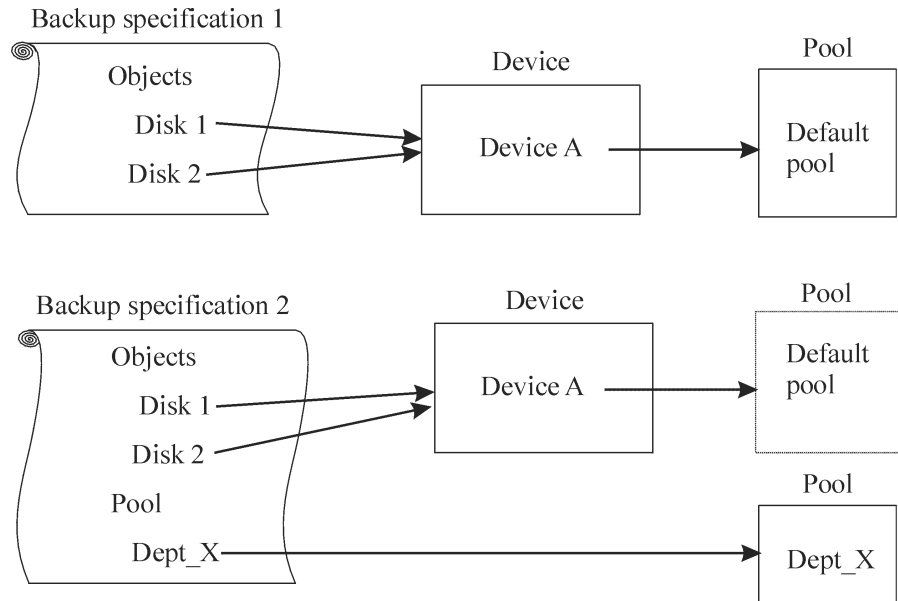


Figure 1-11 shows the relationship among the backup specification, devices, and media pools. The devices are referred to in the backup specification. Each device is linked to a media pool; this media pool can be changed in the backup specification. For example, backup specification 2 references the pool Dept_X instead of the default pool.

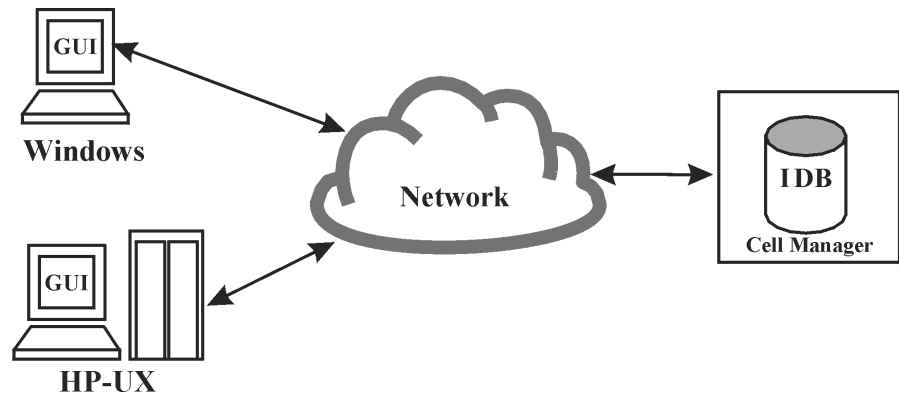
Data Protector supports various devices. Refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References* for more information.

User Interfaces

Data Protector provides easy access to all configuration and administration tasks using the Data Protector GUI provided to run under X11/Motif on UNIX platforms and on the Windows platforms. Additionally, a command-line interface is available on UNIX and Windows platforms.

The Data Protector architecture allows you to flexibly install and use the Data Protector user interface. The user interface does not have to be used from the Cell Manager system; you can install it on your desktop system. As depicted in Figure 1-12, the user interface also allows you to transparently manage Data Protector cells with HP-UX, Solaris or Windows Cell Managers.

Figure 1-12 Using the Data Protector User Interface



TIP

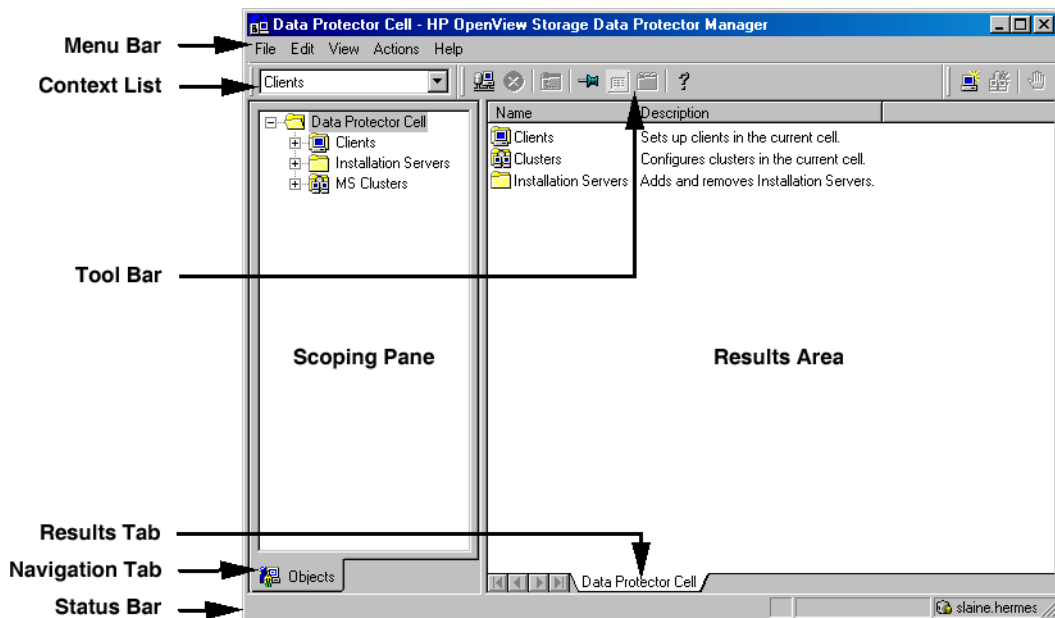
In a typical mixed environment, install the Data Protector user interface on several systems in the environment, thus providing access to Data Protector from several systems.

Data Protector GUI

The Data Protector GUI, depicted in Figure 1-13, is an easy-to-use, powerful interface providing the following functionality:

- A Results Tab with all the configuration wizards, properties and lists.
- Easy configuration and management of the backup of online database applications that run in Windows environments, such as Microsoft SQL Server, Microsoft Exchange Server, SAP R/3, and Oracle or that run in the UNIX environments, such as SAP R/3, Oracle, and Informix Server.
- A comprehensive online Help system called the Help Topics, and context-sensitive Help called the Help Navigator.

Figure 1-13 Data Protector GUI



Overview of Tasks to Set Up Data Protector

This section provides an overview of global tasks to set up your Data Protector backup environment. Depending on the size and complexity of your environment, you may not need to go through all these steps.

1. Analyze your network and organizational structure. Decide which systems need to be backed up.
2. Check if there are any special applications and databases which you want to back up, such as Microsoft Exchange, Oracle, IBM DB2 UDB, SAP R/3, or others. Data Protector provides specific integrations with these products.
3. Decide on the configuration of your Data Protector cell, such as:
 - the system to be your Cell Manager
 - systems on which you want to install the user interface
 - local backup versus network backup
 - systems to control backup devices and libraries
 - type of connections, LAN and/or SAN

4. Purchase the required Data Protector licenses for your setup. This way you obtain the passwords you will need to install.

Alternatively, you can operate Data Protector using an instant-on password. However, this is valid only for 60 days from the date of installation. Refer to the *HP OpenView Storage Data Protector Installation and Licensing Guide* for details.

5. Consider security aspects:
 - Analyze security considerations. Refer to the *HP OpenView Storage Data Protector Installation and Licensing Guide*.
 - Consider which user groups you need to configure.
6. Decide how you want to structure your backups:
 - Which media pools do you want to have, and how will they be used?
 - Which devices will be used, and how?

- How many copies of each backup do you want?
 - How many backup specifications do you need, and how should they be grouped?
 - If you are planning to back up to disk, consider advanced backup strategies such as synthetic backup and disk staging.
7. Install and configure your Data Protector environment.
 - Install the Data Protector Cell Manager system and use the Data Protector user interface to distribute Data Protector components to other systems.
 - Connect devices (tape drives) to the systems that will control them.
 - Configure backup devices.
 - Configure media pools and prepare the media.
 - Configure backup specifications, including backup of the IDB.
 - Configure reports, if needed.
 8. Become familiar with tasks such as:
 - Handling failed backups
 - Performing restores
 - Duplicating backed up data and vaulting media
 - Preparing for disaster recovery
 - Maintaining the IDB

About Backup and Data Protector
Overview of Tasks to Set Up Data Protector

2 Planning Your Backup Strategy

In This Chapter

This chapter describes backup strategy planning. It focuses on planning Data Protector cells, performance, and security, as well as backing up and restoring data. The chapter also discusses basic backup types, automated backup operation, clustering, and disaster recovery.

It is organized as follows:

“Backup Strategy Planning” on page 29

“Planning Cells” on page 35

“Understanding and Planning Performance” on page 41

“Planning Security” on page 48

“Clustering” on page 52

“Full and Incremental Backups” on page 66

“Keeping Backed Up Data and Information About the Data” on page 73

“Backing Up Data” on page 77

“Automated or Unattended Operation” on page 87

“Duplicating Backed Up Data” on page 89

“Restoring Data” on page 100

“Disaster Recovery” on page 104

Backup Strategy Planning

Data Protector is simple to configure and administer. However, if you work in a large environment with diverse client systems and huge amounts of data to back up, plan in advance. Planning simplifies subsequent configuration steps.

What Is Backup Strategy Planning?

Backup strategy planning is a process that includes the following steps:

1. Defining the requirements and constraints for backups, for example, how often your data needs to be backed up or whether you need additional copies of the backed up data on additional media sets.
2. Understanding the factors that influence your backup solution, such as the sustained data transfer rates of the network and of backup devices. These factors can affect how you configure Data Protector and the kind of backup -- network or direct, for example -- that you choose. For instance, if you back up to disk, you can take advantage of advanced backup strategies such as synthetic backup and disk staging.
3. Preparing the backup strategy that shows your backup concept and how it is implemented.

This section provides detailed information on the preceding steps. The rest of this guide provides important information and considerations that help you plan your backup solution.

Defining the Requirements of a Backup Strategy

Defining objectives and constraints of your backup strategy includes answering questions, such as:

- What are your **organizational policies** regarding backups and restores?

Some organizations already have defined policies on archiving and storing data. Your backup strategy should comply with these policies.

- What types of data need to be backed up?

List all types of data existing in your network, such as user files, system files, Web servers, and large relational databases.

- How long is the maximum downtime for recovery?

The allowed downtime has a significant impact on the investments into network infrastructure and equipment needed for backups. For each type of data, list the maximum acceptable downtime for recovery, that is, how long specific data can be unavailable before recovered from a backup. For example, user files may be restored in two days, while some business data in a large database would need to be recovered in two hours.

Recovery time consists mainly of the time needed to access the media and the time required to actually restore data to disks. A full system recovery takes more time, because some additional steps are required. Refer to “Disaster Recovery” on page 104 for more information.

- How long should specific types of data be kept?

For each type of data, list how long the data must be kept. For example, you may only need to keep user files for three weeks, while information about company employees may be kept for five years.

- How should media with backed up data be stored and maintained?

For each type of data, list how long the media with data must be kept in a vault, a safe, external location, if you use one. For example, user files may not be stored in a vault at all, while order information may be kept for five years, with verification of each medium after two years.

- To how many media sets should the data be written during backup?

Consider writing critical data to several media sets during backup to improve the fault tolerance of such backups, or to enable multi-site vaulting. Object mirroring increases the time needed for backup.

- How much data needs to be backed up?

List the estimated amount of data to be backed up, for each type of data. This influences the time needed for backup and helps you to choose the right backup devices and media for backup.

- What is the projected future growth of the amount of data?

Estimate future growth, for each type of data. This will help you to come up with backup solutions that will not be quickly outdated. For example, if your company plans to hire 100 new employees, the amount of users’ data and client systems’ data will grow accordingly.

- How long can a backup take?

Estimate the time needed for each backup. This directly affects the amount of time data is available for use. User files can be backed up at any time when the users are not working on them, while some transactional databases may only have a few hours available for backup.

The time needed for backup depends on the type of backup, full or incremental. Refer to “Full and Incremental Backups” on page 66 for more information. Data Protector also backs up some popular online database applications. Refer to the *HP OpenView Storage Data Protector Integration Guide* for more information.

If you back up to disk, you can take advantage of synthetic backup and disk staging. These advanced backup strategies significantly reduce the time needed for backup. For more information, see Chapter 11, “Synthetic Backup,” on page 261 and “Disk Staging” on page 94.

When there is a very fast and large disk to be backed up on a slower device, consider the possibility of backing up one hard disk through multiple concurrent Disk Agents. Starting multiple Disk Agents on the same disk speeds up the backup performance considerably.

Also, if there is a large volume of information to be backed up and the available time to complete the back up is limited, consider the possibility of using direct backup to take advantage of SAN speed, reduced network traffic, and the lack of the backup-server bottleneck.

- How often does data need to be backed up?

For each type of data, list how often the data needs to be backed up. For example, user working files may be backed up on a daily basis, system data on a weekly basis, and some database transactions twice a day.

Factors Influencing Your Backup Strategy

There are a number of factors that influence how your backup strategy is implemented. Understand these factors before preparing your backup strategy.

- Your company’s backup and storage policies and requirements.
- Your company’s security policies and requirements.

- Your physical network configuration.
- Computer and human resources available at different sites of your company.

Preparing a Backup Strategy Plan

The result of the planning is a backup strategy that must address the following areas:

- How critical system availability (and backup) is to the company
 - The need to keep the backed up data at a remote location in case of a disaster.
 - The level of business continuance

This includes the recovery and restore plan for all critical client systems.
 - The security of backed up data

The need to guard premises to prevent unauthorized people from entering. This also includes safeguarding all relevant data against unauthorized access, using physical access prevention and electronic password protection.
- Types of data that need to be backed up

List the company's types of data and how you want to combine them in backup specifications, including the time frames available for backups. The company's data can be divided into categories like company business data, company resource data, project data, and personal data, each with its own specific requirements.
- Backup policy implementation
 - How backups are done and the backup options that you use

This defines the frequency of full and incremental backups. It also defines the backup options that are used and whether the backups are permanently protected and the backup media stored at a security company.
 - How the client systems are grouped into backup specifications

Consider how best to group backup specifications. This can be done on the basis of departments, data types, or backup frequency.

- How the backups are scheduled

Consider using the staggered approach, whereby full backups are scheduled for different clients (backup specifications) on different days to avoid network load, device load, and time window issues.

- Retaining data on media, and information about backups

Consider protecting data from being overwritten by newer backups for a specified amount of time. This protection, called data protection, is on a session basis.

Define the period of time the Catalog Database should store information about backup versions, the number of backed up files and directories, and messages stored in the database. For as long as this catalog protection has not expired, backed up data is easily accessible.

- Device configuration

Determine devices to use for backups, and the client systems they are connected to. Connect the backup devices to client systems with the largest amount of data, so that as much data as possible is backed up locally and not via the network. This increases the backup speed.

If you need to back up large amounts of data:

- Consider using a library device.
- Consider backing up to a disk-based device. Besides other benefits, backup to disk reduces the time needed for backup and enables the use of advanced backup strategies such as synthetic backup and disk staging.
- Consider configuring your system for direct backup by attaching a library device to the SAN through a fibre channel bridge. This is a solution when the network impairs the backup speed.

- Media management

Determine the type of media to use, how to group the media into media pools, and how to place objects on the media.

Define how media are used for backup policies.

- Vaulting

Decide whether to store media at a safe place (a vault), where they are kept for a specific period of time. Consider duplicating backed up data during or after the backup for this purpose.

Backup Strategy Planning

- Backup administrators and operators

Determine the rights of users that can administer and operate your storage product.

Planning Cells

One of the most important decisions in planning your backup strategy is whether you want to have a single or multiple cell environment. This section describes the following:

- Factors you should consider when planning cells
- How cells relate to a typical network environment
- How cells relate to Windows domains
- How cells relate to Windows workgroup environments

One Cell or Multiple Cells?

When deciding whether to have a single cell or multiple cells in your environment, consider the following items:

- Backup administration issues

The use of multiple cells gives you higher administration freedom within each cell. You can apply completely independent media management policies for each cell. If you have several administrative groups, you may, for security reasons, not want a cell to span across these groups. A disadvantage of having multiple cells is that it can require more administrative work or might even require a separate administrator for each cell.

- Size of each cell

The size of a Data Protector cell affects backup performance and the ability to manage the cell. The recommended maximum size for a Data Protector cell is 100 client systems. Cells with more than 200 client systems are less manageable.

- Network considerations

All client systems of a cell should be on the same LAN for maximum performance. Refer to the sections that follow for more information about other network considerations such as your network configuration.

- Geographical location

If the client systems you want to back up are geographically dispersed, it may be difficult to manage them from a single cell and there may be networking problems between the client systems. Additionally, the security of data may be an issue.

- Time Zones

Each cell should be within one time zone.

- Security of data

Data Protector provides cell level based security. All Data Protector administrative work is done in the context of a single cell: media, backup devices, and backed up data belong to one cell. Note that Data Protector lets you share devices or move media between cells, so physical access to media must be limited to authorized personnel.

- Mixed environments

Data Protector allows you to back up client systems of diverse platforms in a single cell. However, it may be convenient to group client systems in a cell based on the platforms. For example, you may have one cell with the Windows client systems and one with the UNIX client systems. This is especially useful if you have separate administrators and policies for the UNIX and Windows environments.

- Departments and sites

You can group each department or site in a separate cell. For example, you may have one cell for the accounting, one for the IT, and one for the manufacturing department. Even if you choose to have several cells, Data Protector allows you to easily configure common policies among the cells.

Installing and Maintaining Client Systems

If you have several UNIX and Windows client systems, an efficient mechanism for the installation of Data Protector becomes important. Local installation on every client is not feasible in large environments.

Installation Servers and the Cell Manager

The main system in a Data Protector cell is the Cell Manager. To conveniently distribute (push) Data Protector components to client systems from a central location, a system holding the Data Protector

software repository is needed. This system is called the Data Protector Installation Server. The Cell Manager is by default also an Installation Server.

Each time you perform a remote installation, you access the Installation Server. The advantage of using Installation Servers is that the time required for remote installation, update, upgrade, and uninstallation of Data Protector software is greatly reduced, especially in enterprise environments.

There are certain hardware and software requirements that need to be met by Installation Servers and Cell Managers before you start installing the software. A dedicated port, generally port 5555, needs to be available throughout the cell. Refer to the *HP OpenView Storage Data Protector Installation and Licensing Guide* for details.

The Cell Manager and Installation Servers are installed directly from the CD. After you have installed the Cell Manager and Installation Servers you can then install the components on various client systems using the Data Protector Installation GUI.

When you install Data Protector for the first time, it runs with an instant-on license, valid for 60 days, that lets you use Data Protector before you acquire a permanent license. During this time, purchase any required licenses.

Also during this time, you should set up and configure your Data Protector environment and request your permanent license. To request a permanent password string, you need to know which client systems belong in which Data Protector cell, the number of devices connected to the client systems, and whether you need to use any of the Data Protector integrations.

Creating Cells in the UNIX Environment

Creating cells in the UNIX environment is easy. Based on the considerations given in this manual, decide which client systems you want to add to the cell and define the Cell Manager system. During installation, root access is required to every client system. An important prerequisite is to have a clean node name resolving setup, such that each client system is accessible from every other client system using the same fully qualified node name.

Creating Cells in the Windows Environment

Due to the different possible configurations (domain versus workgroup), the various levels of support for Windows Administrators may have some impact on the setup of Data Protector during installation. An important prerequisite is to have a clean node name resolving setup, so that each client system is accessible from every other client system using the same fully qualified node name.

Windows Domains

A Windows domain can easily be mapped to a Data Protector cell. In a single Windows domain, use a one-to-one mapping if the size of the domain does not exceed the recommended size of the Data Protector cell. Otherwise, split it into two or more cells and manage these cells using the Data Protector Manager-of-Managers.

Mapping a Data Protector Cell into a Windows Domain

Mapping a Data Protector cell into a Windows domain also eases administration within Data Protector itself. To ease administration, distribute the software such that all the client systems can be installed using a central Windows account in a domain organization. Other operations, however, are not limited to a Windows domain organization since all operations and security verifications are performed by the Data Protector internal protocol and not by the Windows Security.

In general, there are no limitations on how and where Data Protector can be installed. However, because of the structure of Windows and the most common configurations that are domain environments, some operations are easier when Data Protector is mapped to a single domain or a multiple domain model, where one of the domains is a master domain, to allow a single user to manage all the client systems within the environment (Software Distribution and User Configuration).

In a multiple cell environment with a Manager-of-Managers, this issue is more significant because all the cells that are configured require a central administrator that has access to the entire backup environment. When a single domain or multiple domains with a master domain are configured, the same global master domain user can be the administrator of all the cells and the Manager-of-Managers environment. If multiple independent domains are used, you need to configure multiple users to administer the environment.

Windows Workgroups

Some of the configuration tasks require more steps in some cases, because there are no global users as in a domain. Software distribution requires a unique logon for every client system that you install the software on. This means that to install 100 client systems in a workgroup environment, you are required to enter 100 logons. In such cases, use a domain environment, since installation and many other non-Data Protector related administration tasks are much easier for a large-scale environment.

Using MoM in such an environment requires you to configure the administrators separately for each cell, to manage the MoM environment from any of the cells.

Again, Data Protector is not limited to a Windows domain organization. However, it takes advantage of and simplifies the administration procedures in the areas where user authentication is required (Installation, User Management).

Creating Cells in a Mixed Environment

In a mixed environment, take into account the factors described in Chapter , “Creating Cells in the UNIX Environment,” on page 37. The more the environment is broken into multiple domains and multiple workgroups, the more accounts and steps need to be considered to distribute the software and to prepare the environment for administration.

Geographically Remote Cells

Data Protector allows you to easily administer geographically remote cells. Refer to “Splitting an Environment into Multiple Cells” on page 16 for more information.

Considerations for Geographically Remote Cells

When configuring geographically remote cells, remember the following:

- Data is not sent over a WAN.
The devices and the client systems that you are backing up are configured locally.
- The cells are configured in a MoM.
To manage geographically remote cells centrally, you need to configure the cells in a MoM environment.
- Consider user configurations.
All the considerations that are mentioned regarding single domain, multiple domain, and workgroup configurations need to be taken into account.

You can configure a single cell over geographically remote locations. In this case, you need to ensure that data transfer from each client system to the corresponding device is not done over a WAN. Because a WAN network is not a stable connection, it is possible that connections are lost.

MoM Environment

A MoM environment does not require a reliable network connection from cells to the central MoM cell, because only controls are sent over the long distance connections, and backups are performed locally within each Data Protector cell. However, this is based on the assumption that each cell has its own media management database.

In such a case, use the Data Protector `Reconnect broken connections backup` option so that connections are re-established after they are broken.

Understanding and Planning Performance

In business-critical environments, it is a key requirement to minimize the time needed for data recovery in case of a corrupt database or a disk crash. Therefore, understanding and planning backup performance is extremely important. Optimizing the time required for the backup of a number of client systems and large databases that are all connected on different networks and different platforms is a challenging task.

The following sections give an overview of the most common backup performance factors. Due to the high number of variables it is not possible to give distinct recommendations that fit all user requirements.

The Infrastructure

The infrastructure has a high impact on the backup and restore performance. The most important aspects are the parallelism of data paths and the use of high-speed equipment.

Network Versus Local Backups

Sending data over a network introduces additional overhead, as the network becomes a component of performance consideration. Data Protector handles the data stream differently for the following cases:

**Network
Datastream**

Disk to Memory of Source System to Network to Memory of Destination System to Device

Local Datastream

Disk to Memory to Device

To maximize performance, use local backup configurations for high volume datastreams.

Network/server Versus Direct Backups

Sending data over a network and through a server introduces additional overhead, as the network and the server become performance considerations. Data Protector handles the datastream differently for the following cases:

**Network
Datastream**

Disk to Memory of Source System to Network to Memory of Destination System to Device

Direct Datastream Disk to Device

To maximize performance, use direct backup configurations for high volume datastreams.

Devices

Device Performance

Device types and models impact performance because of the sustained speed at which devices can write data to a tape (or read data from it).

Data transfer rates also depend on the use of hardware compression. The achievable compression ratio depends on the nature of the data being backed up. In most cases, using high speed devices with hardware compression improves performance. This is true, however, only if the devices stream.

At the start and at the end of a backup session backup devices require some time for operations such as rewinding media and mount or unmount media.

Libraries offer additional advantages because of their fast and automated access to a large number of media. At backup time, loading new or reusable media is needed, and at restore time the media which contain the data to be restored need to be accessed quickly.

Data in disk based-devices is accessed faster than that in conventional devices, as there is no need to load and unload media. This reduces the amount of time spent for backup and restore. Additionally, disk-based devices enable the use of advanced backup strategies such as synthetic backup and disk staging, which also reduce the backup and restore time.

High Performance Hardware Other than Devices

Performance of Computer Systems

The speed of computer systems themselves directly impacts performance. The systems are loaded during backups by reading the disks, handling software compression, and so on.

The diskread data rate and CPU usage are important performance criteria for the systems themselves, in addition to I/O performance and network types.

Advanced High Performance Configuration

Data Protector Zero Downtime Backup solution provides a means of shortening the application downtime or backup mode time and reduces the network overhead by using locally attached backup devices instead of network backup devices. The application downtime or backup mode time is limited to the time needed to create a replica of data, which is then backed up on a backup system to a locally attached device.

For more information on Zero Downtime Backup, refer to the *HP OpenView Storage Data Protector Zero Downtime Backup Concepts Guide*.

Using Hardware in Parallel

Using several datapaths in parallel is a fundamental and efficient method to improve performance. This includes the network infrastructure. Parallelism boosts performance in the following situations:

When to Use Parallelism

- Several client systems can be backed up locally, that is, with the disk(s) and the related devices connected on the same client system.
- Several client systems can be backed up over the network. Here the network traffic routing needs to be such that datapaths do not overlap, otherwise the performance is reduced.
- Several objects (disks) can be backed up to one or several (tape) devices.
- An object (disk or files) can be directly backed up to several (tape) devices using several XCOPY engines.
- Several dedicated network links between certain client systems can be used. For example, if system_A has 6 objects (disks) to be backed up, and system_B has 3 fast tape devices, consider using 3 dedicated network links between system_A and system_B.
- Load Balancing

Using this Data Protector feature, Data Protector dynamically determines which object (disk) should be backed up to which device. Enable this feature, especially to back up a large number of filesystems in a dynamic environment. For more information, see “How Load Balancing Works” on page 136.

Note that you cannot predict to which media a particular object is written.

Configuring Backups and Restores

Any given infrastructure must be used efficiently to maximize performance. Data Protector offers high flexibility to adapt to the environment and the desired way to operate backups and restores.

Software Compression

Software compression is done by the client CPU when reading data from a disk. This reduces the data that is sent over the network, but it requires significant CPU resources from the client.

By default, software compression is disabled. Use software compression only for backups of many machines over a slow network, where data can be compressed before sending it over the network. If software compression is used, hardware compression should be disabled since trying to compress data twice actually expands the data.

Hardware Compression

Hardware compression is done by a device that receives original data from a Drive Server and writes it to media in the compressed mode. Hardware compression increases the speed at which a tape drive can receive data, because less data is written to the tape.

By default, hardware compression is enabled. On HP-UX systems, enable hardware compression by selecting a hardware compression device file. On Windows systems, enable hardware compression during device configuration. Use hardware compression with caution, because media written in compressed mode *cannot* be read using a device in uncompressed mode and vice-versa.

Full and Incremental Backups

A basic approach to improve performance is to reduce the amount of data to back up. Carefully plan your full and incremental backups. Note that you may not need to perform all the full backups of all the client systems at the same time.

If you back up to disk, you can use advanced backup strategies such as synthetic backup and disk staging.

Disk Image Versus Filesystem Backups

It used to be more efficient to back up disk images (raw volumes) rather than filesystems. This is still true in some cases, such as heavily-loaded systems or disks containing large numbers of small files. The general recommendation is to use filesystem backups.

Object Distribution to Media

The following are examples of object/media backup configurations provided by Data Protector:

- One object (disk) goes to one medium

The advantage is a known fixed relationship between an object and a medium where the object resides. This can be of benefit for the restore process, since only *one* medium needs to be accessed.

The disadvantage in a network backup configuration is the likely performance limitation due to the network, causing the device not to stream.

- Many objects go to a few media, each medium has data from several objects, one object goes to one device

The advantage here is the flexibility of datastreams at backup time, helping to optimize performance, especially in a network configuration.

The strategy is based on the assumption that the devices receive enough data to be able to stream, since each device receives data from several sources concurrently.

The disadvantage is that data (from other objects) has to be skipped during the restore of a single object. Additionally, there is no precise prediction as to which medium will receive data from which object.

For more information on device streaming and backup concurrency, refer to “Device Streaming and Concurrency” on page 136.

Disk Performance

All data that Data Protector backs up resides on disks in your systems. Therefore, the performance of disks directly influences backup performance. A disk is essentially a sequential device, that is, you can read or write to it, but not both at the same time. Also, you can read or

write one stream of data at a time. Data Protector backs up filesystems sequentially, to reduce disk head movements. It also restores files sequentially.

Sometimes this is not visible because the operating system stores most frequently used data in a **cache memory**.

**Disk
Fragmentation**

Data on a disk is not kept in the logical order that you see when browsing the files and directories, but is fragmented in small blocks all over the physical disk. Therefore, to read or write a file, a disk head must move around the whole disk area. Note that this differs from one operating system to another.

TIP

Backups are most efficient for large files with little fragmentation.

Compression

If data is compressed on a disk, the Windows operating system first decompresses the data before sending it across the network. This reduces the backup speed and uses CPU resources.

**Disk Image
Backups**

Data Protector also allows you to back up UNIX disks as disk images. With a disk image backup, a complete image of the disk is backed up without tracking the filesystem structure. The disk head moves linearly across the surface. Thus a disk image backup can be considerably faster than a filesystem backup.

SAN Performance

If large volumes of data need to be backed up in one session, the time needed to transfer the data becomes significant. This consists of the time required to move the data over a connection (LAN, local, or SAN) to a backup device.

Online Database Application Performance

When you back up databases and applications, such as Oracle, SAP R/3, Sybase, and Informix Server, the performance of the backups also depends on the applications. Database online backups are provided so that backups can occur while the database application remains online.

This helps to maximize database up time but may impact application performance. Data Protector integrates with all popular online database applications to optimize backup performance.

Refer to the *HP OpenView Storage Data Protector Integration Guide* for more information on how Data Protector integrates with various applications and for tips on how to improve backup performance.

Also refer to the documentation that comes with your online database application for more information on how to improve backup performance.

Planning Security

When you plan your backup environment, consider security. A well thought out, implemented, and updated security plan prevents the unauthorized access, duplication, or modification of data.

What Is Security? Security in the backup context typically refers to:

- Who can administer or operate a backup application (Data Protector).
- Who can physically access client systems and backup media.
- Who can restore data.
- Who can view information about backed up data.

Data Protector provides security solutions on all these levels.

Data Protector Security Features The following features allow and restrict access to Data Protector and the backed up data. The items in this list are described in detail in the following sections.

- Cells
- Data Protector user accounts
- Data Protector user groups
- Data Protector user rights
- Visibility and access to backed up data

Cells

Starting Sessions Data Protector security is based on cells. Backup and restore sessions can only be started from the Cell Manager unless you have the Data Protector Manager-of-Managers functionality. This ensures that users from other cells cannot back up and restore data from systems in your local cell.

Access from a Specific Cell Manager Additionally, Data Protector allows you to explicitly configure from which Cell Manager a client system can be accessed, that is, configuring a trusted peer.

Restrict Pre- and Post-Execution For security reasons, various levels of restrictions can be configured for pre-exec and post-exec scripts. These optional scripts allow a client system to be prepared for the backup by, for example, shutting down an application to obtain a consistent backup.

Data Protector **Users Accounts**

Data Protector User Accounts Anyone using any Data Protector functionality, administering Data Protector, or restoring personal data, must have a Data Protector user account. This restricts unauthorized access to Data Protector and backed up data.

Who Defines User Accounts? An administrator creates this account specifying a user login name, systems from which a user can log in, and the Data Protector user group membership that defines the user rights

When Is the Account Checked? When a user starts the Data Protector user interface, Data Protector checks user rights. User rights are also checked when specific tasks are performed by a user.

Refer to Chapter 4, “Users and User Groups,” on page 163 for more information.

Data Protector **User Groups**

What Are User Groups? When a new user account is created, the user becomes a member of the specified user group. Each user group contains defined Data Protector user rights. All the members of the group have the user rights set for the group.

Why Use User Groups? Data Protector user groups simplify user configuration. The administrator groups users according to the access they need. For example, an end-user group could allow members to restore personal data to a local system only, while the operator group allows the starting and monitoring of backups, but not the creating of backups.

Refer to Chapter 4, “Users and User Groups,” on page 163 for more information.

Data Protector **User Rights**

What Are User Rights?

Data Protector user rights define the actions that a user can perform with Data Protector. They are applied on the Data Protector user group level and not to each user individually. Users added to a user group automatically gain the user rights assigned to this user group.

Why Use User Rights?

Data Protector provides flexible user and user group functionality, which allows the administrator to selectively define who can use a particular Data Protector functionality. It is important to carefully apply the Data Protector user rights: backing up and restoring data is essentially the same as copying data.

Refer to Chapter 4, “Users and User Groups,” on page 163 for more information.

Visibility of Backed Up Data

Backing up data means creating a new copy. Therefore, when you deal with confidential information, it is important to restrict access to both the original data and to the backup copy itself.

Hiding Data from Other Users

When you configure a backup, you can decide whether during restore the data is visible to everyone (public) or only to the owner of the backup (private). The owner is the user who configured the backup and started (scheduled) the backup session. Refer to “Who Owns a Backup Session?” on page 51 for more information about backup owners.

Data Encoding

Open Systems and public networking make data security in large enterprises essential. Data Protector lets you encode filesystem and disk image data so that it becomes unreadable. Data is encoded before it is transferred over a network and before it is written to media. Data Protector uses a fixed, built-in algorithm for this purpose.

Who Owns a Backup Session?

What Is Backup Ownership?

The Data Protector user who created a backup specification becomes the owner of the running backup session and of the resulting media set. Note that this ownership refers to the Data Protector user and not to the system (platform) user. Therefore, the backup session does not run under the user name of the owner.

Who Can Start a Backup?

You can only run a backup specification that you have created. Therefore, if the backup administrator created a backup specification, other users are not able to start a backup of this backup specification. Refer to online Help, index keyword "ownership" for instructions on how to change a backup owner. Note that if you change a backup owner, you allow this person to access/restore data they may not actually own.

Backup Ownership and Restore

Ownership also affects your ability to restore data. If the private/public option is set to private, only the owner of the media set or the administrators can see the data saved in the media set.

Clustering

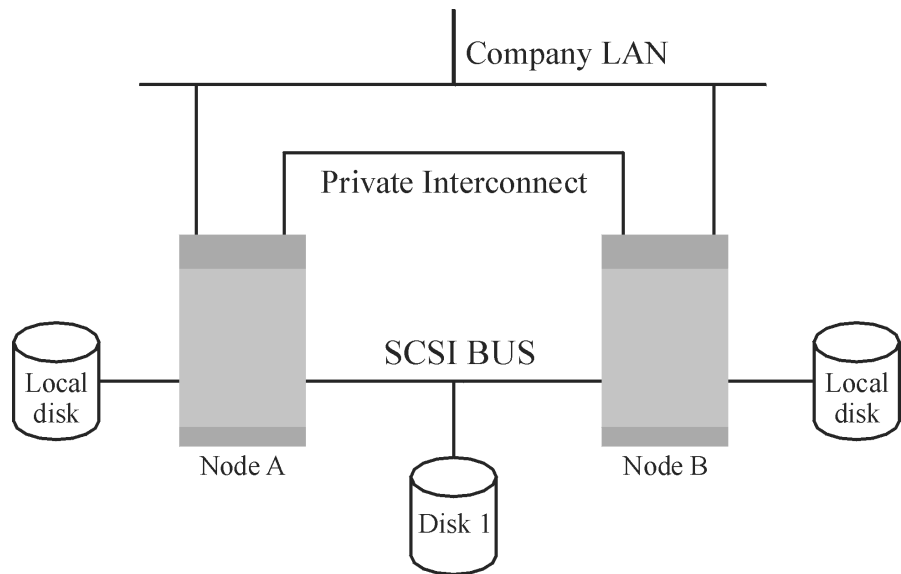
Cluster Concepts

What Is a Cluster? A **cluster** is a group of two or more computers that appear on the network as a single system. This group of computers is managed as a single system and is designed to:

- Ensure that mission-critical applications and resources are as highly-available as possible
- Tolerate component failures
- Support either the addition or subtraction of components

For clustering purposes, Data Protector integrates with Microsoft Cluster Server for Windows Server, with MC/Service Guard for HP-UX, with Veritas Cluster for Solaris and with Novell NetWare Cluster Services. For a list of supported clusters, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

Figure 2-1 Typical Cluster



Components:

- Cluster nodes (two or more)
- Local disks
- Shared disks (shared between nodes)

Cluster Nodes

Cluster nodes are computers that compose a cluster. They are physically connected to one or more shared disks.

Shared Disks

The **shared disks volumes** (MSCS, Novell NetWare Cluster Services) or **shared volume groups** (MC/SG, Veritas Cluster) contain mission-critical application data as well as specific cluster data needed to run the cluster. In MSCS clusters, a shared disk is exclusively active on only one cluster node at a time.

Cluster Network

A cluster network is a private network that connects all cluster nodes. It transfers the internal cluster data called **heartbeat of the cluster**. The heartbeat is a data packet with a time stamp that is distributed among all cluster nodes. Each cluster node compares this packet and determines

the cluster node that is still operational so that you can make an appropriate determination of the ownership of the **package** (MC/SG, Veritas Cluster) or **group** (MSCS).

What is a Package or Group? A package (MC/SG, Veritas Cluster) or a group (MSCS) is a collection of resources that are needed to run a specific **cluster-aware** application. Each cluster-aware application declares its own critical resources. The following resources must be defined in each group or package:

- Shared disk volumes (MSCS, Novell NetWare Cluster Services)
- Shared volume groups (MC/SG, Veritas Cluster)
- Network IP names
- Network IP addresses
- Cluster-aware application services

What Is a Virtual Server? Disk volumes and volume groups represent shared physical disks. A network IP name and a network IP address are resources that define a **virtual server** of a cluster-aware application. Its IP name and address are cached by the cluster software and mapped to the cluster node where the specific package or group is currently running. Since the group or package can switch from one node to another, the virtual server can reside on different machines in different time frames.

What Is a Failover? Each package or group has its own “preferred” node where it normally runs. Such a node is called a **primary node**. A package or group can be moved to another cluster node (one of the **secondary nodes**). The process of transferring a package or group from the primary cluster node to the secondary is called **failover** or switchover. The secondary node accepts the package or group in case of failure of the primary node. A failover can occur for many different reasons:

- Software failures on the primary node
- Hardware failures on the primary node
- The administrator intentionally transfers the ownership because of maintenance on the primary node

In a cluster environment there can be more than one secondary node but only one can be the primary.

A cluster-aware Data Protector Cell Manager that is responsible for running the IDB and managing backup and restore operations has many major benefits over non-cluster versions:

- | | |
|---|---|
| High Availability of the Data Protector Cell Manager | All Cell Manager operations are always available since Data Protector services are defined as cluster resources within the cluster and are automatically restarted when a failover occurs. |
| Automatic Restart of Backups | Data Protector backup specifications that define the backup procedure can easily be configured to be restarted in case of a failover of the Data Protector Cell Manager. Restart parameters can be defined using the Data Protector GUI. |
| Load Balancing at Failover | A special command-line utility is provided for operations that allow backup sessions to be aborted in case applications other than Data Protector perform a failover. The Data Protector Cell Manager allows you to define what should happen in such situations. If the backup is less important than the application, Data Protector can abort running sessions. If the backup is more important or is just ending, Data Protector can continue the sessions. Refer to online Help, index keyword "cluster, managing backups" for more information on how to define the criteria. |

Cluster Support

The Data Protector cluster support means the following:

- The Data Protector Cell Manager is installed in a cluster. Such a Cell Manager is fault tolerant and can *restart operations* in the cell automatically after the failover.

NOTE

If the Cell Manager is installed in the cluster, its cluster critical resources need to be configured in the same cluster package or group as the application being backed up, in order to automatically restart *failed backup sessions* that failed due to a failover. Otherwise, the failed backup sessions must be restarted manually.

- The Data Protector client is installed in a cluster. The Cell Manager (if not installed in the cluster) in such a case is not fault tolerant; the operations in the cell must be restarted manually.

Clustering

The behavior of the Cell Manager after the failover is configurable as far as the *backup session* (failed due to the failover) is concerned - the failed session can be:

- restarted as a whole
- restarted only for the failed objects
- not restarted at all

Refer to online Help, index keyword "cluster, backup specification options" for more information on backup session behavior options on failover of the Data Protector Cell Manager.

Example Cluster Environments

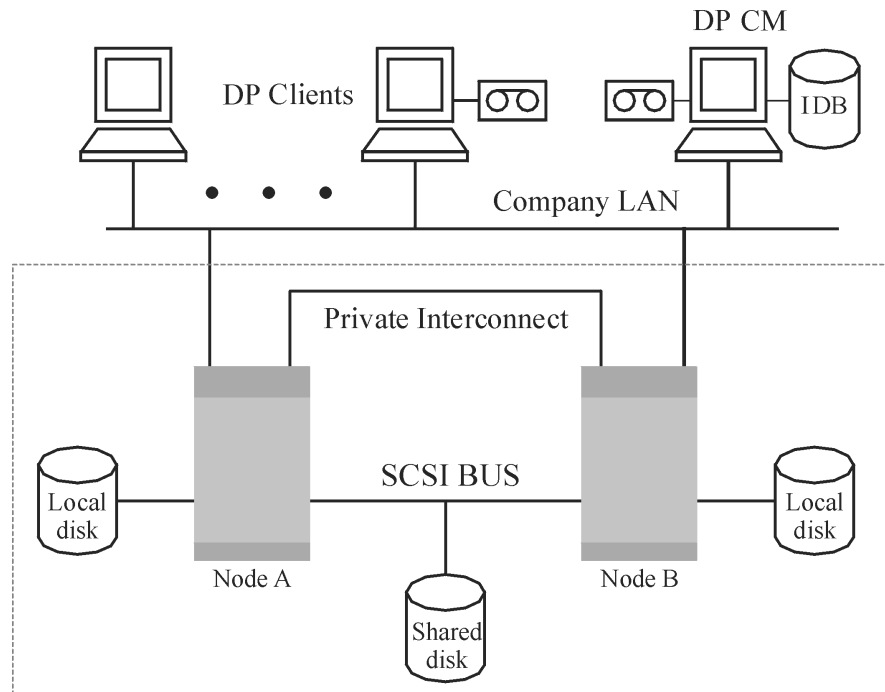
This section gives three example cluster configurations.

Cell Manager Installed Outside a Cluster

In the environment depicted below:

- The Cell Manager installed outside a cluster
- A backup device connected to the Cell Manager or one of the (non-clustered) clients

Figure 2-2 Cell Manager Installed Outside a Cluster



When creating a backup specification, you can see three or more systems that can be backed up in the cluster.

- Physical Node A
- Physical Node B
- Virtual Server

Virtual Server Backup

If you select the virtual server in the backup specification, then the backup session will back up the selected active virtual host/server regardless of the physical node the package or group is currently running on.

Refer to online Help, index keyword "cluster, backup specification options" for more information on how to define these options.

The following is the expected backup behavior under this configuration.

Table 2-1 Backup Behavior

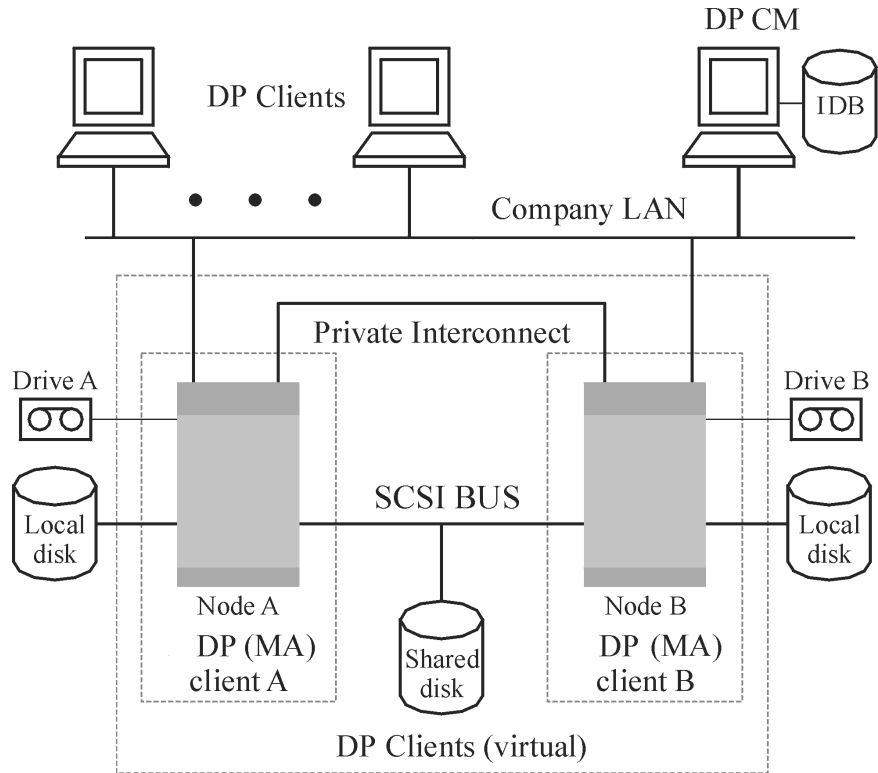
Condition	Result
Failover of the node before a backup starts	Successful backup
Failover of the node during backup activity	Filesystem/disk image backup: The backup session fails. The completed objects from the session can be used for restore, the failed (running and pending) objects need to be backed up again by restarting the session manually.
	Application backup: The backup session fails. The session needs to be restarted manually.

Cell Manager Installed Outside a Cluster, Devices Connected to the Cluster Nodes

In the environment depicted below:

- The Cell Manager installed outside a cluster
- Backup devices connected to the nodes in the cluster

Figure 2-3 Cell Manager Installed Outside a Cluster, Devices Connected to the Cluster Nodes



When creating a backup specification, you can see three or more systems that can be backed up in the cluster.

- Physical Node A
- Physical Node B
- Virtual Server

Virtual Server Backup

If you select the virtual server in the backup specification, then the backup session will back up the selected active virtual host/server regardless of the physical node the package or group is currently running on.

NOTE

The difference with the previous example is that each of the cluster nodes has a Data Protector Media Agent installed. Additionally, you need to use the Data Protector load balancing functionality. Include both devices in the backup specification. With load balancing set to `min=1` and `max=1`, Data Protector will only use the first available device.

The following is the expected backup behavior under this configuration.

Table 2-2 Backup Behavior

Condition	Result
Failover of the node before a backup starts	Successful backup due to automatic device switching (load balancing)
Failover of the node during backup activity	Filesystem/disk image backup: The backup session fails. The completed objects from the session can be used for restore, the failed (running and pending) objects need to be backed up again by restarting the session manually.
	Application backup: The backup session fails. The session needs to be restarted manually.

IMPORTANT

If a failover during backup activity occurs in such a configuration, the MA may not be able to properly abort the session. This results in the corruption of the medium.

Cell Manager Installed in a Cluster, Devices Connected to the Cluster Nodes

In the environment depicted below:

- The Cell Manager installed in a cluster.

With regard to the Data Protector application integrations, there are two possible ways of configuring Data Protector and an application in such a configuration:

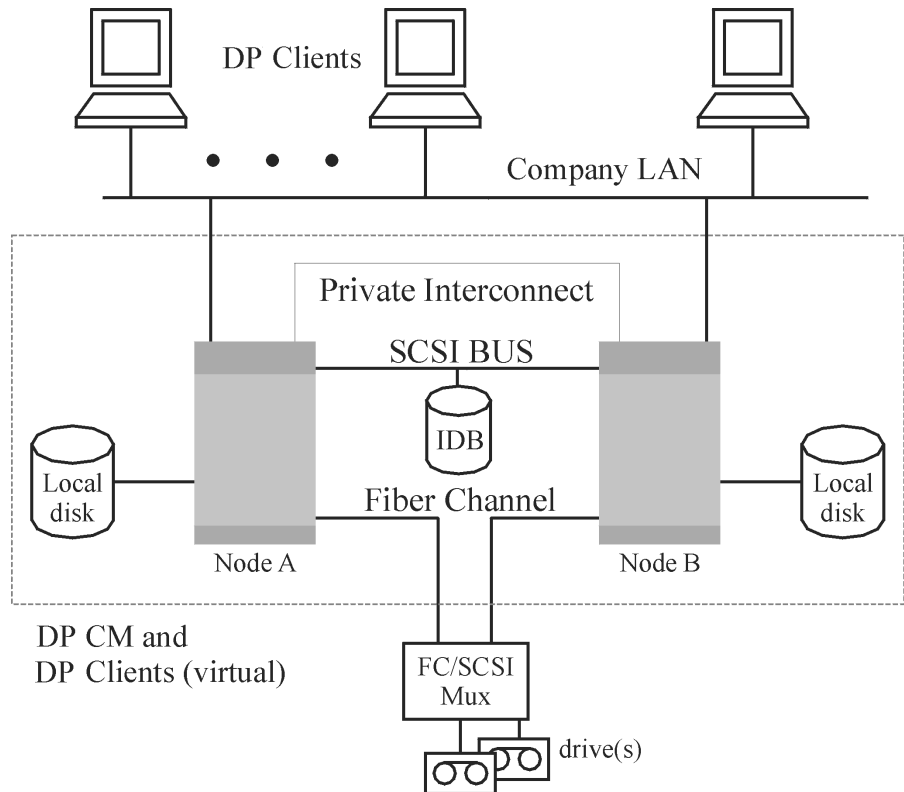
- The Data Protector Cell Manager is configured to run (both during the normal operation and during the failover) on the same node as the application - the Data Protector cluster critical resources are defined in the same package (MC/ServiceGuard) or group (Microsoft Cluster Server) as the application cluster critical resources.

IMPORTANT

Only in such a configuration, it is possible to define the automated action concerning the Data Protector sessions aborted during the failover.

- The Data Protector Cell Manager is configured to run (both during the normal operation and during the failover) on nodes other than the application node - the Data Protector cluster critical resources are defined in some other package (MC/ServiceGuard) or group (Microsoft Cluster Server) as the application cluster critical resources.
- Backup device(s) connected to the cluster shared Fibre Channel bus via an FC/SCSI MUX.

Figure 2-4 Cell Manager Installed in the Cluster, Devices Connected to Cluster Nodes



When creating a backup specification, you can see three or more systems that can be backed up in the cluster.

- Physical Node A
- Physical Node B
- Virtual Server

Virtual Server Backup

If you select the virtual server in the backup specification, then the backup session will back up the selected active virtual host/server regardless of the physical node the package or group is currently running on.

NOTE

Clusters do not support a SCSI bus with shared tapes. To bring high availability also to Media Agents, the Fibre Channel technology can be used as an interface to the device. The device itself is not highly-available in this configuration.

This configuration allows the following features:

- Customizable automatic restart of backups in case of failover of the Cell Manager.

The Data Protector backup specifications can be configured to be restarted in case of failover of the Cell Manager. Restart parameters can be defined using the Data Protector GUI.

- System load control at failover.

Sophisticated control is provided to define Data Protector behavior at failover. A special command, `omniclus`, is provided for this purpose. The Cell Manager allows the administrator to define what should happen in such situations.

- If the backup is less important than the application that just switched to the backup system, Data Protector can abort the running sessions.
- If the backup is more important or it is just pending, Data Protector continues the sessions.

The following is the expected backup behavior under this configuration.

Table 2-3

Backup Behavior

Condition	Result
Failover before a backup starts	Successful backup

Table 2-3 Backup Behavior

Condition	Result	
Failover of the application and the Cell Manager during backup activity (Cell Manager runs on the same node as the application).	Filesystem/disk image backup The backup session fails. The completed objects from the session can be used for restore, the failed (running and pending) objects are backed up again by restarting the session automatically.	IMPORTANT To restart the session, the appropriate Data Protector option must be selected. Refer to online Help, index keyword "cluster, managing backups" for information on defining all possible Data Protector actions in case of failover of the Cell Manager.
	Application backup The backup session fails. The session is restarted automatically.	
Failover of the application during backup activity without Cell Manager failover (Cell Manager runs on other node than the application).	Filesystem/disk image backup The backup session fails at failover of the node where the filesystem is installed. The completed objects from the session can be used for restore, the failed (running and pending) objects need to be backed up again by restarting the session manually.	
	Application backup The backup session fails. The session needs to be restarted manually.	

IMPORTANT

If a failover during backup activity occurs in such a configuration, the MA may not be able to properly abort the session. This results in the corruption of the medium.

Additionally, the Data Protector cluster Cell Manager/client can be integrated with the EMC Symmetrix or HP StorageWorks Disk Array XP environment, producing a very highly-available backup environment. Refer to the *HP OpenView Storage Data Protector Zero Downtime Backup Administrator's Guide* for more information.

Full and Incremental Backups

Data Protector provides two basic types of filesystem backups: full and incremental.

A full backup saves all the files selected for backup in a filesystem. An incremental backup saves only those files that have changed since the last full or incremental backup. This section gives hints on how to choose the backup type and how this influences your backup strategy.

Table 2-4

Comparison of Full and Incremental Backup

	Full backup	Incremental backup
Resources	Takes more time to complete than incremental backup and requires more media space.	Backs up only changes made since a previous backup, which requires less time and media space.
Device handling	If you use a standalone device with a single drive, you need to change the media manually if a backup does not fit on a single medium.	It is less likely that the backup will require additional media.
Restore	Enables simple and quick restore.	A restore takes more time because of the number of media needed.
IDB impact	Occupies more space in the IDB.	Occupies less space in the IDB.

Data Protector can also make incremental backups of online database applications. These vary from application to application. On Sybase, for instance, this type of backup is referred to as a transaction backup (a backup of transaction logs modified since the last backup).

Note that the incremental backup concept is not related to the log level concept, which defines the amount of information written to the IDB.

NOTE

A number of additional backup types (such as direct backup, split mirror backup, snapshot backup, and data mover backup) are available with Data Protector application integrations. Refer to the respective *HP OpenView Storage Data Protector Integration Guides* for more information.

Full Backups

Full backups always back up all selected objects, even if there are no changes since the previous backup.

Synthetic Backup

Synthetic backup is an advanced backup solution that eliminates the need to run regular full backups. Instead, incremental backups are run, and subsequently merged with the full backup into a new, synthetic full backup. For more information, see Chapter 11, “Synthetic Backup,” on page 261.

Incremental Backups

Incremental backups back up changes from a previous still protected (full or incremental) backup. A full backup of an object (with identical client name, mount point, and description) must exist before an incremental backup of this object is possible.

Conventional Incremental Backup

Before running an incremental backup of a specific backup object, Data Protector compares the trees in the backup object with the trees in the valid restore chain of this object. If the trees do not match (for example, an additional directory in the backup object was selected for backup since the last backup or multiple backup specifications with the same backup object and different trees exist), a full backup is automatically performed. This ensures that all files that have changed since the last relevant backup are backed up.

With conventional incremental backup, the criterion for determining whether a file has changed or not since a previous backup is the file's modification time. However, if a file has been renamed, moved to a new

location, or if some of its attributes have changed, its modification time does not change. Consequently, the file is not always backed up in a conventional incremental backup. Such files are backed up in the next full backup.

Enhanced Incremental Backup

Enhanced incremental backup reliably detects and backs up also renamed and moved files, as well as files with changes in attributes.

Enhanced incremental backup also eliminates unnecessary full backups of an entire backup object when some of the trees selected for backup change. For example, if an additional directory is selected for backup since the last backup, a full backup of this directory (tree) is performed, whereas the backup of the rest is incremental.

Using enhanced incremental backup is a prerequisite for synthetic backup.

Types of Incremental Backups

Data Protector provides incremental backups of different types:

- | | |
|--------|--|
| Inc | This type of backup, shown in Figure 2-5, is based on <i>any</i> previous backup, be it a full backup or an incremental backup, which must still be protected. This backup is also called a differential backup because only the changes from the previous backup are backed up. |
| Inc1-9 | A leveled incremental backup , shown in Figure 2-6, depends on the last backup of the next lower level that is still protected. For example, an Inc1 backup saves all changes since the last full backup, while an Inc5 backup saves all changes since the last Inc4 backup. An Inc1-9 backup <i>never</i> references an existing Inc backup. |

Figure 2-5 Differential Backups

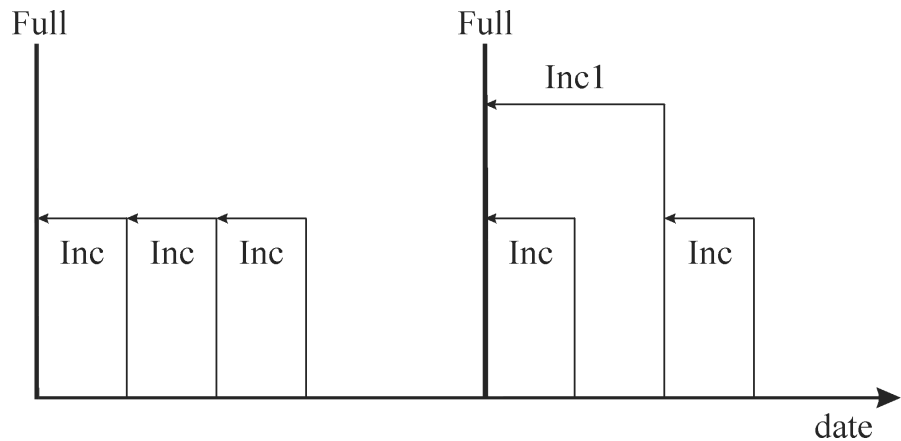


Figure 2-5 and Figure 2-6 indicate various types of incremental backups. Incremental backups depend on the last full backup. Data Protector checks if there is a protected full backup of the backed up data whenever an incremental backup is started. If there is no protected full backup Data Protector starts a full backup instead.

Figure 2-6 Leveled Incremental Backups

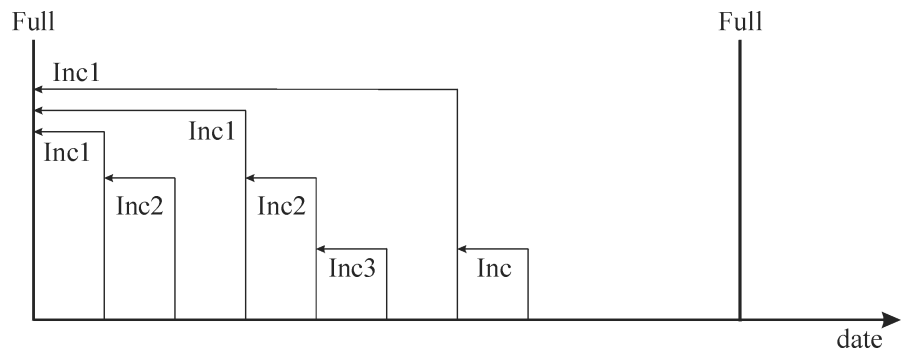


Table 2-5 shows the relative referencing of backup runs with various backup types. See the text following the table for a full explanation.

Table 2-5 Relative Referencing of Backup Runs

1 Full <---- Incr1

Table 2-5 **Relative Referencing of Backup Runs**

2	Full	<----	<----	<----	Incr2		
3	Full	<----	Incr1	<----	Incr2		
4	Full	<----	Incr				
5	Full	<----	Incr1	<----	Incr		
6	Full	<----	Incr1	<----	Incr2	<----	Incr
7	Full	<----	Incr1	<----	Incr	<----	Incr
8	Full	<----	Incr1	<----	Incr3		
9	Full	<----	Incr1	<----	Incr2	<----	Incr3
10	Full	<----	<----	<----	Incr2	<----	Incr3
11	Full	<----	<----	<----	<----	<----	Incr3

How to Read Table 2-5

- The rows in Table 2-5 are independent of each other and show different situations.
- The age of the backups increases from right to left, so that the far left is the oldest and the far right is the most recent backup.
- The full and IncrX represent still-protected objects of the same owner. Any existing IncrX that is not protected can be used for restore, but is not considered for referencing on subsequent backup runs.

Examples

- In the second row, there is a full, still protected backup and an Incr2 is running. There is no Incr1, so the backup is executed as an Incr1.
- In the fifth row, there is a full backup, an Incr1 and another incremental is running. Data Protector references the currently running backup to the previous incremental, that is Incr1.
- In the eighth row, the Incr3 is executed as Incr2, and in the eleventh row, the Incr3 is executed as Incr1.

Considering Restore

To restore the latest data, you need media from your last full backup and subsequent incremental backups. Therefore, the more incremental backups you have, the more media you need to handle. This is inconvenient if you use standalone devices, and the restore can last long.

Media Sets

Using differential and leveled incremental backups, as indicated in Figure 2-7, will require access to all five previously completed **media sets**, up to and including the full backup. The space needed on the media is lowest here, but the restore is rather complex. The series of required media sets is also called a **restore chain**.

TIP

Use the Data Protector Appendable on Incrementals Only option to keep data from full and incremental backups (of the same backup specification) on the same media set.

Another common use of the incremental backup concept is indicated in Figure 2-8. Here the required space on the media is slightly larger. Only two media sets need to be accessed to restore to the desired point in time. Note that there is no dependency on any previous Inc1 media set for this restore, unless the desired restore point in time would be moved.

Figure 2-7

Media Needed to Restore from Differential and Leveled Backups

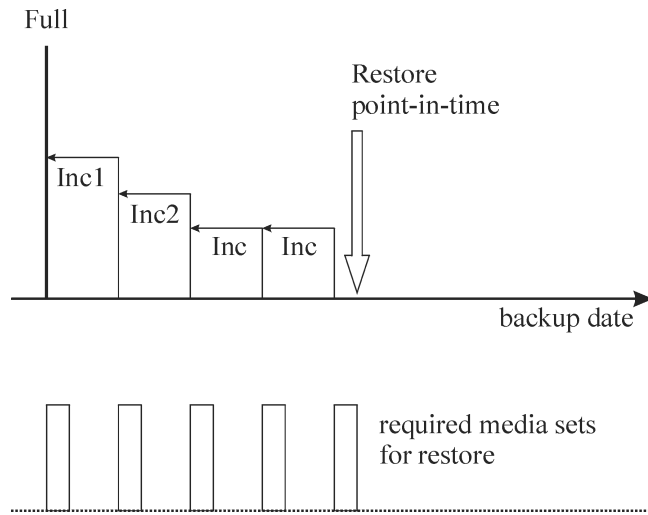
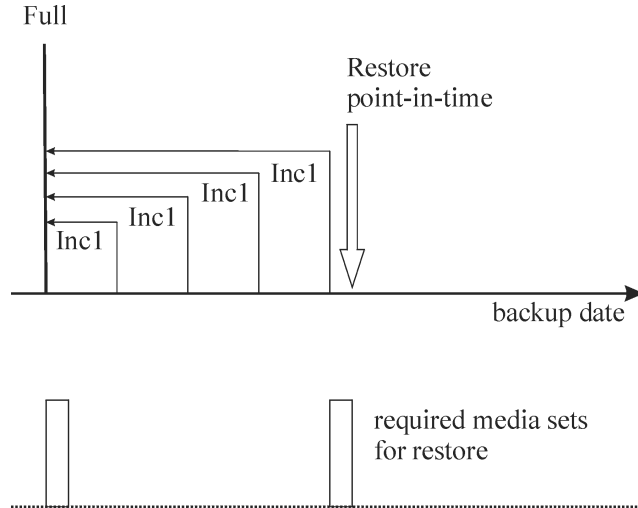


Figure 2-8

Media Needed to Restore from Leveled Incremental Backups



Note that you must set the appropriate data protection in order to get all needed full and incremental backups for restore. If the data protection is not properly set, you can get a broken restore chain. Refer to Appendix B for more information.

Keeping Backed Up Data and Information About the Data

Data Protector lets you specify how long to keep your backed up data on the media itself (data protection), how long to keep information about the backed up data in the IDB (catalog protection), and what level of information to keep in the IDB (logging level).

You can set the protection independently for backed up data and for backup information about this data in the IDB. When copying media, you can specify a different protection period for the copies than the protection of the original media.

Data Protector Internal Database

Restore performance depends, in part, on how fast the media required for a restore can be found. By default, this information is stored in the IDB to enable the highest restore performance as well as the convenience of being able to browse the files and directories to be restored. However, putting all file names of all backups in the IDB and keeping them for a long time can cause the IDB to grow to unmanageable levels.

Data Protector allows you to trade off IDB growth with the convenience of restore, by letting you specify catalog protection independently of data protection. For example you can implement a policy that enables an easy and fast restore within four weeks after the backup, by setting catalog protection to four weeks. From then on restores can still be done in a less convenient way until the data protection expires, say after one year. This would considerably reduce the space requirements in the IDB.

Data Protection

What Is Data Protection?

Data Protector allows you to specify the amount of time data on media is protected from being overwritten by Data Protector. You can specify the protection in absolute or relative dates.

You can specify data protection in different parts of Data Protector. For details, refer to online Help, index keyword "data protection".

If you do not change the Data Protection backup option when configuring a backup, it is permanently protected. Note that if you do not change this protection, the number of media needed for backup grows constantly.

Catalog Protection

What Is Catalog Protection?

Data Protector saves information about backed up data in the IDB. Since the information about the backed up data is written to the IDB each time a backup is done, the IDB grows with the number and the size of backups. Catalog protection tells Data Protector how long the information about backed up data is available to users browsing data during restore. Once catalog protection has expired, Data Protector will overwrite this information in the IDB (not on the media) in one of the subsequent backups.

You can specify the protection using absolute or relative dates.

If you do not change the `Catalog Protection` backup option when configuring your backup, information about backed up data has the same protection duration as data protection. Note that if you do not change this, the IDB grows constantly as new information is added with each backup.

Refer to “Catalog Protection as an IDB Key Tunable Parameter” on page 185 for more information on how catalog protection settings influence the IDB growth and performance.

The protection model used by Data Protector can be mapped to the concept of backup generations, which is elaborated in “Further Information” on page B-1

Logging Level

What Is Logging Level?

Logging level determines the amount of details on files and directories written to the IDB during backup. You can always restore your data, regardless of the logging level used during the backup.

Data Protector provides four logging levels that control the amount of details on files and directories written to the IDB. For more information, refer to “Logging Level as an IDB Key Tunable Parameter” on page 183.

Browsing Files for Restore

The IDB keeps information about the backed up data. This information allows you to browse, select and start the restore of files using the Data Protector user interface. You can also restore data without this

information as long as the media are still available, but you must know which media to use and what needs to be restored, for example, the exact file name.

The IDB also keeps information on how long the actual data on the media will not be overwritten.

Data protection, catalog protection and logging level policies influence the availability of data and access time to data during restore.

Enabling the Browsing of Files and Quick Restore

To restore files quickly, both information about backed up data in the catalog and protected data on the media, must exist. Information in the catalog allows you to browse, select, and start the restore of files using the Data Protector user interface and allows Data Protector to quickly locate data on backup media.

Enabling the Restore of Files, but not Browsing

Once catalog protection has expired and data protection is still valid, you cannot browse files in the Data Protector user interface, but you can still restore data if you know the file name and the media. The restore is slower as Data Protector does not know where on the media the desired data is located. You can also import the media back into the IDB, thus re-establishing the information about backed up data in the catalog, and then start restoring.

Overwriting Backed Up Files with New Data

Once data protection has expired, data on the media is overwritten in one of the subsequent backups. Before this happens, you can still restore the data from the media.

TIP

Set data protection to the amount of time that you must keep the data, for example, one year.

Set the catalog protection to the amount of time you want to be able to browse, select, and restore files quickly using the Data Protector user interface.

Exporting Media from a Cell

Exporting media from a Data Protector cell removes all the information about backed up data on the media and the media themselves from the IDB. You cannot browse, select or restore files from exported media using the Data Protector user interface. You need to re-read (or add) the media back into the Data Protector cell. This functionality is needed to move media to a different cell.

Backing Up Data

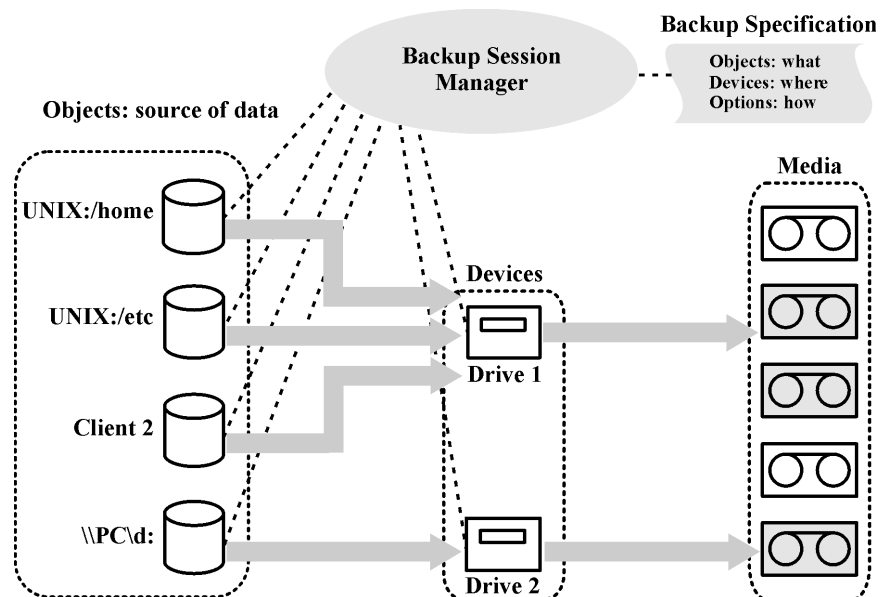
Backing up your data consists of some or all of the following steps:

- Selecting what to back up, from which client system - the source of data.
- Selecting where to back up - the destination.
- Selecting to write the same data to additional media sets - mirroring.
- Selecting how to back up - backup options.
- Scheduling a backup for automated operation.

You can specify all these when creating a backup specification.

Figure 2-9

Backup Session



At the specified time, Data Protector starts a backup session based upon a backup specification. The source of data is specified as a list of objects (such as a filesystem on UNIX or disk drives on Windows systems) and the destinations are specified (tape) devices. During the backup session, Data Protector reads the objects, transfers data through the network,

and writes it to the media residing in the devices. The backup specification names the devices to use. It also can specify a media pool. If no media pool is specified, the default media pool is used.

A backup specification can be a simple definition of the backup of a disk to a standalone DDS drive, or a complex definition of the backup of 40 large servers to a Silo tape library with eight drives.

Creating a Backup Specification

What Is a Backup Specification?

A backup specification allows you to group objects that you want to back up in a group with common characteristics, such as scheduling, used devices, type of backup, and backup session options.

How to Create a Backup Specification

You configure a backup specification using the Data Protector user interface. You need to know what you want to back up, how many mirrors you want to create, which media and which devices you want to use for the backup, and optionally, some desired specific behavior for the backup. Data Protector provides default behavior that is suitable for most cases. You can customize backup behavior using Data Protector backup options.

Data Protector can back up a client with all the disks connected to it by discovering the disks at backup time. Refer to “Backing Up with Disk Discovery” on page 217.

Selecting Backup Objects

What Is a Backup Object?

Data Protector uses the term **backup object** for a backup unit that contains all items selected for backup from one disk volume (logical disk or mount point). The selected items can be any number of files, directories, or the entire disk or mount point. Additionally, a backup object can be a database entity or a disk image (rawdisk).

A backup object is defined by:

- Client name: a hostname of the Data Protector client where the backup object resides.
- Mount point: an access point in a directory structure (drive on Windows and mount point on UNIX) on the client where the backup object is located.

- Description: uniquely defines the backup objects with identical client name and mount point.
- Type: backup object type, for example filesystem or Oracle.

The way in which a backup object is defined is important to understand how incremental backups are done. For example, if the description of a backup object changes, it is considered as a new backup object, therefore a full backup will be automatically performed instead of incremental.

Examples of Backup Options

You can customize the backup behavior for each individual backup object by specifying the backup options for this object. The following are examples of the backup options you can specify:

- Logging level of information going to the IDB.

Data Protector provides four levels that control the amount of details on files and directories stored in the IDB:

- Log All
- Log Files
- Log Directories
- No Log

Note that changing the level of stored information affects the ability to browse the files using the Data Protector user interface when restoring. Refer to “Logging Level as an IDB Key Tunable Parameter” on page 183 for more information on logging levels.

- Automatic load balancing

Dynamic device allocation from a specified list. For more information, see “How Load Balancing Works” on page 136.

Data Protector dynamically determines which object (disk) should be backed up to which device.

- Pre-exec and post-exec scripts

Processing to prepare a client for a consistent backup. For more information, refer to “Pre-Exec and Post-Exec Commands” on page 215.

You can also specify the directories to exclude from a backup, or back up specific directories only. You can also back up disks as they are added. Thus, your backup is fully configurable and dynamic.

Backup Sessions

What Is a Backup Session?

A backup session is a process that backs up data from a client system to media. A backup session always runs on the Cell Manager system. A backup session is based on a backup specification and is started when a backup is run.

During a backup session, Data Protector backs up data using default or customized behavior.

Refer to Chapter 7, “How Data Protector Operates,” on page 209 for advanced information on backup sessions, and how to control sessions.

Object Mirrors

What Is an Object Mirror?

An object mirror is an additional copy of a backup object created during a backup session. When creating a backup specification, you can choose to create one or several mirrors of specific objects. The use of object mirroring improves the fault tolerance of backups and enables multi-site vaulting. However, object mirroring during a backup session increases the time needed for backup.

For more information, see “Object Mirroring” on page 96.

Media Sets

What Is a Media Set?

The result of a backup session is backed up data on a medium or a media set. Each backup session results in one or several media sets, depending on whether you perform backup with object mirroring. Depending on the pool usage, several sessions can share the same media. When you restore data, you need to know the media from which to restore. Data Protector keeps this information in the Catalog Database.

Backup Types and Scheduled Backups

A scheduling policy defines when backups start and the backup types (full or incremental). Consider the differences between full and incremental backups. See Table 2-4 on page 66.

You can combine full and incremental backups when you configure scheduled backups. For example, you may run a full backup on Sundays and incremental backups every working day. To back up a large amount of data and avoid the high volume peak for the full backups, use the staggered approach. Refer to “Staggering Full Backups” on page 82.

Scheduling, Backup Configurations, and Sessions

Backup Configuration

When you schedule a backup, all the objects specified in that backup specification are backed up in the scheduled backup session(s).

For each individual or periodic scheduled backup, you can specify the following options: Backup type (full or incremental), Network load, and Backup protection. With split mirror or snapshot backup, in the case of ZDB to disk or ZDB to disk+tape (instant recovery enabled), you specify the Split mirror/snapshot backup option. For split mirror and snapshot backups, the backup type is ignored (a full backup is performed).

Within one backup specification, you can schedule both ZDB to disk and ZDB to disk+tape, and specify a different data protection period for each individual or periodic scheduled backup.

Backup Session

When a backup session is started, Data Protector tries to allocate all needed resources, such as devices. The session is queued for as long as the required minimum resources are not yet available. Data Protector tries to allocate the resources for a specific period of time, the timeout. Timeout is user configurable. If the resources are still unavailable after the timeout, the session is aborted.

Optimizing Backup Performance

To optimize the load on the Cell Manager, Data Protector by default starts five backup sessions at the same time. If more are scheduled at the same time, the excessive sessions are queued and started subsequently as the others are finished.

Scheduling Tips and Tricks

The sections “Full and Incremental Backups” on page 66 and “Keeping Backed Up Data and Information About the Data” on page 73 describe the concept of backup generations, data protection, and catalog protection.

This section combines all these concepts by giving some examples of backup schedules and some tips for efficient scheduling.

When to Schedule Backups

Typically, you schedule backups to run during lowest user activity, usually at night. Full backups take the most time, so schedule them at weekends.

Consider scheduling full backups for different clients (backup specifications) on different days, as shown in “Staggering Full Backups”.

NOTE

Data Protector offers reports that show available time slots from a device-usage point of view. This allows you to pick a time where the devices to use are not likely to be occupied by serving already existing backups.

Staggering Full Backups

Performing a full backup of all systems during the same day may cause network load and time window problems. To avoid these problems, use the staggered approach for full backups.

Table 2-6

The Staggered Approach

	Mon	Tue	Wed	...
system_grp_a	FULL	Inc1	Inc1	...
system_grp_b	Inc1	FULL	Inc1	...
system_grp_c	Inc1	Inc1	FULL	...

Optimizing for Restore

The combination of your scheduling policy with full and incremental backups highly influences the time needed to restore your data. This is illustrated in three examples in this section.

For a point-in-time restore, you need a full backup plus all the incremental backups to the desired point in time. Since full and incremental backups are typically not on the same media, you may need

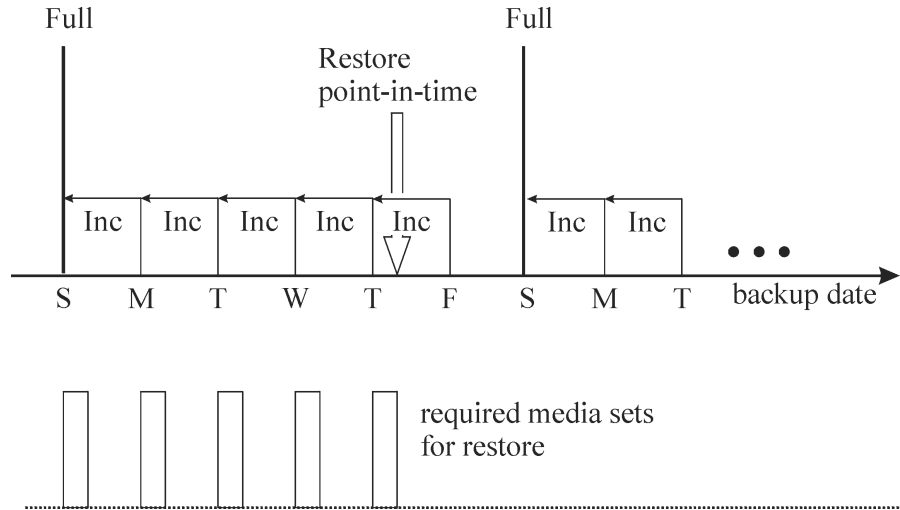
to load different media for the full and each incremental backup. Refer to “Selecting Media for Backups” on page 126 for more information on how Data Protector selects media for backups.

Example 1

Figure 2-10 depicts a scheduling policy based on a full backup plus differential backups.

Figure 2-10

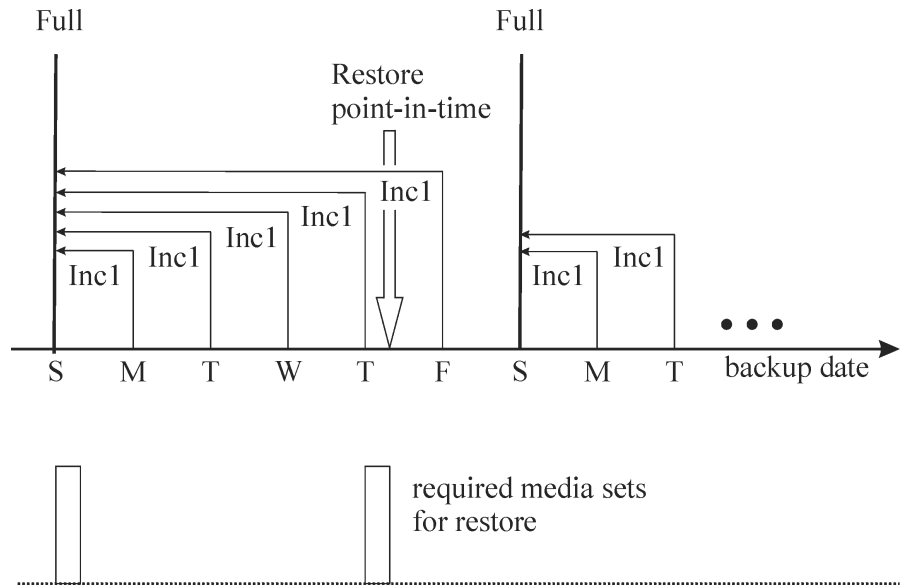
Full Backup with Daily Differential Backups



This policy reduces the media space and time needed for backing up, because you only back up changes from the previous day. However, to restore files from a Thursday backup, you need to provide the media for the full and each of the incremental backups until Thursday, that is five media sets. This complicates and slows down the restore.

Example 2 Figure 2-11 depicts a scheduling policy based on a full backup plus level one incremental backups.

Figure 2-11 Full Backup with Daily Level 1 Incremental Backups



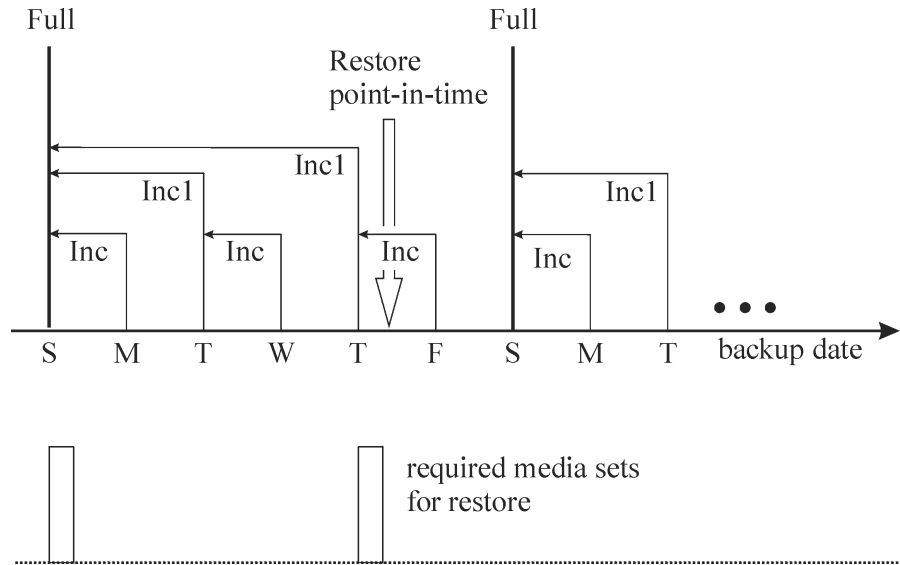
This policy requires slightly more time for backups and also requires a little more media since you back up all the changes from the last full backup every day. To restore files from Thursday's backup, you need to provide media for the full and for Thursday's incremental backup, that is, two media sets only. This considerably simplifies and speeds up the restore.

Example 3

Depending on your environment and requirements, the best solution could lie somewhere in between. For example, you may have the following scheduling policy:

Figure 2-12

Full Backup with Mixed Incremental Backups



This policy takes into account the fact that there are not many changes during weekends. Data is backed up using a combination of differential backups and level one incremental backups, to optimize backup performance. To restore files from Thursday's backup, you need to provide media for the full backup and the second incremental level one backups, that is, two media sets.

Automated or Unattended Operation

To simplify operation and the operator's involvement in the backup process, Data Protector provides extensive functionality supporting unattended or automatic backup during lights-out time. This section describes how to plan your scheduling policies, how these policies influence the behavior of backup, and provides examples of scheduling policies. This section focuses on longer periods of unattended operation spanning from several days to weeks, rather than the unattended operation during a single backup.

Considerations for Unattended Backups

Data Protector provides simple ways of scheduling your backups. Since the effectiveness of scheduling policies depends on your environment, you need to plan before finding the best scheduling policy.

- When is the lowest system usage and user activity?
Typically, this is at night and most backups are scheduled to run during the night. Data Protector can generate reports about devices used for backup.
- What kind of data do you have and how often do you want to schedule backups of this data?
Data that changes often and is important to the company, such as user files, transactions, and databases must be backed up regularly. System-specific data, such as program files that do not change often, do not need to be backed up so often.
- How much do you want to simplify restore?
Depending on how you schedule your full and incremental backups, you will need media from the full and incremental backups to restore the latest version of files. This may take longer or even require manual media handling if you do not have an automatic library device.
- How much data do you need to back up?
Full backups take longer than incremental backups. Backups must typically be done in a limited time-frame.

Planning Your Backup Strategy

Automated or Unattended Operation

- How many media are required?

Define a media rotation policy. Refer to “Implementing a Media Rotation Policy” on page 121. This will show if you can keep enough media inside the planned library to operate for the desired period without having to handle media manually.

- What about mount prompt handling?

Consider whether to use one or several libraries. This enables automatic operation, since Data Protector can have access to all or most of the media, hence significantly reducing the need to manually handle media. If the data volume is too large for a library, then consider using more libraries. Refer to “Large Libraries” on page 143 for more information.

- How do I handle unavailable devices?

Use dynamic load balancing or device chaining, and provide several devices when creating a backup specification. This way you avoid the failure of a backup if a device is not turned on or the system to which the device is connected is not functioning.

- How long can a backup of all data take?

Since backups must finish during a period of low network usage and when users do not use their systems, consider scheduling backups appropriately to distribute the network load caused by the backups, and to maximize the efficiency of backup sessions. This may require using the staggered approach.

If you need to back up large amounts of data and the backup window presents a problem, consider backing up to disk-based devices and using advanced backup strategies such as synthetic backup and disk staging.

- How can I prepare running applications for backups?

Many applications keep enough files open in a way that running a backup would produce an inconsistent backup. This can be avoided by using the pre-exec and post-exec scripts that can be used to synchronize the status of applications with the backup activities.

Duplicating Backed Up Data

Duplicating backed up data brings several benefits. You can copy data to improve its security and availability, or for operational reasons.

Data Protector provides the following methods of duplicating backed up data: object copy, object mirror, and media copy. See Table 2-7 for an overview of the main characteristics of these methods.

Table 2-7 Data Protector Data Duplication Methods

	Object Copy	Object Mirror	Media Copy
What is duplicated	Any combination of object versions from one or several backup sessions	A set of objects from a backup session	An entire medium
Time of duplication	Any time after the completion of a backup	During backup	Any time after the completion of a backup
Media type of source and target media	Can be different	Can be different	Must be the same
Size of source and target media	Can be different	Can be different	Must be the same
Appendability of target media	Yes	Yes	No ^a
Result of the operation	Media containing the selected object versions	Media containing the selected object versions	Media identical to the source media

- a. You can use only unformatted media, empty media, or media with expired protection as target media. After the operation, both the source and the target media become non-appendable.

Copying Objects

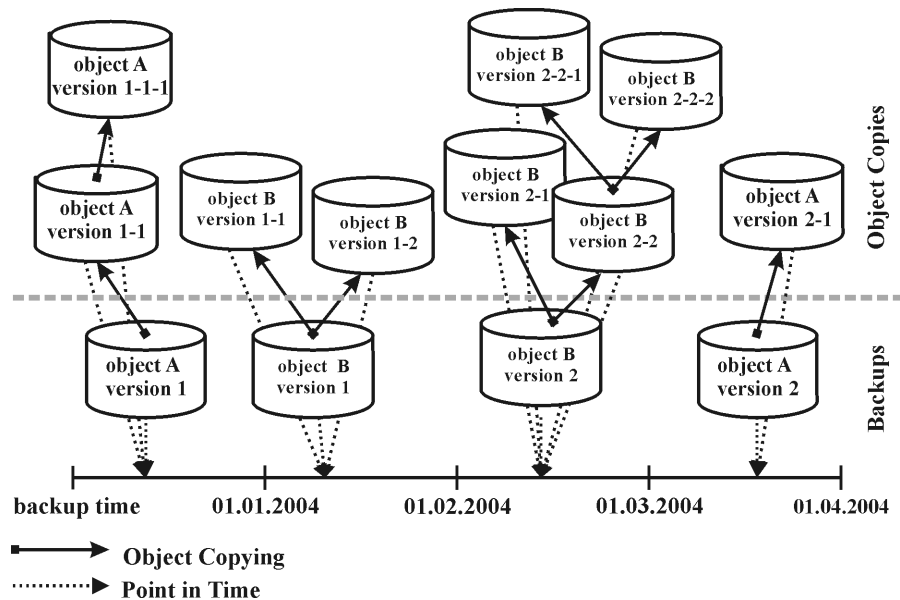
What Is Object Copy?

The Data Protector object copy functionality enables you to copy selected object versions to a specific media set. You can select object versions from one or several backup sessions or object consolidation sessions. During the object copy session, Data Protector reads the backed up data from the source media, transfers the data, and writes it to the target media.

The result of an object copy session is a media set that contains copies of the object versions you specified.

Figure 2-13 on page 90 shows how data backed up at a specific point in time can be copied afterwards. You can copy any backup object from a medium containing a backup or a medium containing a copy of the object.

Figure 2-13 Object Copy Concept



In the figure, there is an object version resulting from a backup of object A, version 1, and two additional copies of the same object version. Version 1-1 has been obtained by copying the object version resulting from the backup, and version 1-1-1 by copying a copy of the object version. Any of these object versions can be used for a restore of the same object version.

Start of Object Copy Session

You can start an object copy session interactively or specify an automated start of the session. Data Protector offers two types of automated object copying: **post-backup object copying** and **scheduled object copying**.

Post-Backup Object Copying

Post-backup object copying takes place after the completion of a session that is specified in the automated object copy specification. It copies objects selected according to the automated object copy specification that were written in that particular session.

Scheduled Object Copying

Scheduled object copying takes place at a user-defined time. Objects from different sessions can be copied in a single scheduled object copy session.

Selection of Devices

You need separate devices to be used with the source media and the target media. The destination devices can have a larger block size than the source devices. However, to avoid impact on performance, it is recommended that the devices have the same block size and are connected to the same system or to a SAN environment.

Object copying is load balanced by default. Data Protector makes optimum use of the available devices by utilizing as many devices as possible.

If you do not specify the source devices to be used in the object copy specification, Data Protector uses the default devices. By default, the devices that were used for writing the objects are used as source devices. If destination devices are not specified per object, Data Protector selects them automatically from those you selected in the object copy specification according to the following criteria in the order of priority:

- destination devices of the same block size as source devices are selected before those with a different block size
- locally attached devices are selected before network attached devices

Devices are locked at the beginning of the session. Devices that are not available at that time cannot be used in the session, as device locking after the beginning of the session is not possible. If a media error occurs, the device with errors will be avoided within that copy session.

Selection of the Media Set to Copy From

If an object version that you want to copy exists on more than one media set, which has been created using one of the Data Protector data duplication methods, any of the media sets can be used as a source for copying. You can influence the media set selection by specifying the media location priority.

The overall process of media selection is the same as for restore. For details, see “Selection of the Media Set” on page 101.

Object Copy Session Performance

An impact on object copy performance can be caused by factors such as device block sizes and the connection of devices. If the devices used in the object copy session have different block sizes, the data will be repackaged during the session, which takes additional time and resources. If the data is transferred over the network, there will be additional network load and time consumption. This impact can be minimized if the operation is load balanced.

Why Use Object Copy?

Additional copies of backed up data are created for multiple purposes:

- Vaulting
You can make copies of backed up objects and keep them in several locations.
- Freeing media
To keep only protected object versions on media, you can copy such object versions, and then leave the medium for overwriting.
- Demultiplexing of media
You can copy objects to eliminate interleaving of data.
- Consolidating a restore chain
You can copy all object versions needed for a restore to one media set.
- Migration to another media type
You can copy your backups to media of a different type.

- Support of advanced backup concepts

You can use backup concepts such as disk staging.

Vaulting

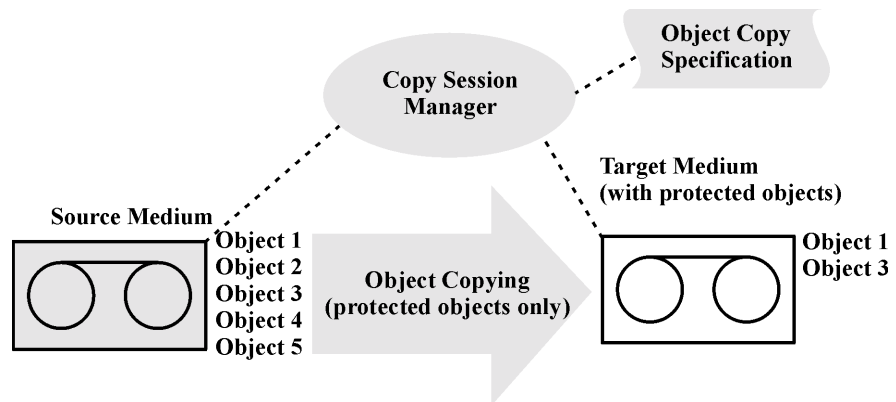
Vaulting is a process of storing media in a safe place, often called a vault, where they are kept for a specific period of time. For details, refer to “Vaulting” on page 131.

It is recommended to keep a copy of the backed up data on site for restore purposes. To obtain additional copies, you can use the object copy, object mirror, or media copy functionality, depending on your needs.

Freeing Media

You can minimize the media space consumption by keeping only protected backups and overwriting unprotected ones. As a single medium may contain both, you can copy protected objects to a new media set and leave the medium for overwriting. See Figure 2-14 on page 93.

Figure 2-14 Freeing Media

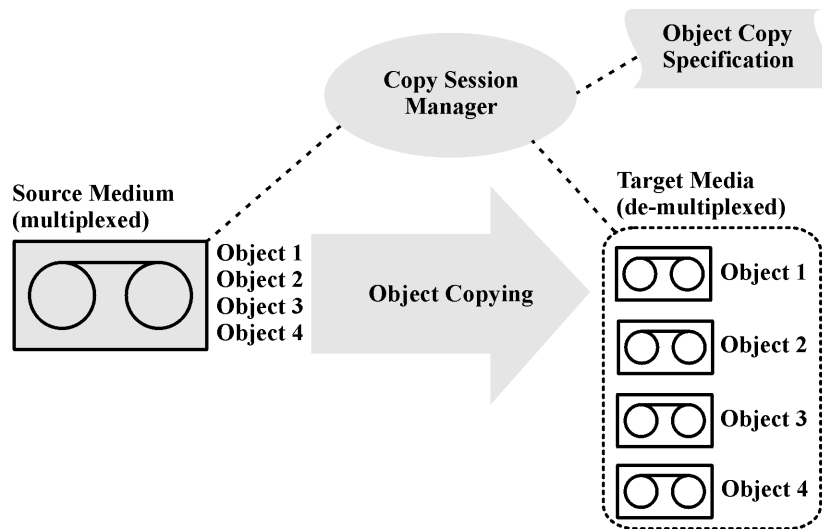


Demultiplexing of Media

Multiplexed media contain interleaved data of multiple objects. Such media may arise from backup sessions with the device concurrency more than 1. Multiplexed media may compromise the privacy of backups and require more time for restore.

Data Protector offers a possibility of demultiplexing of media. Objects from a multiplexed medium are copied to several media that you specify. See Figure 2-15 on page 94.

Figure 2-15 Demultiplexing a Medium



Consolidating a Restore Chain

You can copy a restore chain (all backups that are necessary for a restore) of an object version to a new media set. A restore from such a media set is faster and more convenient, as there is no need to load several media and seek for the needed object versions.

Migration to Another Media Type

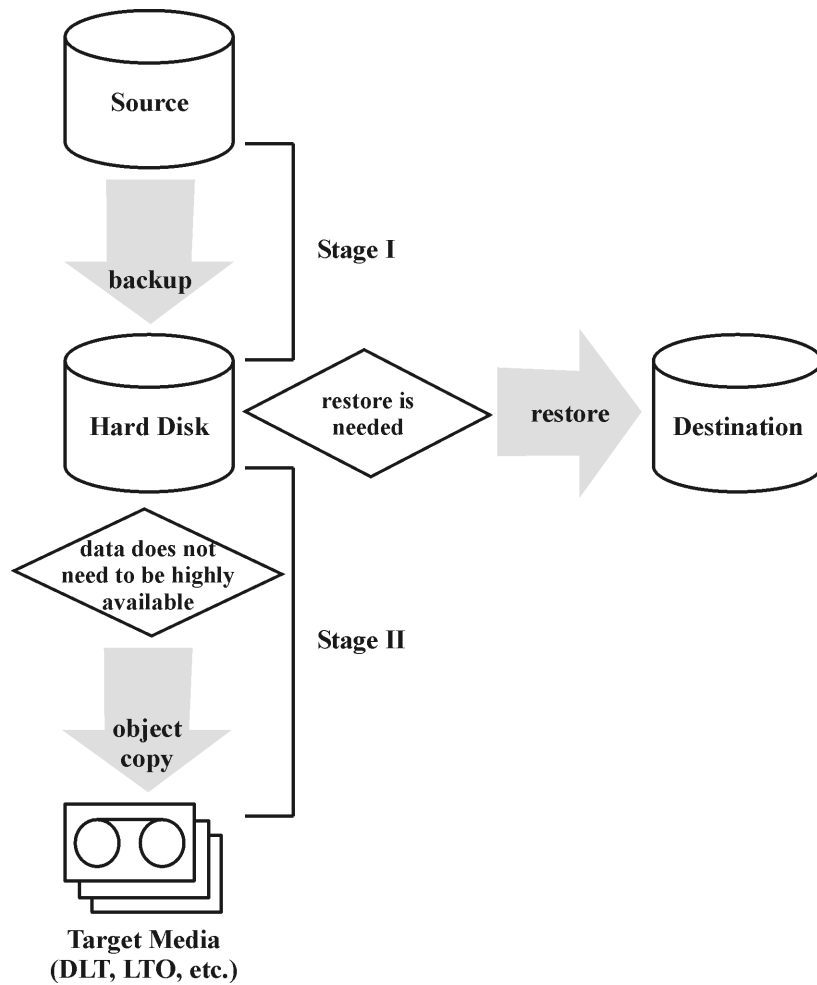
You can migrate backed up data to another media type. For example, you can copy objects from file devices to LTO devices or from DLT devices to LTO devices.

Disk Staging

The concept of disk staging is based on backing up data in several stages to improve the performance of backups and restores, reduce costs of storing the backed up data, and increase the data availability and accessibility for restore.

The backup stages consist of backing up data to media of one type and later moving the data to media of a different type. The data is backed up to media with high performance and accessibility, but limited capacity (for example, system disks). These backups are usually kept accessible for restore for a period of time when a restore is the most probable. After a certain period of time, the data is moved to media with lower performance and accessibility, but high capacity for storage, using the object copy functionality. See Figure 2-16 on page 95.

Figure 2-16 Disk Staging Concept



Disk staging also eliminates the need for frequent backups of numerous small objects to tape. Such backups are inconvenient due to frequent loading and unloading of media. The use of disk staging reduces backup time and prevents media deterioration.

Object Mirroring

What Is Object Mirroring?

The Data Protector object mirror functionality enables writing the same data to several media sets simultaneously during a backup session. You can mirror all or some backup objects to one or more additional media sets.

The result of a successful backup session with object mirroring is one media set containing the backed up objects and additional media sets containing the mirrored objects. The mirrored objects on these media sets are treated as object copies.

Benefits of Object Mirroring

The use of the object mirror functionality serves the following purposes:

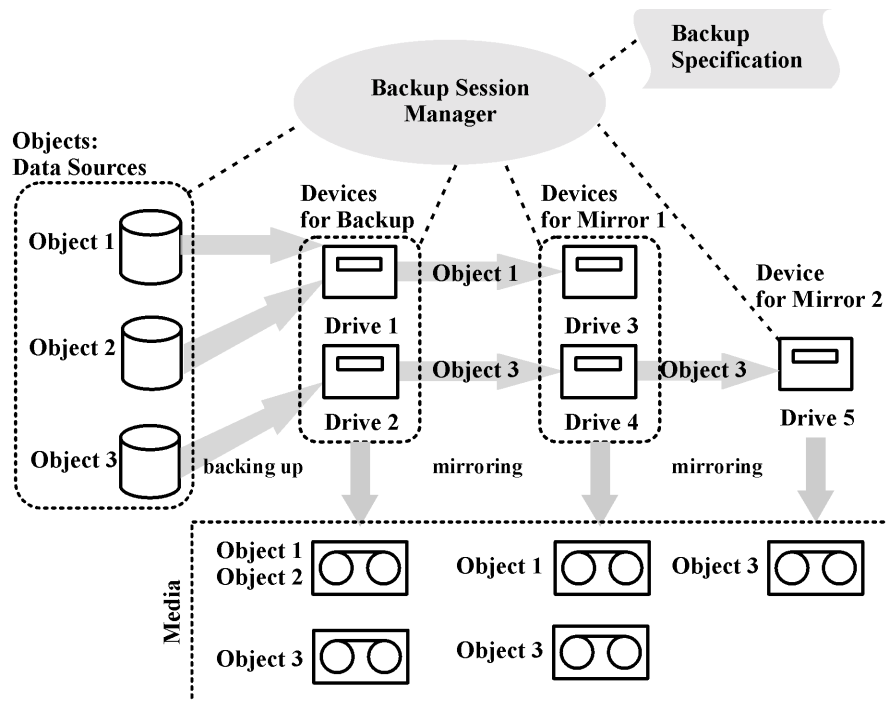
- It increases the availability of backed up data due to the existence of multiple copies.
- It enables easy multi-site vaulting, as the backed up data can be mirrored to remote sites.
- It improves the fault tolerance of backups, as the same data is written to several media. A media failure on one medium does not affect the creation of the other mirrors.

Object Mirror Operation

In a backup session with object mirroring, each selected object is backed up and at the same time mirrored as many times as specified in the backup specification. See Figure 2-17 on page 97.

Let us take Object 3 in the figure as an example. The Disk Agent reads a block of data from the disk and sends it to the Media Agent that is responsible for the backup of the object. This Media Agent then writes the data to the medium in Drive 2 and forwards it to the Media Agent that is responsible for mirror 1. This Media Agent in turn writes the data to the medium in Drive 4 and forwards it to the Media Agent that is responsible for mirror 2. This Media Agent writes the data to the medium in Drive 5. At the end of the session, Object 3 is available on three media.

Figure 2-17 Object Mirroring



Selection of Devices

Object mirroring is load balanced by default. Data Protector makes optimum use of the available devices by utilizing as many devices as possible. Devices are selected according to the following criteria in the order of priority:

- devices of the same block size are selected, if available
- locally attached devices are selected before network attached devices

When you perform an object mirror operation from the command line, load balancing is not available.

Backup Performance

Object mirroring has an impact on backup performance. On the Cell Manager and Media Agent clients, the impact of writing mirrors is the same as if additional objects were backed up. On these systems, the backup performance will decrease depending on the number of mirrors.

On the Disk Agent clients, there is no impact caused by mirroring, as backup objects are read only once.

Backup performance also depends on factors such as device block sizes and the connection of devices. If the devices used for backup and object mirroring have different block sizes, the mirrored data will be repackaged during the session, which takes additional time and resources. If the data is transferred over the network, there will be additional network load and time consumption.

Copying Media

What Is Media Copying?

The Data Protector media copy functionality enables you to copy media after a backup has been performed. Media copying is a process that creates an exact copy of a medium containing a backup. You can use it to duplicate media for archiving or vaulting purposes. After the media have been copied, you can move either the original media or the copies to an off-site vault.

Besides manually started media copying, Data Protector also offers automated media copying. For more information, refer to “Automated Media Copying” on page 99.

How to Copy Media

You need two devices of the same media type, one for the source medium and one for the target medium. The source medium is the medium being copied while the target medium is the medium to which data is copied.

When you copy media within a library that has multiple drives, you can use one drive for the source and one for the copy.

What Is the Result?

The result of copying media is two identical sets of media, the original media set and the copy. Either of them can be used for restore.

After the source medium has been copied, Data Protector marks it as non-appendable to prevent appending new backups (this would result in the original being different from its copy.) The copy is also marked as non-appendable. The default protection of the copy is the same as for the original.

You can make multiple copies of the original media. You cannot, however, make copies of copies, also known as second generation copies.

Automated Media Copying

What Is Automated Media Copying? Automated media copying is an automated process that creates copies of the media containing backups. This functionality is available with library devices.

Data Protector offers two types of automated media copying: post-backup media copying and scheduled media copying.

Post-Backup Media Copying Post-backup media copying takes place after the completion of a backup session. It copies the media used in that particular session.

Scheduled Media Copying Scheduled media copying takes place at a user-defined time. Media used in different backup specifications can be copied in a single session. You create an automated media copy specification to define which media will be copied.

How Does Automated Media Copying Operate? First you create an automated media copy specification. When the automated media copy session begins, Data Protector generates a list of media, referred to as source media, based on the parameters specified in the automated media copy specification. For each source medium, a target medium is selected to which the data will be copied. The target media are selected from the same media pool as the source media, from a free pool, or from the blank media in a library.

For each source medium, Data Protector selects a pair of devices from the devices that you specified in the automated media copy specification. The automated media copy functionality provides its own load balancing. Data Protector tries to make optimum use of the available devices by utilizing as many devices as possible and selecting local devices, if they are available.

The automated media copy functionality does not handle mount or cleanme requests. If a mount request is received, the media pair concerned is aborted, but the session continues.

For examples of use, refer to “Examples of Automated Media Copying” on page B-5.

Restoring Data

Policies for restoring data are an essential part of the overall backup strategy in the company. Keep the following in mind:

- Backing up and restoring files is essentially the same as copying files. Therefore, ensure that only authorized people have the rights to restore confidential data.
- Ensure that unauthorized people cannot restore files of other people.

This section describes some possible implementations of the restore policy using Data Protector. You can restore your filesystem data by browsing through restore objects or restore sessions. By default, data is restored to its original location. However, you can specify any location to be the destination of restored data.

Restore Duration

After data loss, access to data is possible only after the recovery process is finished. It is often critical to minimize restore duration so that users can do their regular work. Therefore, plan for the time needed to restore specific data.

Factors Affecting Restore Duration

The restore duration depends on a number of factors, such as:

- The amount of data to be restored. This also directly influences all the following items.
- A combination of full and incremental backups. Refer to “Full and Incremental Backups” on page 66 for more information.
- Media and devices used for backup. Refer to Chapter 3, “Media Management and Devices,” on page 109 for more information.
- Speed of networks and systems. Refer to “Understanding and Planning Performance” on page 41 for more information.
- The application you are recovering, for example, Oracle database files. Refer to the appropriate *HP OpenView Storage Data Protector Integration Guide* for more information.

- The use of parallel restore. Several objects can be restored with a single read operation, depending on how the data was backed up. Refer to “Parallel Restores” on page 220.
- Speed and ease of selecting the data to be restored, which depends on the logging level settings used during the backup and on catalog protection time. Refer to “Logging Level as an IDB Key Tunable Parameter” on page 183.

Selection of the Media Set

If an object version that you want to restore exists on more than one media set, which has been created using one of the Data Protector data duplication methods, any of the media sets can be used for the restore. By default, Data Protector automatically selects the media set that will be used. You can influence the media set selection by specifying the media location priority. You can also manually select the media set you want to use for the restore, except when restoring integration objects.

Media Set Selection Algorithm

By default, Data Protector selects the media set with the best availability and quality. For example, Data Protector avoids media sets with missing media or poor media; it considers the completion status of the objects, the availability and locality of the device to be used with a certain media set, and so on. A media set located in a library is used before one in a standalone device.

Selection of Restore Chain

If you use synthetic backup, there is often more than one restore chain for the same point in time of an object. By default, Data Protector selects the most convenient restore chain and the most appropriate media within the selected restore chain.

Media Location Priority

To influence the selection of the media set, specify the media location priority. This is important if you use the concept of multi-site storage. If you keep media at different sites, you can specify which location is preferable for a specific restore. Data Protector will use the media set with the highest priority if more than one media set matches the conditions of the selection algorithm.

You can set the media location priority globally or for a specific restore session.

Operators Are Allowed to Restore

A popular restore policy is that only dedicated backup operators or network administrators have the right to restore files or perform disaster recovery.

When to Use This Policy

Use this policy in the following cases:

- In a large network environment where it is best to have a dedicated person to do such jobs.
- In an environment where end users do not have the necessary computer knowledge to restore files, operators can be trusted to restore sensitive data.

What Needs to Be Done

You need to do the following, to implement this policy:

- Add the backup operators or network administrators that will restore data for other people to the Data Protector operators or admin user group.

You do not need to add other people (such as users who want to perform restores to their own systems) to any Data Protector user group.

- During installation, do not install the Data Protector user interface on end-user systems. Install the Disk Agent that allows Data Protector to back up these systems.
- Establish a policy of handling requests for restore. This policy should cover how end users request the restore of files, for example, via email containing all the details necessary for the operator to locate and restore the files back to the end-user system. The end users should also have a way of knowing when the files have been restored.

End Users Are Allowed to Restore

Another possible restore policy is to allow all or just selected end users to restore their own data. This policy provides sufficient security and may relieve the backup operator from doing a number of restore operations.

When to Use This Policy

Use this policy in the following cases:

- When the end users have sufficient knowledge to handle restores. You may need to provide some training for the users on basic backup concepts and restore operations.
- You use library backup devices with media of most recent backups. The end user Data Protector user group, by default, does not allow end users to handle mount requests for needed media. The end users will still need the assistance of the backup operator in case of mount requests. This can be avoided by using large libraries.

What Needs to Be Done

You need to do the following to implement this policy:

- Add the end users that are allowed to restore their own data to the Data Protector `end_users` user group. For additional security, you may limit the Data Protector access of these users, to a specific system only.
- Install the Data Protector user interface on the systems the end users are using. Data Protector automatically checks the user rights and allows restore functionality only.
- When you configure backups of the end-user systems, make backups visible to the end users by setting the Data Protector `public` option.

Disaster Recovery

This section provides only a short overview of the disaster recovery concepts. Detailed disaster recovery concepts, planning, preparation, and procedures are described in the *HP OpenView Storage Data Protector Disaster Recovery Guide*.

A **computer disaster** refers to any event that renders a computer system unbootable, whether due to human error, hardware or software failure, natural disaster, etc. In these cases it is most likely that the boot or system partition of the system is not available and the environment needs to be recovered before the standard restore operation can begin. This includes repartitioning and/or reformatting the boot partition and recovery of the operating system with all the configuration information that defines the environment. *This has to be completed in order to recover other user data.*

After a computer disaster has occurred, the system (referred as **target system**) is typically in a non-bootable state and the goal of Data Protector disaster recovery is to restore this system to the original system configuration. The difference between the crashed and the target system is that the target system has all faulty hardware replaced.

A disaster is always serious, however the following factors can exacerbate the situation:

- The system needs to be returned to online status as quickly and efficiently as possible.
- Administrators are not familiar with the required steps to perform the disaster recovery procedure.
- The available personnel to perform the recovery have only fundamental system knowledge.

Disaster recovery is a complex task that involves extensive planning and preparation before execution. You need to have a well-defined, step-by-step process in place to prepare for, and recover from, disastrous situations.

The **disaster recovery process** consists of 4 phases:

1. **Phase 0** (planning/preparation) is the prerequisite for a successful disaster recovery.

CAUTION

It is too late to prepare for a disaster recovery once a disaster has occurred.

2. In **Phase 1**, DR OS is installed and configured, which usually includes repartitioning and reformatting of the boot partition, since the boot or system partition of the system are not always available and the environment needs to be recovered before normal restore operations can resume.
3. In **Phase 2**, the operating system with all the configuration information that defines the environment with Data Protector (as it was) is restored.
4. Only after phase 2 is completed, is the restore of applications and user data possible (**Phase 3**). A well-defined, step-by-step process has to be followed to ensure a fast and efficient restore.

Disaster Recovery Methods

Data Protector supports the following disaster recovery methods:

- Manual disaster recovery

This is a basic and very flexible disaster recovery method. You need to install and configure the DR OS. Then use Data Protector to restore data (including the operating system files), replacing the operating system files with the restored operating system files.

- Automated disaster recovery

Automated System Recovery (ASR) is an automated system on Windows systems, which reconfigures a disk to its original state (or resizes the partitions if the new disk is larger than the original disk) in the case of a disaster. ASR thus enables the Data Protector `drstart.exe` command to install the active DR OS that provides Data Protector disk, network, tape and file system access.

- Disk delivery Disaster recovery

On Windows clients, the disk of the crashed system (or the replacement disk for the physically damaged disk) is temporarily connected to a hosting system. After being restored, it can be connected to the faulty system and booted. On UNIX systems, the

auxiliary disk with a minimal operating system, networking, and Data Protector agent installed is used to perform Disk Delivery Disaster Recovery.

- **Enhanced Automated Disaster Recovery (EADR)**

Enhanced Automated Disaster Recovery (EADR) is a fully automated Data Protector recovery method for Windows clients and Cell Manager, where user intervention is reduced to minimum. The system is booted from the disaster recovery CD ISO image and Data Protector automatically installs and configures DR OS, formats and partitions the disks, and finally recovers the original system with Data Protector as it was at the time of backup.

- **One Button Disaster Recovery (OBDR)** is a fully automated Data Protector recovery method for Windows clients and Cell Manager, where user intervention is reduced to a minimum. The system is booted from the OBDR tape and automatically recovered.

For a list of supported disaster recovery methods for a particular operating system, see the Support Matrices in the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References* or on the Web:

http://www.openview.hp.com/products/datapro/spec_0001.html

Alternative Disaster Recovery Methods

This section compares the Data Protector disaster recovery concept with concepts of other vendors. This section points out only significant aspects of alternative recovery concepts. Two alternative recovery approaches are discussed:

Recovery Methods Supported by Operating System Vendors

Most vendors provide their own methods, but when it comes to restore, they typically require the following steps:

1. Re-install the operating system from scratch
2. Re-install the application(s)
3. Restore application(s) data

Excessive manual reconfiguration and customization of the operating system and the application(s) is required to reconstruct the status before the disaster. This is a very complicated, time consuming, and error-prone process using different tools that are not integrated with each other. It does not benefit from a backup of the operating system, the application(s), and their configurations as a whole set.

Recovery Using Third-Party Tools (for Windows)

This often consists of a special tool that backs up the system partition as a snapshot, which can be restored rapidly. The method conceptually requires the following steps:

1. Restore the system partition (using the third-party tool)
2. Restore any other partition (perhaps selective) if required using the standard backup tool

It is obvious that one has to work from two different backups with different tools. This is a difficult task to perform on a regular basis. If this concept is implemented for a large organization, the administrative overhead to manage the different versions (weekly backup) for the data from two tools must be addressed.

Data Protector on the other hand represents a powerful all-in-one cross-platform enterprise solution for fast and efficient disaster recovery that includes backup and restore and supports clustering. It provides

easy central administration, easy restore, high availability support, monitoring, reporting and notifications to aid administration of systems in a large organization.

3 Media Management and Devices

In This Chapter

This chapter describes Data Protector concepts of media and device management. It discusses media pools, devices, and large libraries.

It is organized as follows:

“Media Management” on page 111

“Media Life Cycle” on page 113

“Media Pools” on page 114

“Media Management Before Backups Begin” on page 124

“Media Management During Backup Sessions” on page 126

“Media Management After Backup Sessions” on page 131

“Devices” on page 134

“Standalone Devices” on page 141

“Small Magazine Devices” on page 142

“Large Libraries” on page 143

“Data Protector and Storage Area Networks” on page 152

Media Management

Serious challenges can arise when administrating large quantities of media in an enterprise environment. Data Protector media management functionality allows for a flexible and efficient allocation of backup data to media. This can be done in many ways by defining methods of automatic or strict media allocation.

Media Management Functionality

Data Protector provides the following media management functionality that allows simple and efficient management of a large number of media:

- Grouping media into logical groups, media pools, that enable you to think about large sets of media without having to worry about each medium individually.
- Data Protector keeps track of all media and the state of each medium, the data protection expiration time, the availability of media for backups, and a catalog of what has been backed up to each medium.
- Automated media rotation policies so that you do not need to take care of tape rotation manually.
- The possibility to explicitly define which media and which devices you want to use for backup.
- Optimized media management for specific device types, such as standalone, magazine, library devices and large silo devices.
- Fully automated operation. If Data Protector has control of enough media in the library devices, the media management functionality enables the running of backups without the need for an operator to handle media for weeks.
- Recognition and support of barcodes on large libraries with barcode support and silo devices.
- Automatic recognition of Data Protector media format and other popular tape formats.
- Data Protector only writes to blank media initialized (formatted) by Data Protector. You cannot force Data Protector to overwrite foreign tape formats during a backup, thus you avoid accidental overwrites of media that belong to other applications.

Media Management

- Recognition, tracking, viewing, and handling of media used by Data Protector and separating it from media used by other applications in library and silo devices.
- Keeping information about the media used in a central place and sharing this information among several Data Protector cells.
- Support for media vaulting.
- Interactive or automated creation of additional copies of the data on the media.

This chapter describes the above functionality in more detail.

Media Life Cycle

A typical media life cycle consists of the following steps:

1. Preparing media for backup.

This includes initializing (formatting) media for use with Data Protector and assigning media to media pools, which are used to track the media.

Refer to “Media Management Before Backups Begin” on page 124 for more information.

2. Using media for backup.

This defines how media are selected for backup, how the condition of the media is checked, how new backups are added to the media, and when data on the media is overwritten.

Refer to “Media Management During Backup Sessions” on page 126 for more information.

3. Vaulting media for long-term data storage. You can use one of Data Protector’s data duplication methods to make copies of the backed up data for vaulting purposes.

Refer to “Media Management After Backup Sessions” on page 131 for more information on vaulting.

4. Recycling media for new backups once the data on the media is no longer needed.

5. Retiring media.

Once a medium has expired, it is marked poor and will no longer be used by Data Protector.

Refer to “Calculating Media Condition” on page 130.

Media Pools

Data Protector media pools manage large numbers of media, hence reducing the management effort for the administrators to a minimum.

What Is a Media Pool?

A pool is a logical set, or group, of media with a common usage pattern and media properties. It can only have media of the same physical type. DLT and DAT/DDS media cannot be in the same pool for instance.

The current location of a medium has no influence on its relation to the pool. Whether the medium is in a drive, in a repository slot of a library, in the vault or somewhere else, does not matter; it always belongs to its pool until it is recycled and exported from the cell.

Several devices can use media from the same pool.

Media Pool Property Examples

Examples of pool properties are:

- appendable

This allows Data Protector to append data to the media in this pool when performing subsequent backup sessions.

If this option is not selected, then the media will contain data from a single session only.

- append incrementals only

A backup session appends to a medium only if an incremental backup is performed. This allows you to have a complete set of full and incremental backups on the same medium, if there is enough space.

- media allocation policy

There are several levels of strictness as to which media can be used for backup. They range from strict, where Data Protector requires a specific medium, to loose, where Data Protector accepts any suitable medium in the pool, including new (blank) media.

Every device is linked to a default pool. This pool can be changed in the backup specification.

For information on other media pool properties, refer to online Help, index keyword “media pools, properties of“.

**Media Pools and
dcbf Directories**

Data Protector allows you to set a target dcbf directory for a media pool. This means that information about all media from the media pool is stored in the specified dcbf directory.

For information on the DCBF part of the IDB and dcbf directories, refer to “IDB Architecture” on page 173.

**How to Use Media
Pools**

The usage of pools depends mainly on your preferences. For example, pools can be defined using criteria like:

- system platform (one pool for UNIX systems, one for Windows 2000 systems, and one for Windows XP systems)
- per system (every system has its own pool)
- organizational structure (all systems in department_A have a pool, and systems in department_B have another pool)
- systems categories (running large databases, or business critical applications)
- backup type (all full backups use one pool, and all incremental backups use another pool)
- combinations of the above criteria, and more.

A simplified way to think about media pools is to view them as a destination for your backup while you look at the devices as a transfer mechanism between the data and the media pools.

The relationship of a pool to a system category is defined by putting certain systems into the same backup specification and also specifying the pool(s). The options used (when defining the devices, pools, and backup specifications) determine how the data of the objects will end up on the media.

Grouping such media used for a similar kind of backup to media pools allows you to apply common media handling policies on a group level while not bothering with each medium individually. All media in a pool are tracked as one set and have the same media allocation policy.

**Default Media
Pools**

Data Protector provides default media pools for various media types. These default media pools allow you to quickly run backups without having to create your own media pools. However, to efficiently manage your large environment, create different media pools for specific needs. When you run a backup, specify which media pool to use.

Free Pools

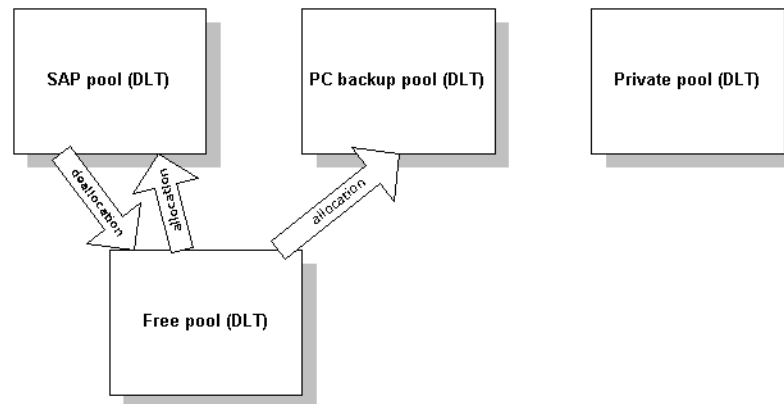
If media allocated to a specific media pool run out, you cannot use the media in another pool, even if the media are of the same type. This can result in unnecessary mount requests and operator intervention. To solve this problem, you can use the single pool model, at which all media are in the same pool. While this allows you to share free media, it compromises the benefits of using media pools in the first place: easier media management, separation of important from not so important data, etc. To alleviate this drawback, free pools are used.

What Is a Free Pool?

A free pool is an auxiliary source of media of the same type (for example, DLT) for use when all free media in a regular pool run out. It helps to avoid failed backups due to missing (free) media.

Figure 3-1

Free Pools



When Is a Free Pool Used?

Media are moved between regular and free pools on two events (Figure 3-1):

- Allocation. Media are moved from a free pool to a regular pool
- Deallocation. Media are moved from a regular pool to a free pool. You can specify in the GUI whether deallocation is done automatically. Media from the PC backup pool in Figure 3-1, for example, are not automatically deallocated.

Protected (allocated, used) media belong to a specific regular pool (like the SAP pool), while free Data Protector media can be (automatically) moved to a free pool. This free pool is later used for allocation of free media for all pools that are configured to use this free pool.

Some regular pools, for example the Private pool in Figure 3-1, can also be configured not to share any media with free pools.

Free Pool Benefits A free pool has the following benefits:

- Sharing of free media between pools

All free (unprotected, empty) media can be grouped in a free pool and shared between all media pools that support free pool usage.

- Reduced operator intervention for backup

Assuming that all free media are shared, the need for mount requests is reduced.

Free Pool Properties

A free pool:

- can be created manually or automatically when you configure the use of one. You cannot delete free pools if they are linked to a normal pool or are not empty.
- is different from a regular pool in that it does not provide allocation policy options.
- contains only Data Protector media (no unknown or blank media).

Media Quality Calculation

Media quality is calculated equally between pools. That means that medium condition factors will be configurable for a free pool only and will be inherited by all pools using the free pool.

Free Pool Limitations

Free pools have the following limitations:

- You cannot select different condition factors for each pool. Instead, all pools that use a free pool use condition factors configured for this free pool.
- You cannot move protected media to a free pool and unprotected media to a regular pool that has automatic deallocation configured.
- You cannot use some operations such as Import, Copy and Recycle on media in a free pool.
- Pools with magazine support cannot use a free pool.

- You may experience some temporary inconsistencies in pools when using free pools, for example, when there is an unprotected medium in a regular pool waiting for the de-allocation process.
- If you change the protection of media after its expiry (for example to Permanent), though the media may be in a free pool, they are not allocated for backup.
- When allocated from a free pool, media with different data format type can be used and are automatically reformatted, for example NDMP media are reformatted to normal media.

For further information on free pools, refer to Data Protector online Help index keyword “free pools, characteristics”.

Media Pool Usage Examples

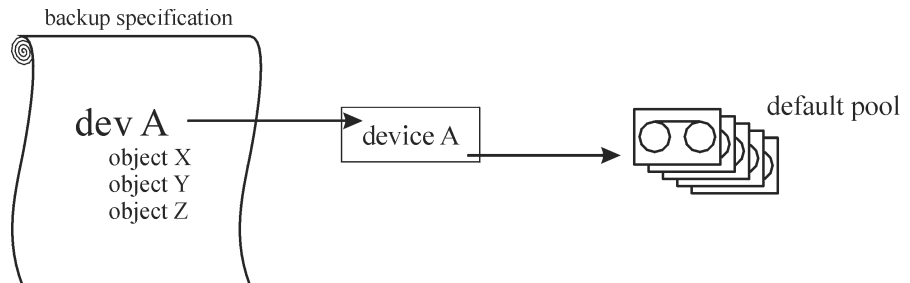
The examples below show some configurations you may want to consider when choosing the appropriate strategy for a particular backup environment.

Example 1

In the model shown in Figure 3-2, all objects are backed up to the same media pool. The backup specification does not reference a pool, so the default pool is used, which is part of the device definition.

Figure 3-2

A Simple One Device/One Media Pool Relation



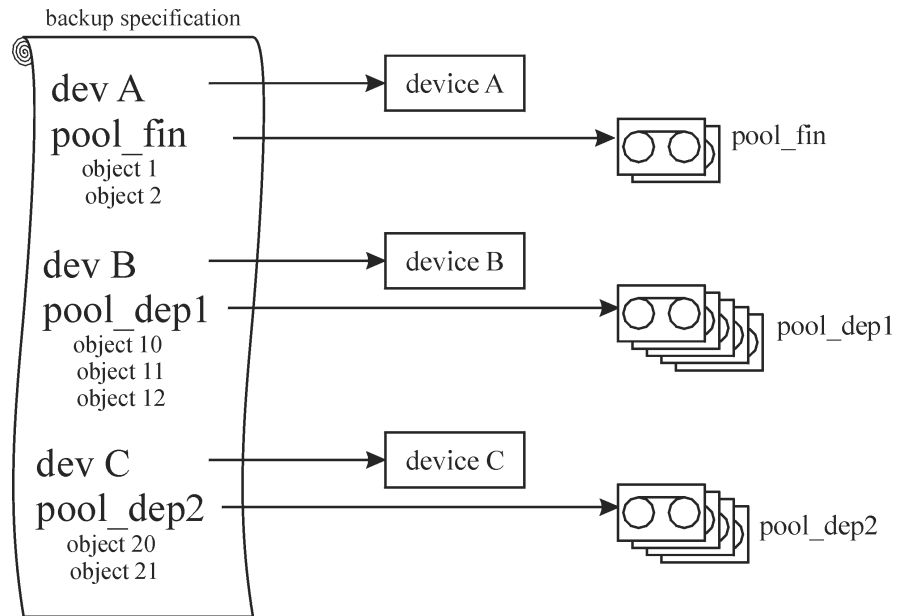
Example 2

Large library devices contain a number of physical drives and media used by different departments or applications. You can configure a media pool for each department, as shown in Figure 3-3, and decide which drive in the library will handle the actual data transfer. The arrow pointing from a backup specification to a media pool indicates that you defined a

target media pool in a backup specification. If you do not specify a media pool in the backup specification, the default pool, specified in the device definition, is used.

Refer to “Large Libraries” on page 143 for details about the relation between media pools and large library devices.

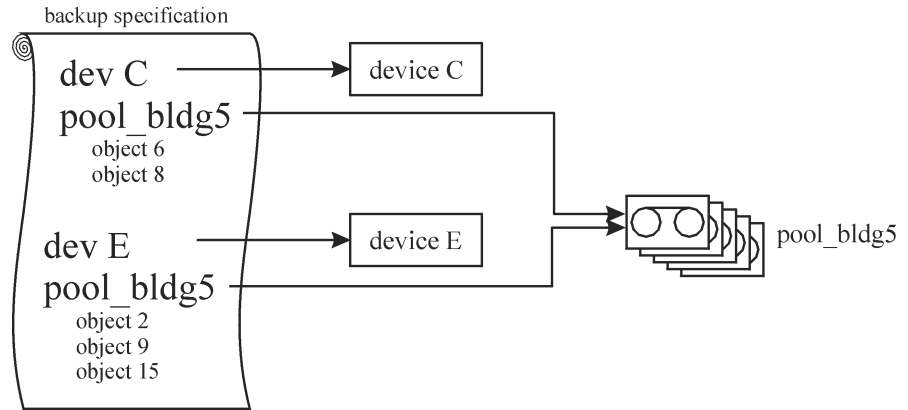
Figure 3-3 Configuration of Media Pools for Large Libraries



Example 3 Figure 3-4 shows an example when data is backed up to media in a media pool with multiple devices simultaneously. Higher performance is achieved due to the use of several devices in parallel, regardless of which pool is used.

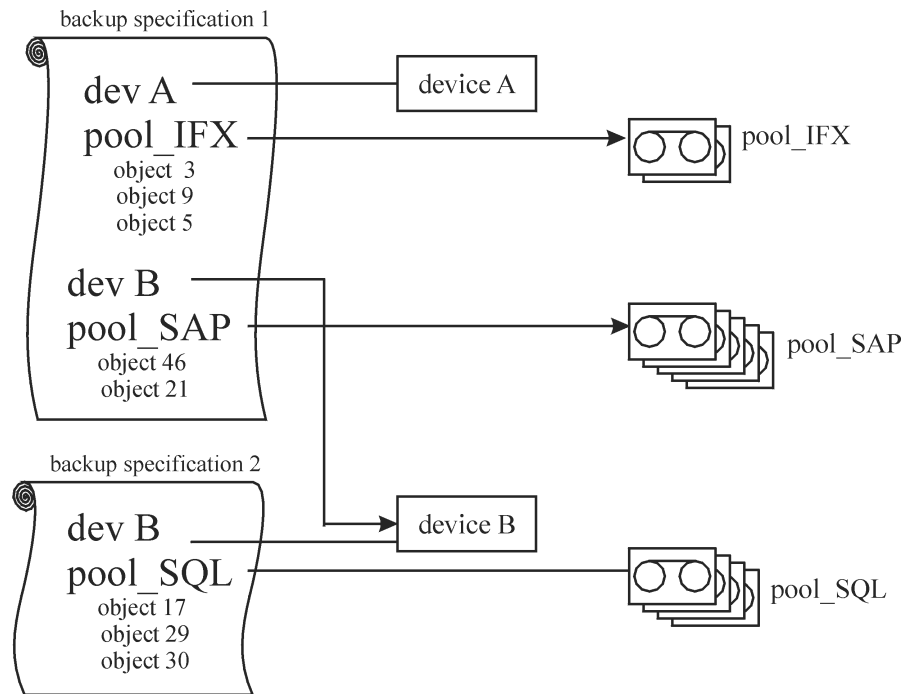
Refer to “Device Lists and Load Balancing” on page 135 for more information.

Figure 3-4 Multiple Devices, Single Media Pool



Example 4 Data is backed up to media in multiple media pools on multiple devices simultaneously. If you want to use the same device with different pools, you need to create several backup specifications. In the example below, a separate media pool is dedicated to each database application.

Figure 3-5 Multiple Devices, Multiple Media Pools



Implementing a Media Rotation Policy

What Is a Media Rotation Policy?

A media rotation policy defines how media are used during backup, including the following. In defining a media rotation policy, answer the following questions:

- How many backup generations are needed?
- Where are media stored?
- How often media are used?
- When can media be overwritten and re-used for new backups?
- When are media old enough to be replaced?

Traditional backup strategies used with older backup tools required a thoroughly planned and well defined media rotation policy controlled by the administrator rather than a backup application. With Data

Protector, you can implement a rotation policy by specifying usage options such that media selection for subsequent backups is done automatically.

Media Rotation and Data Protector

Data Protector automates media rotation and media handling as follows:

Automatic Media Rotation and Media Handling

- Because media are grouped into media pools, you no longer need to manage single media. Data Protector automatically tracks and manages each single medium in the media pools.
- You do not need to decide to which media the backed up data is to be written to; Data Protector does that for you. You back up to a media pool.
- Data Protector automatically selects media from a media pool according to the media allocation policy and usage options you specified. You can also disable the automatic selection and perform manual media selection.
- The location of media is tracked and displayed in the Data Protector user interface as long as the media are configured in Data Protector.
- Data Protector automatically tracks the number of overwrites on the media and the age of the media and thus tracks the condition of the media.
- Data Protector provides a security mechanism so that media with protected data do not get overwritten accidentally by Data Protector.

Media Needed for Rotation

Estimating the Quantity of Needed Media

The following helps to estimate the quantity of media you might need for a full rotation:

- Determine if the media capacity can be used fully or if some media are non-appendable and can only be used partially.
- Determine the systems that will be backed up and the media space required for the related data. For example, you can use backup preview.
- Determine the backup frequency, such as the number of incremental backups between two full backups.
- Determine the quantity of media needed for one backup generation, where a backup generation contains a full backup and a sequence of incremental backups up to the next full backup. Consider also hardware compression if you have planned to use it with the devices.
- Determine for how long the media will remain protected.
- Calculate the number of backup generations that will have been created before the first backup generation can be overwritten.

By now you should be able to estimate the quantity of media required for a full media rotation. Additional media will be required in case you:

- Assume 10% overhead added by Data Protector to the data on the media for directory and file information. This information is already calculated in the backup preview size.
- After the media no longer fulfill the usage criteria, they need to be replaced.
- Expect some growth in the volume of data to be backed up.

Media Management Before Backups Begin

Before you can use media for backup, media must be initialized, or formatted, for use with Data Protector. You can either initialize (format) media manually, or you can let Data Protector automatically initialize (format) media when the media are selected for backup. Refer to “Selecting Media for Backups” on page 126.

Initializing or Formatting Media

What Is Initializing (Formatting) Media?

Before Data Protector uses media for backup, it initializes (formats) the media. This saves the information about each medium (medium ID, description and location) in the IDB and also writes this information on the medium itself (to the medium header). When you initialize (format) media, you also specify to which media pool the media belong.

If media are not initialized (formatted) before backup, Data Protector can initialize (format) blank media during backup with the default labels, if the pool policy is set accordingly. The first backup to such media will take more time. Refer to “Selecting Media for Backups” on page 126 for more information.

Labeling Data Protector Media

How Data Protector Labels Media?

When you add media for use with Data Protector by initializing (formatting) media, you must specify the media label which helps you identify the media later. If a device has a barcode reader, the barcode is automatically displayed as a prefix of the medium description. A barcode provides a unique ID for each medium in the IDB. You can optionally use the barcode as medium label during the initialization of the medium.

Data Protector also assigns each medium a media ID that uniquely identifies this medium.

An ANSI X3.27 label is also written on the tape for identification on other systems. Data Protector writes these labels with other information to a medium header and to the IDB.

If you change the medium label, Data Protector modifies the medium label in the IDB and not on the medium itself. Therefore, if you export and import media that have not been updated, the medium label in the IDB is replaced with the medium label from the media. The media label on the tape can be changed only by re-initializing (formatting) the media.

How Are Labels Used?

These labels identify the medium as a Data Protector medium. When loading a medium for backup or restore, Data Protector checks the medium for the medium ID. The media management system maintains the information about this medium, which tells Data Protector whether the requested action is allowed for this medium. For example, if you try to write a new backup to this medium, the media management system checks whether the data protection for the data already contained on this medium has expired. The user defined label is used to identify a specific medium.

Location Field

Backup media are usually stored in different locations. For example, a backup needs to be available on site for fast restore access, whereas a medium containing a copy of the backed up data is often stored off-site for safety reasons.

Data Protector provides a location field for each medium, which can be used freely by the operator(s). This field can help to track the location of the media. Examples of meaningful location fields would be: In Library, off-site, and vault_1.

The media location setting is also useful if an object version that you want to restore exists on more than one media set. You can set the media location priority, which influences the selection of the media set that will be used for the restore. For more information on the selection of media for restore, see “Selection of the Media Set” on page 101.

Media Management During Backup Sessions

What Happens During Backup?

During a backup session, Data Protector automatically selects media for backup and keeps track of which data is backed up to which media. This simplifies management of media so that the operator does not need to know exactly which data was backed up to which media. Backup objects that have been backed up within the same backup session represent a media set.

This section provides the following information:

- How Data Protector selects media for backup
- How full and incremental backups are added to the media
- How the condition of media is calculated

Refer to the following sections for related information:

- “Full and Incremental Backups” on page 66
- “Media Pools” on page 114

Selecting Media for Backups

Data Protector automatically selects media for backup based on media allocation policies. This simplifies media management and media handling; a backup operator does not need to manually administer the media for backup.

Media Allocation Policy

You can influence how media are selected for backup using the media allocation policy. You can specify a loose policy, where any suitable medium is used for backup, including new, blank media or a strict policy, where media must be available in a predefined order to facilitate balanced media usage. Additionally, you can use a pre-allocation list.

Pre-Allocating Media

Data Protector allows you to explicitly specify media from a media pool that you want to use for a backup using a pre-allocation list. Combine this list with the strict media allocation policy. In this case, the media are used in the exact order as specified. If media are not found in this order, Data Protector issues a mount request.

Media Condition The condition of the media also influences which media are selected for backup, for example, media in good condition are used for backup before media in fair condition. Refer to “Calculating Media Condition” on page 130 for more information.

Adding Data to Media During Backup Sessions

To maximize space usage of media as well as backup and restore efficiency, you can select how Data Protector treats the space on the medium left over from the previous backup. This is defined with a media usage policy.

Media Usage Policy

The available media usage policies are listed below:

Appendable A backup session starts writing data to the space remaining on the last medium used from a previous backup session. Subsequent media needed in this session are written from the beginning of the tape, hence only unprotected or new tapes can be used. Appending media conserves media space but can add complexity to vaulting, because one medium can contain data from several media sets.

Non Appendable A backup session starts writing data at the beginning of the first available medium for backup. Each medium contains data from a single session only. This simplifies vaulting.

Appendable of Incrementals Only A backup session appends to a medium only if an incremental backup is performed. This allows you to have a complete set of full and incremental backups on the same medium, if there is enough space.

Distributing Objects Over Media

The following figures show some examples of how objects can be distributed over media:

Figure 3-6 Multiple Objects and Sessions per Medium, Sequential Writes

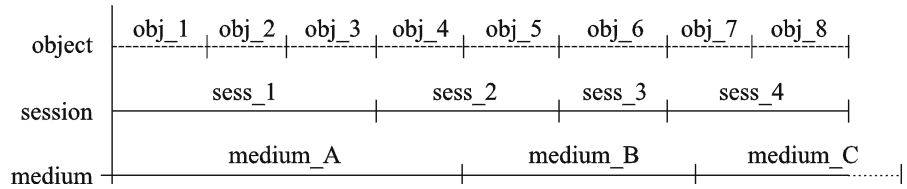


Figure 3-6 shows an example of eight sequential writes over four sessions, using the appendable media usage policy. The data was written in four sessions, one object at a time. The three media belong to the same media pool. *Medium_A* and *medium_B* are already full, while *medium_C* has still some space left.

Figure 3-7 Multiple Objects and Sessions per Medium, Concurrent Writes

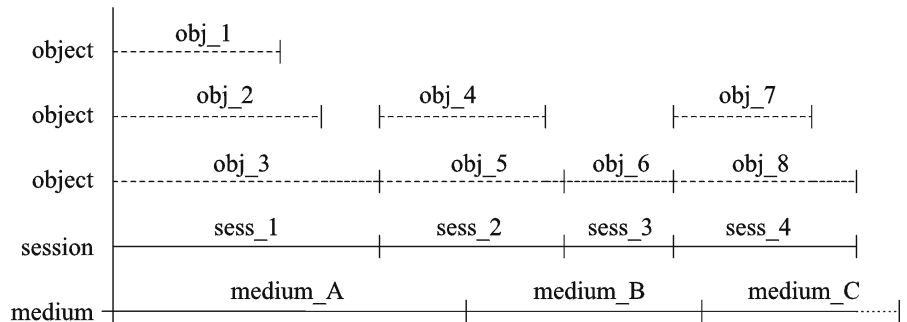


Figure 3-7 shows an example of eight objects that have been written during four sessions with the concurrency settings that allow for simultaneous writes. In this case, *obj_1*, *obj_2*, and *obj_3* have been backed up concurrently in *sess_1*; *obj_4* and *obj_5* have been backed up concurrently in *sess_2*, and so on. *Obj_1* could come from *system_A* and *obj_2* from *system_B*, or they could come from different disks on the same system. The media usage policy is appendable.

Figure 3-8 Multiple Media per Session, Multiple Media per Object

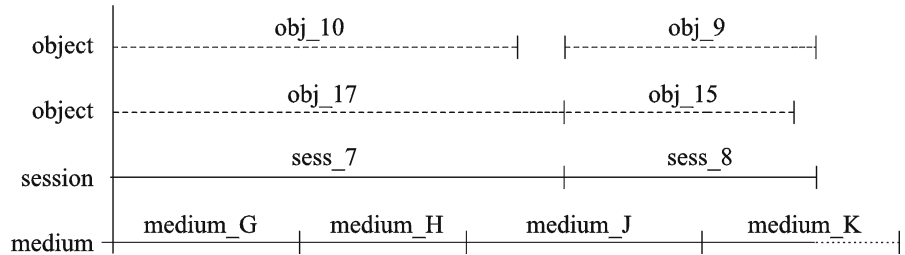


Figure 3-8 shows an example of four backup objects that have been backed up during two sessions, so that the first pair of backup objects has been concurrently written in *sess_7* and the second one in *sess_8*. Note that one object can be stretched over several media. The media usage policy is appendable.

Figure 3-9 Each Object Written on a Separate Medium

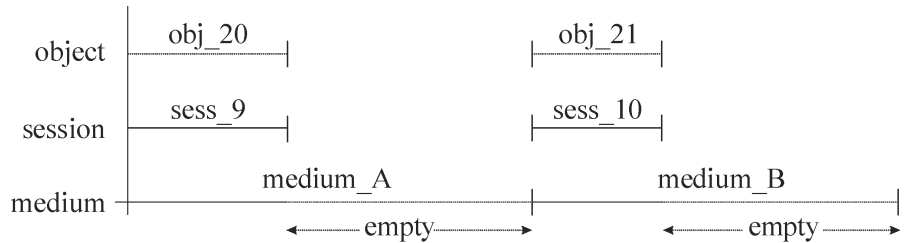


Figure 3-9 shows an example of using one backup specification per object with the non-appendable media usage policy. The result is higher media consumption. You could combine this with the append incrementals only policy, to get the incremental backups of the object on the same medium.

For more information on how full and incremental backup policies influence restore performance and media usage, refer to “Full and Incremental Backups” on page 66.

Writing Data to Several Media Sets During Backup

During a backup session, you can write all or some objects to several media sets simultaneously, using the Data Protector object mirror functionality. For more information, see “Object Mirroring” on page 96.

Calculating Media Condition

Media Condition Factors

Data Protector calculates the state of used media using **media condition factors**. The state of the poorest medium in a pool determines the state of the entire pool. For example, as soon as the state of one medium in a media pool is poor, the state of the pool becomes poor. When that particular medium is removed from the pool, the state reverts to either fair or good.

Media can have three states: good, fair, or poor.

On a per medium basis, the following is used for calculating the condition:

- number of overwrites

The usage of a medium is defined as the number of overwrites from the beginning of the medium. Once the medium has more than the threshold number of overwrites, it is marked as poor

- media age

The age of a medium is calculated as the number of months that have elapsed since you formatted, or initialized, the medium. Once a medium is older than the threshold number of months, it is marked as poor.

- device errors

Some device errors result in the medium being marked as poor. If a device fails during a backup, the medium used for the backup in this device is marked as poor.

Media Management After Backup Sessions

Once the data is stored on the media, you must take the right precautions to protect the media and the data on the media. Consider the following:

- Protecting media from overwrites.
You have specified this when you configured a backup of data, but you can change this after the backup is done. Refer to “Keeping Backed Up Data and Information About the Data” on page 73 for more information on data and catalog protection.
- Protecting media from physical damage.
Media with permanent data may be stored to a safe place.
- Copying backed up data and keeping the copies at a safe place.
Refer to “Duplicating Backed Up Data” on page 89.

The following sections describe how to vault media and restore from such media.

Vaulting

What Is Vaulting? Vaulting is a process of storing media with important information to a safe place, where they are kept for a specific period of time. The safe place for media is often called a **vault**.

Data Protector supports vaulting with the following features:

- Data protection and catalog protection policies.
- Easy selecting and ejecting of media from a library.
- The field `media_location` tells you the physical location where the media are stored.
- A report showing media used for backup within a specified time-frame.
- A report showing which backup specifications have used specified media during the backup.

- A report showing media stored at a specific location with data protection expiring in a specific time.
- Displaying a list of media needed for a restore and the physical locations where the media are stored.
- Filtering of media from the media view based on specific criteria.

Implementing Vaulting

The implementation of vaulting depends on your company's backup strategy and policies for handling data and media. Generally, it consists of the following steps:

1. Specifying the desired data protection and catalog protection policies when configuring backup specifications.
2. Configuring a vault in Data Protector. Essentially, this means specifying a name for the vault you will use for media, for example: Vault_1.
3. Establishing the appropriate media maintenance policy for media in the vault.
4. Optionally, creating additional copies of the backed up data for vaulting purposes, using the object mirror functionality during backup, or the object copy or media copy functionality after backup.
5. Selecting the media you want to store in a vault, ejecting the media and storing it in the vault.
6. Selecting the media with expired data which is in a vault and inserting the media in a library.

Vaulting Usage Example

Your company backup policy, for example, says that you must back up data daily. Each week a full backup must be stored in a vault where it must be available for the next five years. You must be able to easily restore data from all the previous year's backups stored in the vault. After five years, media from the vault can be re-used.

This implies the following Data Protector settings: a full backup once a week with daily incrementals. Data protection is set to five years. Catalog protection is set to one year. Therefore, you will be able to simply browse and restore data for one year and the data will be available for restore from media for five years. Media from the full backup are copied and stored to a vault. After one year, Data Protector automatically deletes detailed information from the IDB about the data on the media, thus creating more space in the database for new information.

Restoring from Media in a Vault

Restoring media from a vault is no different than restoring from any other media. Depending on how your data and catalog protection policies are defined, you may need to do some additional steps:

1. Bring media from a vault and insert the media into a device.
2. If the catalog protection for the media is still valid, restore data simply by selecting what you want to restore using the Data Protector user interface.

If the catalog protection for the media has expired, Data Protector does not have detailed information about the backed up data. You must restore by manually specifying the files or directories you want to restore. You can also restore the complete object to a spare disk and then search for files and directories in the restored filesystem.

TIP

To re-read detailed information about the files and directories backed up on the media once the catalog protection has expired, export the media and import them back. Then specify that you want to read the detailed catalog data from those media. Now you will be able to select files and directories in the Data Protector user interface again.

For more information on how data protection and catalog protection policies influence restores, refer to “Keeping Backed Up Data and Information About the Data” on page 73.

Devices

Data Protector supports a number of devices available on the market. Refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References* for up-to-date list of supported devices.

Using Devices with Data Protector

To use a device with Data Protector, you must configure the device in the Data Protector cell. When you configure a device, you specify a name for the device, some device specific options, such as barcode or cleaning tape support, and a media pool. The process of configuring devices is simplified with a wizard that leads you through all the steps and can even detect and configure devices automatically. The same physical device can be defined multiple times with different usage properties in Data Protector using different (logical) device names, for example, one without hardware data compression and another one with hardware data compression.

The following sections describe some specific device functionality and how Data Protector operates with various devices.

Library Management Console Support

Many modern tape libraries provide a management console that allows libraries to be configured, managed, or monitored from a remote system. The scope of tasks that can be performed remotely depends on the management console implementation, which is independent of Data Protector.

Data Protector eases access to the library management console interface. The URL (web address) of the management console can be specified during the library configuration or re-configuration process. By selecting a dedicated menu item in the GUI, a web browser is invoked and the console interface is automatically loaded into it.

For a list of device types for which this feature is available, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

IMPORTANT

Before using the library management console, consider that some operations which you can perform through the console may interfere with your media management operations and/or your backup and restore sessions.

TapeAlert

TapeAlert is a tape device status monitoring and messaging utility that makes it easy to detect problems that could have an impact on backup quality. From the use of worn-out tapes to defects in the device hardware TapeAlert provides easy-to-understand warnings or errors as they arise, and suggests a course of action to remedy the problem.

Data Protector fully supports TapeAlert 2.0, as long as the connected device also provides this functionality.

Device Lists and Load Balancing

Multiple Devices for Backup

When configuring a backup specification, you can specify several standalone devices or multiple drives in a library device that will be used for the operation. In this case, the operation is faster because data is backed up in parallel to multiple devices (drives).

Balancing the Use of Devices

By default, Data Protector automatically balances the load (the usage) of devices so that they are used evenly. This is called **load balancing**. Load balancing optimizes the usage by balancing the number of the objects backed up to each device. Since load balancing is done automatically during backup time, you do not have to manage the allocation of objects to devices used in the session; you just specify the devices to be used.

When to Use Load Balancing

Use load balancing when:

- You back up a large number of objects.
- You use library (autochanger) devices with several drives.
- You do not need to know on which media objects will be backed up.
- You have a good network connection.
- You want to increase the robustness of the backup. Data Protector automatically redirects the backup operation from failed devices to other devices in a device list.

When Not to Use Load Balancing

Do not use load balancing when:

- You want to back up a small number of large objects. In this case Data Protector often cannot effectively balance the load among devices.
- You want to explicitly select to which device each object will be backed up.

Device Chaining

Data Protector allows you to configure several standalone devices of the same type, connected to the same system, as a device chain. When a medium in one device gets full, the backup automatically continues on the medium in the next device in the device chain.

How Load Balancing Works

For example, assume that there are 100 objects configured for backup to four devices with concurrency set to three and with load balancing parameters MIN and MAX both configured at two. If at least two devices are available, the session will start with three objects being backed up in parallel to each of the first two available devices. The other 94 objects will be pending and will not be assigned to a particular device at that time.

Once a backup of a particular object is done, the next pending object is started and assigned to the device that has less than three concurrent objects being backed up. Load balancing ensures that the two devices are running in parallel as long as there are still pending objects to be backed up. If a device fails during backup, one of the two devices in reserve is used. The objects that were being backed up to the failed device are aborted, while the next three pending objects are assigned to the new device. This means that each failure of a device can cause a maximum of three objects to be aborted, provided that other devices are available for the backup session to continue.

Device Streaming and Concurrency

What Is Device Streaming?

To maximize a device performance, it must be kept streaming. A device is streaming if it can feed enough data to the medium to keep the medium moving forward continuously. Otherwise, the medium tape has to be stopped while the device waits for more data. In other words, if the rate at which data is written to the tape is less than or equal to the rate which data can be delivered to the device by the computer system, then the device is streaming. In network-focused backup infrastructures, this

deserves attention. For local backups, where disks and devices are connected to the same system, a concurrency of 1 may suffice if your disks are fast enough.

How to Configure Device Streaming

To allow the device to stream, a sufficient amount of data must be sent to the device. Data Protector accomplishes this by starting multiple Disk Agents for each Media Agent that writes data to the device.

Disk Agent Concurrency

The number of Disk Agents started for each Media Agent is called **Disk Agent (backup) concurrency** and can be modified using the Advanced options for the device or when configuring a backup. Data Protector provides default numbers that are sufficient for most cases. For example, on a standard DDS device, two Disk Agents send enough data for the device to stream. For library devices with multiple drives where each drive is controlled by one Media Agent, you can set the concurrency for each drive independently.

Increased Performance

If properly set, backup concurrency increases backup performance. For example, if you have a library device with four drives, each controlled by a Media Agent and each Media Agent receives data from two Disk Agents concurrently, data from eight disks is backed up simultaneously.

Device streaming is also dependent on other factors, such as network load and the block size of the data written to the device.

Refer to “Backup Sessions” on page 212 for related information.

Multiple Data Streams

Data Protector allows you to concurrently back up parts of a disk to multiple devices. This feature is useful for backing up very large and fast disks to relatively slow devices. Multiple Disk Agents read data from the disk in parallel and send the data to multiple Media Agents. This method speeds up the backup, but requires that you take into account the following:

If one mount point was backed up through many Disk Agents, data is contained in multiple objects. To restore the whole mount point define all parts of the mount point in a single backup specification and then restore the entire session.

Segment Size

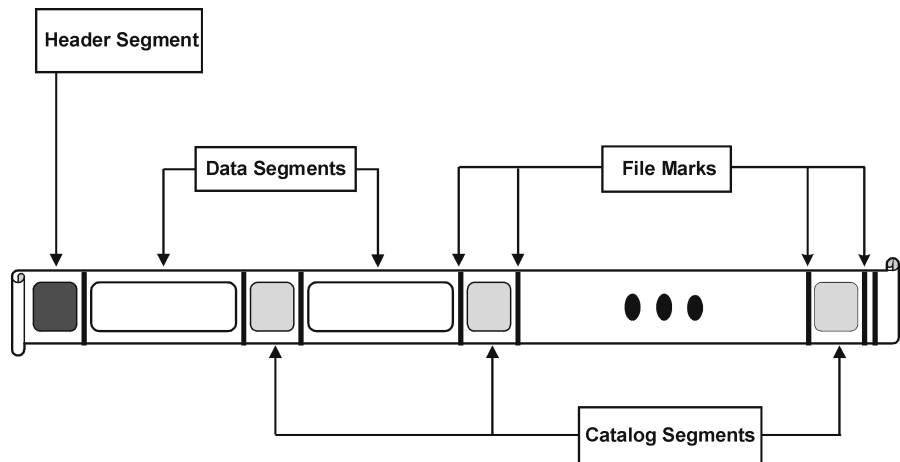
A medium is divided into data segments, catalog segments and a header segment. Header information is stored in the header segment, which is the same size as the block size. Data is stored in data blocks of data

segments. Information about each data segment is stored in the corresponding catalog segment. This information is first stored in the Media Agent memory and then written to a catalog segment on the medium as well as to the IDB. All segments are divided by file marks as shown in Figure 3-10.

NOTE

Some tape technologies place limitations on the number of file marks per medium. Ensure that your segment size is not too low.

Figure 3-10 Data Format



Segment size, measured in megabytes, is the maximum size of data segments. If you back up a large number of small files, the actual segment size can be limited by the maximum size of catalog segments. Segment size is user configurable for each device. It affects the speed of a restore. A smaller segment size leaves less space on the medium for data, because each segment has a file mark that takes up media space. However, a larger number of file marks results in faster restores, because a Media Agent can more quickly locate the segment containing the data to be restored. Optimal segment size depends on the type of media used in the device and the kind of data to be backed up. For example, by default the segment size for DLT medium is 150 MB.

Block Size

Segments are not written as a whole unit, but rather in smaller subunits called blocks. The hardware of a device processes data in units of a device-type specific block size. Data Protector allows you to adjust the size of the blocks it sends to the device. The default block size value for all devices is 64 KB.

Increasing the block size can improve performance. Changing the block size should be done *before* formatting tapes. For example, a tape written with the default block size cannot be appended to using a different block size.

NOTE

Use the same block size for media that can be used with different device types. Data Protector can only append data to media using the same block size.

Number of Disk Agent Buffers

Data Protector Media Agents and Disk Agents use memory buffers to hold data waiting to be transferred. This memory is divided into a number of buffer areas (one for each Disk Agent, depending on device concurrency). Each buffer area consists of 8 Disk Agent buffers (of the same size as the block size configured for the device). You can change this value to be anything between 1 and 32, although this is rarely necessary. There are two basic reasons to change this setting:

- Shortage of memory

The shared memory required for a Media Agent can be calculated as follows:

$DAConcurrency * NumberOfBuffers * BlockSize$

Reducing the number of buffers from 8 to 4, for instance, results in a 50% reduction in memory consumption, with performance implications.

- Streaming

If the available network bandwidth varies significantly during backup, then it becomes more important that a Media Agent has enough data ready for writing to keep the device in the streaming mode. In this case, increase the number of buffers.

Device Locking and Lock Names

Device Names

When configuring devices for use with Data Protector, you can configure the same physical device many times with different characteristics simply by configuring the same physical device in Data Protector with different device names. For example, a simple standalone DDS device can be configured as a compressed device and then as an uncompressed device, although this is not recommended.

Physical Device Collision

When specifying a device used for backup, you may specify one device name in one backup specification and another device name of the same physical device in a different backup specification. Depending on the backup schedule, this may result in Data Protector trying to use the same physical device in several backup sessions at the same time, thus creating a collision.

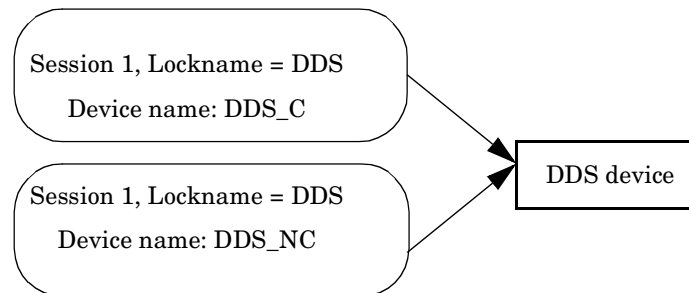
Preventing Collision

To prevent this collision, specify a virtual lockname in both device configurations. Data Protector checks if the devices have the same lockname and prevents collision.

For example, a DDS device is configured as a compressed device named DDS_C, and as a non-compressed device DDS_NC as shown in Figure 3-11. Specify the same lockname, DDS, for both devices.

Figure 3-11

Device Locking and Device Names



Standalone Devices

What Are Standalone Devices?

Standalone devices are devices with one drive that reads/writes to one medium at time.

Standalone devices are used for small scale backups or special backups. When the medium is full, the operator must manually replace it with a new medium for the backup to proceed.

Data Protector and Standalone Devices

Once you have connected a device to the system, you use the Data Protector user interface to configure the device for use with Data Protector. To do this, you must first install a Data Protector Media Agent on the system with the device connected. Data Protector can detect and automatically configure most standalone devices.

During a backup, Data Protector issues a mount request when the medium in a device is full. The operator must replace the medium for the backup to continue.

What Are Device Chains?

Data Protector allows you to configure multiple standalone devices to a device chain. When a medium in one device gets full, the backup automatically continues on the medium in the next device in the device chain.

Device chains allow running unattended backups using several standalone devices without having to manually insert/eject media when the media are full.

Stacker Devices

Stacker devices, similar to device chains, contain a number of media that are used in a sequential order. When a medium gets full, the next medium is loaded and used for backup.

Small Magazine Devices

What Are Magazine Devices?

Magazine devices group a number of media into a single unit called a magazine. Data Protector treats the magazine as if it were a single medium. A magazine has a larger capacity than a single medium and is easier to handle than several single media. For a list of supported devices, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

Data Protector and Magazine Devices

Data Protector allows you to perform media management tasks on magazines as sets, emulating single media by providing magazine and media views, or on a single medium.

You can alternatively use magazine devices as normal libraries without using Data Protector magazine support. Data Protector can detect and automatically configure magazine devices.

Cleaning Dirty Drives

Using cleaning tapes, Data Protector can automatically clean magazines and other devices when they get dirty.

Large Libraries

What Are Library Devices?

Library devices are automated devices, also called autoloaders, exchangers or jukeboxes. In Data Protector, most libraries are configured as SCSI libraries. They contain a number of media cartridges in a device's repository and can have multiple drives writing to multiple media at a time.

A typical library device has a SCSI ID for each drive in the device and one for the library robotic mechanism that moves media from slots to drives and back. For example, a library with four drives has five SCSI IDs, four for the drives and one for the robotic mechanism.

Data Protector also supports silo libraries, such as HP StorageWorks Libraries, StorageTek/ACSLs and ADIC/GRAU AML. For a list of supported devices, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

Handling of Media

The Data Protector user interface provides a special library view, which simplifies managing library devices.

Media in a large library device can all belong to one Data Protector media pool, or they can be split into several pools.

Configuring a Library

When configuring a device, you configure the slot range you want to assign to Data Protector. This allows sharing of the library with the other application. The assigned slots may contain blank (new) media, Data Protector or non-Data Protector media. Data Protector checks the media in the slots and displays the information about the media in the library view. This allows you to view all kinds of media, not just the media used by Data Protector.

Size of a Library

The following may help you estimate the size of the library you need:

- Determine if you need to distribute the media to several locations or keep them in a central location.

- Obtain the number of required media. Refer to “Implementing a Media Rotation Policy” on page 121.

Sharing a Library with Other Applications

A library device can be shared with other applications storing data to media in the device.

You can decide which drives from the library you want to use with Data Protector. For example, out of a four-drive library you may choose to use only two drives with Data Protector.

You can decide which slots in the library you want to manage with Data Protector. For example, out of the 60 slots library you might use slots 1-40 with Data Protector. The remaining slots would then be used and controlled by a different application.

Sharing of the library with other applications is especially important with large HP libraries and silo libraries, such as StorageTek/ACSLs or ADIC/GRAU AML devices.

Enter / Eject Mail slots

Library devices provide special enter/eject mail slots an operator uses to enter or eject media to or from the device. Depending on the device, more than one enter/eject slot can be provided. In case of a single mail slot, media are inserted one by one, while in case of multiple mail slots, a particular number of slots can be used in one enter/eject operation.

Data Protector allows you to enter/eject several media in one step. For example, you can select 50 slots in the device and eject all media in one action. Data Protector will automatically eject media in the correct order for the operator to remove the media from the enter/eject mail slot.

Refer to the documentation with your device for more information.

Barcode Support

Data Protector supports library devices with a barcode reader. In these devices, each medium has a barcode that uniquely identifies media.

Advantages of Barcodes

Barcodes enable Data Protector to significantly improve media recognition, labeling, and cleaning tape detection.

- Scanning the barcodes of the media in a device's repository is faster, because Data Protector does not need to actually load the media to a drive and read the medium header.
- A barcode is automatically read by Data Protector and used to identify the media.
- A cleaning tape is automatically detected if it has a CLN barcode prefix.
- A barcode is a unique identifier for media in the IDB. You cannot have duplicate barcodes in your environment.

TIP

You can optionally use the barcode as medium label during the initialization of the medium.

Cleaning Tape Support

HP Data Protector provides automatic cleaning for most devices using a cleaning tape. This medium will be used automatically by Data Protector if a dirty drive event from the device is detected.

- For SCSI libraries it is possible to define which slot holds a cleaning tape.
- For devices with a barcode reader, Data Protector recognizes cleaning tape barcodes automatically if they have the CLN prefix.
- For devices without a cleaning tape, a dirty drive detection will cause a cleaning request to be displayed on the session monitor window. The operator must clean the device manually.

You cannot continue your backup without cleaning the drive, since the backup may fail because data may not be correctly written and stored on the media.

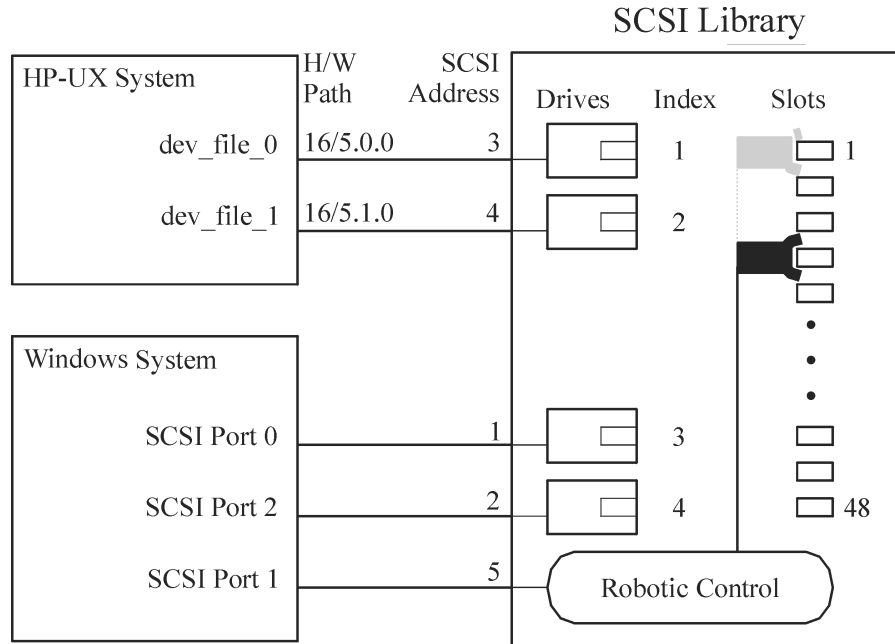
Sharing a Library with Multiple Systems

What Is Library Sharing?

Device sharing allows you to connect different drives of a physical library to different systems. These systems can then perform local backups to the library. The result is significantly higher backup performance and less network traffic. To enable library sharing, the drives in the library

must have the possibility to connect to separate SCSI buses. This is useful with high performance libraries to allow the drive to receive data in a continuous stream from multiple systems, further enhancing performance. Data Protector internally redirects the robotic commands to the system that manages the robotics.

Figure 3-12 Connecting Drives to Multiple Systems



Control Protocols and Data Protector Media Agents

The drives in the library must be able to physically connect to different systems that have a Data Protector Media Agent (the General Media Agent or the NDMP Media Agent) installed.

With Data Protector, there are two types of protocols used for drive control:

- SCSI—for SCSI or Fibre Channel connected drives.

This protocol is implemented in both the General Media Agent and in the NDMP Media Agent.

- NDMP—for NDMP dedicated drives.

This protocol is implemented in the NDMP Media Agent only.

On the other hand, there are four types of protocols used for library robotic control:

- ADIC/GRAU—for ADIC/GRAU library robotics
- StorageTek ACS—for StorageTek ACS library robotics
- SCSI—for robotics other libraries
- NDMP—for NDMP robotics

All four library robotic control protocols are implemented in both the General Media Agent and in the NDMP Media Agent.

Drive Control

Any Data Protector client system configured to control a drive in a library (regardless of the drive control protocol and platform used) can communicate with any Data Protector client system configured to control the robotics in the library (regardless of the robotics control protocol and platform used). Thus, it is possible to share drives in any supported library among Data Protector clients systems on various platforms using various robotic and drive protocols. The NDMP Media Agent is needed only on client systems controlling the backup of an NDMP server (on client systems configured for NDMP dedicated drives). In all other cases the two Data Protector Media Agents are interchangeable.

Table 3-1 show the Data Protector Media Agent (the General Media Agent or the NDMP Media Agent) required on client systems configured for drive control of a library with drives shared among multiple client systems.

Table 3-1 Required Data Protector Media Agent for Drive Control

	Drive Control Protocol	
	NDMP	SCSI
Robotic Control Protocol (ADIC/GRAU, StorageTek ACS, SCSI, or NDMP)	NDMP Media Agent	NDMP Media Agent or General Media Agent

Robotic Control

A Data Protector client system controlling the library robotics can have either the General Media Agent or the NDMP Media Agent installed, regardless of the type of drive protocol (NDMP or SCSI) used with the drives in the library.

Table 3-2 show the Data Protector Media Agent (the General Media Agent or the NDMP Media Agent) required on a client system configured for robotic control of a library with drives shared among multiple client systems.

Table 3-2 Required Data Protector Media Agent for Robotic Control

	Robotic Control Protocol			
	ADIC/GRAU	StorageTek ACS	SCSI	NDMP
Drive Control Protocol (NDMP or SCSI)	NDMP Media Agent or General Media Agent	NDMP Media Agent or General Media Agent	NDMP Media Agent or General Media Agent	NDMP Media Agent or General Media Agent

Exemplary Configurations

Figures 3-13 to 3-15 show exemplary configurations of shared drives in libraries and Data Protector Media Agents distributions in such configurations.

Figure 3-13 Sharing a SCSI Library (Robotics Attached to a Data Protector Client System)

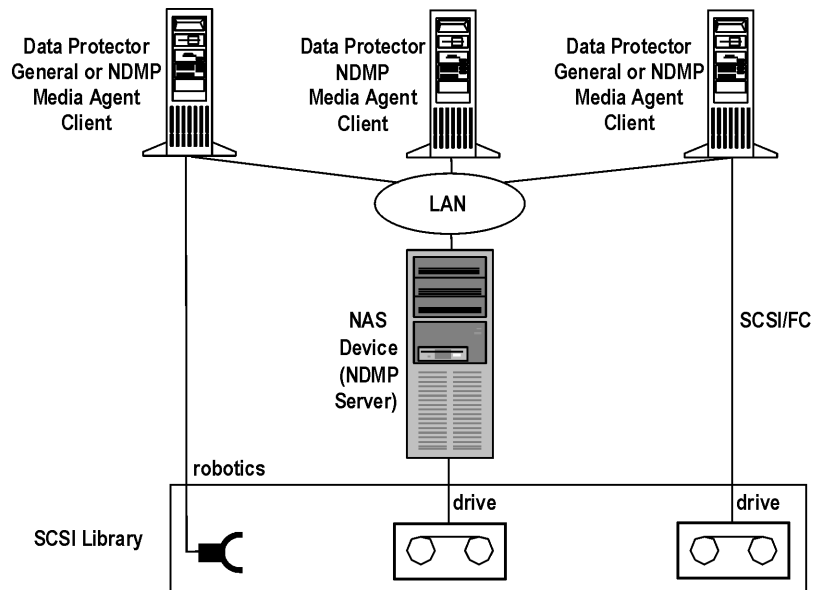


Figure 3-13 shows a SCSI library, with its robotics attached to and configured on the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI robotic control protocol is used by the General Media Agent or the NDMP Media Agent on the client. The Data Protector client system with the attached robotics can also have one or more drives attached.

The NDMP dedicated drive in the library is configured on the Data Protector client system with the NDMP Media Agent installed. The NDMP drive control protocol is used by the NDMP Media Agent on the client.

Another drive in the library is configured on and attached to the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI drive control protocol is used by the General Media Agent or the NDMP Media Agent on the client.

Figure 3-14 Sharing a SCSI Library (Robotics Attached to an NDMP Server)

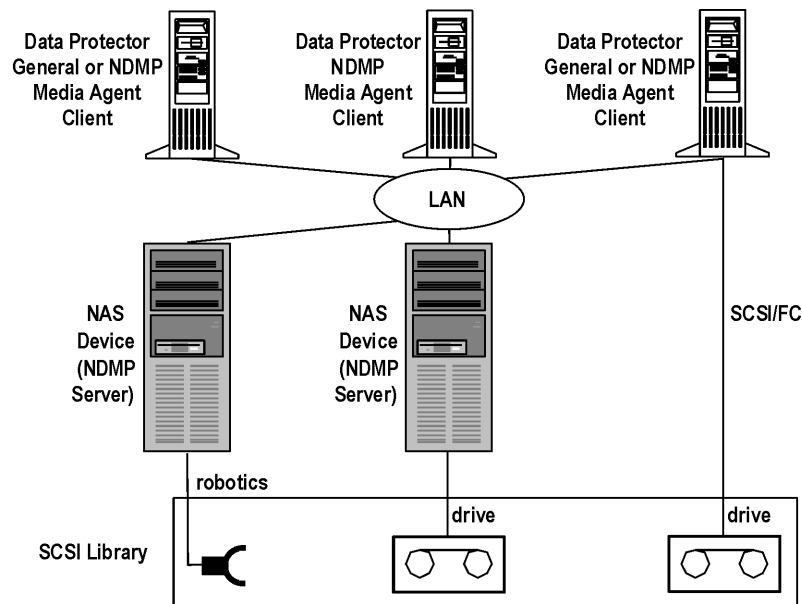


Figure 3-14 shows a SCSI library, with its robotics attached to an NDMP Server and configured on the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI robotic control protocol is used by the General Media Agent or the NDMP Media Agent on the client. The NDMP Server with the attached robotics can also have one or more drives attached.

IMPORTANT

If the NDMP Server with the attached robotics also have an NDMP dedicated drive attached, the Data Protector client system on which the robotics and the NDMP dedicated drive are configured, can only have the NDMP Media Agent installed, since the NDMP drive control protocol is used for the NDMP dedicated drive.

The NDMP dedicated drive in the library is configured on the Data Protector client system with the NDMP Media Agent installed. The NDMP drive control protocol is used by the NDMP Media Agent on the client.

Another drive in the library is configured on and attached to the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI drive control protocol is used by the General Media Agent or the NDMP Media Agent on the client.

Figure 3-15 Sharing an ADIC/GRAU or StorageTek ACS Library

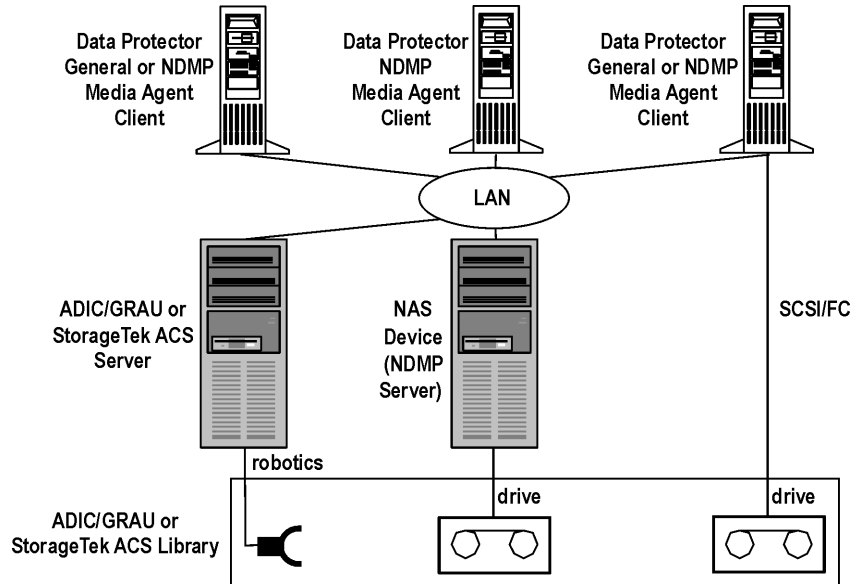


Figure 3-15 shows an ADIC/GRAU or StorageTek ACS library, with its robotics attached to an ADIC/GRAU or StorageTek ACS Server and configured on the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The ADIC/GRAU robotic control protocol is used by the General Media Agent or the NDMP Media Agent on the client. The ADIC/GRAU or the StorageTek ACS Server can also have one or more drives attached.

The NDMP dedicated drive in the library is configured on the Data Protector client system with the NDMP Media Agent installed. The NDMP drive control protocol is used by the NDMP Media Agent on the client.

Another drive in the library is configured on and attached to the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI drive control protocol is used by the General Media Agent or the NDMP Media Agent on the client.

Data Protector and Storage Area Networks

Where and how you store data in your enterprise may have a serious impact on your business. Information is becoming increasingly mission-critical to most companies. Today, terabytes of data must be accessible to users across the network. The Data Protector implementation of SAN-based Fibre Channel technology provides you with the data storage solution you need.

Storage Area Networks

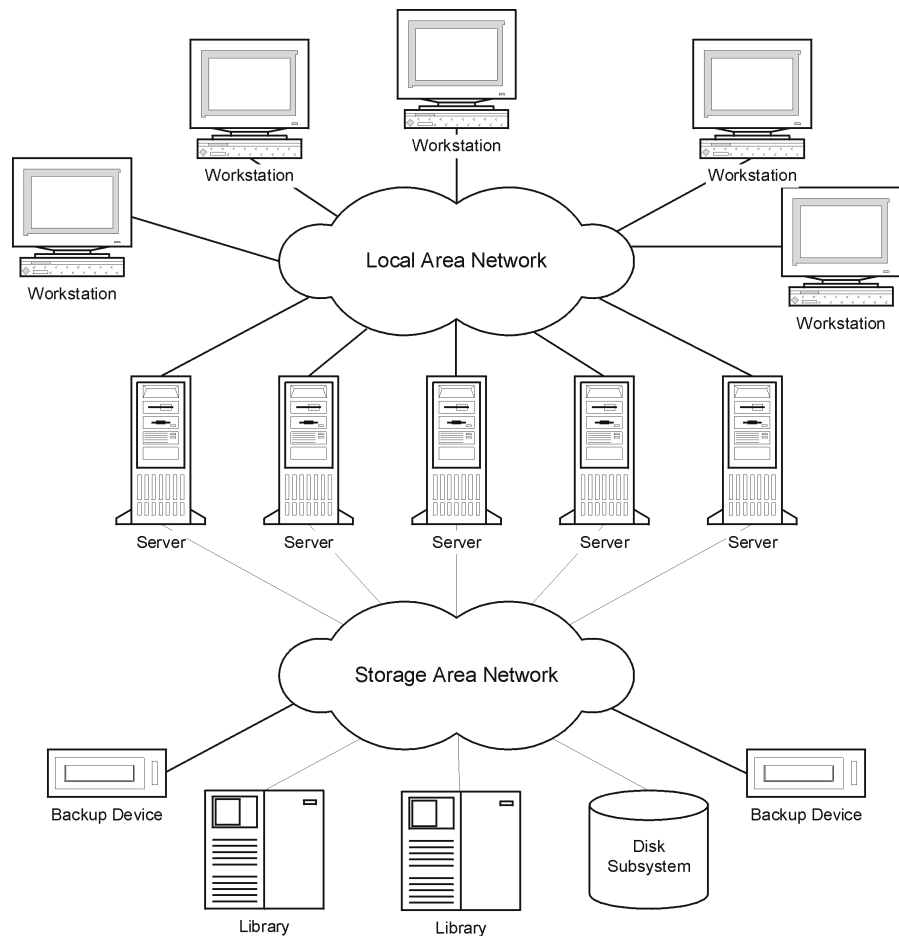
A Storage Area Network (SAN), depicted in Figure 3-16, is a new approach to network storage that separates storage management from server management with a network devoted to storage.

A SAN provides *any-to-any* connectivity for all network resources, thus enabling device sharing between multiple client systems and increasing data traffic performance as well as the availability of devices.

The SAN concept allows the exchange of information between multiple data storage devices and servers. The servers can access data directly from any device and do not need to transfer data over the conventional LAN. A SAN consists of servers, backup devices, disk arrays, and other nodes, all connected with a fast network connection, typically Fibre Channel. This additional network provides off-loading storage operations from the conventional LAN to a separate network.

Data Protector's direct backup functionality is a productive application of SAN and fibre channel technology.

Figure 3-16 Storage Area Network



Fibre Channel

Fibre Channel is an ANSI standard for high-speed computer interconnection. Using either optical or copper cables, it allows the bidirectional transmission of large data files at up to 4.25 gigabits per second, and can be deployed between sites within a 30 kilometer range. Fibre Channel is the most reliable, highest performance solution for information storage, transfer, and retrieval available today.

Fibre Channel connects nodes using three physical topologies that can have variants:

- Point-to-point
- Loop
- Switched

Point-to-point, loop, and switched Fibre Channel topologies can be mixed to best suit your connectivity and growth requirements.

Refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References* or to http://www.openview.hp.com/products/datapro/spec_0001.html for a list of supported configurations.

Point-to-Point Topology

This topology allows the connecting of two nodes, typically a server and a backup device. It provides the basic benefit of improved performance and longer distances between nodes.

Loop Topology

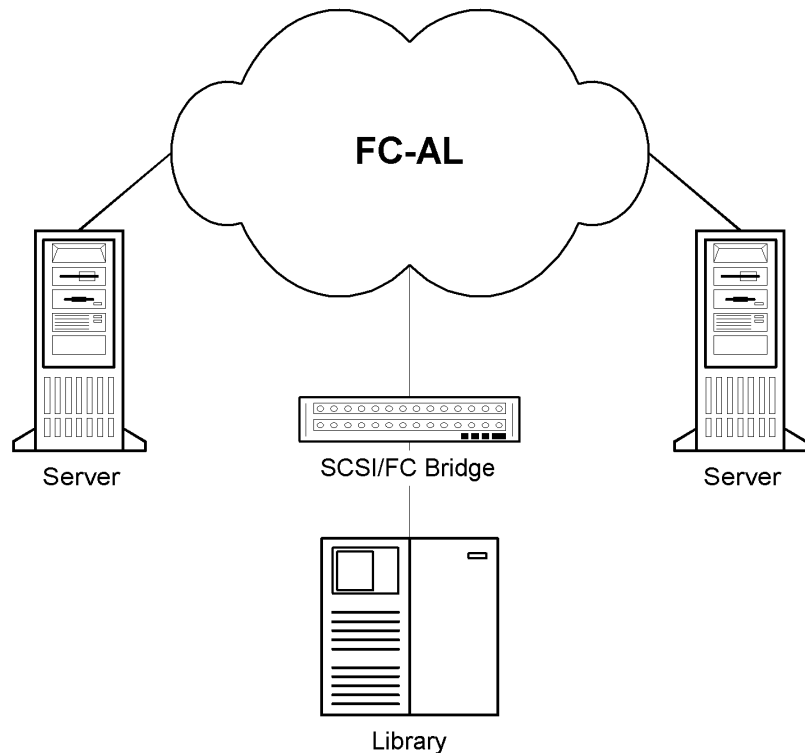
The loop topology is based on the Fibre Channel Arbitrated Loop (FC-AL) standard, which allows the connecting of up to 126 nodes. Nodes include servers, backup devices, hubs, and switches. Any node in a loop can communicate with any other node in the loop, and all nodes share the same bandwidth. An FC-AL loop is typically implemented using an FC-AL hub with automatic port by-pass. Automatic port by-pass allows the hot-plug of nodes into the loop.

LIP

A Loop Initialization Primitive (Protocol) (LIP) may be triggered by a number of causes, most common being the introduction of a new device. The new device could be a former participant that has been powered on or an active device that has been moved from one switch port to another. A LIP occurrence can cause an undesirable disruption of an ongoing process on the SAN, for example, a tape backup operation. It resets the SCSI bus connecting the SCSI/FC Bridge and the node (SCSI device). Refer to Figure 3-17.

In the case of a backup or restore, a SCSI bus reset is registered as a write error. Data Protector aborts all operations upon write errors. In the case of backups, it is recommended to (copy the information already backed up on the medium and then) reformat the medium and restart the backup.

Figure 3-17 Loop Initialization Protocol



Switched Topology

The switched topology provides any-to-any connectivity between all nodes connected to a switch. Switches are easy to install and use, because the Fibre Channel protocol provides self-configuration and self-management. Switches automatically detect what is connected (nodes, FC-AL Hubs or other FC switches), and configure themselves accordingly. Switches provide scaled bandwidth to connected nodes. The switched topology provides real hot-plug of nodes.

NOTE

Hot-plug refers to protocol capabilities such as reset, re-establish communication, and so on. Take into account that ongoing data transfers are interrupted during hot-plug and that some devices, such as tape devices, cannot handle this behavior. Connecting nodes to or disconnecting nodes from a loop is likely to interrupt your backup or restore process and cause the operation to fail. Connect or disconnect nodes from loops only when there are no running backups or restores using the related hardware.

Device Sharing in SAN

Data Protector supports the SAN concept by enabling multiple systems to share backup devices in the SAN environment. The same physical device can be accessed from multiple systems. Thus, any system can perform a local backup on some device or any other device. Because data is transferred over the SAN, backups do not need any bandwidth on your conventional LAN. This type of backup is sometimes referred to as a “LAN-free” backup. Backup performance is also improved, because SAN-based Fibre Channel technology typically provides an order of magnitude higher throughput than LAN technologies.

You need to prevent several computer-systems from writing to the same device at the same time. This can become even more complex when devices are used from several applications. Access to the devices needs to be synchronized between all systems involved. This is done using locking mechanisms.

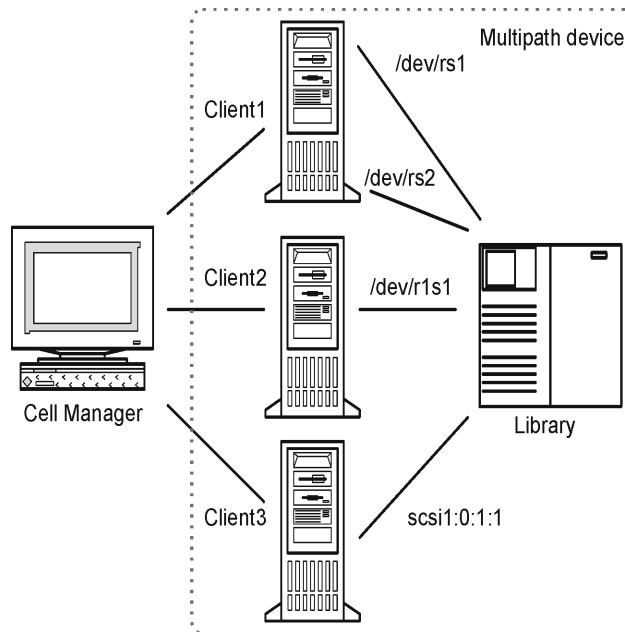
SAN technology provides an excellent way to manage the robotics of a library from multiple systems. This allows the option to manage the robotics from one system (classic) or allow each system that uses the library to access the robotics directly, provided the requests to the robotics are synchronized between all the systems involved.

Configuring Multiple Paths to Physical Devices

A device in a SAN environment is usually connected to several clients and can thus be accessed through several paths, that is client names and SCSI addresses (device files on UNIX). Data Protector can use any of these paths. You can configure all paths to a physical device as a single logical device - **multipath device**.

For example, a device is connected to `client1` and configured as `/dev/rs1` and `/dev/rs2`, on `client2` as `/dev/r1s1` and on `client3` as `scsi1:0:1:1`. Thus, it can be accessed through four different paths: `client1:/dev/rs1`, `client1:/dev/rs2`, `client2:/dev/r1s1` and `client3:scsi1:0:1:1`. A multipath device therefore contains all four paths to this tape device.

Figure 3-18 Example Multipath Configuration



Why Use Multiple Paths

With previous versions of Data Protector, a device could be accessed from only one client. To overcome this problem, several logical devices had to be configured for a physical device using a lock name. Thus, if you were using lock names for configuring access from different systems to a single physical device, you had to configure all devices on every system. For example, if there were 10 clients which were connected to a single device, you had to configure 10 devices with the same lock name. With this version of Data Protector, you can simplify the configuration by configuring a single multipath device for all paths.

Multipath devices increase system resilience. Data Protector will try to use the first defined path. If all paths on a client are inaccessible, Data Protector will try to use paths on the next client. Only when none of the listed paths is available, the session aborts.

Path Selection

During a backup session, the device paths are selected in the order defined during the device configuration, except if a preferred client is selected in the backup specification. In this case, the preferred client is used first.

During a restore session, the device paths are selected in the following order:

1. Paths that are on the client to which the objects are restored, if *all* objects are restored to the same target client
2. Paths that were used for backup
3. Other available paths

If direct library access is enabled, local paths (paths on the destination client) are used for library control first, regardless of the configured order.

Backward Compatibility

Devices configured with previous versions of Data Protector are not reconfigured during the upgrade and can be used as in previous releases of Data Protector without any changes. To utilize the new multipath functionality, you must reconfigure devices as multipath devices.

Device Locking

Locking devices must cover the possibility of several applications using the same device, as well as only Data Protector using a device by sending data and commands to it from several systems. The purpose of locking is to ensure that only one system at a time communicates with a device that is shared between several systems.

Device Locking with Multiple Applications

If Data Protector and at least one other application want to use the same device from several systems, the same (generic) device locking mechanism has to be used by each application. This mechanism needs to work across several applications. This mode is not currently supported by Data Protector. Should this be required, operational rules must ensure exclusive access to all devices from only one application at a time.

Device Locking within Data Protector

If Data Protector is the only application that uses a drive, but that same drive needs to be used by several systems, Device Locking has to be used.

If Data Protector is the only application that uses a robotics control from several systems, Data Protector handles this internally, provided the library control is in the same cell as all the systems that need to control it. In such a case, all synchronization of access to the device is managed by Data Protector internal control.

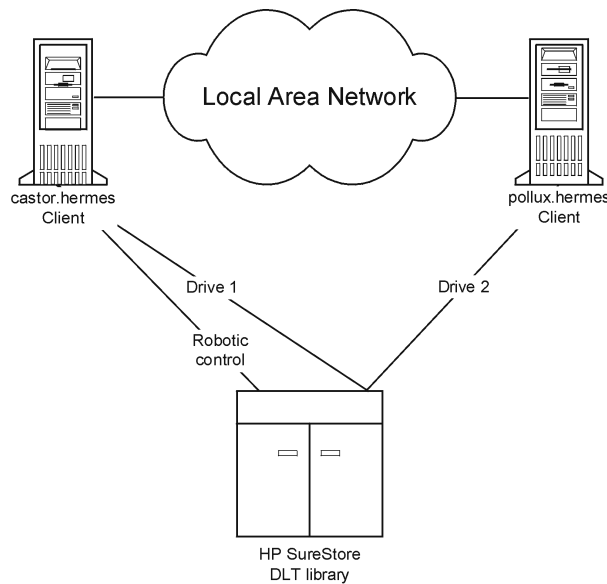
Indirect and Direct Library Access

Upon configuring Data Protector with a SCSI Library device, there are two ways in which client systems can access library robotics: Indirect Library Access and Direct Library Access.

Indirect Library Access

This configuration can be used in SAN as well as conventional SCSI direct connect environments. Several systems can access the library robotics by forwarding their requests to a client system that has direct access to the library robotics. This is called Indirect Library Access. In the example depicted in Figure 3-19, two client systems are attached to an HP StorageWorks DLT multidrive library. The client system `castor` controls the robotics and the first drive, while the client system `pollux` controls the second drive. A Data Protector Media Agent on `pollux` communicates with a process running on `castor` to operate the robotics. This Data Protector library sharing feature is used automatically when the hostnames of the library and drive are different.

Figure 3-19 Indirect Library Access



Note that you cannot use a shared library if the client system that controls the robotics, *castor*, in our example, fails.

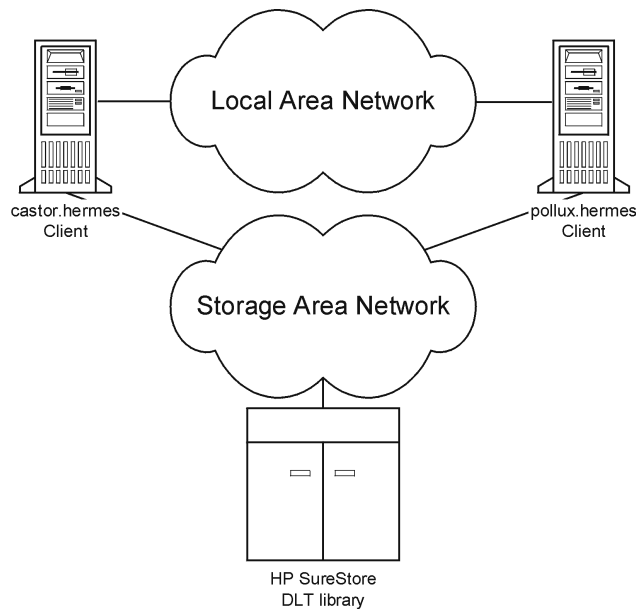
Direct Library Access

When the SAN concept is used, Data Protector can be configured with a SCSI Library so that each client system has its own access to library robotics and drives. This is called Direct Library Access

There is no single “controlling client system” for the robotics: a failure of the system controlling the robotics does not exclude any other system from using the library. This is performed without reconfiguration. Several client systems can be used to control the robotics.

Figure 3-20 shows an HP StorageWorks DLT multidrive library attached via a SAN to two client systems. Both client systems have access to the library and to both drives. The SCSI protocol is used for communication with the library.

Figure 3-20 Direct Library Access



Device Sharing in Clusters

Clustering, which is often used in combination with the SAN concept, is based on sharing network resources (for example network names, disks, and tapes devices) between nodes.

Cluster-aware applications can at any time run on any node in a cluster (they run on virtual hosts). To perform a local backup of such an application, you need to configure devices with virtual hostnames instead of real node names. Configure as many devices for each physical device as you need, using the Lock Name device locking mechanism. Refer to “Device Locking” on page 158 for details.

Static Drives

Static drives are devices that are configured on a real node in a cluster. They can be used to back up data from systems with disks that are not shared. However, they are not useful for backing up cluster-aware applications, because such application can run on any node in the cluster.

Floating Drives

Floating drives are device that are configured on a virtual host, using virtual system names. Floating drives should be configured for the backup of cluster-aware applications. This ensures that no matter on which node in the cluster the application is currently running, Data Protector always starts a Media Agent on that same node.

4 **Users and User Groups**

In This Chapter

This chapter discusses Data Protector security, users, user groups, and user rights.

It is organized as follows:

“Increased Security for Data Protector Users” on page 165

“Users and User Groups” on page 166

Increased Security for Data Protector Users

Data Protector provides advanced security functionality that prevents unauthorized backing up or restoring of data. Data Protector security involves hiding data from unauthorized users, data encoding, and restricted grouping of users according to their responsibilities.

This section describes security issues related to using Data Protector for backing up data, restoring data, or monitoring the progress of backup sessions.

Access to Backed Up Data

Backing up and then restoring data is essentially the same as copying data. Therefore, it is important to restrict access to this data to authorized users only.

Data Protector provides the following user-related security:

- All users intent on using any of the Data Protector functionality must be configured as Data Protector users.
- Backed up data is hidden from other users, except the backup owner. Other users do not even see that data was backed up. For example, if the backup operator has configured a backup, only the backup operator or the system administrator can see and restore the backed up data. You can make data visible to other users using the Data Protector `Public` option. Refer to the Data Protector online Help for instructions.

Visibility of Backed Up Data

Users and User Groups

To use Data Protector, you must be added to the Data Protector configuration as a Data Protector user with certain privileges. Note that adding a new user is not a prerequisite for backing up the system this user is using.

Users are grouped into user groups with specific user rights, for example, to monitor sessions in the cell, configure backups, and restore files.

Predefined User Groups

To simplify the configuration of your backup, Data Protector provides predefined user groups with specific rights to access Data Protector functionality. For example, only members of the admin user group can access all Data Protector functionality. Operators can, by default, start and monitor backups.

TIP

In small environments, only one person is required to perform all backup tasks. This person must be a member of the Data Protector admin user group. In this case, there is no need to add other users to the Data Protector configuration.

Depending on your environment, you may decide to use the default Data Protector user groups, modify them, or create new ones.

Default Administrators

During installation, the following users are automatically added to the Data Protector admin user group:

- UNIX root user on the UNIX Cell Manager system
- User installing Data Protector on the Windows Cell Manager system

This allows them to configure and use the complete Data Protector functionality. For more information refer to online Help, index keyword “user groups, admin“.

Using Predefined User Groups

The following default groups are provided by Data Protector:

Table 4-1

Data Protector Predefined User Groups

User Group	Access Rights
Admin	Allowed to configure Data Protector and perform backup, restore, and all other available operations.
Operator	Allowed to start backups and respond to mount requests.
End-user	Allowed to perform restore of their own objects. In addition, users can monitor and respond to mount requests for their own restore sessions.

NOTE

Admin capabilities are powerful. A member of the Data Protector admin user group has system administrator privileges on all the clients in the Data Protector cell.

Data Protector User Rights

Data Protector users have the Data Protector user rights of the user group they belong to. For example, all members of the admin user group have the rights of the Data Protector admin user group.

When configuring a user from the Windows domain in Data Protector running on the UNIX Cell Manager, the user must be configured with the Domain Name or the wildcard group `"*"`.

For a detailed description of the Data Protector user rights for each user group, refer to online Help.

Users and User Groups
Users and User Groups

5

The Data Protector Internal Database

In This Chapter

This chapter describes the Data Protector internal database (IDB) architecture, as well as its usage and operation. Explanations of the database parts and their records are presented, along with recommendations on how to manage database growth and performance, including formulas for calculating its size. This information is needed to effectively administer the database configuration and maintenance.

It is organized as follows:

“About the IDB” on page 171

“IDB Architecture” on page 173

“IDB Operation” on page 178

“Overview of IDB Management” on page 180

“IDB Growth and Performance” on page 181

About the IDB

What Is the Data Protector Internal Database (IDB)?

The IDB is an embedded database, located on the Cell Manager, which keeps information regarding what data is backed up, on which media it resides, the result of backup, restore, copy, object consolidation, and media management sessions, and what devices and libraries are configured.

Why Is the IDB Used?

There are three key reasons for using the IDB:

- **Fast and convenient restore**
The information stored in the IDB enables you to quickly find the media required for a restore, and therefore makes the restore much faster. It also offers you the convenience of being able to browse for files and directories to be restored.
- **Backup management**
The information stored in the IDB enables you to verify how backups were done. You can also configure various reports using the Data Protector reporting functionality.
- **Media management**
The information stored in the IDB enables Data Protector to allocate media during backup, copy, and object consolidation sessions, track media attributes, group media in different media pools, and track media locations in tape libraries.

IDB Size and Growth Consideration

The IDB can grow very big and have a significant impact on backup performance and the Cell Manager system. Therefore, the Data Protector administrator must understand the IDB and, according to needs, decide which information to keep in the IDB and for how long. It is the administrator's task to balance between restore time and functionality on the one hand, and the size and growth of the IDB on the other. Data Protector offers two key parameters to assist in balancing your needs: **logging level** and **catalog protection**. Refer also to "IDB Growth and Performance" on page 181.

The IDB on the Windows Cell Manager

IDB Location

The IDB on the Windows Cell Manager is located in the `<Data_Protector_home>\db40` directory.

IDB Format

The IDB on the Windows Cell Manager stores all text information in UNICODE, double-byte format. Therefore, the IDB grows slightly faster than the IDB on the UNIX Cell Manager, which stores information in the ASCII format.

The UNICODE format allows for full support of filenames and messages localized to other languages.

The IDB on the UNIX Cell Manager

IDB Location

The IDB on the UNIX Cell Manager is located in the `/var/opt/omni/server/db40` directory.

IDB Format

The IDB on the HP-UX and Solaris Cell Manager stores all text information in ASCII single- and multi-byte formats.

The ASCII format limits the support of filenames and messages localized to other languages. When backing up files with filenames in a double-byte format, such as UNICODE, the filenames are converted to the ASCII format and may not appear correctly in the Data Protector user interface. However, the files and filenames will be restored correctly.

For more information, refer to “Internationalization” on page B-14.

The IDB in the Manager-of-Managers Environment

In the Manager-of-Managers (MoM) environment, you can use the Centralized Media Management Database (CMMDB), which allows you to share devices and media across several cells. Refer to “Enterprise Environments” on page 15 for more information on the MoM functionality.

IDB Architecture

The IDB consists of the following parts:

- MMDB (Media Management Database)
- CDB (Catalog Database), divided into two parts: filenames and other CDB records
- DCBF (Detail Catalog Binary Files)
- SMBF (Session Messages Binary Files)
- SIBF (Serverless Integrations Binary Files for the NDMP integration)

Each of the IDB parts stores certain specific Data Protector information (records), influences IDB size and growth in different ways, and is located in a separate directory on the Cell Manager. Refer to Figure 5-1.

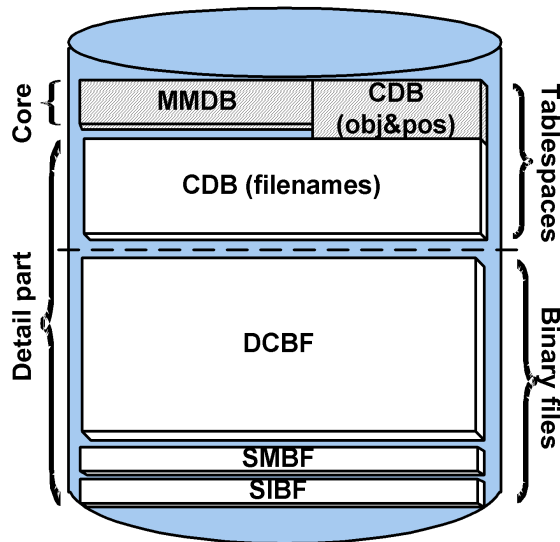
For robustness considerations and recommendations for optimizing robustness by relocating some IDB directories, refer to online Help, index keyword “robustness of IDB“.

Underlying Technology

The MMDB and CDB parts are implemented using an embedded database consisting of tablespaces. This database is controlled by the RDS database server process. All changes to the MMDB and CDB are updated using transaction logs. The transaction logs are stored in the `db40\logfiles\syslog` directory. The CDB (objects and positions) and the MMDB parts represent the core part of the IDB.

The DCBF, SMBF and SIBF parts of the IDB consist of binary files. Updates are direct (no transactions).

Figure 5-1 IDB Parts



Media Management Database (MMDB)

MMDB Records

The Media Management Database stores information about the following:

- Configured devices, libraries, library drives, and slots
- Data Protector media
- Configured media pools and media magazines

MMDB Size and Growth

The MMDB does not grow very big in size. The largest portion of the MMDB is typically occupied by information about the Data Protector media. Space consumption is in the range of 30 MB. Refer to “IDB Size Estimation” on page 187 for more details.

MMDB Location

The MMDB is located in the following directory:

- On Windows: `<Data_Protector_home>\db40\datafiles\mmdb`
- On UNIX: `/var/opt/omni/server/db40/datafiles/mmdb`

Catalog Database (CDB)

CDB Records

The Catalog Database stores information about the following:

- Backup, restore, copy, object consolidation, and media management sessions. This is a copy of the information sent to the Data Protector Monitor window.
- Backed up objects, their versions, and object copies.
- Positions of backed up objects on media. For each backed up object, Data Protector stores information about the media and data segments used for the backup. The same is done for object copies and object mirrors.
- Pathnames of backed up files (filenames) together with client system names. Filenames are stored only once per client system. The filenames created between backups are added to the CDB.

Filename Size and Growth

The biggest and fastest growing part of the CDB is the filenames part. It typically occupies 20% of the entire database.

The growth of the filenames part is proportional to the growth and dynamics of the backup environment, and not to the number of backups.

A file or directory on the HP-UX or Solaris Cell Manager occupies approximately 50-70 bytes, and a file or directory on the Windows Cell Manager occupies 70-100 bytes in the IDB.

Filenames are stored in the `fnames.dat` file and in some other files, depending on the filename length. The maximum size of each of these files is 2 GB. You are notified when one of these files starts running out of space, so that you can add new files to extend the size of the filenames part of the IDB.

Size and Growth for CDB (Objects and Positions)

The CDB records other than filenames occupy a minor share of space in the IDB. Space consumption is in the range of 100 MB for a medium size backup environment. Refer to “IDB Size Estimation” on page 187 for more details.

CDB Location

The CDB is located in the following directory:

- On Windows: `<Data_Protector_home>\db40\datafiles\cdb`
- On UNIX: `/var/opt/omni/server/db40/datafiles/cdb`

Detail Catalog Binary Files (DCBF)

DCBF Information The Detail Catalog Binary Files part stores file version information. This is information about backed up files, such as file size, modification time, attributes/protection, and so on.

One DC (Detail Catalog) binary file is created for each Data Protector medium used for backup. When the medium is overwritten, the old binary file is removed and a new one is created.

DCBF Size and Growth

In an environment where filesystem backups using the `Log all` option are typical, the DCBF occupies the largest part (typically 80%) of the IDB. Approximately 30 bytes are used for each version of each backed up file. Logging level and catalog protection can be used to specify what is actually stored in the IDB and for how long. Refer to “IDB Growth and Performance: Key Tunable Parameters” on page 182.

By default, the `db40\dcbf` DC directory is configured for the DC binary files. Its default maximum size is 4 GB. You can create more DC directories and have them on different disks on the Cell Manager, thus extending IDB size. The maximum number of supported directories per cell is 10.

DCBF Location

By default, the DCBF is located in the following directory:

- On Windows: `<Data_Protector_home>\db40\dcbf`
- On UNIX: `/var/opt/omni/server/db40/dcbf`

Consider the disk space on the Cell Manager and relocate the DC directory, if necessary. You can create more DC directories and locate them to different disks. Create several DC directories only if the number of media/DC binary files grows very large (several thousand) or if you have space problems. Refer to online Help, index keyword “DC directories“ for more information.

Session Messages Binary Files (SMBF)

SMBF Records

The Session Messages Binary Files stores session messages generated during backup, restore, copy, and media management sessions. One binary file is created per session. The files are grouped by year and month.

SMBF Size and Growth

The SMBF size depends on the following:

- The number of sessions performed, since one binary file is created per session.
- The number of messages in a session. One session message occupies approximately 200 bytes on Windows and 130 bytes on UNIX systems. You can change the amount of messages displayed when backup, restore, and media management operations are performed by specifying the `Report level` option. This also influences the amount of messages stored in the IDB. Refer to online Help for more details.

SMBF Location

The SMBF is located in the following directory:

- On Windows: `<Data_Protector_home>\db40\msg`
- On UNIX: `/var/opt/omni/server/db40/msg`

You can relocate the directory by editing the `SessionMessageDir` global option. Refer to the *HP OpenView Storage Data Protector Troubleshooting Guide* for more information on the Data Protector global options file.

Serverless Integrations Binary Files (SIBF)

SIBF Records

The Serverless Integrations Binary Files stores raw NDMP restore data. This data is necessary for restore NDMP objects.

SIBF Size and Growth

The SIBF does not grow very big in size. Refer to “IDB Size Estimation” on page 187 for more details. For NDMP backups, the SMBF grows proportionally to the number of objects backed up. Approximately 3 KB are used for each backed up object.

SIBF Location

The SIBF is located in the following directory:

- On Windows: `<Data_Protector_home>\db40\meta`
- On UNIX: `/var/opt/omni/server/db40/meta`

IDB Operation

During Backup

When a backup session is started, a session record is created in the IDB. Also, for each object and each object mirror in the session, an object version record is created. All these records are stored in the CDB and have several attributes. The Backup Session Manager updates media during a backup. All media records are stored in the MMDB and are allocated for a backup depending on policies.

When a data segment is written to the tape and then to a catalog segment, then for each object version that was part of this data segment, a media position record is stored in the CDB. In addition, the catalog is stored in the DC (Detail Catalog) binary file. One DC binary file is maintained per Data Protector medium. A DC binary file is named `<MediumID>_<TimeStamp>.dat`. If a medium is overwritten during a backup, its old DC binary file is removed and a new one is created.

All session messages generated during backups are stored in session messages binary files (the SMBF part).

If transaction logging is enabled, an IDB backup removes old transaction logs and starts creating new ones, which are necessary for an IDB recovery.

During Restore

When configuring a restore, Data Protector performs a set of queries in the CDB and DCBF parts to enable users to browse virtual filesystems of backed up data. These browse queries are done in two steps. The first step is to select a specific object (filesystem or logical drive). If this object has many backup versions and/or copies stored, this can take some time because Data Protector scans the DCBF to build a lookup cache for later browsing. The second step is browsing the directories.

After specific versions of files are selected, Data Protector determines the required media and locates media position records that are used by the selected files. These media are then read by Media Agents and data is sent to the Disk Agents that restore the selected files.

During Object Copying or Object Consolidation

During an object copy or object consolidation session, the same processes run as during a backup and a restore session. Basically, data is read from source media as if it was restored and written to target media as if it was

backed up. An object copy or object consolidation session has the same effect on the IDB operation as backup and restore. For details, see “During Backup” on page 178 and “During Restore” on page 178.

Exporting Media

When a medium is exported, the following is removed:

- All the media position records from that medium are removed from the CDB.
- All objects and object copies that now have no positions on any other media are removed from the CDB part.
- Obsolete sessions (whose media have either been overwritten or exported) older than 30 days are removed (this can be modified using the `KeepSession` variable from the global option file). Session messages of such sessions are also removed.
- The medium record is removed from the MMDB part, and the DC binary file for that medium is removed from the DCBF.

Removing the Detail Catalog

When the detail catalog is removed for a specific medium, its DC binary file is removed. The same result is achieved by removing the catalog protection for all object versions and object copies on that medium (the next daily maintenance of DC binary files removes the binary file). All other records stay in the CDB and MMDB and it is possible to run a restore from such media (however, browsing is not possible).

Filenames Purge

DC binary files show whether a given file is backed up on a related medium or not, but the filenames are actually stored in the CDB. A filename is considered “used” if it is marked as backed up in at least one DC binary file. Over time, it can happen that a large number of filenames are not used. To remove such filenames, Data Protector scans all DC binary files and then removes unused filenames.

File Versions Purge

File versions purge used to be a major maintenance task in OmniBack II A.03.50 and earlier versions. With OmniBack II A.04.10 and later, it is reduced to a minor automatic daily maintenance task.

When the catalog protection of all object versions stored on a specific medium expires, automatic daily maintenance of DC binary files removes the respective binary file.

Overview of IDB Management

IDB Configuration One of the most important steps in setting up your Data Protector backup environment is to configure the IDB. The initial configuration enables you to set your internal policies regarding IDB size, the location of IDB directories, the IDB backup necessary in case of IDB corruption or a disaster, and the configuration of IDB reports and notifications.

IMPORTANT It is highly recommended to schedule an IDB backup to be performed on a daily basis. Creating a backup specification for the IDB backup is part of the IDB configuration.

IDB Maintenance Once you configure the IDB, its maintenance is reduced to a minimum, mainly acting on notifications and reports.

IDB Recovery An IDB recovery is needed if some of the IDB files are missing or corrupted. The recovery procedure depends on the level of corruption. Refer to online Help, index keyword “IDB, recovery“ for detailed information.

IDB Growth and Performance

For proper IDB configuration and maintenance it is necessary to understand the key factors that influence the IDB growth and performance, as well as the key tunable parameters that you can adapt to your needs, and thus handle the growth and performance of the IDB as efficiently as possible.

Key IDB Growth and Performance Factors

The key factors for IDB growth and performance are the following:

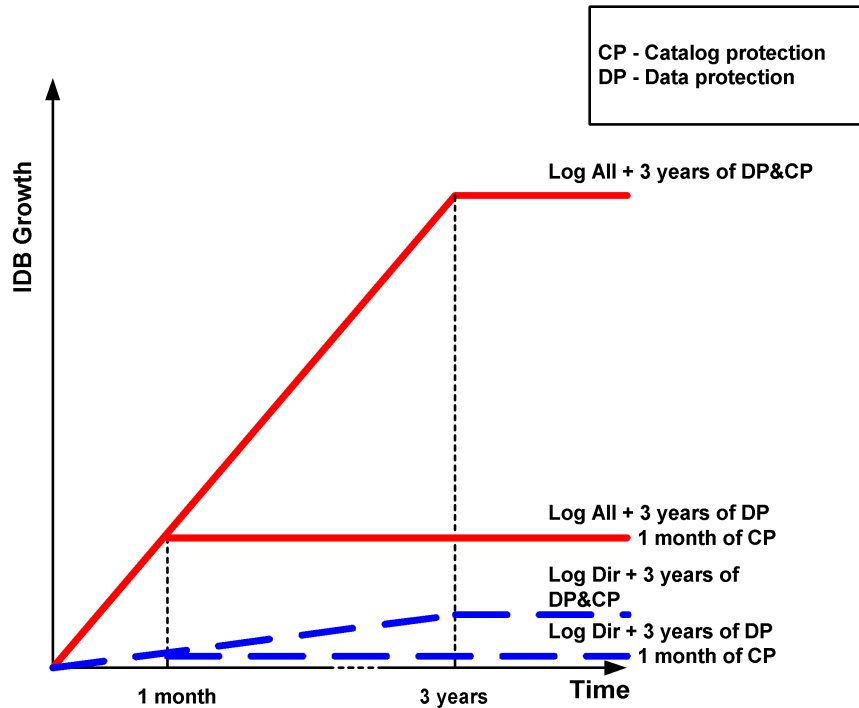
- **Logging level settings**
Logging level defines the amount of detail written to the IDB during backup. The more detailed logging level you use, the greater influence it has on the IDB. For details, see “IDB Growth and Performance: Key Tunable Parameters” on page 182.
- **Catalog protection settings**
Catalog protection determines how long the information about backed up data is available in the IDB. The longer period of catalog protection you set, the greater influence it has on the IDB. For details, see “IDB Growth and Performance: Key Tunable Parameters” on page 182.
- **Number of backed up files**
Data Protector keeps track of each file and each version of that file. Different backup types impact the IDB differently. For information on backup types, see “Full and Incremental Backups” on page 66.
- **Number of backups**
The more often you perform a backup, the more information is stored in the IDB.
- **Filesystem dynamics**
The number of files created and removed between backups can have a significant impact on the growth of the filenames part of the IDB. The Report on System Dynamics gives you information about the system dynamics. You can avoid the IDB growth due to filesystem dynamics by using the Log Directories logging level.

- **Growth of your backup environment**
The number of systems being backed up in the cell influences the IDB growth. Plan for the growth of your backup environment.
- **Character encoding used for your filenames (applicable for UNIX only)**
Depending on the filename encoding, a character in the filename can take up from one to three bytes in the IDB. Shift-JIS encoded filenames, for example, take up to three bytes in the IDB, while pure ASCII filenames take up only one byte. The character encoding is relevant for growth of filename part of IDB on UNIX (on Windows, all characters take up two bytes in the IDB).
- **Number of object copies and object mirrors**
The more object copies and object mirrors you create, the more information is stored in the IDB. For object copies and object mirrors, the IDB stores the same information as for backed up objects, except for filenames.

IDB Growth and Performance: Key Tunable Parameters

The logging level and catalog protection are the main factors of the IDB growth and performance. Their impact on the IDB depends on the settings you use. For a graphic representation of the impact of different logging level and catalog protection settings, see Figure 5-2 on page 183.

Figure 5-2 The Influence of Logging Level and Catalog Protection on IDB Growth



Logging Level as an IDB Key Tunable Parameter

What Is Logging Level?

Logging level determines the amount of details about backed up files and directories written to the IDB. You can always restore your data, regardless of the logging level used during backup.

Data Protector provides four logging levels that control the amount of details about files and directories written to the IDB:

Table 5-1

Log All	Logs all detailed information about backed up files and directories (names, versions, and attributes).
----------------	--

Table 5-1

Log Files	Logs all detailed information about backed up files and directories (names and versions). This represents approximately 30% of all detailed information about backed up files and directories.
Log Directories	Logs all detailed information about backed up directories (names, versions, and attributes). This represents approximately 10% of all detailed information about backed up files and directories.
No Log	No information about backed up files and directories is logged to the IDB.

The different settings influence the IDB growth, the backup speed, and the convenience of browsing for data to be restored.

Impact on Performance

The logging level defines the amount of data written to the IDB during a backup. This also influences the IDB speed, and therefore the backup process.

Logging Level and Browsing for Restore

Changing the level of stored information affects your ability to browse files using the Data Protector GUI during a restore. If the `No Log` option is set, browsing is not possible; if the `Log Directories` option is set, browsing of directories is possible; if the `Log Files` option is set, full browsing is possible but file attributes (size, creation, and modification dates and so on) are not displayed.

Regardless of the logging level set, it is always possible to restore your data:

- Instead of browsing for your data, you can always manually select a file to restore (if you know the name of the file).
- You can retrieve information about backed up data from the media.

Logging Level and Restore Speed

The restore speed is approximately the same when the `Log All`, `Log Directories`, or `Log Files` options are set.

If the `No Log` option is set, the restore speed can be slower when restoring single files. This is because Data Protector has to read all data from the beginning of an object before finding a file to be restored.

In case of a full system restore, the whole object should be read anyway, so the logging level settings do not play an important role.

Catalog Protection as an IDB Key Tunable Parameter

What Is Catalog Protection?

Catalog protection determines how long the information about backed up data is available in the IDB. This is different from data protection, which determines how long the backed up data is available on the medium itself. If there is no catalog protection, you can still restore your data, but you cannot browse for it in the Data Protector GUI.

Catalog protection is based on the fact that the data stored last is most important and accessed most frequently. Old files are seldom searched for, and therefore it is allowable for their search to take more time.

Expired Catalog Protection

Once the catalog protection expires, the information is not immediately removed from the IDB. Data Protector removes it automatically once per day. Since the information in the IDB is organized on a per-medium basis, it is removed completely when catalog protection expires for all objects on the medium.

Impact on Performance

Catalog protection settings do not have any impact on the backup performance.

Catalog Protection and Restore

When catalog protection expires, data is restored as if it were backed up using the `No Log` option. Refer to “Logging Level as an IDB Key Tunable Parameter” on page 183.

Recommended Usage of Logging Level and Catalog Protection

Always Use Catalog Protection

Always set a reasonable level of catalog protection. The only exception is if the `Log None` option is set (in this case catalog protection does not apply anyway).

If you set the catalog protection to `Permanent`, the information in the IDB is removed only when media are exported or deleted. In this case, the size of the IDB grows linearly until the data protection period is reached, even if the number of files in the cell does not change. For example, if the data protection period is one year and media are recycled, then significant growth of the IDB stops after one year. The addition of new catalogs is approximately equal to the removal of old ones. If catalog

protection is set for 4 weeks, then significant growth of the IDB stops after 4 weeks. Therefore, in this case, the IDB is 13 times larger if the catalog protection is set to Permanent.

It is recommended that catalog protection includes at least the last full backup. For example, you can set a catalog protection of 8 weeks for full backups and one week for incremental backups.

Use Different Logging Levels in the Same Cell

A cell often consists of mail (or similar) servers that generate a large number of files on a daily basis, database servers that store all information in a handful of files, and some user workstations. Since the dynamics of these systems are rather different, it is very difficult to prescribe one setting that suits them all. Therefore, it is recommended to create several backup specifications with the following logging level settings:

- For mail servers, use the `Log Directories` option.
- For database servers, no logging is necessary as they have their own restore policies. Therefore, use the `No Log` option.
- For workstations, the `Log All` or `Log Files` options allow for searching and restoring different versions of files. For backups with the `Log Directories` or `No Log` options set, you can import catalogs from the media, which, in a reasonably short time, allows the possibility to browse for the selected object. For information on importing catalogs from media, refer to online Help, index keyword “importing, catalogs from media“.

Different Logging Levels for Object Copies

Backed up objects and object copies or mirrors of these objects can have the same or different logging levels. Depending on your backup policy, the selected logging level of object copies can be more or less detailed than that of the source objects.

For example, you can specify the `No Log` option for object mirrors if you create these mirrors just to ensure a successful completion of a backup session. Or, you can specify the `No Log` option for a backup object to increase the backup performance, and then specify the `Log All` option for this object in a subsequent object copy session.

Specifics for Small Cells

If the number of files in a cell is small and will remain small (a million files or less) and the systems in the cell perform usual business activities, you can always use the `Log All` option, which is the Data Protector default. However, you need to take care of IDB growth and set a reasonable level of catalog protection.

Specifics for Large Cells If the number of files grows into the tens of millions, or there are tens of thousands of files generated on a daily basis, and you use the `Log All` option, then backup speed and IDB growth will become a problem in a relatively short period of time.

In this situation, you have the following options:

- Reduce the logging level to the smallest acceptable level. Setting the `Log Files` option can reduce the IDB size to a third, and setting the `Log Directories` option to almost a tenth. This, of course, depends on the nature of the file systems in the cell.
- Reduce the catalog protection to a minimum.
- Split the cell in two. As a final solution, you can always introduce another IDB and redirect half of the systems into it.

You can configure `Report on System Dynamics`, which informs you about dynamics of the growth of filenames on a particular client.

IDB Size Estimation

If you mainly perform filesystem backups, the IDB can grow to a significant size (larger than 16 GB) under certain conditions. If you perform disk image or online database, it is very likely that your IDB will not grow beyond 2 GB.

To estimate the size of the IDB use the Internal Database Capacity Planning Tool located at:

- On the UNIX Cell Managers:
`/opt/omni/doc/C/IDB_capacity_planning.xls`
- On the Windows Cell Manager:
`<Data_Protector_home>\docs\IDB_capacity_planning.xls`

You can also use this tool to estimate the size of the IDB in environments with online databases (Oracle, SAP R/3).

6 **Service Management**

In This Chapter

Service Management, reporting, and monitoring help administrators manage their backup environments more effectively. This chapter describes the concepts behind the service management features and benefits available in both a standalone Data Protector installation and through its integration with HP OpenView service management products.

It is organized as follows:

“Overview” on page 191

“Native Data Protector Functionality” on page 194

“Service Management Integrations” on page 203

Overview

Enterprise information technology (IT) departments are increasingly using service management tools, techniques, and methods to set service level expectations, measure service delivery against those expectations, and to justify future service expansion.

Because IT groups must manage the risk of data loss, data backup and recovery are critical elements in IT service delivery and management. Threats ranging from user error to viruses or other unauthorized data access and modification, or the occasional failure of the storage device itself put data at risk constantly. Business-critical data loss can cost the enterprise thousands, even millions of dollars per hour of downtime.

Users, however, may perceive data backup as something that can slow down or deny access to services while the backup is being conducted. But without this key activity, the continued availability and timeliness of services can be compromised and placed at significant risk.

While all data is at risk, not all data justifies equal recovery ability. IT departments must protect the business-critical data to a higher level of protection than the less valuable data - and do so cost effectively.

Service management measures and reports are a key tool IT managers can use to demonstrate value delivered to the organization and also to maintain competitive cost structures. Service providers use Service Level Agreements (SLAs), that typically establish availability and performance objectives, to document provider-customer contractual expectations.

Demonstrating SLA compliance requires constant monitoring and periodic reporting to show whether SLA expectations have been met. Data Protector, out of the box has monitoring, notification, and reporting tools to document backup and recovery operations. Integration with other OpenView service management products consolidates service views, service performance data, and other capabilities into one console, giving you better information and insight into overall IT service delivery.

Data Protector provides IT service managers with key data to enable operative monitoring and planning of backup and data recovery operations. This data can be used in service availability and recovery

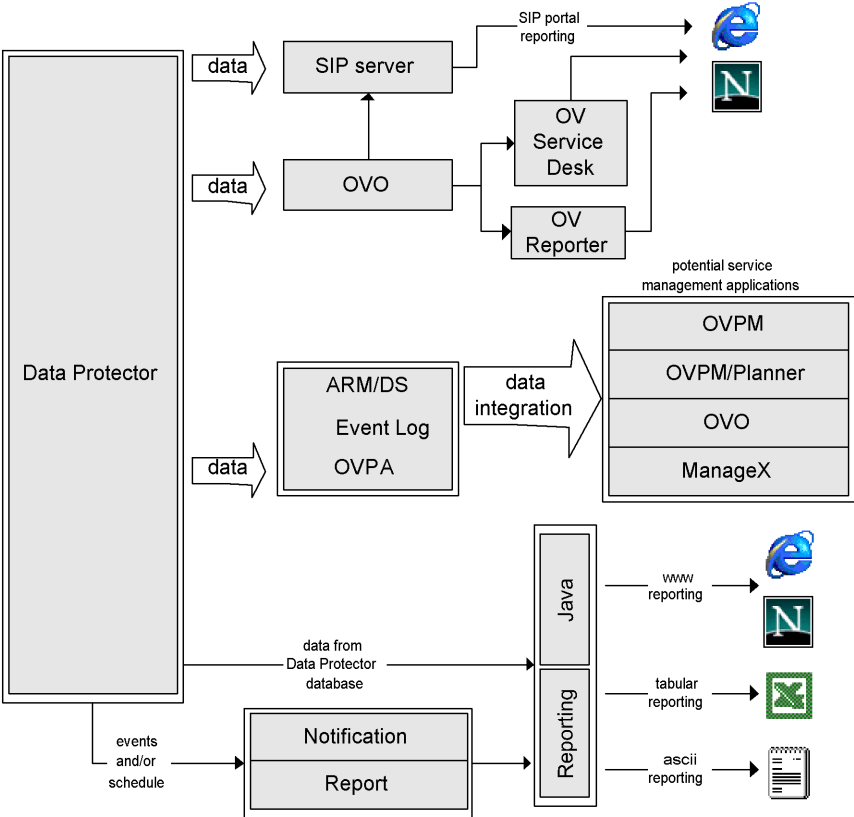
planning activities that are key if service agreements are to be adhered to. In addition, Data Protector information can be used to implement cost management and chargeback models for true IT financial management.

Data Protector and Service Management

Data Protector provides service management supports and can be integrated with service management applications, such as ManageX, OpenView Performance Agent (formerly MeasureWare Agent), OpenView Reporter, OpenView Service Desk, and OpenView Service Information Portal.

Data Protector service management falls into two categories: native (or out-of-the-box) and application integrations. The items in each category are described in more detail later in this chapter.

Figure 6-1 Service Management Information Flow



Native Data Protector Functionality

The functionality described in the following sections comes with Data Protector “out of the box.”

Key Functions

- Data Protector has been equipped to track the elapsed times of key operations and to register this data as well as volume data using the Application Response Measurement Version 2.0 API (ARM 2.0 API). Registration of this data can be performed with HP OpenView Performance Agent (OVPA).
- Built-in monitoring of running sessions allows you to instantly react to occurrences in your backup environment.
- The Data Protector built-in notification and reporting engine allows you to receive concise reports as well as immediate alerts in many different formats (such as ASCII, HTML, and spreadsheet compatible format) and delivered in various ways (such as e-mail, SNMP, broadcast (available on Windows only), write to file, and send to external command). As the Data Protector built-in notification engine can send alerts via SNMP, it is possible to integrate virtually any application that can receive SNMP traps.
- The integration of Data Protector with HP OpenView Operations allows you to receive alerts from Data Protector on the OVO console and have automatic actions performed.
- The Data Protector capability to send major and critical events into the Windows Event Log opens up a variety of interesting integration possibilities.
- The integration with HP OpenView ManageX automatically forwards Data Protector major and critical events to the ManageX operator console. Automatic actions can be set up to react upon failures in the backup environment.
- The Data Protector built-in Java-based online reporting allows you to do online reporting from wherever you are in your network (even from a remote location) without the need to have the Data Protector user interface installed on your local system. This functionality requires a Web browser.

Application Response Measurement Version 2.0 (ARM 2.0 API)

What Is ARM?

The ARM API is an emerging standard for measuring end-to-end response times of transactions in distributed environments. Application programs that use the ARM API act as sources of response time information (and also user supplied information that may be relevant to a particular transaction) for ARM compliant system management and monitoring tools such as HP OpenView Performance Agent (OVPA). OVPA will log ARM transaction information in its repository for subsequent analysis and reporting. It can also raise real time alerts (or “alarms”) when the elapsed time of a specific transaction, such as a backup operation, exceeds a predefined threshold. When a real time alert is raised, a number of actions are possible, including, but not limited to, informing a central operations console, such as HP OpenView Operations, paging a system operator, or taking automated remedial action to resolve the problem.

Table 6-1 ARM Functionality

Transaction Description (ARM 1.0)	Additional Data Logged to ARM (ARM 2.0)	Usage
Backup specification session duration	Processed data [MB]	Availability and recovery planning. Chargeback.
Object backup session duration	Processed data [MB]	Availability and recovery planning. Chargeback.
Restore session duration	Recovered data [MB]	Availability and recovery planning
IDB check duration	IDB size [MB]	Data Protector architecture management
IDB purge duration	IDB size after purge and number of purged records	Data Protector architecture management

As Data Protector is already ARM equipped, it is a fairly simple task to integrate Data Protector with an application like OVPA that supports the ARM API. On Windows platforms, this is completely automatic. If Data Protector is installed on a system where OVPA is already present or vice versa, the transaction data will immediately show up in OVPA and HP OpenView Performance Manager (OVPM). On HP-UX, the only required task is to create a link from an OVPA library to a Data Protector directory. For more information, refer to online Help, index keyword “ARM integration, installing”.

Another way to interface OVPA with Data Protector is the Data Source Integration (DSI). This is important if the application that you use for transaction tracking is not ARM 2.0 compliant. ARM 1.0 permits you to log time-specific data only, such as the duration of a backup session. With DSI you can add the capability to report any data which can be retrieved from the command line to tools like OVPA. This enables highly customized reporting.

Integration with HP OpenView Operations

Functionality of the Data Protector OVO Integration

Data Protector integrates with HP OpenView Operations (OVO). OVO simplifies management of large networks by allowing the operator to monitor and administer the network and the applications from a single point. Once Data Protector is integrated in the OVO environment, the network administrator can immediately see if anything is wrong during backup and react upon the information given. Data Protector messages can be displayed in the OVO message window.

Integration with ManageX

Functionality of the Data Protector ManageX Integration

The Data Protector ManageX integration is available on Windows platforms only. It provides the following functionality:

- Data Protector writes all major and critical messages that occur during backup, restore or any other operation to the Windows Event Log. ManageX then uses these events and forwards them to the ManageX console, so that an operator can react to them.

- Service monitoring

ManageX monitors all Data Protector services running on the Cell Manager as well as any Data Protector client system. In case of failure of any of these services, ManageX immediately alerts the operator. ManageX can also be configured in such a way that it automatically attempts to restart the failed service.

ManageX 3.5 and higher has these policies already built in. To make use of this integration, simply distribute the respective ManageX policies to Data Protector systems.

SNMP Traps

SNMP traps allow a Service Management application to receive and process an SNMP trap message when a Data Protector event occurs or when an SNMP trap is sent as a result of Data Protector's checking and maintenance mechanism. For more information on Data Protector on configuring SNMP traps, refer to online Help, index keyword "SNMP, reports send methods".

The Monitor

The Data Protector monitor is a part of the Data Protector user interface and allows you to supervise and to take corrective action on currently running backup, restore, and media management sessions.

Monitoring lets you view all sessions in a cell and shows you detailed messages and the current status of these sessions. In a multi-cell environment, you can view the sessions that run on computer systems in other cells. From the monitor's user interface, you can abort a backup, restore, or media management session or respond to "mount" requests.

If you make use of the Manager-of-Managers, you can monitor sessions of multiple cells simultaneously from one user interface.

Reporting and Notification

Data Protector reporting represents a powerful, customizable, and flexible tool for managing and planning your backup environment. Data Protector has always had a rich set of built-in reports that system administrators have relied upon to manage Cell Managers. IT Service

Providers now can use these same reports to demonstrate data protection SLA compliance. Built-in reports that are especially relevant to service level management include:

- Inventory/Status Reports such as the `host_not_conf` report, which contains information about unprotected systems, the `dl_sched` report, which lists all scheduled backups, and the `media_list` report, which is a media inventory report.
- Capacity Utilization Reports such as the licensing report, which is a Data Protector license utilization report and the `dev_unused` report, which lists devices currently not used for backup (available).
- Problem Reports such as the `backup_statistics` report, which consists of failed backup-related information. An administrator can receive an hourly, daily, or weekly E-mail report on failed jobs and the reasons for failure.

The notification and reporting capabilities that have always been part of the Cell Manager (and that have been extended significantly from earlier versions) also allow you to

- Choose from approximately 30 pre-configured reports (including, but not limited to, reports such as sessions in a specific time frame, IDB reports, and device usage report)
- Specify your own parameters for those reports (such as time frames, backup specifications, and groups of backups)
- Select from various different output formats (such as ASCII, HTML, and spreadsheet compatible formats)
- Schedule those reports with the Data Protector built-in scheduler
- Trigger report sending based on events (such as device failure, mount requests, and end of sessions)
- Select from many delivery methods used to deliver reports (such as e-mail, SNMP, broadcast (available on Windows only), write to file, and send to external command)

You can combine most of these different formats, delivery methods, schedules, and triggers.

Some examples are shown below:

Reporting and Notification Examples

- Every morning at 7:00, a report about all backup sessions in the last 24 hours is created and sent by e-mail in the ASCII format to the backup administrator's mailbox. Additionally, the same report is written to a file on your Web server in the HTML format so that others can also access this information.
- In event of a device failure or a mount request, a broadcast message is immediately sent to the backup administrator's Windows workstation, and an external command is triggered, which activates the backup administrator's pager.
- At the end of a backup session, every end user whose system has been backed up receives an e-mail in ASCII format that contains a backup status report.

Event Logging and Notification

The Data Protector Event Log is a central repository of all Data Protector-related notifications. The Data Protector built-in notification engine sends alerts or activates the Data Protector reporting mechanism based on the log entries. The event log is the information source for SLA-compliance reports in Data Protector or in OpenView management applications. In addition to reports, log entries feed OpenView management applications via the Data Protector SPI (smart plug-in) so that they can trigger preventive or corrective actions (for details, refer to example under 3.1).

Since the Data Protector built-in notification engine can send alerts via SNMP, virtually any application that can receive SNMP traps can integrate with Data Protector. Integration with OpenView Operations and OpenView Reporter is an example of SNMP trap-based implementation.

The Event Log is accessible only for Data Protector users in the Admin group and for Data Protector users that are granted the Reporting and notifications user rights. You can view or delete all events in the Event Log.

Data Protector Log Files

Some Service Management applications, such as HP OpenView Operations, allow you to specify when and which log files should be monitored for a specific log entry. If the specified entry is detected in the file, an action can be specified. In OVO this is called *Log file encapsulation*.

You can configure such a Service Management application to monitor Data Protector log files for specific log entries (Data Protector events) and define an action that is to be executed in case a particular Data Protector event is detected.

For more information on Data Protector log files refer to the *HP OpenView Storage Data Protector Troubleshooting Guide*. Note that there is no log files formatting specification provided.

Windows Application Log

Some Service Management applications, such as ManageX, monitor the Windows Application Log.

To enable automatic forwarding of all Data Protector messages and messages about the Data Protector services (if they are stopped) to Windows Application Log, set the `EventLogMessages` variable in the Data Protector global options file to 1. For more information on the Data Protector global options file, refer to the *HP OpenView Storage Data Protector Troubleshooting Guide*.

Java-Based Online Reporting

Data Protector comes with a Java-based online reporting capability that lets you configure, run, and print all Data Protector built-in reports, live and interactive. During reporting operations, Data Protector Java reporting directly accesses the Cell Manager to retrieve current data. You can make this Java applet available through a Web server, copy it to the client machine for direct access, or use it locally.

Using this facility only requires a supported Web browser; there is no need to have the Data Protector GUI installed on the system.

Not only can you use the Java reporting facility to get online access to your reports, but you can also configure your reporting structure through it, such as adding new reports to a schedule or changing a report's parameters.

Data Protector Checking and Maintenance Mechanism

Data Protector has a rich automated daily self-check and maintenance mechanism, which improves its operational reliability and predictability. Data Protector's self-check and maintenance tasks include:

- “Not Enough Free Media” check
- “Data Protector License Expiration” check

For more information, refer to online Help, index keyword “checks performed by Data Protector”.

Central Management, Distributed Environment

The Data Protector MoM enables administrators to centrally manage an enterprise environment consisting of several Data Protector Cell Managers. The MoM system administrator performs configuration, media management, monitoring, and status reporting tasks for the whole enterprise from a single console. With MoM, managing many Data Protector Cell Managers is as convenient as managing just one. IT service providers can administer larger clients' environments without adding employees. For more information on MoM, refer to online Help, index keyword “MoM environment”.

Using the Data Provided by Data Protector

Here are some examples of what you can do with the data that Data Protector provides:

What Can I Do with the Data?

- Real-time alerting of backup or restore sessions that exceed the specified time window (OVPA).
- Creating graphs of backup duration of key systems in your environment to detect trends in operation time (OVPM).
- Forecasting of the IDB growth to be able to spot points in time where certain limits will be reached (OVPM Planner).
- Regular e-mail reports to back up operators, end users, and management (Data Protector built-in reporting with the capability to send e-mails).

Native Data Protector Functionality

- Backup reports written to a Web server to make them available on an on-demand basis (built-in Data Protector reporting with the capability to write HTML).
- Sending major and critical Data Protector events to your network management solution, such as HP OpenView Network Node Manager (Data Protector built-in notification engine sending SNMP traps).

Service Management Integrations

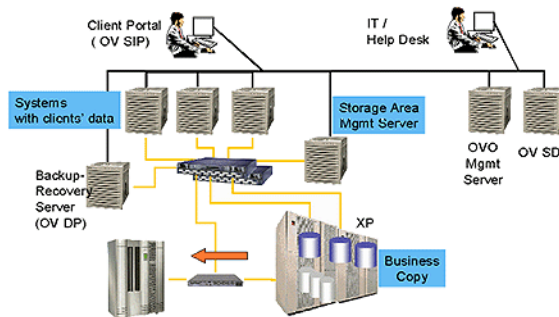
The following Data Protector integrations can be installed to help facilitate service management and to give you single-point access to rich service management functionality.

Key Functions

- Standard and custom report formats
- Availability of a “trouble ticket” interface for Data Protector
- Facilitates a specific, consistent, measurable level of service
- Availability of Data Protector information through a web interface
- Graphical representation of data

Figure 6-2

Example of an IT service provider environment with service management access through the client portal.



Data Protector OVO-OVR Integration

The integration of Data Protector with HP OpenView Operations (OVO) is extended by adding HP OpenView Reporter 3.0 (English version). With Reporter, service providers can generate reports from the OVO console as a central management point. Integration with Reporter adds a variety of new reports in the following categories:

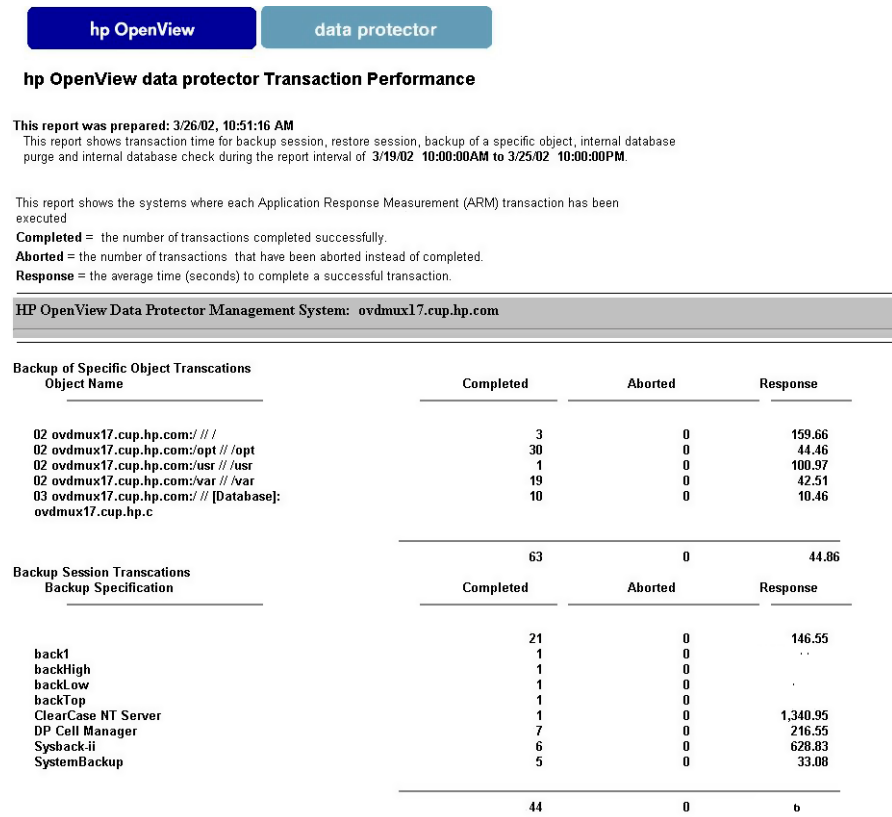
- Backup Session Reports
- Administration Reports
- Media Pool Reports

Service Management
Service Management Integrations

- Performance

An IT Service Provider can use these reports to demonstrate its SLA compliance to a customer. For example, the “Data Protector Transaction Performance” Report consists of service performance metrics (one of the IT SLA parameters):

Figure 6-3 Data Protector OVR example



In addition to SLA compliance reports, An IT Service Provider can generate monthly operational reports for the Data Protector environment. For example, the “Data Protector Operational Error Status” report aggregates the “problem” data and can be used by an IT service provider for operational planning.

Figure 6-4 Operational Error Status report

HP OpenView Storage Data Protector

hp OpenView storage data protector operational error status

This report was prepared on: 7/11/2002, 2:04:39 AM

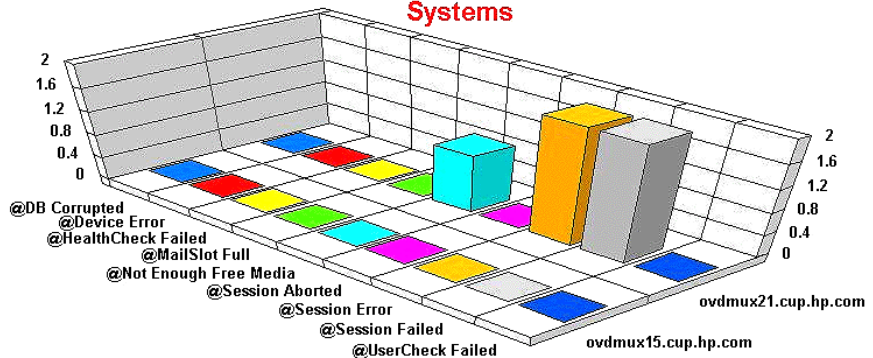
This report shows the number of operational errors that occurred on the Data Protector management systems (cell managers). Data is collected for the reporting interval of 7/10/2002 2:04:39AM - 7/11/2002 2:04:39AM. The "Operational Error Status for All Data Protector Management Systems" graph shows the sum of various errors on each Data Protector Management System. For details of the errors relating to each Data Protector management system, see the graphs titled: "Combined Error Status for All Data Protector Management Systems", "Error Status For Unacknowledged Error Messages" and "Error Status For Acknowledged Error Messages".

HP OpenView Operations Management Server: ovdmux24.cup.hp.com

Application: HP OpenView Storage Data Protector

The "Operational Error Status for All Data Protector Management Systems" graph shows the combined operational error status based on acknowledged and unacknowledged error messages for all the Data Protector Management Systems.

Operational Error Status for All Data Protector Management Systems

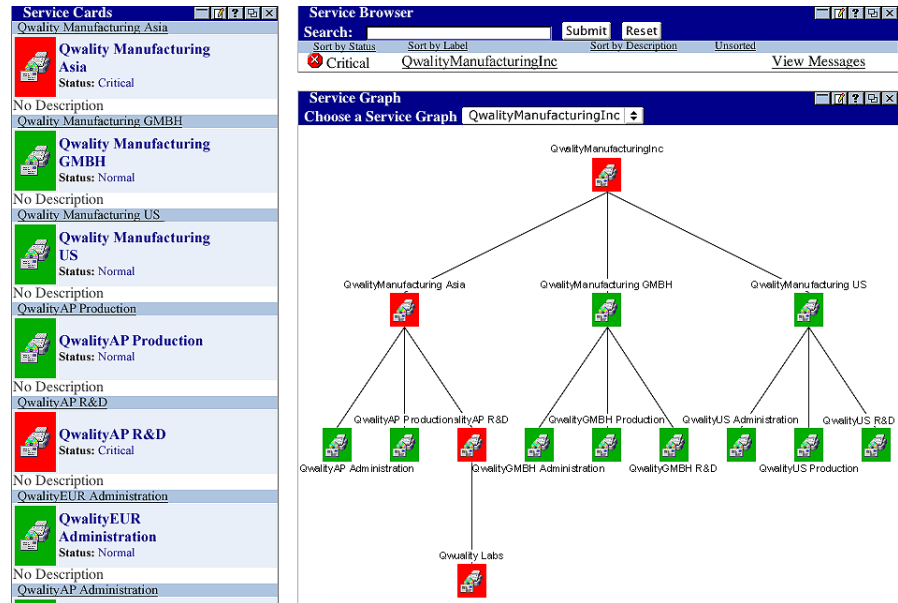


Data Protector OVO-SIP

SIP can give you visibility into services that you are providing. Rather than giving you a generalized view of your infrastructure, SIP personalizes the information for each of your customers and shows status and business information specific to the customer's environment.

With this integration, you can give your customers a view into the status of their outsourced data protection operations. This integration uses OVO components to present a graphical representation of the storage network.

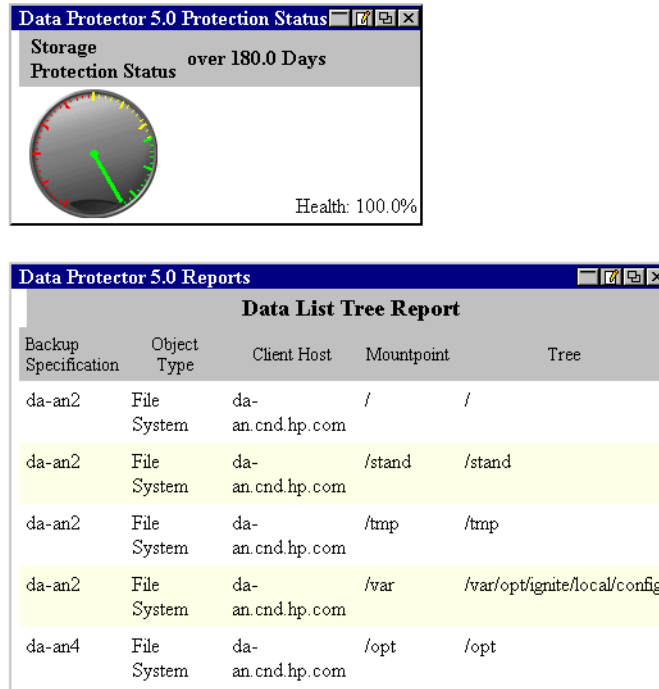
Figure 6-5 Data Protector OVO-SIP example



Data Protector SIP

This integration also uses SIP to provide Data Protector information through a web-based interface. It does not require OVO to be installed. The integration provides information through tables and gauges.

Figure 6-6 Direct SIP integration example



Integrating Data Protector with HP OpenView Service Desk

Service Desk is a help desk solution. It enables the IT support organization to implement configuration, help desk, incident resolution, problem resolution, and change management processes into a single workflow. Service Desk automates and regulates IT troubleshooting processes. It stores SLAs and monitors support-service compliance to them.

When integrated with Data Protector, Service Desk (without human involvement) monitors the time taken to resolve backup-related problems, such as adding media or restarting a failed backup, thus enhancing Data Protector's monitoring and measuring capabilities.

Service Desk manages service help desk workflow, measures service quality levels, and generates reports demonstrating SLA compliance. Data Protector A.06.00 integration with Service Desk gives support

personnel access to Data Protector data for a timely response and resolution of operational problems before they affect vital data protection services.

7**How Data Protector Operates**

In This Chapter

This chapter describes the operation of Data Protector. It explains Data Protector processes (on UNIX) and services (on Windows), backup and restore sessions, and media management sessions.

It is organized as follows:

“Data Protector Processes or Services” on page 211

“Backup Sessions” on page 212

“Restore Sessions” on page 218

“Object Copy Sessions” on page 223

“Object Consolidation Sessions” on page 226

“Media Management Sessions” on page 229

Data Protector Processes or Services

Data Protector runs several background processes (on UNIX) and services (on Windows) that enables it to run backup and restore sessions. It provides the necessary communication paths, activates backup and restore sessions, starts Disk Agents and Media Agents, stores information about what was backed up, manages media, and performs similar functions.

Inet	The Data Protector Inet service runs on each Windows system in the Data Protector cell. Inet is responsible for communication between systems in the cell and starts other processes needed for backups and restores. The Data Protector Inet service is started when Data Protector is installed on a system. On UNIX systems, the system inet daemon (INETD) starts the Data Protector Inet process.
CRS	The CRS (Cell Request Server) process (service) runs on the Data Protector Cell Manager. It starts and controls backup and restore sessions. The service is started when Data Protector is installed on the Cell Manager system and is restarted each time the system is restarted.
MMD	The MMD (Media Management Daemon) process (service) runs on the Data Protector Cell Manager and controls media management and device operations. The process is started by the Cell Request Server process (service).
RDS	The RDS (Raima Database Server) process (service) runs on the Data Protector Cell Manager and manages the IDB. The process is started when Data Protector is installed on the Cell Manager.

For instructions on how to manually start or stop the Data Protector processes and services, see online Help.

Backup Sessions

This section describes how a backup session is started, what happens during a backup session, and the processes and services involved.

What Is a Backup Session?

When a backup specification is started it is called a backup session. The backup session copies data from a source, typically a hard disk, to a destination, typically tape media. The result of a backup session is a copy of data on the backup media, the media set.

Scheduled and Interactive Backup Sessions

Scheduled Backup Session

A scheduled backup session is started by the Data Protector Scheduler at the time you have specified. You can view the progress of the scheduled backup session in the Data Protector monitor.

Interactive Backup Session

An interactive backup session is started from the Data Protector user interface directly. The Data Protector monitor starts immediately and you can view the progress of the backup session. Note that multiple users can monitor the same backup session. You may want to stop monitoring by disconnecting the user interface from the session. The session will then continue in the background.

Backup Session Data Flow and Processes

What Happens in a Backup Session?

The information flow of a backup session is shown in Figure 7-1 on page 214. Note that the data flow and processes described here are for a standard network backup. For data flow and processes specific to other types of backup, such as direct backup, refer to the related chapter.

When a backup session is started, the following happens:

1. The Backup Session Manager (BSM) process is started on the Cell Manager system and controls the backup session. This process reads the backup specification for information on what to back up, and which options, media, and devices to use for the backup.
2. The BSM opens the IDB and writes to the IDB information about the backup session, such as generated messages, details about the backed up data, and the devices and media that were used for the session.

3. The BSM starts Media Agents (MAs) on the systems with devices configured for backup. A new Media Agent is started for each drive used in parallel. The number of Media Agents that can be started in the cell is limited by the cell configuration and the number of licenses you have purchased.

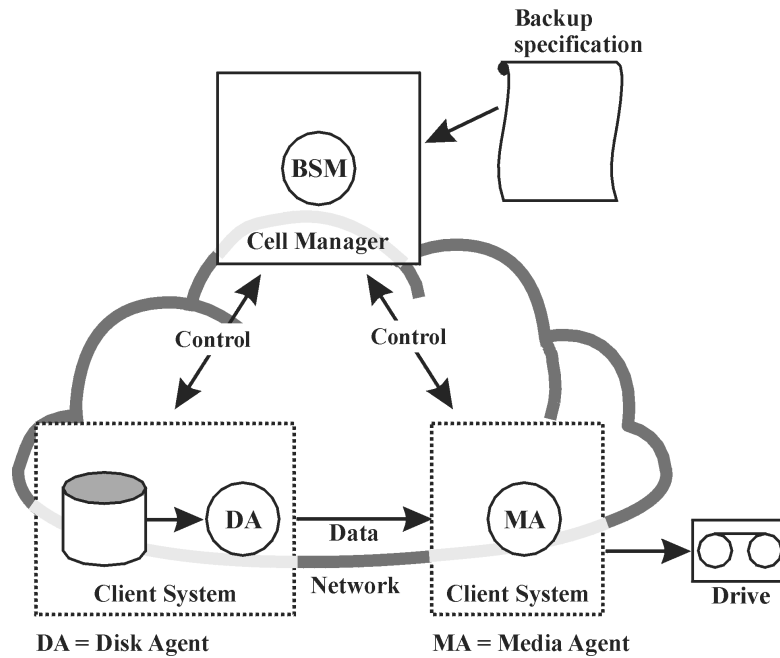
In a backup session with object mirroring, the BSM also starts Media Agents that will be used for mirroring.

4. The BSM starts Disk Agents (DAs) for each disk to be backed up in parallel. The actual number of Disk Agents started depends on the concurrency of Disk Agents configured in the backup specification. This is the number of Disk Agents that can be started to send data in parallel to a Media Agent, thus allowing a device to stream.
5. Disk Agents read data from disks and send it to the Media Agents that write data to media.

In a backup session with object mirroring, Media Agents used for writing mirrored objects are daisy-chained. Each Media Agent writes the received data to media and forwards it to the next Media Agent in the chain.

6. The BSM monitors the progress of the session and starts new Disk Agents and new Media Agents as necessary.
7. When the backup session is completed, the BSM closes the session.

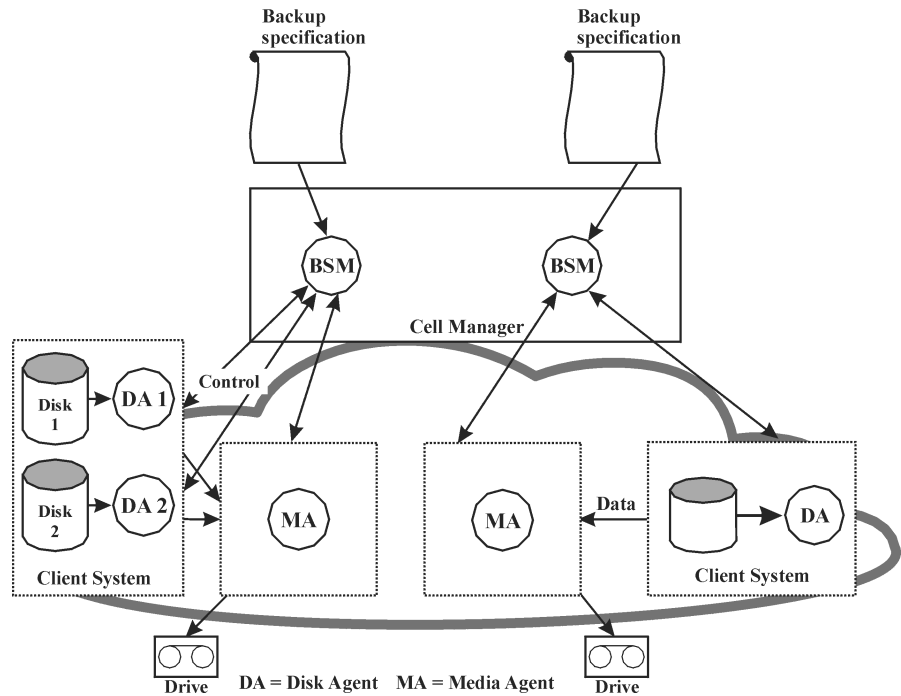
Figure 7-1 Backup Session Information Flow (1)



How Many Sessions Can Run Concurrently?

A number of backup sessions can run in the cell at the same time, as shown in Figure 7-2. This number is limited by resources in the cell, such as the configuration of the Cell Manager, for instance, processor speed, main memory size, disk space, and similar. The maximum number of backup sessions running concurrently can be configured.

Figure 7-2 Backup Session Information Flow - Multiple Sessions



Pre-Exec and Post-Exec Commands

Data Protector pre-exec commands enable you to execute some actions before a backup or a restore session. Data Protector post-exec commands enable you to execute some actions after a backup or a restore session. A typical pre-exec action would be to shut down a database to put data in a consistent state.

The pre-exec and post-exec commands can be set for a backup specification and, as such, executed on the Cell Manager system, or they can be specified as a backup object option and thus executed on the client system where the respective Disk Agent is running.

Pre-exec and post-exec script commands can be written as executables or shell scripts. These are not supplied by Data Protector and must be written separately by, for example, the backup operator.

Queuing of Backup Sessions

- Timeout** When a backup session is started, Data Protector tries to allocate all needed resources, such as devices. The session is queued until the required minimum resources are available. If the resources are still unavailable after the timeout, the session is aborted. The timeout period can be set using the `SmWaitForDevice` global option.
- Optimizing the Load** To optimize the load on the Cell Manager, Data Protector can, by default, start up to five backup sessions at the same time. The default value can be modified in the global options file. If more are scheduled at the same time, the extra sessions are queued and started subsequently as others are finished.

Mount Requests in Backup Sessions

- What Is a Mount Request?** A mount request in a backup session appears when Data Protector needs a new medium for backup and the medium is not available. Data Protector issues a mount request for one of the following reasons:
- Issuing a Mount Request**
- There is not enough space on the backup media and there are no new media available.
 - Data Protector media allocation policy for backup requires a medium that is not available in the device.
 - The order of media used for backup is defined in the pre allocation list and media are not available in this order.
- Refer to “Adding Data to Media During Backup Sessions” on page 127 and “Selecting Media for Backups” on page 126 for more information.
- Responding to a Mount Request** Responding to a mount request includes providing the required media and telling Data Protector to proceed with the backup. Data Protector allows you to configure what happens when a mount request is issued:
- Sending Notification to an Operator** You can configure a Data Protector notification to send an e-mail to the operator with information about the mount request. The operator can take the appropriate actions, such as manually loading the needed media or aborting the session. Refer to “Reporting and Notification” on page 197 for more information.

Automating a Mount Request

You can configure automated actions for the handling of mount requests. To do this, write a script or a batch program that performs the desired action.

Backing Up with Disk Discovery

What Is Disk Discovery?

In backing up with disk discovery, Data Protector creates a detailed list of disks on the target system when the backup session is started, and backs up all disks. Therefore, all local disks on the system are backed up even though they were not present on the system when the backup was configured. Backup with disk discovery is particularly useful in dynamic environments, where configurations change rapidly. It enables you to select or exclude specific directories in the backup.

How Does It Compare to a Standard Backup?

In a standard backup, you explicitly configure specific disks, directories or other objects for backup by configuring them in the backup specification. Therefore, only these objects are backed up. If you add new disks to the system or want to back up some other objects, you must manually edit the backup specification and these new objects. You can select, as you configure the backup, the method you want to use - disk discovery or standard backup.

Restore Sessions

This section describes how a restore session is started, what happens during a restore session, and the processes and services involved.

What Is a Restore Session?

In a restore session, data is copied from a backup copy, typically on a tape medium, back to a disk.

A restore session is started interactively. You tell Data Protector what to restore, let Data Protector determine the needed media, select some options and start the restore. You and other users can monitor the progress of the session.

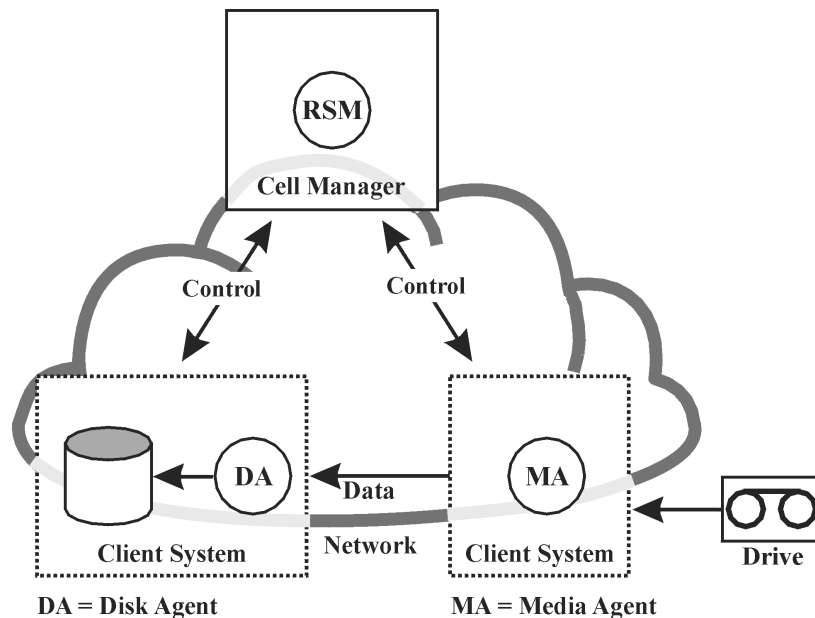
Restore Session Data Flow and Processes

What Happens in a Restore Session?

When a restore session, as shown in Figure 7-3, is started, the following happens:

1. The Restore Session Manager (RSM) process is started on the Cell Manager system. This process controls the restore session.
2. The RSM opens the IDB, reads the information about media needed for restore, and writes the information about the restore session to the IDB, such as generated messages.
3. The RSM starts Media Agents (MAs) on the systems with devices used for restore. For each drive used in parallel, a new Media Agent is started.
4. The RSM starts Disk Agents (DAs) for each disk restored in parallel. The actual number of Disk Agents started depends on the objects you selected for restore. Refer to “Parallel Restores” on page 220 for more information.
5. Media Agents read data from media and send it to the Disk Agents that write the data to disks. The RSM monitors the progress of the session and starts new Disk Agents and new Media Agents as necessary.
6. When the restore session is completed, the RSM closes the session.

Figure 7-3 Restore Session Information Flow



How Many Restore Sessions Can Run Concurrently? A number of restore sessions can run in the cell at the same time. This number is limited by resources in the cell, such as the Cell Manager and systems with connected devices.

Queuing of Restore Sessions

Timeout When a restore session is started, Data Protector tries to allocate all needed resources, such as backup devices. The session is queued for as long as the required minimum resources are not yet available. Data Protector tries to allocate the resources for a specific period of time, the timeout. Timeout is user configurable. If the resources are still unavailable after the timeout, the session is aborted.

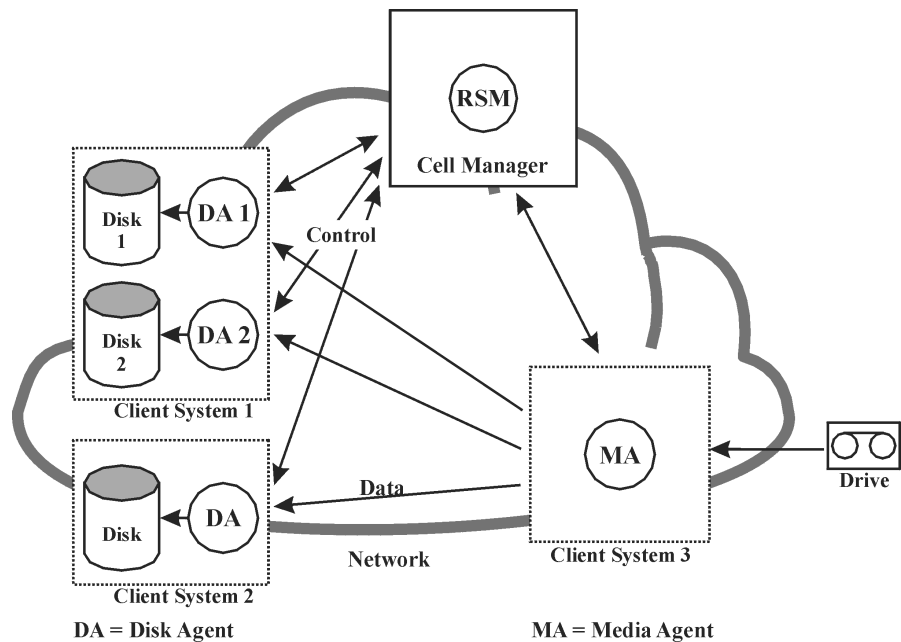
Mount Requests in a Restore Session

- What Is a Mount Request?** A mount request appears in a restore session when the media needed for restore are not available in the device. Data Protector allows you to configure a desired action that should happen when a mount request appears.
- Responding to a Mount Request** Responding to a mount request includes providing the required media or any copy of media and telling Data Protector to proceed with the restore.

Parallel Restores

- What Is a Parallel Restore?** In a parallel restore, interleaved data from multiple objects is read concurrently from media in a single path and restored. A parallel restore significantly improves restore performance when restoring multiple objects from the same media. Refer to Figure 7-4 for more information.
- How Does It Compare to a Standard Restore?** Data from multiple Disk Agents is (most of the time) multiplexed and stored on the media. Refer to “Multiple Objects and Sessions per Medium, Concurrent Writes” on page 128. In a standard restore, Data Protector reads multiplexed data from the media and assembles only the parts needed for the selected object. When the next object is restored, Data Protector must rewind the media and read the parts for the other object, assuming both objects are on the same medium and written using multiplexing.

Figure 7-4 Parallel Restore Session Flow



In a parallel restore, Data Protector reads multiplexed data for all selected objects and assembles the parts needed for all the objects on the fly, sending the right data to the right Disk Agents. This improves performance when reading from the media. The performance is additionally improved if the selected objects are written to different physical disks. In this case, data is copied to multiple disks at the same time.

Fast Multiple Single File Restore

Data Protector uses discontinuous object restore to improve restore performance. After restoring a specific file or tree, Data Protector repositions itself directly on the next file or tree on the medium, if there's at least a single segment between the files or trees, and continues the restore.

Restore Sessions

Within an individual restore object you can start multiple Disk Agents. This way the restoring of multiple single files that are located all over the medium is much faster than if Data Protector were to traverse the medium.

Object Copy Sessions

This section describes how an object copy session is started, what happens during the session, and the processes and services involved.

What Is an Object Copy Session?

An object copy session is a process that creates an additional copy of the backed up data on a different media set. During an object copy session, the selected backed up objects are copied from the source to the target media.

Automated and Interactive Object Copy Sessions

Automated Object Copy Session

An automated object copy session can either be scheduled or started immediately after a backup. A scheduled object copy session is started at the time you have specified using the Data Protector Scheduler. A post-backup object copy session is started after the specified session finishes. You can view the progress of the automated object copy session in the Data Protector monitor.

Interactive Object Copy Session

An interactive object copy session is started from the Data Protector user interface directly. The Data Protector monitor starts immediately and you can view the progress of the session. Multiple users can monitor the same object copy session. You may want to stop monitoring by disconnecting the user interface from the session. The session will then continue in the background.

Object Copy Session Data Flow and Processes

What Happens in an Object Copy Session?

The information flow of an object copy session is shown in Figure 7-5 on page 225. When an object copy session is started, the following happens:

1. The Copy and Consolidation Session Manager (CSM) process is started on the Cell Manager system. This process reads the object copy specification for information on what to copy and which options, media, and devices to use. It also controls the object copy session.
2. The CSM opens the IDB, reads the information about the media needed for copying, and writes the information about the object copy session, such as generated messages, to the IDB.

Object Copy Sessions

3. The CSM locks the Media Agents. The session is queued until all read Media Agents and the minimum required write Media Agents are locked, with the same timeout as for backup. If the resources are still unavailable after the timeout, the session is aborted.
4. The CSM starts the Media Agents on the systems with devices configured for copying. Media Agents are loaded with source and target media allocated according to the backup policies.
5. Media Agents read the data from the source media and connect to the Media Agents loaded with the target media.

If destination devices are not specified per object, Data Protector selects them automatically from those you selected in the object copy specification according to the following criteria in the order of priority:

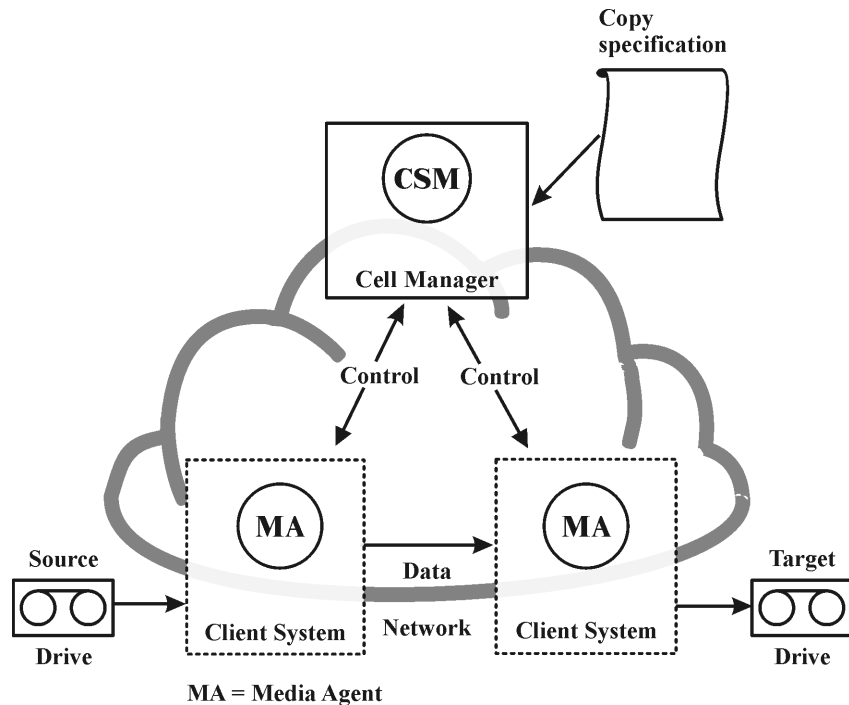
- destination devices with the same block size as source devices are selected before those with a different block size
 - locally attached devices are selected before network attached devices
6. Media Agents loaded with the target media accept connections from the Media Agents loaded with the source media and start writing object copies to the target media.

If the block size of the source device is smaller than the block size of the destination device, blocks are repackaged at this stage of the object copy session.
 7. When the object copy session is completed, the CSM closes the session.

How Many Sessions Can Run Concurrently?

A number of object copy sessions can run in the cell at the same time. This number is limited by the resources in the cell, such as the Cell Manager and the systems with connected devices.

Figure 7-5 Object Copy Session Information Flow



Queuing of Object Copy Sessions

Timeout

When an object copy session is started, Data Protector tries to allocate all needed resources. The session is queued until the required minimum resources are available. If the resources are still unavailable after the timeout, the session is aborted. The timeout period can be set using the `SmWaitForDevice` global option.

Mount Requests in an Object Copy Session

What Is a Mount Request?

A mount request in an object copy session is issued when a source or a target medium needed for the object copy operation is not available.

Responding to a Mount Request

Responding to a mount request includes providing the required medium and confirming the mount request. If the required source medium has media copies, you can provide a copy instead of the original medium.

Object Consolidation Sessions

This section describes how an object consolidation session is started, what happens during the session, and the processes and services involved.

What Is an Object Consolidation Session?

An object consolidation session is a process that merges a restore chain of a backup object, consisting of a full backup and at least one incremental backup, into a new, consolidated version of this object. During an object consolidation session, Data Protector reads the backed up data from the source media, merges the data, and writes the consolidated version to the target media.

For more information, see Chapter 11, “Synthetic Backup,” on page 261.

Automated and Interactive Object Consolidation Sessions

Automated Object Consolidation Session

An automated object consolidation session can either be scheduled or started immediately after a backup. A scheduled object consolidation session is started at the time you have specified using the Data Protector Scheduler. A post-backup object consolidation session is started after the specified backup session finishes. You can view the progress of an automated object consolidation session in the Data Protector monitor.

Interactive Object Consolidation Session

An interactive object consolidation session is started from the Data Protector user interface directly. The Data Protector monitor starts immediately and you can view the progress of the session. Multiple users can monitor the same object consolidation session. You may want to stop monitoring by disconnecting the user interface from the session. The session will then continue in the background.

Object Consolidation Session Data Flow and Processes

When an object consolidation session is started, the following happens:

1. The Copy and Consolidation Session Manager (CSM) process is started on the Cell Manager system. This process reads the object consolidation specification for information on what to consolidate and which options, media, and devices to use. It controls the object consolidation session.
2. The CSM opens the IDB, reads the information about the needed media, and writes the information about the object consolidation session, such as generated messages, to the IDB.
3. The CSM locks the Media Agents. The session is queued until all read Media Agents and the minimum required write Media Agents are locked, with the same timeout as for backup. If the resources are still unavailable after the timeout, the session is aborted.
4. The CSM starts the Media Agents on the systems with devices that will be used in the session. Media Agents are loaded with source and target media allocated according to the backup policies.
5. Some Media Agents read full object versions. They send data to other Media Agents that read incremental object versions. The latter Media Agents do the actual consolidation and send data to the Media Agents loaded with the target media.

If destination devices are not specified per object, Data Protector selects them automatically from those you selected in the object consolidation specification according to the following criteria in the order of priority:

- destination devices with the same block size as source devices are selected before those with a different one
 - locally attached devices are selected before network attached devices
6. Media Agents loaded with the target media write consolidated objects to the target media.

If the block size of the source device is smaller than that of the destination device, blocks are repackaged.

7. When the object consolidation session is completed, the CSM closes the session.

How Many Sessions Can Run Concurrently?

A number of object consolidation sessions can run in the cell at the same time. This number is limited by the resources in the cell, such as the Cell Manager and the systems with connected devices.

Queuing of Object Consolidation Sessions

Timeout

When an object consolidation session is started, Data Protector tries to allocate all needed resources. The session is queued until the required minimum resources are available. If the resources are still unavailable after the timeout, the session is aborted. The timeout period can be set using the `SmWaitForDevice` global option.

Mount Requests in an Object Consolidation Session

What Is a Mount Request?

A mount request in an object consolidation session is issued when a source or a target medium needed for the object consolidation operation is not available.

Responding to a Mount Request

Responding to a mount request includes providing the required medium and confirming the mount request. If the required source medium has media copies, you can provide a copy instead of the original medium.

Media Management Sessions

What Is a Media Management Session? A media management session is used to perform a certain action on the media, such as initializing media, scanning the content, verifying data on the media, and copying media.

Logging to the IDB Information about a media management session, such as generated messages, is stored in the IDB.

Data Protector Monitor and Media Management Session A media management session can be viewed in the monitor window. If you close the Data Protector GUI, the session will continue in the background.

Media Management Session Data Flow

What Happens in a Media Management Session? When a media management session is started, the following happens:

1. The Media Session Manager (MSM) process is started on the Cell Manager system. This process controls the media session.
2. The MSM starts the Media Agents (MAs) on the system that has devices used for the media management session.
3. Media Agents perform the requested operation and send generated messages to the Data Protector user interface, where you can track the progress. The session is also stored in the IDB.
4. When the session is complete, the MSM closes the session.

How Many Sessions Can Run? A number of media management sessions can run in the cell at the same time if they do not use the same resources, such as devices or media.

How Data Protector Operates
Media Management Sessions

8 Integration with Database Applications

In This Chapter

This chapter gives a brief description of the integration of Data Protector with database applications, such as Microsoft Exchange Server, Oracle Server, IBM DB2 UDB, and Informix Server.

It is organized as follows:

“Overview of Database Operation” on page 233

“Filesystem Backup of Databases and Applications” on page 236

“Online Backup of Databases and Applications” on page 237

For a detailed list of supported integrations, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.

Overview of Database Operation

From the user's perspective, a **database** is a set of data. Data in a database is stored in **tables**. Relational tables are defined by their columns and are given a name. Data is stored in rows in the table. Tables can be related to each other, and the database can be used to enforce these relationships. Data can thus be stored in **relational format** or as **object-oriented** structures such as abstract data types and methods. Objects can be related to other objects, and objects can contain other objects. A database is usually managed by the server (manager) process that maintains data integrity and consistency.

Whether you use relational structures or object-oriented structures, databases store data in **files**. Internally, these are database structures that provide a logical mapping of data to files, allowing different types of data to be stored separately. These logical divisions are called **tablespaces** in Oracle, **dbspaces** in Informix Server, and **segments** in Sybase.

Figure 8-1 **Relational Database**

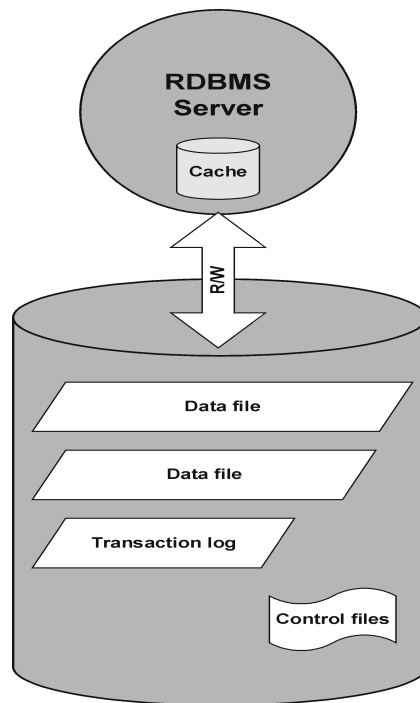


Figure 8-1 shows a typical relational database with the structures described below.

Data files are physical files that contain all of a database's data. They change randomly and can be very large. They are internally divided into pages.

Transaction logs record all database transactions before they are further processed. Should a failure prevent modified data from being permanently written to data files, the changes can be obtained from log files. Any kind of recovery is done in two parts: **roll forward**, which applies transaction changes into the main database and **roll back**, which removes uncommitted transactions.

Control files hold information about the physical structure of the database, such as, database names, names and locations of a database's data files and log files, and the time stamp of the database's creation. This control data is kept in control files. These files are critical for the operation of the database-

The **cache** of the database server process contains the most-often used pages of the data files.

The following is the standard flow of transaction processing:

1. A transaction is first recorded into the transaction log.
2. Changes required in the transaction are then applied to cached pages.
3. From time to time sets of modified pages are flushed to data files on disk.

Filesystem Backup of Databases and Applications

Databases are constantly changing while they are online. Database servers consist of multiple components that minimize response time for connected users and increase performance. Some data is kept in the internal cache memory and some in temporary log files, which are flushed at **checkpoints**.

Because data in a database can change during a backup, a filesystem backup of database files makes no sense without putting the database server into a special mode or even offline. Saved database files have to be in a consistent state, otherwise the data is of no use.

The following steps are required to configure a filesystem backup of the database or application:

- identify all data files
- prepare two programs that are able to shut down and start up the database, respectively
- configure the filesystem **backup specification** with all the data files included and specify the shut-down program as a **pre-exec command** and the start-up program as a **post-exec command**

This method is relatively simple to understand and configure but has one key disadvantage: *the database is not accessible during the backup*, which is unacceptable for most business environments.

Online Backup of Databases and Applications

To overcome the necessity to shut down the database during a backup, database vendors have prepared interfaces that can be used to put databases temporarily into special modes to save the data to tapes. Server applications are thus online and available to users during the backup or restore process. These application-specific interfaces allow backup products, like Data Protector, to back up or restore logical units of the database application. The functionality of the backup APIs varies depending on the database vendor. Data Protector integrations are available for major databases and applications. For a detailed list of supported integrations, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*

The essence of the backup interface is that it provides the backup application with consistent data (even if it may not be consistent on the disk) while at the same time keeping the database operational.

Figure 8-2 Data Protector Integration with Databases

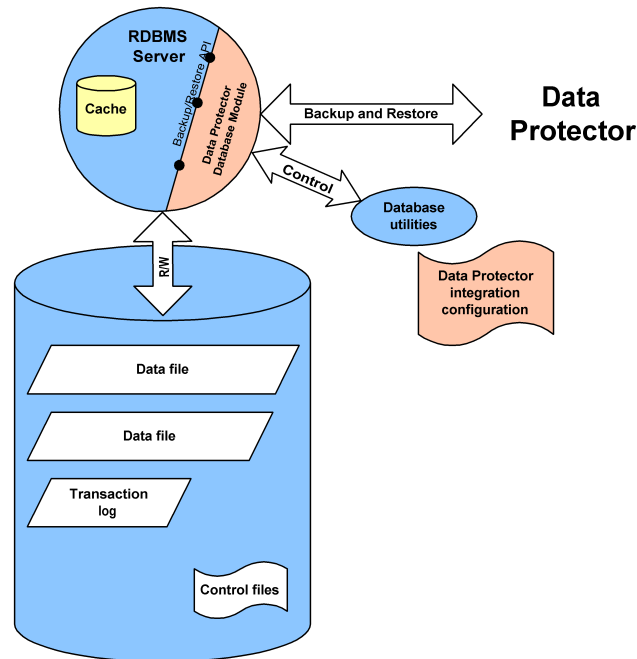


Figure 8-2 shows how a relational database is integrated with Data Protector. Data Protector provides a **Database Library** that is linked in to the database server. The database server sends data to Data Protector and requests data from it. Database utilities are used to trigger backup and restore operations.

A typical procedure to configure the backup of a database through the Data Protector integration is as follows:

1. A database/application-specific agent is installed on the database system
2. The Data Protector integration is configured for each database. Data needed for Data Protector to work with this database are stored on the database system (into configuration files or registry entries). Typically, this includes pathnames and user names/passwords.
3. The backup specification is prepared using the Data Protector user interface.

Besides the key advantage of the database being **online** all the time there are also other benefits of using the Data Protector integrations with the databases:

- There is no need to specify the location of data files. These can be located on different disks.
- The logical structure of the database can be browsed. It is possible to select only a subset of the database.
- Applications are aware of backup operation and keep track of which parts are backed up.
- Several modes of backup are possible. Besides **full** backups, users can select (block level) **incremental** backups or only the backup of transaction logs.
- Several modes of restore are possible and after the restore of data files, the database can automatically restore transaction logs and apply them as configured.

9 Direct Backup

In This Chapter

This chapter introduces the direct backup concept and the technologies that enable it. The chapter also discusses the direct backup configurations that are supported by Data Protector.

It is organized as follows:

“Overview” on page 243

“Requirements and Support” on page 250

“Supported Configurations” on page 251

Overview

The storage industry's demand for backup solutions that minimize application downtime and system loads while maximizing backup speeds is growing. Data volume is also growing; it has doubled every 1.5 years over the last 20 years and continues to grow even faster.

Applications and services need to be online nearly all the time with maximum performance. Backup windows are narrow and performance degradation due to backup (or anything else) is no longer acceptable.

In addition, the requirement for solutions that do not demand substantial investment in solution-specific equipment is also growing.

This multi-pronged requirement has led to the development and introduction of new direct or “serverless” backup technologies.

For enterprises and service providers that manage mission critical Oracle environments, Data Protector's direct backup feature is a non-intrusive, serverless backup extension to HP's family of network backup solutions.

Direct backup extends the benefits of HP's ZDB solution by moving data directly from disk to tape and making the load on the backup server negligible or even making the backup server optional.

It minimizes the impact on database production servers by using hardware-based mirroring technologies rather than intrusive software-based snapshots.

Additionally, the direct backup solution is fully integrated with the industry-standard XCopy (ANSI T10 SCP-2 Extended Copy Standard) command embedded in HP StorageWorks tape libraries (as well as external fiber channel SCSI bridges), removing the need for a separate “data mover” appliance.

NOTE

Refer to “Supported Configurations” on page 251 for an explanation of which applications, operating systems, and devices that direct backup supports in HP OpenView Storage Data Protector A.06.00.

Direct Backup

What does it mean to make a direct backup? This backup solution is “serverless,” meaning that it does not use a dedicated backup server to move data and data is not moved across the LAN. The data moves *directly* from the client system to be backed up to a tape device without moving data through a backup server.

Direct backup supports backup of application data files and control files and of disk images -- either raw disk or raw logical volume.

Direct backup uses existing split-mirror and Storage Area Network (SAN) technology to:

- access application data with minimal impact to the application -- there is minimal utilization of the application server (providing little to no application down time)
- move data without encountering bottlenecks associated with network traffic and LAN speed.

In order to support direct/serverless backups, Data Protector also incorporates new technology to resolve target file systems and move data across the SAN. This new technology, based on the XCopy standard, provides a method for moving data from the target system to the tape device without moving data through a server. Refer to “About XCopy” on page 247 for a brief explanation of XCopy.

This direct disk to tape data path (over SAN) helps to lower the need for equipment investment and boost the utilization of existing equipment.

Backup Types

Direct backup supports backup of application data files and control files and of disk images -- either raw disk or raw logical volume.

Direct Backup Benefits

Because the data mover is in the SAN bridge and the technology that interprets the target system is built into the General Media Agent, direct backup users can use a low-cost management server to drive the backup and can avoid having to invest in multiple servers to perform block identification.

Additionally, direct backup is designed to let you leverage hardware capabilities to increase uptime and instant recovery capabilities to reduce restore time.

Direct backup is not limited to proprietary file systems and logical volume managers (LVMs).

Direct backup adds value to your backup solution both by addition and omission. For example, direct backup:

- Takes advantage of the most advanced XCopy features to speed backups
- Enables maximum uptime by leveraging existing hardware mirroring and snapshot capabilities
- Lets you take advantage of Data Protector's industry-leading instant recovery capability to speed recovery
- Requires very little CPU and memory resource from the XCopy host device

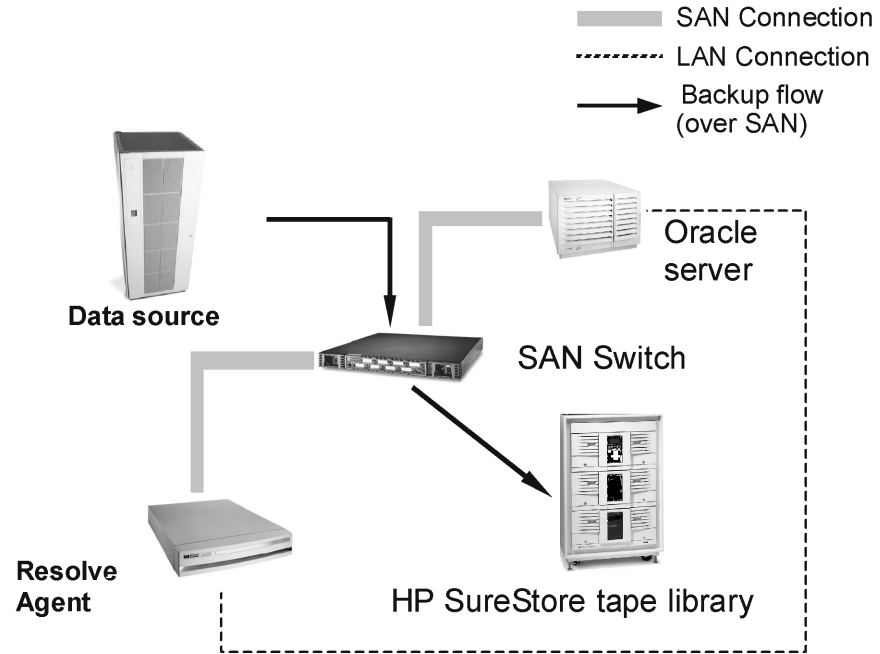
How Direct Backup Works

As with any other Data Protector backup, you will make a backup specification to control how and when the backup happens.

- The General Media Agent on the application server quiescens the application
- The split mirror agent on the application server and backup host splits the mirror
- The General Media Agent on the backup host:
 - resolves the disk of the target system
 - calculates the resolve information
 - calls XCopy
- XCopy, in turn, retrieves the target data and moves it across the bridge to the tape device.

Figure 9-1 shows a basic direct backup configuration. The Resolve Agent is on a separate backup host in this configuration. Data, however, does not move through this host.

Figure 9-1 Direct Backup Architecture



Environment

This section describes the direct backup environment in terms of the devices that need to be connected and what they need to be connected to. It also discusses the required agents and where they are installed.

For information about platform, tape drive, and library support, refer to “Supported Configurations” on page 251.

Direct backup requires the General Media Agent to be off the application server. Also, the Resolve Media Agent must be on the application server or other host and must have access to the XCopy engine. Refer to “Supported Configurations” on page 251 for discussion of Resolve Agent placement.

Direct backup requires that:

- The disk array, XCopy engine, application server, and tape drive or library are attached to the SAN.
- The Resolve host and application server are attached to the LAN.

- The HP StorageWorks Disk Array XP (XP) is configured as Business Copy (BC) with mirrors configured with enough disk space.
- The SAN is properly configured in order to provide access to source (disks) and target (tape) devices from both the XCopy engine and the host where the Data Protector General Media Agent is running. This means that LUN masking and SAN zoning must be configured so that:
 - The General Media Agent host has access to the XCopy engine
 - The General Media Agent host has access to the target tape drive or library
 - The SSEA host has access to the source disk
 - The XCopy engine has access to the source disk
 - The XCopy engine has access to the tape drive or library

About Resolve

The Resolve program is a proprietary component of Data Protector that understands the native disk layout of many different file systems. Resolve allows Data Protector direct backup to back up data written by many kinds of operating systems without needing many servers running those operating systems.

Resolve examines the raw information on the disk and chooses the appropriate method of interpreting the disk's file system. Note that Resolve does not read the data itself; it reads only information related to disk location. It then returns information suitable for direct input to the XCopy engine.

About XCopy

XCopy is a National Committee for Information Technology (NCITS) standard that allows two devices to communicate with one another without the help of another computer/server in between.

XCopy specifies a set of SCSI commands that, when given to an XCopy engine, allow the transfer of data from one device to another without the need to have a computer/server in between. The data flows from the source device (either block or streaming, i.e., disk or tape) to the destination device (either block or streaming) through XCopy.

It assumes that the streaming (tape) device is set up and the device is ready to write/read the data (that is, the drive is online, there is a tape in the drive, and the tape is properly positioned at the starting point for the write/read.) This frees the controlling server from reading the data from one device into its memory and then writing the information to the destination device. With XCopy, all the server does is send the XCopy commands to the XCopy engine and then wait for the results.

XCopy + Resolve

Prior to Resolve, you needed a server with a matching file system to get this information. This was because even with the right server, getting this information could be difficult as the operating system might have converted the actual physical sectors into its logical view before returning the information. Resolve removes the need for multiple servers to support multiple file systems and the difficulties presented by file system-specific information formats.

Direct Backup Process Flow

The process flow for direct backup is listed below. These are the basic steps -- start to finish -- of a direct backup.

- Read backup specification
- Determine what to backup
- Quiesce application
- Split mirror
- Release application
- Resolve blocks
- Move data -- The XCopy engine
- Reconnect and resynchronize mirror

Backup Stages for Data Files

Original data files that are to be backed up go through several stages before they end up as copies that are used later to perform restores. The direct backup process (generally) follows these steps:

1. Obtain data-file consistency (quiesce the application)
2. Read meta data (file attributes) and group files into objects

3. Obtain data-file stability (use Split-Mirror technology for point-in-time data stability)
4. Map data files to list of disk blocks (using Resolve technology)
5. Move disk blocks to tape (using XCopy technology)

Typically each stage is managed by one Data Protector agent. Agents are spawned by the Backup Session Manager (BSM). All errors that cannot be handled internally by agents are reported through the BSM to the user and stored in the Internal Database. The Backup Media Agent (BMA) writes catalog segments and delimiters between data and catalog segments known as file-marks.

Restore

There are two restore options when using direct backup:

- If you are using the HP StorageWorks XP disk array and you have instant recovery capability, you can use it to restore the data. Refer to the *HP OpenView Storage Data Protector Zero Downtime Backup Administrator's Guide* for an explanation of using instant recovery.
- Restore of information backed up using direct backup can also be accomplished using the standard Data Protector network restore.

Note that in either case it is important to verify that the application server can handle the load of the restore. This is not an issue on the backup end, as the data does not pass through the server during this process. On restore, however, the data does impact the server.

Requirements and Support

This section lists the requirements for using direct backup successfully and the file systems and applications that direct backup supports.

- Data Protector Cell Manager running on any supported operating system
- Resolve Agent running on HP-UX 11.11
- Support of application servers running HP-UX 11.11
- Support of HP LVM on HP-UX 11.11
- XCopy host, source disk, destination device, and XCopy engine must be in the same SAN zone.
- File System support:
 - Veritas' VxFS 3.1, 3.3
- Application support:
 - Oracle 9.i
- Raw volume support
- Support of ServiceGuard environments for the application server
- Restore through the standard Data Protector restore interface
- Support of instant recovery for the XP
- XCopy engine in the bridge

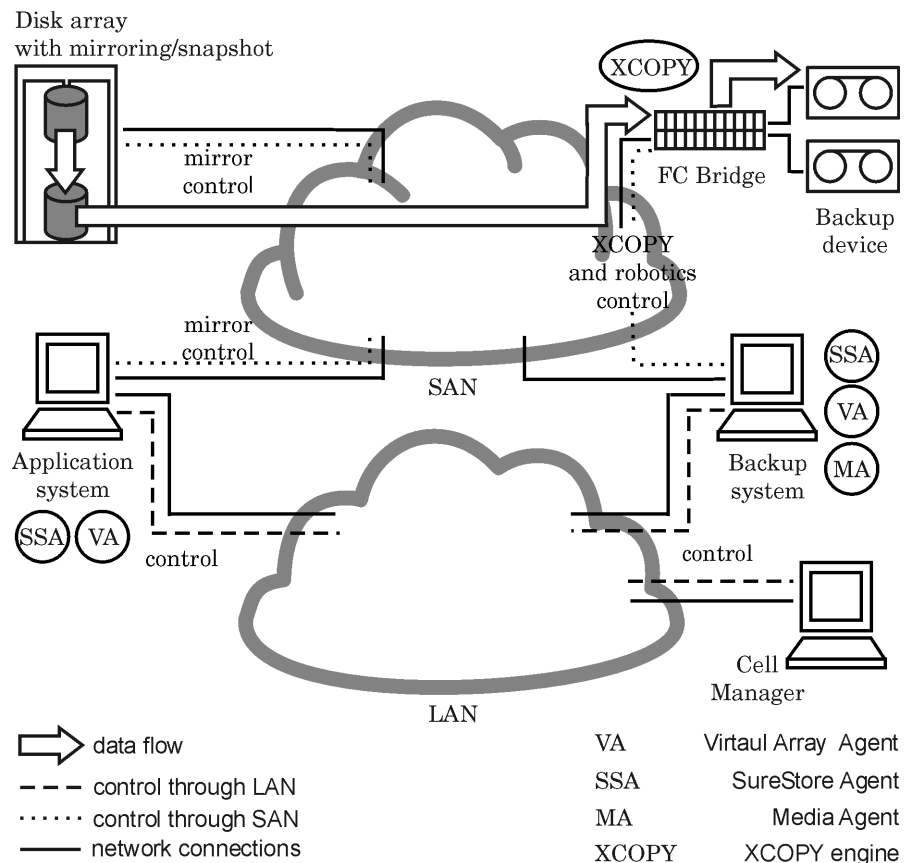
Supported Configurations

Three Hosts: CM, Application, Resolve

This solution uses three hosts: one each for the Cell Manager, Resolve Agent, and application. Although this configuration requires three machines, the Resolve host can be a less expensive host and the resource load is split, helping you avoid performance impact on the application.

Note that in this configuration, the Cell Manager host can be running *any* of the operating systems supported by Data Protector. The application and Resolve Agent hosts must be running HP-UX 11.11.

Figure 9-2 Basic Three Host Configuration



Two Hosts: Cell Manager/Resolve Agent and Application

This solution uses two hosts: one for the Cell Manager and Resolve Agent and one for the application. Although this configuration does require two machines, the resource load is split; this can help you avoid performance impact on the application. Also, the machine that hosts the Cell Manager and the Resolve Agent can be of minimal processing capacity.

Note that in this configuration, both hosts must be running HP-UX 11.11.

Basic Configuration: Single Host

This solution uses a single host on which the Cell Manager, application, and Resolve Agent are all installed. As all three components are running on the same physical machine, they are sharing resources (I/O channels, CPUs, memory, etc.) for their activities. This configuration minimizes the equipment required to facilitate direct backup. However, because resources are shared, the Cell Manager and General Media Agent may impact the application database's performance (XCOPY's processing requirements are negligible).

Note that in this configuration, the host must be running HP-UX 11.11.

10 Disk Backup

In This Chapter

This chapter introduces the concepts associated with backing up data to disk and the technologies that enable it. It also discusses the disk-to-disk backup configurations that are supported by Data Protector.

It is organized as follows:

“Overview” on page 255

“Disk Backup Benefits” on page 256

“Data Protector Disk-Based Devices” on page 258

Overview

Industry has requirements for increasingly faster methods of backing up and restoring data. In addition, it has become more and more important that the time required for data backup and restore be reduced to a minimum so as not to interrupt the day-to-day running of company applications.

Many applications and databases frequently make small changes to existing files or produce many new files containing business-critical data throughout the working day. These files need to be backed up immediately to guarantee the data in them will not be lost. This requirement means that a fast medium that can store large amounts of data that works without interruption is necessary for storing data.

Disk-based storage media have become increasingly cheaper in recent years. At the same time, the storage capacity of disks has risen. This has led to the availability of low-cost, high-performance single disks and disk arrays for storing data.

Disk backup (also known as disk-to-disk backup) is becoming ever more important. In the past, tape storage was the favored medium for backup and restore because of its price and effectiveness in meeting disaster recovery requirements. Today, more and more businesses are augmenting their tape storage backup solutions with faster disk-based backup solutions. This ensures faster data backup and recovery.

Disk Backup Benefits

There are many situations in which it is advantageous to use disk-based devices when performing backups. Disk-based devices are, in fact, specific files in specified directories, to which you can back up data instead of or in addition to backing it up to tape. The following list indicates some situations in which disk-based devices are particularly useful:

- Many applications and databases continuously generate or change a large number of files, which contain business-critical data. Under these circumstances, it is necessary to continuously back up the files concerned, in order to guarantee the capability of restoring them without data loss.

In these environments, tape devices typically have to operate in stop/start mode, because they do not receive a constant data stream. This may result in the tape device limiting access to the files concerned. In addition, the lifetime of the backup device may be greatly reduced.

Alternatively backups can be performed to any disk-based device, overcoming the limitations described. As a short-term backup solution, this is adequate in itself. If a longer term backup solution is required, the data in the disk-based devices can be moved periodically to tape to free up the disk space. This process is known as **disk staging**.

- In environments that have fast, high-capacity disk drives and slow tape drives, you can shrink the backup window by performing backup to disk-based devices first and moving the data to tape later.
- Using disk-based devices for backup enables you to take advantage of advanced backup strategies such as synthetic backup.
- Disk-based devices are useful for providing fast restore capability for recently backed up data. For example, backup data could be kept in a disk-based device for 24 hours to enable fast, convenient restore.
- Mechanically, a disk-based device is quicker to use than a tape. When using a disk-based device there is no need to mount and unmount a tape. When backing up or restoring a small amount of data, a disk-based device is quicker because it does not need the initialization time that a tape drive requires. With a disk-based device there is no

need to load or unload media, which consumes a significant amount of time in a small backup or restore. The advantages of using a disk-based device are even more evident when restoring from an incremental backup.

- The risk of media problems such as faulty tapes and tape mounting failures are reduced to a minimum. The availability of RAID disk configurations provides protection of data in cases where a disk fails.
- Overhead costs are reduced because there is no need for tape handling.
- Overall, disk-based storage space is becoming increasingly cheaper even if compared to tape-based storage.

Data Protector Disk-Based Devices

Data Protector has the following disk-based devices:

- Standalone file device
- File jukebox device
- File library device

Standalone File Device

The standalone file device is the simplest disk-based backup device. It consists of a single slot to which data can be backed up. Once configured, its properties cannot be changed. The file device has a maximum capacity of 2 TB, if this file size is supported by the operating system on which the device is running.

File Jukebox Device

The file jukebox device is a special version of the Data Protector jukebox device. The jukebox device can be configured to back up either optical or file media. The jukebox device used to back up file media is referred to as the file jukebox device. The type of media to be backed up by the jukebox is specified during device configuration.

The file jukebox device consists of multiple slots to which you can back up data. Configuration is a two phase process, firstly a file jukebox device is created and then one or several drive(s) is configured for it. Once the device has been configured it is possible to change its properties. Each slot in the file jukebox device has a maximum capacity of 2 TB. The device's maximum capacity is equal to:

Number of slots X 2 TB

File Library Device

The file library device is the most sophisticated disk-based backup device. It has multiple slots called **file depots** to which you can back up data. The configuration of the file library device is completed in a single stage. It is possible to change the properties of the file library device at any time. The device's maximum capacity is the same as the maximum that can be saved on the filesystem on which the device resides. Each file depot has a maximum capacity of up to 2 TB. File depots are created automatically as required.

The file library device has intelligent disk space management. It anticipates potential problems saving data to it. A warning message is written in the event log if the amount of free disk space approaches the

configured minimum amount required for the device to work. This enables you to free more disk space in good time for the device to continue saving data. If all the space allocated to the file library device is ever completely used, a warning message appears on the screen with instructions as to how to solve the problem.

The file library device automatically creates more file depots if a particular backup requires more space than is available in a single file depot.

**Recommended
Disk-Backup
Device**

Hewlett-Packard recommends using the file library device as the preferred disk-based backup device. The file library device is the most flexible and intelligent of the set of disk-based backup devices. It can be re-configured at any time during use and is capable of performing more sophisticated disk space handling than any other disk-based backup devices. Furthermore, it enables the use of advanced backup strategies such as synthetic backup.

For description of the file library device functionality, refer to online Help, index keyword “file library devices“.

Data Format

The data format of the disk-based devices is based on the tape data format. Data Protector converts the data to be backed up into tape format before it writes the data to the disk-based device.

Configuration

It is possible to set properties for the all disk devices both during the initial device setup and after the devices are in operation. The degree of changes that can be made to the properties of each device vary according to the device.

**Backing Up to a
Disk Device**

A backup can be made to a disk-based device by creating a normal Data Protector backup specification.

Disk Backup
Data Protector Disk-Based Devices

| 11 **Synthetic Backup**

In This Chapter

This chapter introduces the concept of synthetic backup and explains the synthetic backup solution provided by Data Protector.

It is organized as follows:

“Overview” on page 263

“Synthetic Backup Benefits” on page 264

“How Data Protector Synthetic Backup Works” on page 265

“Restore and Synthetic Backup” on page 267

Overview

With the volume of data increasing and backup windows shrinking, performing a full backup often presents a problem in terms of time and storage space. On the other hand, having many incremental backups can be problematic because each incremental increases the time needed to perform a restore.

As backup to disk is gaining popularity due to the high performance and capacity as well as increasingly lower price of disks, new opportunities have arisen. The industry's requirements are to minimize the backup window, minimize the load on production servers and the network, and enable a quick restore. These requirements are met by synthetic backup.

Synthetic backup is an advanced backup solution that produces a **synthetic full backup**, an equivalent to a conventional full backup in terms of data, without putting stress on the production servers or the network. A synthetic full backup is created from a previous full backup and any number of incremental backups.

Performing synthetic backup eliminates the need to run regular full backups. Instead, incremental backups are run, and subsequently merged with the full backup into a new, synthetic full backup. This can be repeated indefinitely, with no need to run a full backup again.

In terms of restore speed, a synthetic full backup is equivalent to a conventional full backup. The restore chain consists of only one element, so a restore is as quick and simple as possible.

Synthetic Backup Benefits

Synthetic backup brings the following benefits:

- It eliminates the need for full backups. After the initial full backup, only incrementals are performed, which significantly reduces the time needed for the backup.
- Consolidation of backed up objects is performed on the device server, putting no stress on either the production servers or the network.
- A type of synthetic backup, called virtual full backup, is even more efficient. Virtual full backup consolidates data using pointers, which eliminates unnecessary duplication of data.
- A restore from a synthetic full backup is as fast as from a conventional full backup, as there is no need to retrieve data from incremental backups. This eliminates the reading of each incremental backup in the restore chain, and if tape devices are used, also loading and unloading of several media and seeking for object versions.

How Data Protector Synthetic Backup Works

Data Protector synthetic backup enables you to merge a full backup and any number of incremental backups into a new, synthetic full backup.

A synthetic full backup can be created from a full backup that is written to a disk or tape device and incremental backups that are written to a disk-based device, a Data Protector file library. The synthetic full backup can, again, be written to a disk or tape device.

If all the backups, full and incremental, are written to the same file library, an even more efficient type of synthetic backup is available, called **virtual full backup**. This solution uses pointers to consolidate data rather than copy the data. As a result, the consolidation takes less time and avoids unnecessary duplication of data.

The following figures explain the concept of synthetic backup and virtual full backup. They show how a synthetic full backup or a virtual full backup is created from a full backup and any number of incremental backups.

Figure 11-1 Synthetic Backup

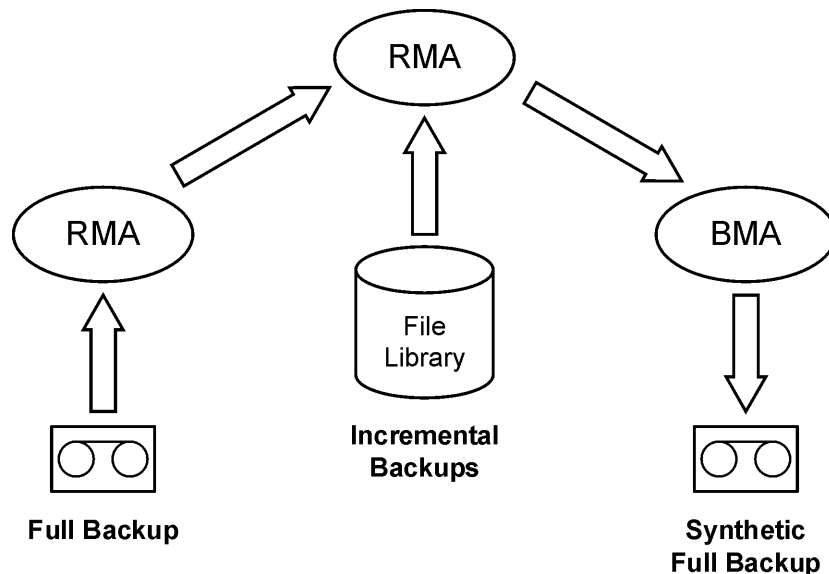


Figure 11-1 shows how a synthetic full backup is created. The Restore Media Agent (RMA) reads the full backup from the backup medium, which can be a tape or a disk. The data is sent to another RMA, which reads the incremental backups from the file library and consolidates the data. The consolidated data is then sent to the Backup Media Agent (BMA), which writes the synthetic full backup to the backup medium, which can, again, be a tape or a disk.

Later on, the synthetic full backup is typically merged with subsequent incremental backups into a new synthetic backup. The procedure can be repeated indefinitely, either after each incremental backup, or at a desired interval.

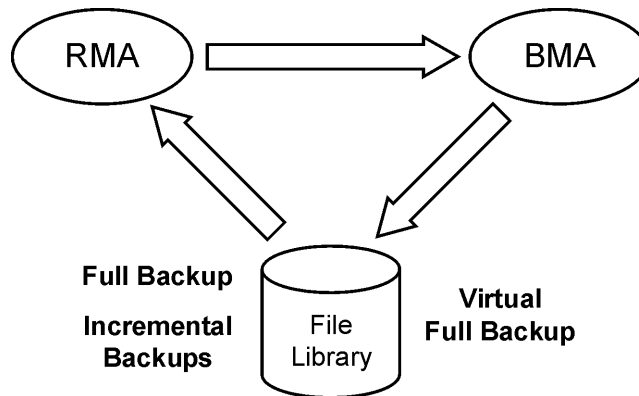
Figure 11-2**Virtual Full Backup**

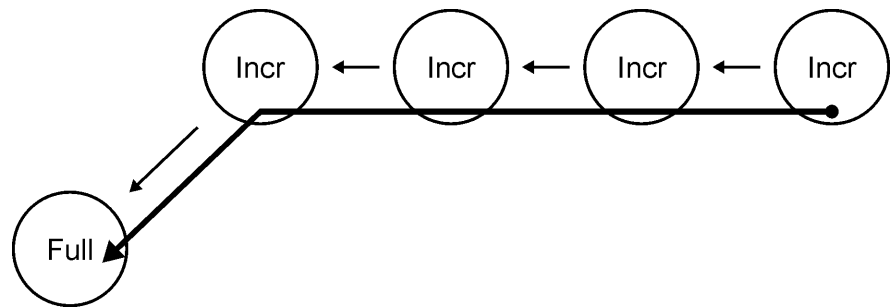
Figure 11-2 shows how a virtual full backup is created. With this type of backup, all the backups reside in a single file library. The Restore Media Agent (RMA) reads the information about the full backup and the incremental backups, and generates the data for the virtual full backup. The generated data is sent to the Backup Media Agent (BMA), which creates the virtual full backup in the file library.

Restore and Synthetic Backup

Restore from a synthetic full backup is equivalent to restore from a conventional full backup. The following figures present different situations, supposing you need to restore your data to the latest possible state. In all examples, a full backup and four incremental backups of the backup object exist. The difference is in the use of synthetic backup.

Figure 11-3

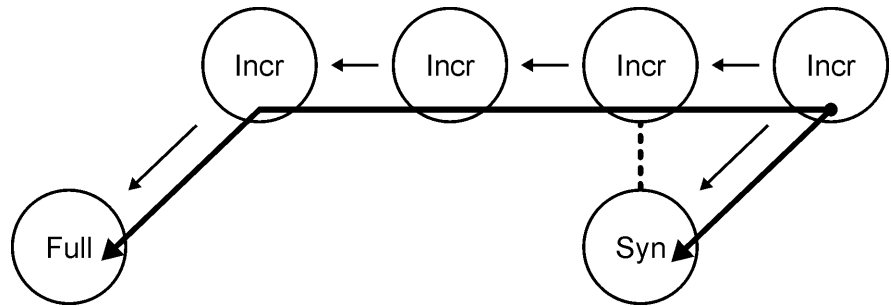
Full and Incremental Backups



In Figure 11-3, conventional backups were performed. To restore to the latest possible state, you need the full backup and all four incremental backups. The restore chain consists of five elements, which often reside on different media.

Such a restore can take a considerable amount of time, as each incremental backup must be read. If tape devices are used, time is spent for loading and unloading of several media and seeking for object versions to restore.

Figure 11-4 Synthetic Backup



In Figure 11-4, a synthetic full backup exists, which is used for restore by default. The restore chain consists of only two elements, namely the synthetic full backup and the subsequent incremental backup. The restore is significantly simpler and quicker than that without the synthetic full backup. In the figure, both possible restore chains are shown.

Figure 11-5 Regular Synthetic Backup

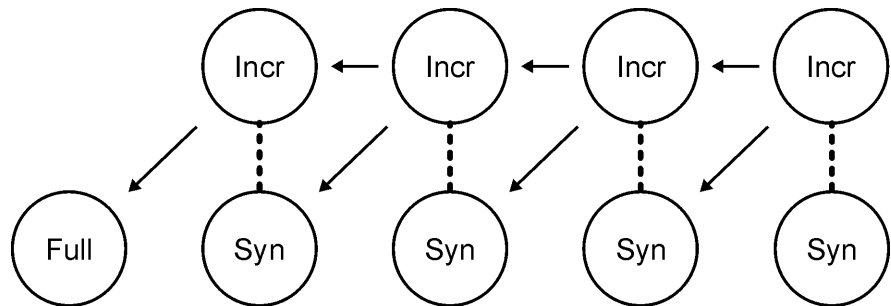
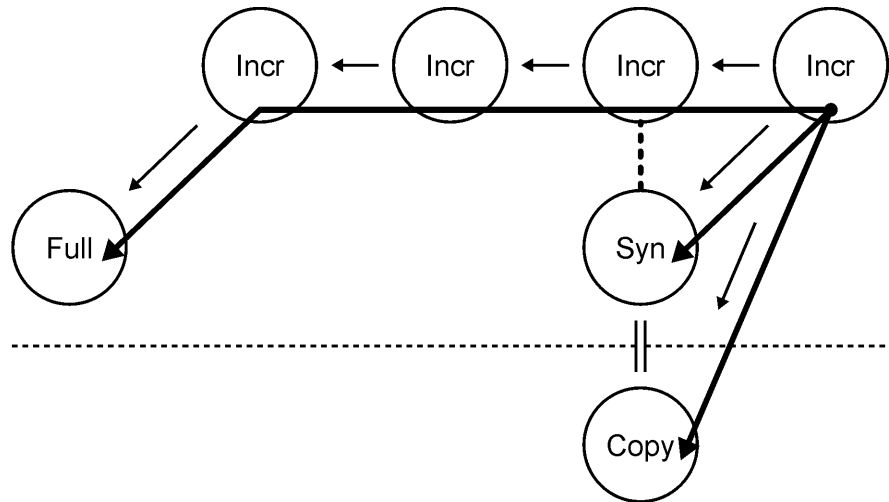


Figure 11-5 shows a situation where a synthetic backup was performed after each incremental backup. This strategy enables the simplest and quickest restore to the latest possible state, or to any earlier point in time that was backed up. Only one element is required for restore, namely the synthetic full backup of the desired point in time.

Figure 11-6 Synthetic Backup and Object Copy



In Figure 11-6, a synthetic backup was performed and then copied. This provides additional safety. The restore to the latest possible state can use any of the three different restore chains shown. By default, Data Protector selects the optimum restore chain, which normally includes the synthetic full backup or its copy. In case of missing media, a media error, or similar, an alternative restore chain is used.

Synthetic Backup
Restore and Synthetic Backup

12**Split Mirror Concepts**

In This Chapter

This chapter introduces the split mirror backup concept and discusses the configurations that are supported by HP.

It is organized as follows:

“Overview” on page 273

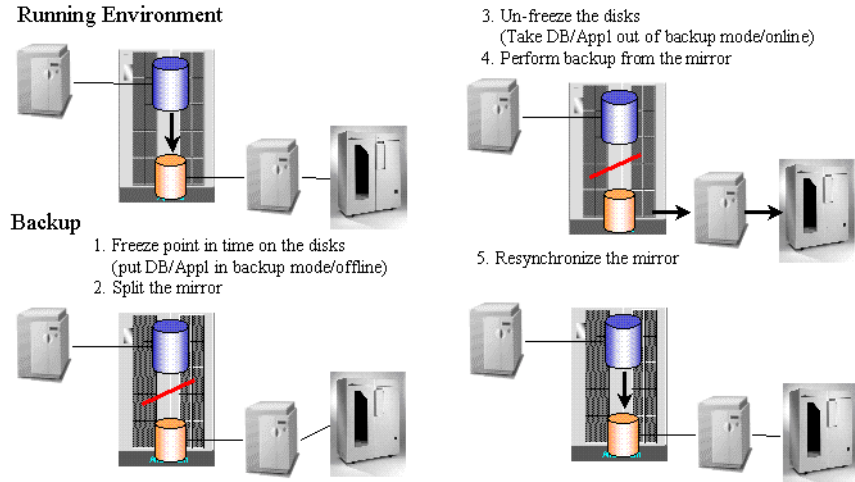
“Supported Configurations” on page 277

Overview

Modern high availability (HA) storage configurations introduce new demands on backup concepts. The configuration consists of one of numerous variations of single or multiple mirror structures.

The usual approach is to use one **replica** (mirror copy) for the backup task, while the **source volumes** still serve the application. Refer to Figure 12-1.

Figure 12-1 Split Mirror Backup Concept



The **target volumes** in replica are typically connected to a separate client, which also has tape devices connected to allow for local backup. Usually, hardware mirror technologies such as HP StorageWorks Disk Array XP or EMC Symmetrix are used to create a replica such as:

- HP StorageWorks ContinuousAccess XP or
- HP StorageWorks BusinessCopy XP

The availability of the application remains almost permanent, with the exception of a short period of time (lasting several seconds to a few minutes). This time is needed to make the data on the disk consistent and perform the actual split of the mirrors. The data must be consistent so that the application can make use of the data after a restore.

Normally, the replica is not created at the time of backup, but is already available and synchronized to provide high availability to the application. The backup and the resyncing of the replica does not affect the application performance, since this occurs in parallel on separate hardware.

As the application client and backup client are different (in most cases), it is very important that all cached information (database cache, filesystem cache) on the client is flushed to the disk before the backup mirror is split off. One of the following options can achieve this:

- Databases can be put into backup mode
- Databases can be taken offline
- A mount point can be unmounted

Only when this is carried out *before* a replica is consistent. However, if a database is running on a filesystem or a rawdisk, there is no need to unmount the filesystem or rawdisk as the database ensures that data is really written to the disk and not to the filesystem cache.

For an online database backup, a replica alone cannot be restored. The archive log files from the application client are also needed. An archive log backup can be started right after a split, when the database is taken out of backup mode.

The use of one replica in combination with the HP StorageWorks ContinuousAccess XP technology to perform the backup does take away high availability of storage for the duration of the backup. Additional mirrors retain full high availability of storage and allow for the same backup approach.

The backup client can be a centralized backup client for multiple application clients running different applications. In such cases, the backup client must run on the same operating system as the application client, so as to access mirrored resources in a native way.

The backup client should be capable of performing backups in a reasonable amount of time. Though, theoretically, almost 24 hours may be required to perform a backup, the restore time must be considered as well. It is thus recommended to have a backup client that can perform the backup in 2 to 4 hours. It is recommended to perform the restore through the application client.

In this approach the bulk of the data transfer happens via the backup client and its access to the replica. The LAN connection between the backup client and application client is only used to coordinate processes that are involved in the backup. There are processes running on each client to allow the automation of the split.

Instant Recovery

Data Protector instant recovery takes advantage of the split mirror technology to provide instant data restore. The solution is based on zero downtime backup (ZDB) solutions like the HP StorageWorks Disk Array XP integration, which uses split mirror technology.

During a split mirror backup session, a replica is used for the purpose of moving the data to a backup medium (tape). After a backup is completed, the replica can be discarded and disk pair prepared for the next backup session by resynchronization, or the replica can be left unchanged for the purpose of instant recovery. Several replicas can exist at the same time. For example, HP StorageWorks Disk Array XP allows up to three replicas, and each can have an additional two copies if cascading is used.

During the instant recovery, the data on the specified replica (left unchanged for the purpose of instant recovery) is synchronized to the application client source volumes without restoring from a backup medium.

Data Protector will only use the first three replicas because secondary mirrors cannot perform fast-resynchronization, which is critical for ensuring minimal restore time. Instant recovery is only possible using the HP StorageWorks BusinessCopy XP configuration (local mirror - dual host and local mirror - single host configurations).

ZDB to Tape and ZDB to Disk+Tape

During ZDB-to-tape and ZDB-to-disk+tape sessions, a replica of the application data is streamed to a tape device, which is connected to a separate backup system, using Data Protector Disk Agent and General Media Agent, with minimal impact on the application system. After the backup is completed, the replica is either:

- discarded - ZDB to tape
- retained and can be used for instant recovery - ZDB to disk+tape

ZDB to Disk

During a ZDB-to-disk session, the original data is not moved to a backup medium (tape) from the replica. The replicas (up to three) can be used for various purposes, such as offline data processing or instant recovery; the

latter is possible only if HP StorageWorks BusinessCopy XP configuration was used. It is only possible to restore objects from a ZDB-to-disk session by using the instant recovery functionality.

Replica Set Rotation

Several replicas can exist at the same time. HP StorageWorks Disk Array XP allows up to three replicas, and each can have an additional two copies if cascading is used. Data Protector can use only disks from the first three replicas (**first level mirrors** or **MUs**) for backup and instant recovery purposes. The additional six copies (cascading mirrors) are not supported.

When configuring a ZDB backup specification for a source volume (LDEV) with first level mirrors configured or when restoring to such a source volume, it is, using Data Protector, possible to define a **replica set** from which this integration selects one replica for the current session.

Backup Clients and Clusters

The backup client should not be used as a failover server for the application client. It is recommended to have application and backup services on separate clusters.

Supported Configurations

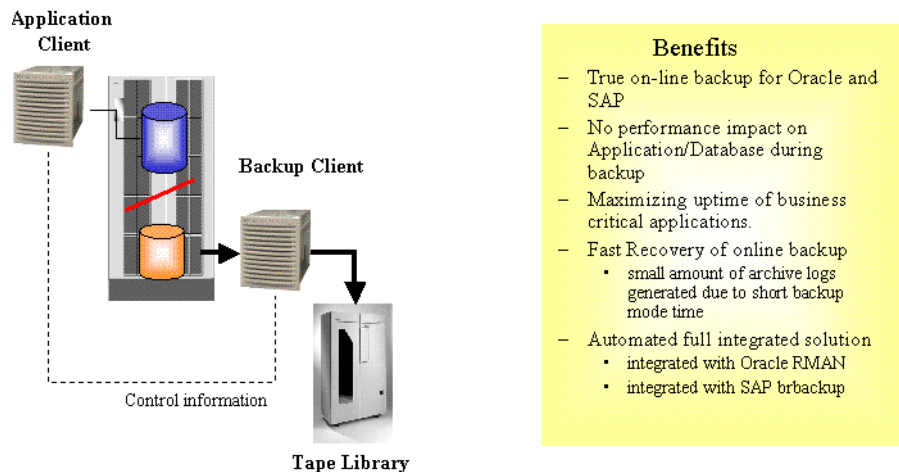
Local Mirror - Dual Host

This solution uses a local mirroring functionality such as Business Copy XP. Both disks are in the same disk array, which means the I/O infrastructure of the RAID system is actually shared between the application client (or host) and the backup client.

As the application client and the backup client are two physically different systems, they can use their own resources (I/O channels, CPUs, memory, etc.) for their dedicated activities, such as backup, without interfering with each other. In this way, the backup performance does not impact the database performance.

Figure 12-2

Local Mirror - Dual Host (Full Performance, Zero Downtime Backup)



The Data Protector split mirror backup integration allows automatic handling of mirror status as well as tight integration with applications such as SAP R/3 and Oracle (to ensure data consistency and application/database-aware backups). Only if the application/database is aware of a backup can a secure operation be guaranteed and native application tools be used for restore (sapdba, for example). The impact of

a backup on the application is reduced to the time needed to perform a split of the mirror and put the database into a consistent mode that permits the split, as well as to take it out of this mode again.

This configuration enables an offline backup of a very large database in a short time, as well as an online backup that creates very few archive log files, since the backup mode time of the database is kept to a minimum.

A small number of archive logs reduces the space needed for the archive logs in total, as well as speeds up the recovery process of the database. After a restore of an online database, a recovery is needed to return the database to a consistent state. All archive logs that have been created during the backup must be applied. In a split mirror backup, only the archive log files created during the split are applied.

Local Mirror - Single Host

In cases where no dedicated backup server is available, both functions (application and backup) are performed on the same client (or host). Offline backups of mail applications, for instance, could reduce the downtime of the application to minutes instead of hours in this way.

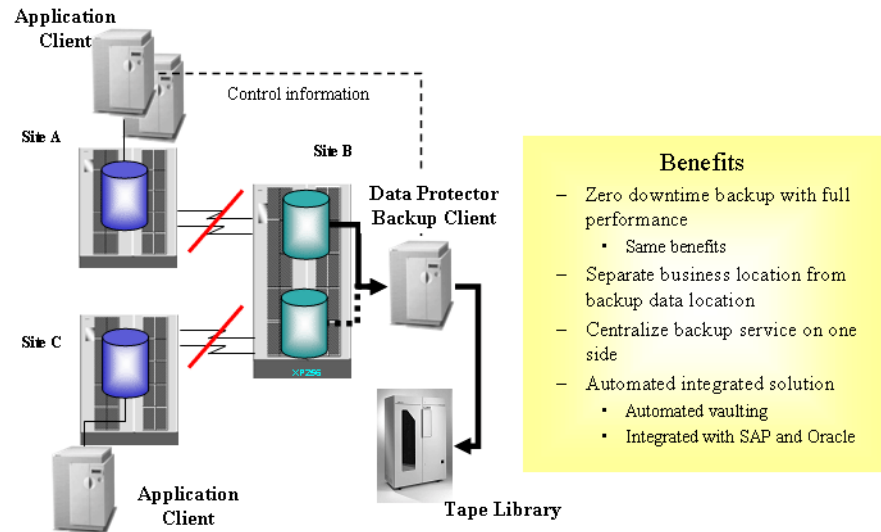
In this type of configuration, only **disk image** (raw disk) and **filesystem** backups are supported. Database and application backups, like Oracle and SAP R/3, cannot be supported, since the database has to be mounted on the backup server, which would not be possible on the same server that has the database already mounted.

Remote Mirror

Remote mirror technology, such as Continuous Access XP, enhances the configurations shown earlier due to the fact that the backup and application processes utilize different disk array resources at different locations.

Figure 12-3

Split Mirror - Remote Mirror (LAN-Free Remote Backup - Data HA)



The remote mirror transfers data to a physically separate site where it can be backed up to locally available tapes. This allows the separation of production data from backup data, eliminating the risk of a fire or other disaster damaging both the production and the backup environment at the same time.

No network resources are required to sync the mirrors during a backup. Although data is not transferred through the network, Data Protector still needs the communication between the Cell Manager and its clients.

This solution allows you to centralize a backup service by mirroring the application data from several production sites (A and C in this case) to a central location or central disk array. In this way, your investment in a backup service (server and tape library) can be consolidated and combined with the high availability of a remote mirror configuration.

The remote site cannot be used as an automatic disaster recovery site during the time of the backup, as the link between the two sites is split for the duration of the backup (and both disks are out of sync). This means that in case of a site A failure, site B cannot take over automatically (as it normally would) for x hours (x being the time the data takes to stream to the tape). This problem applies to local mirroring

as well. However, it is particularly important for the remote solution, as the concept of a remote disaster recovery site using hardware mirror concepts is widely accepted in the industry.

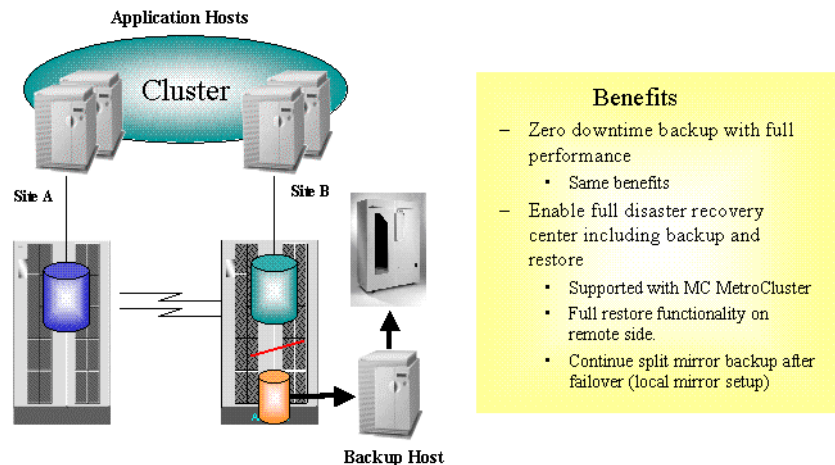
Local/Remote Mirror Combination

If the customer has a need for a permanently available recovery site (provided, for example, by a MetroCluster) in addition to a zero downtime backup solution, the combination of a remote mirror and a local mirror can be used.

This solution allows for full split mirror advantage together with a full recovery solution at the remote site. In this example, the remote mirror is constantly maintained with only the local link split for backup purposes. This gives the cluster the continuous ability to fail over to the remote site (site B).

Figure 12-4

Local/Remote Mirror Combination (Disaster Recovery Integrated Backup [Service HA - HP-UX only])



In order to have the failover functionality independent of the backup operation, the backup client must be a separate additional client outside the cluster. If a MetroCluster solution is implemented, the cluster arbitration client could be the backup client.

Other Configurations

There are many other possible split mirror configurations that provide some particular advantage or fulfill a specific user need. However, each configuration has its specific behavioral pattern that imposes specific requirements on the control functions in order to guarantee backup and recovery. It is important to control and specify which configurations are supported.

All the configurations shown above are supported by HP. For an updated list of supported configurations, refer to the following URL:
http://www.openview.hp.com/products/datapro/spec_0001.html.

In the event that you want to back up data in a configuration not listed, this does not mean that it cannot be supported. Please contact your local HP representative or HP Consulting to investigate the supportability of additional configurations.

Split Mirror Concepts
Supported Configurations

13 **Snapshot Concepts**

In This Chapter

This chapter introduces the snapshot backup concepts and discusses the configurations that are supported by HP.

It is organized as follows:

“Overview” on page 285

“Supported Configurations” on page 291

Overview

The rapidly expanding requirement for high availability storage configurations has led to the introduction of new zero downtime backup (ZDB) technologies. The advances in storage virtualization technology have provided the opportunity for an alternative to conventional split mirror technology.

Within the Data Protector ZDB solution, different disk array technologies are combined with the latest developments in the snapshot technology, to create snapshots of application or database data stored on a disk array. These snapshots can subsequently be kept on a disk array as point-in-time copies of the original data for **instant recovery** purposes or can be used to produce ZDB-to-tape sessions on a backup system. The processes concerned have minimal impact on the application server, providing an effective ZDB solution.

Storage Virtualization

The term “storage virtualization” is used to describe the technology that separates the logical representation of storage from the actual physical storage components. This means the creation of logical volumes out of a pool of physical disks residing in a disk array. A logical volume is limited by the boundaries of the pool, but may span over any number of physical disks within the disk array. Logical volumes can be presented to one or multiple host systems. You cannot have control over the exact allocation of logical volumes on physical disks, but you can influence it with a choice of protection characteristics.

RAID

Redundant Array of Inexpensive Disks (RAID) technology is used to control the way in which the data is distributed across the physical disks within a disk array. Various levels of RAID are available, providing different levels of data redundancy and data security, speed and access time. For example, RAID0 provides no duplication of data, RAID1 provides duplication of all data, RAID5 provides protection of data by parity.

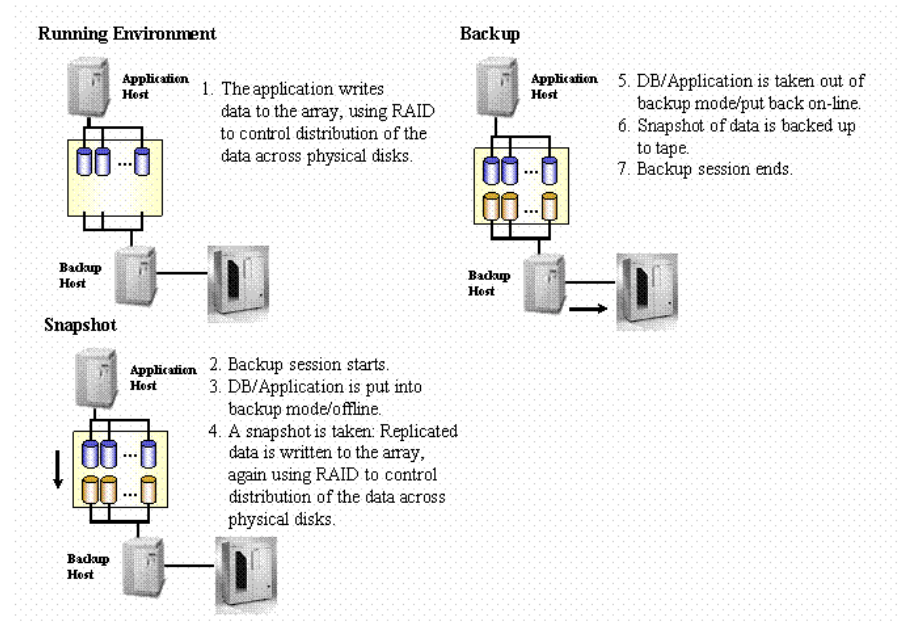
The snapshot integrations for Data Protector are designed to work with disk arrays that use the snapshot technology, such as HP StorageWorks Virtual Array and HP StorageWorks Enterprise Virtual Array.

Snapshot Concepts

In a typical basic setup using the snapshot technology, a single disk array might be connected to separate application and backup systems. The disk array can be used as a storage device by both the application system and the backup system and logical volumes can be mounted on either. Using this arrangement, the application system uses logical volumes within the disk array to store its data during its normal operation. The logical volumes storing the application system data are for the needs of Data Protector snapshot integrations also referred to as **source volumes**. When a snapshot backup is performed, the application data residing on the source volumes is replicated and written to other logical volumes of the same disk array, also referred to as **target volumes**. This replicated data is also referred to as snapshot data and presents an almost instantaneous point-in-time copy of a given filesystem or volume. The set of thus created target volumes is referred to as **replica**. Once the replica for snapshot data is created, the primary data can continue being modified without affecting the backup operation.

Figure 13-1

Snapshot Backup



The backup client is set up as a Data Protector client with tape devices connected, to allow a local backup to be performed.

When a backup session begins, the application client enters the backup mode of operation while the backup client is being prepared for the backup process; a snapshot of the application data is produced.

Once the backup client is ready and the replica for the snapshot data is created, the application is returned to normal operation.

During the time that the application client is in backup mode (or the application may be stopped for a brief period, depending on the application), the impact on application availability is minimal.

If a ZDB to tape is specified, the snapshot data is then streamed to tape media on the backup client. During the tape media streaming operation, the application client can run undisturbed.

Since the application client and backup client are different (in most cases), it is very important that all cached information (database cache, filesystem cache) on the application client is flushed to the array before the snapshot is made. One of the following options can achieve this:

- Databases can be put into backup mode
- Databases can be taken offline
- A mount point can be unmounted

For an online database backup, snapshot data alone does not suffice for a restore. The archive log files from the application client are also needed. An archive log files backup utilizing the standard Data Protector backup procedure can be started immediately after creating snapshots, when the database is taken out of backup mode.

Snapshot data of the application data is produced using the virtual disk array technologies, such as:

- HP StorageWorks Business Copy Virtual Array
- HP StorageWorks Enterprise Virtual Array

Snapshot Backup Types

Within the Data Protector snapshot integrations, the following types of snapshot backups are available:

- ZDB to tape
- ZDB to disk
- ZDB to disk+tape

ZDB to Tape and ZDB to Disk+Tape

During ZDB-to-tape and ZDB-to-disk+tape sessions, a point-in-time snapshot data of the application data is streamed to a tape device, which is connected to a separate backup system, using Data Protector Disk Agent and General Media Agent, with minimal impact on the application system. After the backup is completed, the snapshot data is either:

- discarded - ZDB to tape
- retained and can be used for instant recovery - ZDB to disk+tape

ZDB to Disk

During a ZDB-to-disk session, the same standard snapshot technology is used as in ZDB to tape and ZDB to disk+tape, however, the snapshot data is not streamed to a backup medium (tape device) from the snapshot copy and is retained on a disk array. It can be used for instant recovery. The session effectively ends after the snapshot data is created.

Instant Recovery

During snapshot backup sessions, several snapshot copies of data can be produced and can be retained on a disk array, each point-in-time copy in its own replica. The retained snapshot copies of data can then be used for various purposes, such as offline data processing or instant recovery. Only the point-in-time copies produced during ZDB-to-disk and ZDB-to-disk+tape sessions can be restored using the instant recovery functionality.

Using the instant recovery functionality, the point-in-time copy from a selected replica is restored within a disk array and returned to its state at the point in time that the snapshot data was produced. This process does not involve any restore of data from tape media, dramatically reducing the overall restore time.

Application archive log files are not included in snapshot backup, therefore to restore and apply them, they need to be restored from tape media.

Replica Set and Replica Set Rotation

The maximum number of replicas that can be kept concurrently on a disk array is dependant on the disk array used. The replicas kept on the disk array for the same backup specification form the **replica set** for that backup specification. The replica set is defined by the maximum number of replicas that are to be kept on a disk array for a particular backup specification. When during a snapshot backup session, this

number is reached, the snapshot data in the oldest replica in the replica set is overwritten; if the number is not reached yet, a new replica is created - these two actions are referred to as **replica set rotation**.

Types of Snapshots

Depending on a disk array used, different types of snapshots can be created during a Data Protector snapshot backup session. The Data Protector snapshot integrations utilize the following types of snapshots:

- copy-on-write snapshots with the preallocation of disk space
- copy-on-write snapshots without the preallocation of disk space
- snapclones

Snapshots with the Preallocation of Disk Space

The creation of copy-on-write snapshots with the preallocation of disk space requires the same amount of disk capacity to be allocated as for the source volume. Data is not written to that reserved space until necessary. As the data changes on the source volume, the snapshot data on the target volume is updated with the original data.

Since this snapshot technique caches only the difference between the ever-changing original data content against the point-in-time state, copy-on-write snapshots with the preallocation of disk space are depended on their source volumes; if the data on source volumes is lost, the associated snapshots are useless.

Snapshots without the Preallocation of Disk Space

Copy-on-write snapshots without the preallocation of disk space also represents a point-in-time copy of the original data but it does not require preallocation of disk capacity. The disk capacity is allocated dynamically on as-needed basis. As the data on source volume changes, free space in a disk array is used for the creation of the snapshot. Copy-on-write snapshots without the preallocation of disk space are intended to be short-lived snapshots. Note that their size grows dynamically and may eventually run out of storage capacity if they are not deleted regularly.

The main benefit of copy-on-write snapshots without the preallocation of disk space over copy-on-write snapshots with the preallocation of disk space is in significant reduction of costs. Considerably less additional storage capacity for replication space is needed, if the snapshots are deleted regularly, than with a standard snapshot technology.

Since this snapshot technique caches only the difference between the ever-changing original data content against the point-in-time state, copy-on-write snapshots without the preallocation of disk space are depended on their source volumes; if the data on source volumes is lost, the associated snapshots are useless.

Snapclones

The first part of the snapclone creation is similar to the creation of a copy-on-write snapshot with the preallocation of disk space, which is followed by the cloning process. During this process, all data from the source volume is copied to the target volume. A snapclone enables immediate access to the replicated data while the cloning process runs in the background using the disk array idle time. When the cloning process is finished, the snapclone becomes a full data copy that represents a point-in-time state of the source volume; if the data on source volume is lost, you can always revert to the snapclone.

Supported Configurations

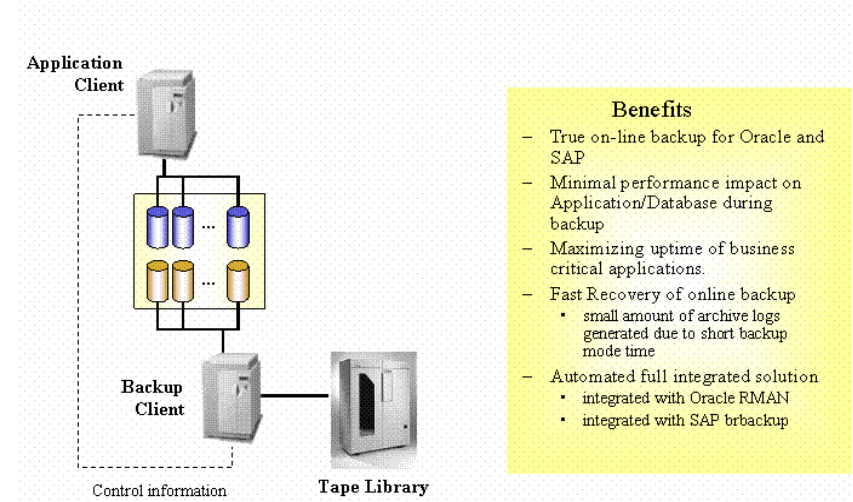
Basic Configuration: Single Disk Array - Dual Host

Both hosts are connected to the same disk array, so that the I/O infrastructure of the RAID system is actually shared between the application client and the backup client.

As the application client and the backup client are two physically different systems, they can use their own resources (I/O channels, CPUs, memory, etc.) for their dedicated activities, such as backup, without interfering with each other. In this way, the impact of the backup on the database performance is minimal.

Figure 13-2

Single Disk Array - Dual Host (Full Performance, Zero Downtime Backup)



The Data Protector snapshot integrations allows automatic handling of disk array status as well as tight integration with applications such as SAP R/3, Oracle or Microsoft SQL or Exchange Server (to ensure data consistency and application/database-aware backups). Only if the application/database is aware of a backup can a secure operation be

Snapshot Concepts

Supported Configurations

guaranteed and native application tools be used for restore (sapdba, for example). The impact of a backup on the application is reduced to the time to perform the following steps:

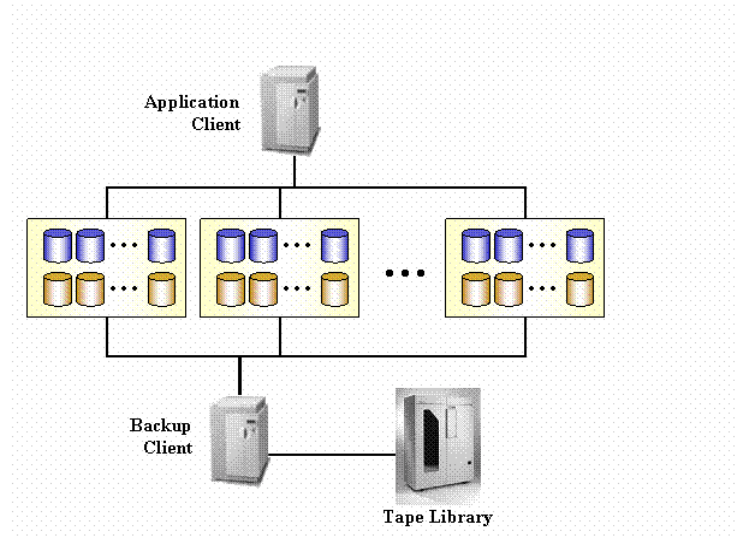
1. Put the database into a consistent mode that permits a snapshot to be taken.
2. Perform a snapshot of the application data.
3. Return the database to normal operating mode.

This configuration enables an offline backup of a very large database in a short time, as well as an online backup that creates very few archive log files, since the backup mode time of the database is kept to a minimum.

The small number of archive logs reduces the space needed for the archive logs in total, as well as speeding up the recovery process of the database. After a restore of an online database, a recovery is needed to return the database to a consistent state. All archive logs that have been created during the backup must be applied. In a snapshot backup, only the archive log files created during the snapshot are applied.

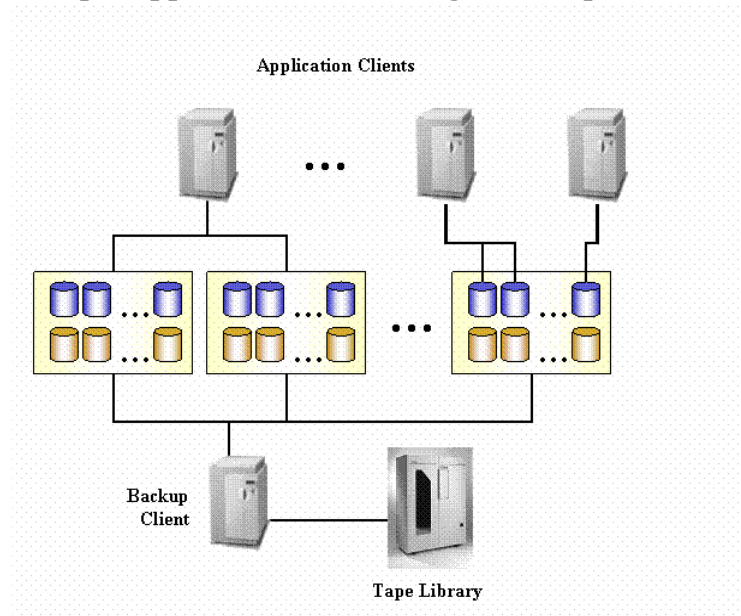
Other Supported Configurations

Figure 13-3 Multiple Disk Arrays - Dual Host



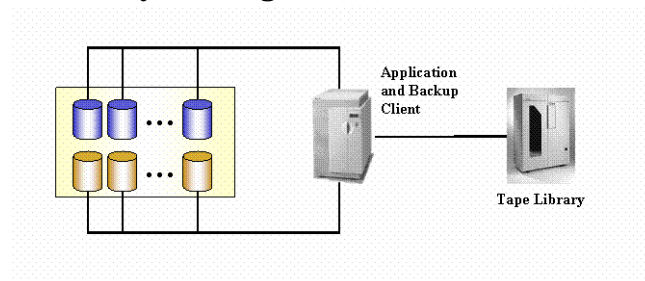
With this solution, both hosts are connected to multiple disk arrays. The I/O infrastructure of the RAID systems is shared between the application client and the backup client.

Figure 13-4 Multiple Application Hosts - Single Backup Host



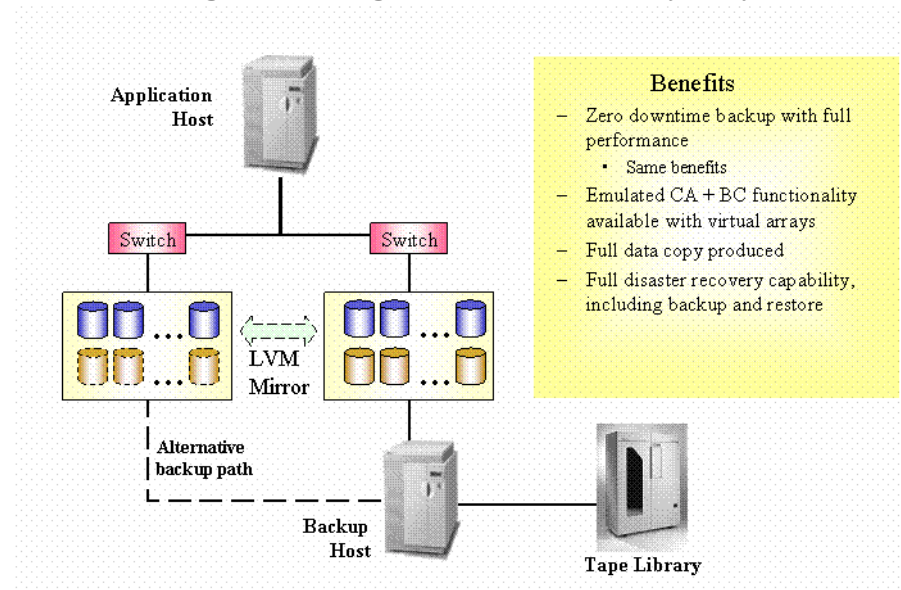
With this solution, multiple application hosts may be connected to a single or multiple disk arrays, which are, in turn connected to a single dedicated backup host. The I/O infrastructure of the RAID systems is shared between the application clients and the backup client.

Figure 13-5 Disk Array(s) - Single Host



In cases where no dedicated backup server is available, both functions (application and backup) can be performed on the same client (or host). Offline backups of mail applications, for instance, could reduce the downtime of the application to minutes instead of hours in this way.

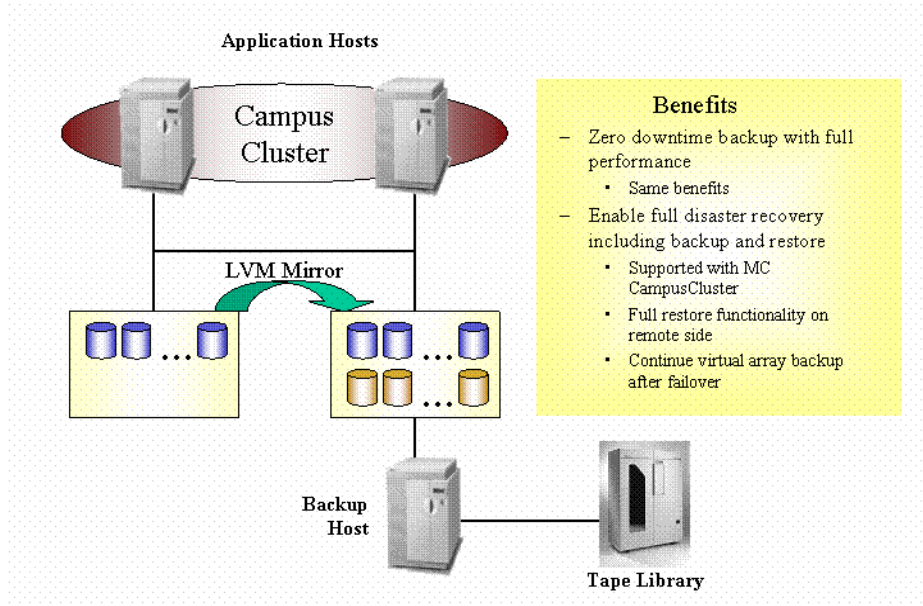
Figure 13-6 LVM Mirroring - HP StorageWorks Virtual Array Only



In the supported configurations described previously, only Business Copy functionality is possible with the HP StorageWorks Virtual Array integration. However, by using LVM mirroring, it is possible to create snapshot copies of data between different virtual arrays, writing to both at the same time. This allows the emulation of the Continuous Access plus Business Copy functionality that is available with HP StorageWorks Disk Array XP.

Figure 13-7

Campus Cluster with LVM Mirroring - HP StorageWorks Virtual Array Only



With this configuration, it is possible to emulate Continuous Access plus Business Copy functionality, with standard cluster failover functionality. This is often required for mission critical applications.

Backup Clients and Clusters

The backup client should not be used as a failover server for the application client. It is recommended to have application and backup services on separate clusters.

Other Configurations

There are many other possible disk array configurations that provide some particular advantage or fulfill a specific user needs. However, each configuration has its specific behavioral pattern that imposes specific requirements on the control functions in order to guarantee backup and recovery. It is important to control and specify which configurations are supported.

Only the configurations shown are supported by HP. For an updated list of supported configurations, refer to the following URL:
http://www.openview.hp.com/products/datapro/spec_0001.html.

In the event that you want to back up data in a configuration not listed, this does not mean that it cannot be supported. Please contact your local HP representative or HP Consulting to investigate the supportability of additional configurations.

Snapshot Concepts
Supported Configurations

14

Microsoft Volume Shadow Copy Service

In This Chapter

This chapter introduces the Microsoft Volume Shadow Copy service (VSS) concept and its role in the backup and restore process. It also outlines the backup and restore flow when using this feature.

The chapter is organized as follows:

“Overview” on page 301

“Data Protector Volume Shadow Copy Integration” on page 306

“VSS Filesystem Backup and Restore” on page 308

For detailed information on the integration, refer to the *HP OpenView Storage Data Protector Integration Guide*. For detailed information on the filesystem backup and restore, refer to the Data Protector online Help.

Overview

A traditional backup process is based on the direct communication between the backup application (application, which initiates and performs backup) and an application to be backed up. This backup method requires from the backup application an individual interface for each application it backs up.

The number of applications on the market is constantly increasing. The necessity of handling application specific features can cause difficulties in backup, restore, and storage activities. An effective solution to this problem is introducing a coordinator among the actors of the backup and restore process.

VSS

Volume Shadow Copy service (VSS) is a software service introduced by Microsoft on Windows operating systems. This service collaborates with the backup application, applications to be backed up, shadow copy providers, and the operating system kernel to implement the management of volume shadow copies and shadow copy sets.

The idea of the Volume Shadow Copy service is to provide a unified communication interface that can coordinate backup and restore of any application regardless of their specific features. With this approach, a backup application does not need to handle each application to be backed up specifically. However, this approach is applicable to a backup application only in case it conforms to the VSS specification.

What Is a Shadow Copy?

A **shadow copy** refers to a volume that represents a duplicate of the original volume at a particular moment in time. The volume shadow copy technology provides a copy of the original volume at a certain point in time. The data is then backed up from the shadow copy, not from the original volume. The original volume continues to change as the backup process continues, but the shadow copy of the volume remains constant.

Shadow copy is basically a snapshot backup, which allows applications and users to continue writing to data volumes, even if they are in the middle of a backup process, while the backup is getting data from a shadow copy of the original volume.

A shadow copy set is a collection of shadow copies created in the same point in time.

What Is a Writer? A **writer** refers to any process that initiates change of data on the original volume. Writers are typically applications (for example, MSDE Writer for MS SQL Server) or system services (for example, System Writer and Registry Writer) that write persistent information on a volume. Writers participate in the shadow copy synchronization process by assuring data consistency.

What Is a Shadow Copy Provider? A **shadow copy provider** refers to some entity that performs the work involved in creating and representing the volume shadow copies. Shadow copy providers own the shadow copy data and expose the shadow copies. Shadow copy providers can be software (including a system provider, MS Software Shadow Copy Provider) or hardware (local disks, disk arrays).

The example of the hardware provider is disk array, which has its hardware mechanism of providing point-in-time state of a disk. A software provider operates on physical disks and uses software mechanism for providing point-in-time state on a disk. The system provider, MS Software Shadow Copy Provider, is a software mechanism, which is a part of the Windows Server 2003 operating system.

The VSS mechanism guarantees that all hardware providers will be offered for creating shadow copy before all software providers. If none of them is able to create a shadow copy, VSS will use the MS Software Shadow Copy Provider for the shadow copy creation, which is always available.

Data Protector and VSS The Volume Shadow Copy service enables coordination among the backup application, writers, and shadow copy providers during the backup and restore process.

Figure 14-1 and Figure 14-2 show differences between the traditional backup model and the model with the VSS coordinator.

Figure 14-1 Actors of the Traditional Backup Model

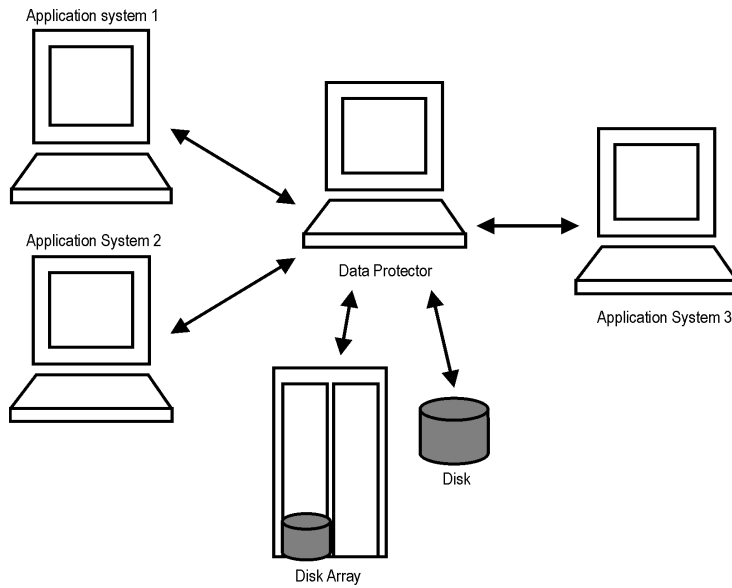
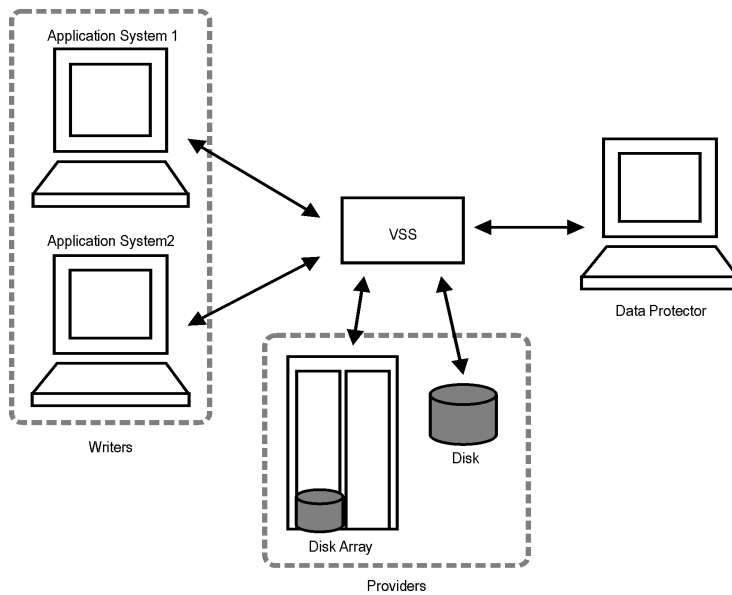


Figure 14-2 Actors of the VSS Backup Model



Overview

In the traditional model, the backup application had to communicate with each application it backed up individually. In the VSS model, the backup application communicates with the VSS only, and the VSS coordinates the whole backup process.

VSS Benefits

The advantages of using Volume Shadow Copy service are the following:

- Unified backup interface is provided for all writers.
- Unified backup interface is provided for all shadow copy providers.
- Data integrity is provided on application level by the writers. No interference is needed from the backup application.

Data Protector supports the Microsoft Volume Shadow Copy service on two levels:

- Within the integration with the Microsoft Volume Shadow Copy service, Data Protector provides a shadow copy backup and restore of the VSS-aware writers, including the ZDB and instant recovery functionality.
- Within the Disk Agent functionality, Data Protector provides VSS filesystem backup.

Data Protector VSS integration supports a consistent shadow copy backup only for the VSS-aware writers. Consistency in this case is provided on the writer's side. When the applications are not VSS-aware, the shadow copy is created, but the consistency of the shadow copy data is not guaranteed on the level of the application. However, the consistency is improved in comparison with the non-VSS filesystem backup.

The table below outlines the differences when using Data Protector VSS integration backup, VSS filesystem backup, or non-VSS filesystem backup:

Table 14-1

Benefits of Using VSS

	Data Protector VSS Integration Backup	VSS Filesystem Backup	Non-VSS Filesystem Backup
Open Files	No open files.	No open files.	If files are open, backup may fail.

Table 14-1 **Benefits of Using VSS**

	Data Protector VSS Integration Backup	VSS Filesystem Backup	Non-VSS Filesystem Backup
Locked Files	No locked files.	No locked files.	If files are locked, backup skips them.
Data Integrity	Provided by the writer.	Crash consistency state (like in case of power failure).	None (inherent).

Data Protector Volume Shadow Copy Integration

Data Protector integration with Microsoft Volume Shadow Copy service provides full support of the VSS-aware writers. This includes automatic detection of the VSS-aware writers, backup and restore functionality. For detailed information on the integration, refer to the *HP OpenView Storage Data Protector Integration Guide*.

VSS Backup

In case of the VSS-aware writers' backup, the consistency of data is provided on the level of writer and is not dependent on the backup application. Data Protector follows the requirements provided by the writers when selecting what to back up.

During the backup of the VSS-aware writers, Data Protector does not communicate with each writer individually, but through the VSS interface. It uses the VSS integration agent to connect the Volume Shadow Copy service, which coordinates the backup process. VSS provides Data Protector with the writers-related metadata necessary for performing a consistent backup and restore. Data Protector examines this data and identifies the volumes to be backed up. Then Data Protector requests VSS for creating a shadow copy of the specified volumes.

NOTE

A **Writer Metadata Document** (WMD) is metadata provided by each writer. Writers identify themselves by the metadata and instruct backup application what to back up and how to restore the data. Thus, Data Protector follows the requirements provided by the writer when selecting the volumes to be backed up and the restore method.

Volume Shadow Copy service synchronizes the writers and providers. After a backup shadow copy is created, VSS communicates this to Data Protector. Data Protector performs a backup from the shadow copy volume to the media and then notifies VSS that the shadow copy can be released.

VSS Restore

VSS integration restore refers to the restore of data, which have been backed up with the coordination of the Volume Shadow Copy service and with the cooperation of a writer. During the restore procedure, Volume Shadow Copy service coordinates communication between Data Protector and writers.

When restoring VSS-aware writers, Data Protector first restores all the relevant metadata to identify the backup components and determine the restore method. Then it connects to the Volume Shadow Copy service and declares that restore is about to start. VSS coordinates writers' activities during the restore time. After Data Protector successfully restored the data, VSS informs the writers that the restore is completed and the writers can access the restored data and start their internal processing.

VSS Filesystem Backup and Restore

There are applications that are not aware of the Volume Shadow Copy service. Such applications cannot guarantee consistency of data in the time of creating shadow copy. The VSS mechanism cannot coordinate activities of these applications in order to perform a consistent backup.

However, you can still benefit from the VSS functionality. The cooperation between the backup application and a shadow copy provider can be still used to assure the higher level of data consistency. Microsoft calls this state of data consistency “crash consistent data”. This means, that the VSS mechanism commits all pending I/O operations and holds incoming writing requests during preparation of a shadow copy volume. In this way all files on the filesystem are closed and unlocked when the shadow copy is being created.

Microsoft Volume Shadow Copy functionality allows the creation of volume shadow copy without the participation of the applications being backed up. In this case, the shadow copy volume is created and then backed up by Data Protector. This approach can be used on the applications that are not aware of the VSS mechanism.

IMPORTANT

In case of backing up applications that are not aware of the VSS mechanism, the data consistency from the applications’ point of view cannot be assured. The data are as consistent as in the case of power failure. Data Protector cannot guarantee any data consistency when applications are not actively participating in shadow copy creation.

During the VSS filesystem backup the consistency of data is improved in comparison with the non-VSS filesystem backup. VSS allows you to create shadow copy backups of volumes, exact point-in-time copies of files, including all open files. For example, databases that are held open exclusively and files that are open due to operator or system activity are backed up during a VSS filesystem backup. In this way, files that have changed during the backup procedure are copied correctly.

The advantages of the VSS filesystem backup are the following:

- A computer can be backed up while applications and services are running. Therefore, applications can continue to write data to the volume during a backup.
- Files that are open are no longer skipped during the backup process, because they appear closed on the shadow copy volume at the time of the shadow copy creation.
- Backups can be performed at any time without locking out users.
- There is little or no impact on performance of the application system during the backup process.

Backup and Restore

The VSS backup is implemented as an additional Windows filesystem backup on Windows Server 2003. The level of data integrity is slightly increased compared to traditional backup of active volume. For detailed information on the Windows filesystem backup and restore, refer to online Help.

During the VSS filesystem backup, applications cannot effectively contribute to the data consistency, as they are not aware of the VSS mechanism. However, Data Protector and a provider can still cooperate in creating volume shadow copies. The VSS filesystem backup offers an opportunity to back up data as they appear at a certain point-in-time regardless of the system I/O activity during the backup.

When Data Protector requests the backup of the volumes specified in the backup specification, the VSS mechanism commits all pending I/O operations, holds incoming writing requests, and prepares a shadow copy volume.

When the shadow copy is created, Data Protector starts its normal backup procedure, except that the source volume is replaced by the newly created shadow copy. If the shadow copy creation fails, Data Protector can proceed with the normal filesystem backup, if this behavior was specified in the backup specification.

A computer is backed up while files are open and services are running. Files are not skipped during such backup. The VSS allows services and applications to continue running uninterrupted on the actual volumes while a shadow copy is being made. After the backup is completed, the shadow copy is deleted.

The restore of data backed up using the VSS filesystem backup does not differ from the standard restore procedure.



A Backup Scenarios

In This Appendix

This Appendix describes two scenarios: one for company XYZ and one for company ABC. Both companies plan to enhance their data storage systems. Their current backup solutions are described along with the inherent problems. Solutions are then proposed to alleviate the problems and to meet the future data storage needs of both companies.

Considerations

In both cases, the following considerations must be taken into account when formulating a company's backup strategy:

- How critical system availability (and backup) is to the company
 - The need to keep the backed up data at a remote location in case of disaster.
 - The level of business continuance. This includes the recovery and restore plan for all critical systems.
 - The security of backed up data
 - The need to guard premises to prevent unauthorized people from entering. This also includes safeguarding all relevant data against unauthorized access with physical access prevention and electronic password protection.
- The type of data that needs to be backed up
 - The company's data can be divided into categories like company business data, company resource data, project data, and personal data, each with its own specific requirements.
- Performance aspects for backups and restores
 - Network and system topology
 - Determine which systems can use what network links and what transfer rates are possible.
 - Time window
 - Define the periods of time during which backups of specific systems can be done.

- Local versus network backups

Determine which systems, that the backup devices are connected to, are backed up locally and which are backed up over the network.
- Backup policy implementation
 - How backups are done, and which backup options are used

This defines the frequency of full and incremental backups. It also defines the backup options that are used, and whether the backups are permanently protected with the backup media stored at a remote site.
 - How the systems are grouped into backup specifications

Consider how best to group backup specifications. This can be on the basis of departments, data types, or backup frequency.
 - How the backups are scheduled

Consider using the staggered approach, whereby full backups are scheduled for different clients (backup specifications) on different days to avoid network load, device load, and time window issues.
 - Retaining data on media and information about backups

Consider protecting data from being overwritten by newer backups for a specified amount of time.

Define the period of time that the Data Protector Catalog Database should store information about backups.
- Device configuration

Determine the devices to be used for backups and the systems they are connected to. Connect the backup devices to systems with the greatest amount of data so that as much data as possible is backed up locally and not through the network. This increases backup speed.

If you have large amounts of data to back up, consider using a library device
- Media management

Determine the type of media to be used, how to group the media into media pools, and how to position objects on the media.

- Vaulting
Decide whether to store media to a safe place, where it is kept for a specific period of time.
- Backup administrators and operators
Determine the administration and operations rights for the backup systems users.

Company XYZ

XYZ is a translation agency providing the following services:

- Translation, localization, language editing, and proof-reading
- Certification of translated documents
- Simultaneous and consecutive interpretation
- Desktop publishing and graphic design
- Rental of conference interpreting equipment

XYZ is currently growing at 20-25 percent per year. Their current backup solution is not able to keep pace with this growth. The backup process is very labor intensive because of the manual process in handling backup tapes.

Environment

This section describes the present-day hardware and software environment of XYZ and how the data storage policy is implemented.

XYZ is divided into three departments, which are connected to a Corporate Network backbone:

- English Department
- Other Languages Department
- Admin Department

The hardware and software environment of XYZ is depicted in Table A-1 on page A-5 and the current backup topology in Figure A-1 on page A-7.

Table A-1

Hardware and Software Environment of XYZ

Depart.	#Servers	#Clients	Current Data	Projected Data (in 5 Years)	Current Devices
English	1 Windows 2000	15 Windows	35 GB	107 GB	3 HP StorageWorks DAT24 autoloaders

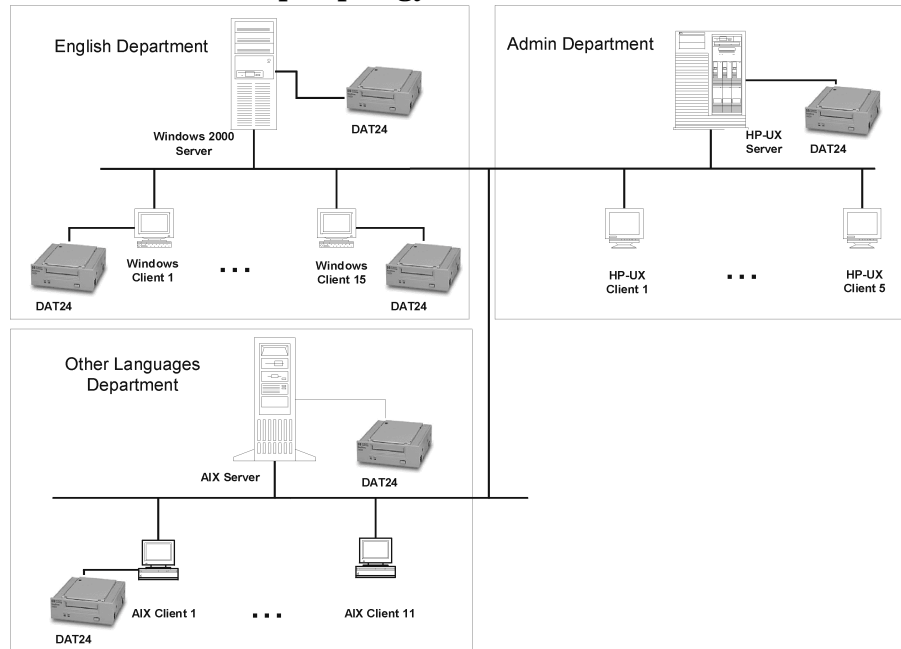
Table A-1 Hardware and Software Environment of XYZ

Depart.	#Servers	#Clients	Current Data	Projected Data (in 5 Years)	Current Devices
Other Languages	1 AIX	11 UX	22 GB	67 GB	2 HP StorageWorks DAT24 autoloaders
Admin	1 HP-UX	5 UX	10 GB	31 GB	1 HP StorageWorks DAT24 autoloader

Figure A-1 on page A-7 shows how the XYZ backup environment is organized.

Figure A-1

Current XYZ Backup Topology



XYZ currently has three servers with an estimated total data volume of 67 GB. In the English Department, data is copied manually by each of the employees to their respective servers at the end of each day. One of the Windows 2000 clients in this department accounts for approximately a third of the data (12 GB).

The backup of clients in the Other Languages Department is done through a Network File System, while the backup of clients in the Admin Department is done through network shares. Employees in the Other Languages Department also work on Saturdays.

Problems with the Current Solution

The current backup solution is not able to keep pace with the growth rate of XYZ. The actual backup process is very labor intensive. The current backup process makes it impossible to consolidate backup management or create an enterprise-wide backup architecture. Each of the backup servers is managed individually. There is no capability for a central backup management. The problems of the current backup solution include the following:

- The backup solution is not automated.

- People must copy their work regularly, which creates a high potential for errors.
- The backup utilities that are used are not the same, resulting in higher training costs.
- The solutions used in the Other Languages and the Admin Departments are less primitive but do have their problems. Network usage has a high impact on backup performance. Moreover, not all data gets backed up. Only Network File System shared files and network shared files are backed up in the Other Languages and Admin Departments, respectively.
- Because there are three independent backup servers for the three departments, there is no central control or administration of the following key areas:
 - Device configuration
 - Media management
 - Backup configuration
 - Scheduling
 - Monitoring
 - Restore operations
- Because each of the backup servers is managed individually, there is no central reporting.
- The current solution does not offer disaster recovery capabilities. This is an increasingly important setback. A disaster may result in the company losing a significant part of its business.

Backup Strategy Requirements

Requirements

After addressing the items under “Considerations” on page A-2, the following requirements have been identified for the backup solution of company XYZ:

- Backup Policy
 - Full, weekly backups will occur and be completed within 12 hours.
 - Daily incremental backups will occur at the end of each workday and will be completed within 8 hours.

- A permanent data protection period will be included.
- Backup media will be stored at a remote site.
- Backup

All backup operations must require less manual intervention than currently.
- Restore
 - Convenient and fast restore of recent data must be provided. Data to be restored must be browsable for the first 3 weeks after backup.
 - Restores of backups of data in the vault must be possible within two days.
- Network Connectivity

The backup servers and the departments will be connected to a 100TX Ethernet LAN.
- Planned Growth

Growth in the current data capacities is projected at 20 to 25% per year in the next five years.
- Software

The backup servers need to be running on one of the supported operating systems. For information on supported operating systems for the Cell Manager, refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*.
- Protection Against Disaster

Upon completion of backups, the media will be stored on-site, where they will be retrieved upon request for file restoration. After 20 days, they will be moved to an off-site storage facility for protection in case of a disaster at the company site, and to make space for new backups.

Proposed Solution

Because of the limitations of the current backup solution for both performance and enterprise-wide management, there is a need to redesign XYZ's backup architecture and strategy to meet its business

objectives. An overview of the proposed solution is given, followed by a detailed account of the solution. Note that this is a proposal and not the only possible solution to XYZ's storage management problems.

Solution Overview All clients and servers should be configured into a single Data Protector cell with the Windows 2000 Server of the English Department as both the Cell Manager and Installation Server for Windows systems. Use the HP-UX backup server of the Admin Department as the Installation Server for UNIX systems. The backup devices consist of an HP StorageWorks DLT 4115w Library, as well as two of the HP StorageWorks DAT24 autoloaders that had been used to date.

This suffices for the next five years at the present data growth rate of 20 to 25% per year. The use of devices that have been used to date provides an added advantage in case of disaster recovery. The Windows 2000 client, which accounts for approximately a third of the data in the English Department (12 GB), should be backed up locally to an HP StorageWorks DAT24 autoloader. The proposed backup solution addresses the following key items:

- Achieving high performing backups
- Media management with minimum human effort
- Simple and effective disaster recovery
- Centralized backup reporting
- Automation of most backup operations

All this is achieved with a single solution in combination with the proposed hardware:

Table A-2 Proposed Environment

Department	Current Data	Projected Data (In 5 Years)	Devices	
English*	35 GB	107 GB	HP DLT 4115 library	2 HP StorageWorks DAT24 autoloaders
Other Languages	22 GB	67 GB		
Admin	10 GB	31 GB		
<p>* One HP StorageWorks DAT24 autoloader is currently used to locally back up the 12 GB of data. The other HP StorageWorks DAT24 autoloader is used to back up the IDB and configuration files. The rest of the data in this department is backed up remotely to the HP StorageWorks DLT 4115 library.</p>				

The remaining 4 HP StorageWorks DAT24 autoloaders are used in a separate R&D system, which is not of our configuration.

The software components proposed for the Enterprise Backup solution include HP OpenView Storage Data Protector A.06.00.

Proposed Solution in Detail

The following is a detailed account of the proposed solution:

- Cell Configuration

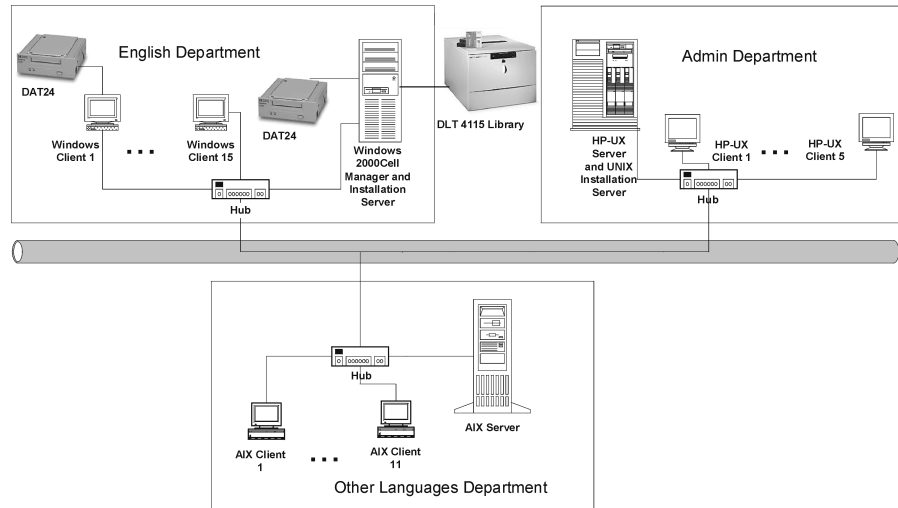
All clients and servers should be configured in a single Data Protector cell. The Data Protector Cell Manager can run on the Windows 2000 Server of the English Department.

All systems in the cell should be on the same LAN for maximum performance. The Cell Manager should also be the Installation Server for Windows. Use the HP-UX backup server of the Admin Department as the Installation Server for UNIX. The HP StorageWorks DLT 4115w Library should be connected to the Cell Manager as well as one HP StorageWorks DAT24 autoloader for

backing up the IDB and configuration files. The Windows 2000 client, which accounts for approximately a third of the data in the English Department (12 GB) should be backed up locally to an HP StorageWorks DAT24 autoloader.

The proposed backup environment is as depicted in Figure A-2 on page A-12:

Figure A-2 Proposed XYZ Backup Topology



The Cell Manager maintains the Catalog Database (CDB). This provides a minimum of 20 days of file and directory detail on the current database.

Estimating the Size of the IDB

The Internal Database Capacity Planning Tool was used to estimate the size of the IDB in a year. The tool is located in the same directory as the rest of the Data Protector online manuals. Input parameters shown in Figure A-3 on page A-13 include the number of files in the environment (2 million), the growth factor (1.2), data protection (52 weeks), catalog protection (3 weeks), the number of full backups per week (1), and the number of incremental backups per week (5).

Figure A-3 Input Parameters

Environment description			
Files:	2	million	
Files per directory:	10		
Data volume:	200	GB	
Growth factor:	1.20		
Device performance:	10.00	MB/second	
Medium capacity:	70.00	GB	
Objects:	50		
Change per incr-bkup:	5.00%		
Backup parameters			
Device concurrency:	2		
Data segment size:	2,048.00	MB	
Log level:	All		
Data protection:	52	Weeks	
Catalog protection:	3	Weeks	
Full backups/week:	1		
Incr backups/week:	5		
Cell Manager parameters			
Insertion speed:	12	million/hour	

The results are shown in Figure A-4 on page A-13. In one year, the database is expected to grow to approximately 419.75 MB.

Figure A-4 Results

Results/calculation			
Avg. file size:	123.36	KB	
Files/segment:	16,931.38		
Catalog size:	1.03	MB	
K devices:	40		
K performance:	1445,647059	GB/hour	
K duration:	0.14	hour	
Protected media:	278.5714286		
Space estimation			
MMDB:	30.00	MB	
CDB:	Fnames:	153.00	MB
	Overs:	5.71	MB
	Mpos:	1.54	MB
DCBF:	229.50	MB	
SMBF:	2.86	MB	
Total:	419.75	MB	

- Hardware
 - Network

All systems should be on the same 100TX network for maximum performance. This network has a sustained data transfer rate of 10 MB/s, or 36 GB/h, of data.

— Backup Devices

The backup devices consist of an HP StorageWorks DLT 4115w Library as well as two HP StorageWorks DAT24 autoloaders.

Why Use the HP StorageWorks DLT 4115w Library?

The HP StorageWorks DLT 4115w Library has a single DLT4000 drive with 15 slots. It has a total compressed storage capacity of 600 GB and a maximum sustained data transfer rate of 3 MB/s, or 10.5 GB/h, with data compression. This is the transfer rate assumed for the remainder of this section. Currently, the total amount of data to be backed up to the HP StorageWorks DLT 4115w Library as a full backup, whether this is a single full backup, or the staggering approach is used, is about 55 GB. Assuming that the size of an incremental backup is approximately 5% of that of a full backup, a backup generation, representing a full backup and all incremental backups based on this full backup, requires $(55+55*5\%*5)$ GB, or **68.75 GB**, of library space. In five years time, this figure is projected to increase to about 210 GB. XYZ's backup policy requires that two backup generations of data be kept. Therefore, $210*2$ GB, or 420 GB, of library space will be required for storage. The HP StorageWorks DLT 4115w Library's 600 GB storage capacity therefore suffices.

Why Use the HP StorageWorks DAT24 Autoloader?

The HP StorageWorks DAT24 autoloader has 6 24-GB data cartridges. It has a total compressed storage capacity of 144 GB and a maximum sustained data transfer rate of 2 MB/s, or 7 GB/h, with data compression. This is the transfer rate assumed for the remainder of this section. Currently, the total amount of data to be backed up to the HP StorageWorks DAT24 autoloader connected to the aforementioned Windows 2000 client in the English Department in a single full backup is 12 GB. Assuming that the size of an incremental backup is approximately 5% of that of a full backup, a backup generation, representing a full backup and all incremental backups based on this full backup, requires $(12+12*5\%*5)$ GB, or **15 GB**, of space. In five years time, this figure is projected to increase to about 45 GB. XYZ's backup policy requires that two backup generations of data be kept. Therefore, $45*2$ GB, or **90 GB**, of

library space will be required for storage. The HP StorageWorks DAT24 autoloader's 144 GB storage capacity therefore suffices.

How Long Does a Full Backup Last?

The Windows 2000 client in the English Department, which accounts for 12 GB of data is backed up locally to an HP StorageWorks DAT24 autoloader. This device has a sustained data transfer rate of 2 MB/s, or approximately 7 GB/h. Therefore, a full backup of this Windows 2000 client takes about **2 hours**. As the amount of data is growing at 20 to 25% per year, this client is projected to hold about 36 GB of data in five years time. This data would then be backed up in **6 hours**.

The Data Protector Catalog Database is approximately 0.4 GB in size. It is backed up locally to an HP StorageWorks DAT24 autoloader, which has a sustained data transfer rate of 2 MB/s or 7 GB/h. Data Protector by default checks the integrity of the database before the database is backed up. It takes less than half an hour to check the integrity of a 0.4 GB database and only a few minutes to back up the database. Therefore, to check the integrity of, and then back up the IDB and configuration files requires less than **1 hour**.

The projected size of the database in five years time is 1.2 GB. It takes less than an hour to check the integrity of a 1.2 GB database and less than half an hour to back it up. Therefore, to check the integrity of, and then back up the IDB and configuration files requires less than **2 hours**.

All the other available data in the system, which is currently about 55 GB) is backed up remotely to the HP StorageWorks DLT 4115w Library, which has a sustained data transfer rate of 3 MB/s, or 10.5 GB/h. Most of this data is via the 100TX network, which has a sustained data transfer rate of 10 MB/s, or 36 GB/h, of data. This does not present a bottleneck. The backup of all these data would therefore take about **5 to 7 hours** to complete. This is well within the allowed 12 hours. The problem would then be that in five years time, when the data is projected to be about 170 GB, the backup would take 15 to 21 hours!

To solve this problem, use the staggering approach. Schedule the full backup of data in the English Department for Fridays at 20:00, and that in the Other Languages Department for Saturdays at 20:00 and that in the Admin Department for Sundays at 20:00.

Table A-3 The Staggering Approach

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
English	Inc1	Inc1	Inc1	Inc1	Full	Inc1	
Other Languages	Inc1	Inc1	Inc1	Inc1	Inc1	Full	
Admin	Inc1	Inc1	Inc1	Inc1	Inc1		Full

Table A-4 on page A-16 shows the size and time requirements for these full backups as of today, as well as the five year projection.

Table A-4 Remote Full Backups to the HP DLT 4115 Library

Department	Current Data/Backup Time	Projected Data/Backup Time
English	23 GB / 3 h	70 GB / 7 h
Other Languages	22 GB / 3 h	67 GB / 7 h
Admin	10 GB / 1 h	31 GB / 3 h

Based on the assumption that the estimated size of an incremental backup is 5% of that of a full backup, a full backup of all data that is remotely backed up in the largest department, the English Department, as well as incremental backups of the other two departments is projected in five years to take $7+5\%(7+3)$ hours, which is less than 8 hours. This is well within the allowed 12 hours.

- Media Pools

Media are grouped into media pools to provide better media tracking and control. Group each of the two media types (DLT and DDS) in its own pool.

- Default DDS

This pool should be used for all DDS media.

- Default DLT
This pool should be used for all DLT media.
- DB_Pool
This pool should be used for the IDB and configuration files. The database should be backed up to two media for security reasons.
- Backup Specifications
Configure five backup specifications, one for each department, and one for the IDB and configuration files:

- ENG1_BS

Backup specification for the Windows 2000 client to be backed up locally in the English Department. Schedule the backup specification such that Data Protector will run a full backup every Friday and a level 1 incremental backup every day, except Friday and Sunday at 20:00.

Why Use Level 1 Incremental Backups?

To restore the latest data, only two media sets need to be accessed, one for the latest full backup and one for the latest level 1 incremental backup prior to the restore point-in-time. This simplifies and speeds up restore considerably.

- ENG2_BS

Backup specification for data in the English Department to be backed up remotely to the HP StorageWorks DLT 4115w Library. Schedule the backup specification such that Data Protector will run a full backup every Friday and level 1 incremental backups every day, except Sunday at 20:00.

- OTH_BS

Backup specification for data in the Other Languages Department to be backed up remotely to the HP StorageWorks DLT 4115w Library. Schedule the backup specification such that Data Protector will run a full backup every Saturday at 20:00 and level 1 incremental backups every day, except Sunday at 20:00.

- ADM_BS

Backup specification for data in the Admin Department to be backed up remotely to the HP StorageWorks DLT 4115w Library. Schedule the backup specification such that Data Protector will

run a full backup every Sunday at 20:00 and level 1 incremental backups every day, except Saturday at 20:00.

— DB_BS

Backup specification for the IDB and configuration files. Schedule the backup specification such that Data Protector will run a full backup every day at 4:00. At this time, other full and incremental backups would be completed and there would be no CPU resource sharing problem between the Cell Manager and other client systems. Two copies of the database should be made.

Backup Options

Use default Data Protector backup options. Set the following options as follows:

— Catalog Protection

Catalog protection sets the amount of time that the Data Protector Catalog Database stores information about backed up versions, information about the number of backed up files and directories, and messages stored in the database. Once catalog protection expires the browsing of files and directories using the Data Protector GUI is no longer possible. Set catalog protection to 20 days.

— Data Protection

Data protection determines the amount of time until each medium can be reused. Set data protection to permanent so that data on the media is not overwritten unintentionally.

— Concurrency

Set to 5 to allow up to five Disk Agents to concurrently write data to the HP StorageWorks DLT 4115w Library. This will increase backup performance.

— Media Pool

For the IDB, select the DB_Pool with the appropriate media to be used for the backup. Other objects use default media pools.

Restore Options

Use default Data Protector restore options. Set the following options as follows:

— List Restored Files

Set to ON to list the pathnames of files and directories that are restored. This option can slow down the restore, if there are too many files to be restored.

— Display Statistical Information

Set to ON to display detailed statistical information about a specific restore session, which includes the number of restored files and directories as well as the amount of restored data.

- Reporting and Notification

Email notifications will be set up for backup administrators for mount requests, low database space, device errors, and end of session events for all backup specifications. Optionally, email or broadcast notifications will be set up for those end users interested in being notified about the success of backups of their systems.

To enable all users to easily determine the status of backup, set up client backup information on the company intranet as follows:

1. Configure a report group with a Client Backup Report for each client. The report should be logged to the file in HTML format.
2. Schedule the report group.
3. Link the logged files to the company intranet page.

- Vaulting

Vaulting is a process of storing media to a safe location for a specified period of time.

Media will be moved to the vault once a week and replaced by new media in the HP StorageWorks DLT 4115w Library and HP StorageWorks DAT24 autoloaders. All actions excluding the actual moving of media to the vault are done by the software solution, including queries done internally in the database to prevent the administrator from having to find media that require ejection.

The second migration of media is done to move media from the vault to a security company. This is done once a month. Data Protector provides a report on what media need to be moved to a security company.

Track the location of media that are moved to a vault. This is important when you want to restore from backups on media that were moved to a security company. Data Protector allows you to perform the following vaulting tasks:

- Generate reports showing media stored at a specific location with data protection expiring in a specified time
- Generate reports showing media used for backup within a specified time frame
- Display a list of backup specifications that have used specified media during the backup.
- Display a list of media needed for restore and the physical locations where the media are stored.
- Filter media from the media view based on specific criteria, such as media with expired protection.
- Restore
 - Restore by Query

Requests for restores by query will be sent to the administrator. If the files were last backed up less than 20 days before the request was placed, then the administrator can use the Restore by Query restore task to select the files and directories to be restored using a specified criteria. The administrator then selects the Overwrite option to replace files and directories on the disk with the versions on the media.
 - Complete Filesystem Restore

Requests for the restore of whole filesystems will be sent to the administrator. If the files were last backed up less than 20 days before the request is placed, then the administrator can select the objects for restore and use the Restore Into option.

With the Restore Into option selected, the object is restored with the exact directory structure to a selected directory. Use a Windows or UNIX utility to compare the restored object with the backed up object.
 - Restore from a Vault

To restore data from a vault, which is, for example, 3 years old, send a request to the administrator who then:

1. Identifies the media needed for restore.
2. Brings the media from a vault, enters the media in the HP StorageWorks DLT 4115w Library or other device and then scans the media.
3. Selects the specific object to be restored using the List From Media option, if the media are not in the IDB.
4. Performs the restore.

Company ABC

ABC is a high growth software engineering company with headquarters in Cape Town, South Africa. As a software engineering outsourcer for multinational partners, ABC transparently sets up multi-site project teams and the accompanying infrastructure to seamlessly execute a wide array of software engineering projects. ABC has been growing at a rate of 30-40% per year. The growth rate is expected to slow down to 15 to 20% in the next five years.

Environment

This section describes the present-day hardware and software environment of ABC and how the data storage policy is implemented.

ABC has offices at three locations. The main hardware data at the three locations is given in Table A-5 on page A-22.

Table A-5

Size of Backup Environment

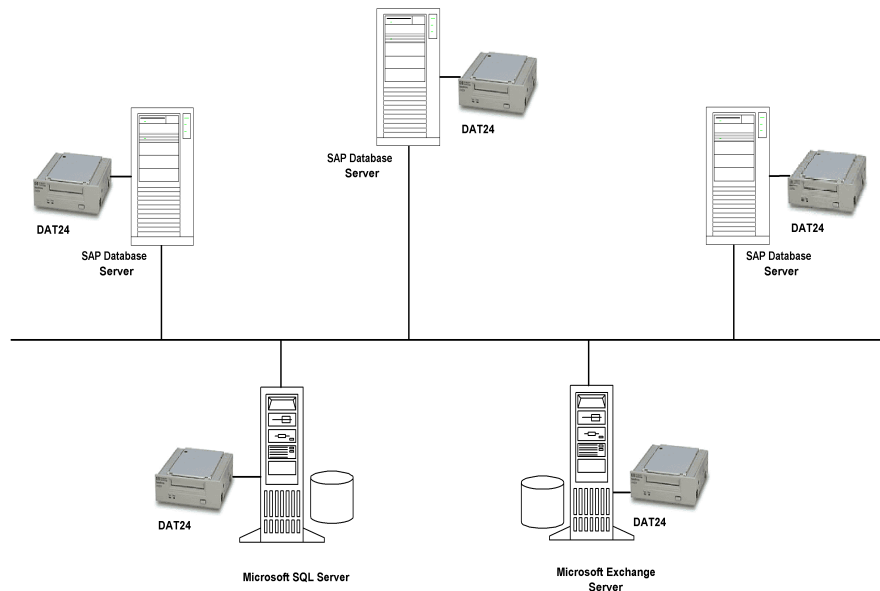
Location	#Win Servers	#Win Clients	#UX Servers	#UX Clients	Current Data	Data (in 5 Years)	Current Devices
ABC Cape Town	7	55	11	40	100	250	5 DAT24*
ABC Pretoria	5	39	5	32	22	55	1 DAT24*
ABC Durban	3	21	6	59	16	40	1 DAT24*
* HP StorageWorks DAT24 autoloader							

Three departments at ABC Cape Town use the Microsoft SQL database to store their data and the company uses Microsoft Exchange Server for mailing services. These databases, currently containing 11 GB and 15 GB of data, respectively, are backed up to 2 HP StorageWorks DAT24 autoloaders.

The system architecture of ABC Cape Town includes the SAP R/3 system using Oracle databases. Three HP T600 servers are used as SAP database servers. ABC Cape Town uses K260 SAP application servers that are configured into application groups, i.e. Sales and Distribution,

Finance, and Production. The application servers are not highly available. The current backup environment of ABC Cape Town is depicted in Table A-5 on page A-23.

Figure A-5 Current ABC Cape Town Backup Topology



Currently, backups of the SAP database servers at ABC Cape Town are performed using the SAP BRBACKUP and BRARCHIVE utilities to 3 HP StorageWorks DAT24 autoloaders. Data is copied manually by employees to their respective servers on a daily basis. The Microsoft Exchange Server and Microsoft SQL database are backed up separately each to an HP StorageWorks DAT24 autoloader by the backup administrator.

The same system is used at both ABC Durban and ABC Pretoria, with the difference that no SAP system is in place at these sites. Employees copy their data to their respective servers. Data is backed up to an HP StorageWorks DAT24 autoloader on a daily basis.

Two of the servers at ABC Pretoria have more than 500 000 files each.

Backup media are denoted by the name of the department, the name of the server and first and last dates on which backups were performed on the media. At the end of each quarter, media are sent for storage to a central offsite location.

Problems with Current Solution

The current backup solution has the following deficiencies:

- There is no online backup solution of the SAP database server.
- The backup solution is not centralized.
- Backup operations are not fully automated.
- Media management requires considerable human effort.
- Disaster recovery is complex.
- Backup operations last longer than the allowed time window.
- The backup solution cannot keep pace with the high growth rate of ABC.
- No reporting and notifications of important events pertaining to the backup.

Backup Strategy Requirements

Before addressing ABC's backup strategy requirements, consider the items under "Considerations" on page A-2.

Requirements

The following section gives a description of ABC backup strategy requirements.

- Organizational policies regarding backups and restores

The company policy on archiving and storing data defines that weekly backups be completed within **12 hours** and that daily incremental or differential backups be completed within **8 hours**.

- Maximum downtime for recovery

The allowed downtime has a significant impact on the investments into the network infrastructure and the equipment needed for the backup. The following table lists, for each type of data, the maximum acceptable downtime for recovery, that is, how long specific data can be unavailable before recovered from the backup.

Table A-6 Maximum Acceptable Downtime for Recovery

Type of Data	Maximum Downtime
Company business data	6 hours
Company resource data	6 hours
Project data	1 day
Personal data	2 days

This recovery time mainly consists of the time needed to access the media and the time required to actually restore data to a disk.

- How long specific types of data should be kept

Table A-7 on page A-25 shows how long data should be kept. This has implications on the amount of backup media required.

Table A-7 How Long Data Should Be Kept

Type of Data	Max Data Storage Time
Company business data	5 years
Company resource data	5 years
Project data	5 years
Personal data	3 months

- How media with backed up data should be stored and maintained

Media should be kept in the tape library in the computer room. All data included in the company backup system should be archived in full every week and incrementally every day. The data should be stored at a security company.

- Amount of data that needs to be backed up

The amount of data that currently needs to be backed up is shown in Table A-8 on page A-26:

Table A-8 **Amount of Data to Be Backed Up**

Location	Data (in GB)
ABC Cape Town	100
ABC Pretoria	22
ABC Durban	16

Plans for future growth of the amount of data

ABC plans to grow at 15 to 20% per year. The amount of data to be backed up is expected to grow accordingly. This has implications not only on the amount of time it takes to run backups and backup devices needed for backup, but also on the size of the IDB.

Table A-9 **Amount of Data to Be Backed Up in Five Years**

Location	Data (in GB)
ABC Cape Town	250
ABC Pretoria	55
ABC Durban	40

- How often data needs to be backed up

Full backups of each type of data are carried out once a week on Fridays, Saturdays, or Sundays. Level one incremental backups are carried out daily on week days. However, if a full backup is carried out on Friday, then the corresponding level one incremental backups are carried out on weekdays and then on Saturday, skipping Friday.

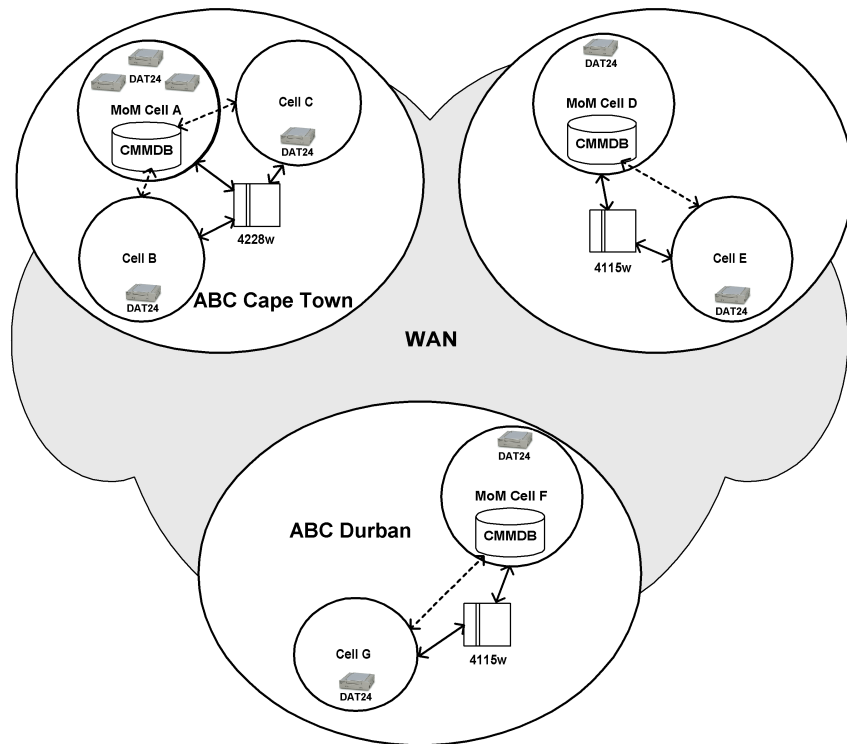
Proposed Solution

Because of the issues presented by the current backup solution, described in “Problems with Current Solution” on page A-24, ABC is undertaking a project to redesign its data storage system.

Solution Overview Each of the three departments at ABC Cape Town must be configured into a Manager-of-Managers (MoM) cell. Additionally, both ABC Durban and ABC Pretoria should be configured into MoM cells, each with two Data Protector cells.

Configure cell A as the MoM cell for the ABC Cape Town environment, cell D as the MoM cell for the ABC Pretoria environment, and cell F as the MoM cell for the ABC Durban environment. This configuration is depicted in Figure A-6 on page A-27.

Figure A-6 ABC Enterprise Environment



The Cell Managers and MoM Managers in all the 7 cells should be Windows systems. Use a Centralized Media Management Database (CMMDB) in one of the cells in each MoM environment and Catalog

Databases in each of the 7 cells. The Centralized Media Management Database allows you to share libraries between cells within each MoM environment.

Each of the three locations should have its own library. Use the HP StorageWorks DLT 4228w Library for the ABC Cape Town environment. Use HP StorageWorks DLT 4115w Libraries for ABC Pretoria and ABC Durban.

The three cells at the ABC Cape Town MoM environment should each have one SAP database server. The SAP database servers share the HP StorageWorks DLT 4228w Library. The Microsoft SQL and Microsoft Exchange databases are backed up locally to HP StorageWorks DAT24 autoloaders.

The two cells at the ABC Pretoria MoM environment should also share a Centralized Media Management Database. This should be configured on the MoM Manager of cell D to enable the sharing of the HP StorageWorks DLT 4115w Library between the cells.

The two cells at the ABC Durban MoM environment should also share a Centralized Media Management Database. This should be configured on the MoM Manager of cell F to enable the sharing of the HP StorageWorks DLT 4115w Library between the cells.

The following is a detailed account of the proposed solution:

Proposed Solution in Detail

- Cell Configuration

Configure the departments into 7 cells, of which three are at ABC Cape Town, and two each at ABC Pretoria and ABC Durban.

Why Configure into Seven Cells?

- Because ABC's departments are geographically dispersed, it would be difficult to manage them from a single cell. Moreover, there may be networking problems between the systems. The configuration also coincides with number of departments, which is an important aspect in terms of security. Each of the cells is also of the recommended size of 30 to 50 client systems. Note, however that this number depends among other things on the number of files and directories in individual client systems.

Then configure each of the three locations as a Manager-of-Managers environment. The MoM allows you to efficiently, transparently and centrally manage your cells from a single point. This then enables you to configure the Centralized Media Management Database (CMMDB) in each MoM environment.

Why Use the CMMDB?

- The Centralized Media Management Database (CMMDB) enables all cells in a MoM environment to share devices and media. Each of the three MoM environments at ABC can then use a single library, shared by client systems in all cells in the environment. Using only one very large library for all ABC’s data would not make much sense, because it would require that huge amounts of data be transferred over WAN for backup purposes.

Use a Catalog Database in each of the 7 cells. The systems in the cells would be as depicted in Table A-10 on page A-29:

Table A-10 ABC Cell Configuration

MoM Environment	Cell	#Windows Servers	#Windows Clients	#UNIX Servers	#UNIX Clients	#SAP
ABC Cape Town	A*	3	24	2	7	1
	B	2	11	5	21	1
	C	2	20	4	12	1
ABC Pretoria	D*	4	33			
	E	1	6	5	32	
ABC Durban	F*	2	10	4	30	
	G	1	11	2	29	
#SAP is the number of SAP database servers						
* represents a MoM cell						

The Cell Managers and MoM Managers in all the 7 cells should be Windows systems.

Why Choose the Windows System?

- Windows systems provide the native Unicode support and therefore require less configuration to properly handle international character in file names.

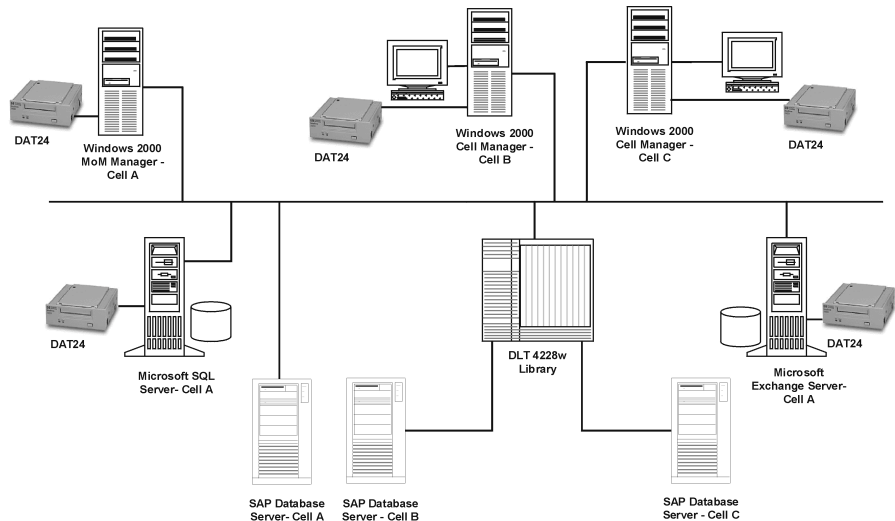
Configure cell A as the Manager-of-Managers cell of the ABC Cape Town environment and import the rest of the cells into the MoM environment. Configure a Centralized Media Management Database in MoM cell A to allow you to share the same library with cells B and C. Share the HP StorageWorks DLT 4228w Library for the ABC Cape

Backup Scenarios
Company ABC

Town environment. With a capacity of 1,1 TB in compressed format, this library should suffice for the company's projected needs in the next five years.

The three cells at ABC Cape Town should each have one SAP database server. The SAP database servers share the HP StorageWorks DLT 4228w Library. The Microsoft SQL and Microsoft Exchange databases are backed up locally to existing HP StorageWorks DAT24 autoloaders. Each of the cells in the environment should have its own Catalog Database. The configuration of the Cape Town environment is depicted in Figure A-7 on page A-30.

Figure A-7 ABC Cape Town Enterprise Backup Environment



The two cells at the ABC Pretoria MoM environment should share a Centralized Media Management Database. This should be configured on the MoM Manager of cell D. The purpose of using the CMMDB is to enable the sharing of the HP StorageWorks DLT 4115w Library between the cells. Each of the cells in the environment should have its own Catalog Database.

The two cells at the ABC Durban MoM environment should, likewise, share a Centralized Media Management Database. This should be configured on the MoM Manager of cell F. Each of the cells in the environment should also have its own Catalog Database.

Use an HP StorageWorks DLT 4115w Library for the ABC Pretoria environment and for the ABC Durban environment. With a capacity of 600 GB in compressed format, this library should suffice for the company's projected needs in the next five years in each of these environments.

Estimating the Size of the IDB

The Internal Database Capacity Planning Tool was used to estimate the size of the IDB in cell F in a year. This tool is located at:

— On the HP-UX and Solaris Cell Managers:

`/opt/omni/doc/C/IDB_capacity_planning.xls`

— On the Windows Cell Manager:

`<Data_Protector_home>\docs\IDB_capacity_planning.xls`

Input parameters shown in Figure A-8 on page A-32 include the number of files in the environment (2 million), the growth factor (1.2), data protection (260 weeks), catalog protection (3 weeks), number of full backups per week (1), and number of incremental backups per week (5).

Figure A-8 **Input Parameters**

Environment description			
Files:		2	million
Files per directory:		10	
Data volume:		16	GB
Growth factor:		1.20	
Device performance:		10.00	MB/second
Medium capacity:		70.00	GB
Objects:		100	
Change per incr.bkup:		10.00%	
Backup parameters			
Device concurrency:		2	
Data segment size:		2,048.00	MB
Log level:		All	
Data protection:		260	Weeks
Catalog protection:		3	Weeks
Full backups/week:		1	
Incr backups/week:		5	
Cell Manager parameters			
Insertion speed:		12	million/hour

The results are shown in Figure A-9 on page A-32. In one year, the database is expected to grow to approximately 667.47 MB.

Figure A-9 **Results**

Results/calculation			
Avg. file size:		7.29	KB
Files/segment:		269,057.37	
Catalog size:		16.42	MB
K devices:		2	
K performance:		85.48173913	GB/hour
K duration:		0.19	hour
Protected media:		133.7142857	
Space estimation			
MMDB:		30.00	MB
CDB:	Fnames:	207.00	MB
	Overs:	57.13	MB
	Mpos:	0.74	MB
DCBF:		372.60	MB
SMBF:		5.72	MB
Total:		667.47	MB

You can also use the Internal Database Capacity Planning Tool to estimate the size of the IDB in environments with online databases (Oracle, SAP R/3).

- Hardware

- Network

All systems in the same location should be on the same LAN for maximum performance. Use the 100TX network to connect all the systems in each of the locations and the WAN to connect the cells in the three locations. The 100TX network has a sustained data transfer rate of 10 MB/s, or 36 GB/h, of data.

- Backup Devices

The backup devices consist of an HP StorageWorks DLT 4228w Library for ABC Cape Town and two HP StorageWorks DLT 4115w Libraries for ABC Pretoria and ABC Durban as well as 7 HP StorageWorks DAT24 autoloaders for backing up the IDB and configuration files in all the cells and 2 HP StorageWorks DAT24 autoloaders for backing up the Microsoft SQL database and the Microsoft Exchange database at ABC Cape Town. The Microsoft Exchange Server and the Microsoft SQL Server currently consist of 15 GB and 11 GB of data, respectively, while the rest of the data (100 GB - 15 GB - 11 GB = 74 GB) is backed up using the three SAP database servers.

Why Use the HP StorageWorks DLT 4228w Library?

The HP StorageWorks DLT 4228w Library has two DLT4000 drive with 28 slots. It has a total compressed storage capacity of 1.1 TB and a maximum sustained data transfer rate of 6 MB/s (2 x 3 MB/s), or 21 GB/h, with data compression. This is the transfer rate assumed for the remainder of this section. Currently, the total amount of data to be backed up to the HP StorageWorks DLT 4228w Library as a full backup, whether this is a single full backup, or the staggering approach is used, is about 74 GB. Assuming that the size of an incremental backup is approximately 5% of that of a full backup, a backup generation, representing a full backup and all incremental backups based on this full backup, requires $(74 + 74 * 5\% * 5)$ GB, or **92.5 GB**, of library space. In five years time, this figure is projected to increase to about 230 GB. ABC's backup policy requires that three backup generations of data be kept. Therefore, $230 * 3$ GB, or 690 GB, of library space will be required for storage. The HP StorageWorks DLT 4228w Library's 1.1 TB GB storage capacity therefore suffices.

The library at ABC Cape Town is shared among the three cells at the location. The library at the ABC Pretoria environment is shared between cells D and E, while that at ABC Durban is shared between cells F and G. Such a configuration requires the use of the Data Protector Centralized Media Management Database in each of the three MoM environments. These databases are configured on the MoM Manager of cells A, D, and F.

Why Use the HP StorageWorks DLT 4115w Library?

The HP StorageWorks DLT 4115w Library has a single DLT4000 drive with 15 slots. It has a total compressed storage capacity of 600 GB and a maximum sustained data transfer rate of 3 MB/s, or 10.5 GB/h, with data compression. This is the transfer rate assumed for the remainder of this section. Currently, the total amount of data to be backed up at ABC Pretoria to the HP StorageWorks DLT 4115w Library as a full backup, whether this is a single full backup, or the staggering approach is used, is about 22 GB. Assuming that the size of an incremental backup is approximately 5% of that of a full backup, a backup generation, representing a full backup and all incremental backups based on this full backup, requires $(22+22*5\%*5)$ GB, or **27.5 GB**, of library space. In five years time, this figure is projected to increase to about 68.75 GB. ABC's backup policy requires that three backup generations of data be kept. Therefore, $68.75*3$ GB, or 206.25 GB, of library space will be required for storage. The HP StorageWorks DLT 4115w Library's 600 GB storage capacity therefore suffices.

HP StorageWorks DAT24 autoloaders are used to back up the Microsoft Exchange Server and Microsoft SQL Server at ABC Cape Town as well as each of the 7 Cell Managers in the 3 MoM environments.

Why Use the HP StorageWorks DAT24 Autoloader?

The HP StorageWorks DAT24 autoloader has 6 24-GB data cartridges. It has a total compressed storage capacity of 144 GB and a maximum sustained data transfer rate of 2 MB/s, or 7 GB/h, with data compression. This is the transfer rate assumed for the remainder of this section. Currently, the total amount of data to be backed up to the HP StorageWorks DAT24 autoloader connected to the aforementioned Microsoft Exchange Server at ABC Cape Town is 15 GB. Assuming that the size of an incremental backup is approximately 5% of that of a full backup, a backup generation, representing a full backup and all incremental backups based on this full backup,

requires $(15+15*5\%*5)$ GB, or **18.75 GB**, of space. In five years time, this figure is projected to increase to about 47 GB. ABC's backup policy requires that two backup generations of data be kept. Therefore, $47*2$ GB, or **94 GB**, of library space will be required for storage. The HP StorageWorks DAT24 autoloader's 144 GB storage capacity therefore suffices.

How Long Does a Full Backup Last?

The SAP database servers in the three cells at ABC Cape Town contain about 74 GB of data to be backed up to an HP StorageWorks DLT 4228w Library. This library has two drives and a sustained data transfer rate of 6 MB/s (2 x 3 MB/s), or 21 GB/h. Therefore, data is backed up to this library in **up to 5 hours**. The projected amount of data in five years, 185 GB, would be backed up in **9 to 10 hours**, which would still be within the acceptable 12 hours.

Cells D and E at ABC Pretoria share an HP StorageWorks DLT 4115w Library. This library has a single drive and a sustained data transfer rate of 3 MB/s, or 10.5 GB/h. The total amount of data to be backed up in these cells is approximately 22 GB. This would be backed up in **2 to 3 hours**. The projected amount of data in five years, 55 GB, would be backed up in **5 to 7 hours**, which would be within the acceptable 12 hours.

Similarly, the 16 GB in cells F and G at ABC Durban would be backed up in **up to 2 hours**. The projected amount of data in five years, 40 GB, would be backed up in **about 4 hours**, which would be within the acceptable 12 hours.

The largest, 1.3 GB, Data Protector Catalog Database at ABC Pretoria should be backed up in a few minutes, when no database integrity checking is performed beforehand. Data Protector by default checks the integrity of the database before the database is backed up. The check operation takes less than an hour for a 1.3 GB database. Therefore, the IDB and configuration files at ABC Pretoria should then be backed up in under **2 hours**.

- Media Pools

Media are grouped into media pools to provide better media tracking and control. Media pools facilitate the management of large numbers of media, reducing the management effort of backup administrators to a minimum. Use the organizational structure and the systems categories criteria to define the following media pools:

Table A-11 **ABC's Media Pool Usage**

Media Pool Name	Location	Description
CT_SAP_Pool	Cape Town	SAP database server
CT_SQL_Pool	Cape Town	Microsoft SQL Server
CT_Exchange_Pool	Cape Town	Microsoft Exchange Server
CT_DB_Pool	Cape Town	IDB
P_DLT_Pool	Pretoria	HP StorageWorks DLT 4115w Library
P_DAT_Pool	Pretoria	HP StorageWorks DAT24 autoloaders
P_DB_Pool	Pretoria	IDB
D_DLT_Pool	Durban	HP StorageWorks DLT 4115w Library
D_DAT_Pool	Durban	HP StorageWorks DAT24 autoloaders
D_DB_Pool	Durban	IDB

- Backup Specifications

Configure backup specifications as follows:

- DB_A...G

Backup specifications for each of the 7 IDBs and configuration files. Schedule the backup specification such that Data Protector will run a weekly full backup and a level one incremental every day, except Sundays at 03.00.

Why Use Level 1 Incremental Backups?

To restore the latest data only two media sets need to be accessed, one for the latest full backup and one for the latest level 1 incremental backup prior to the restore point-in-time. This considerably simplifies and speeds up the restore. Where differential backups are used the number of media sets may increase considerably, making the restore process more complex and slower.

Two copies of the IDB and configuration files should be made, for security reasons.

- SAP_A...C

Backup specification for the SAP database servers in cells A, B and C, respectively. Use the staggering approach to avoid network load, device load, and time window issues as depicted in Table A-12 on page A-37:

Table A-12 The Staggering Approach for ABC Cape Town

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Cell A	Inc1	Inc1	Inc1	Inc1	Full	Inc1	
Cell B	Inc1	Inc1	Inc1	Inc1	Inc1	Full	
Cell C	Inc1	Inc1	Inc1	Inc1	Inc1		Full

- SERVERS_A...G

Backup specifications for the company’s servers to prepare for disaster recovery. Each time a new server is installed, or an existing server is upgraded, this backup specification is updated. Schedule the backup specifications such that Data Protector will run full backups as shown in Table A-13 on page A-38 and level 1 incremental backups every work day.

— USERS_D...G

Backup specifications for user data. This is the main production backup at ABC Pretoria and ABC Durban. Schedule the backup specification such that Data Protector will run a weekly full backup as shown in Table A-13 on page A-38 every Friday and level 1 incremental backups every work day. However, if a full backup is carried out on Friday, then the corresponding level one incremental backups are carried out on weekdays and then on Saturday, skipping Friday.

Table A-13 on page A-38 shows the backup specification configuration in greater detail.

Table A-13 ABC's Backup Specification Configuration

Name	Cell	Description	Backup Day	Time
DB_A	A	IDB	Saturday	03:00
DB_B	B	IDB	Saturday	03:00
DB_C	C	IDB	Saturday	03:00
SQL_A	A	Microsoft SQL database	Friday	20:00
EXCHANGE_A	A	Microsoft Exchange database	Friday	20:00
SAP_A	A	SAP database server	Friday	20:00
SAP_B	B	SAP database server	Saturday	20:00
SAP_C	C	SAP database server	Sunday	20:00
SERVERS_A	A	Servers	Friday	23:00
SERVERS_B	B	Servers	Saturday	23:00
SERVERS_C	C	Servers	Sunday	23:00
DB_D	D	IDB	Saturday	03:00
DB_E	E	IDB	Saturday	03:00
SERVERS_D	D	Servers	Friday	23:00

Table A-13 ABC's Backup Specification Configuration

Name	Cell	Description	Backup Day	Time
SERVERS_E	E	Servers	Saturday	23:00
USERS_D	D	User data	Saturday	0:00
USERS_E	E	User data	Sunday	0:00
DB_F	F	IDB	Saturday	03:00
DB_G	G	IDB	Saturday	03:00
SERVERS_F	F	IDB	Friday	23:00
SERVERS_G	G	Servers	Saturday	23:00
USERS_F	F	User data	Saturday	0:00
USERS_G	G	User data	Sunday	0:00

Backup Options

Use default Data Protector backup options. Set the following options as follows:

— Log Directories

This filesystem backup option ensures that details only on directories are stored in the Catalog Database. This disables the search feature during restore and allows you to browse only directories. Use this option for backing up the two servers with more than 500 000 files each in cell D. Not using this option would result in a large increase in the size of the Data Protector Catalog Database.

— Protection

Data should be easily accessible for a period of three weeks. Since we will have one weekly full backup, we set catalog protection to 27 days (3 weeks*7 days+6 days=27 days).

Set data protection to 5 years for all backup specifications except for Exchange_A, which is sued to back up personal mail. Set data protection for this backup specification to 3 months.

- Concurrency

Set to 5 to allow up to five Disk Agents to concurrently write data to the library. This will increase backup performance.
- Media Pool

Select appropriate media pools and media to be used for backup.
- Reporting and Notification

Email notifications will be set up for backup administrators for mount requests, low database space, device errors, and on end of session events for all the backup specifications. Optionally, email or broadcast notifications will be set up for those end users interested in being notified about the success of backups of their systems.

To enable all users to easily determine the status of backup, set up client backup information on the company home page as follows:

 1. Configure a report group with a Client Backup Report for each client. The report should be logged to the file in HTML format.
 2. Schedule the report group.
 3. Link the logged files to the company home page.
- Vaulting

Vaulting is a process of storing media to a safe location for a specified period of time.

Media will be moved to the vault once a week and replaced by new media in the HP StorageWorks DLT 4228w Library, the HP StorageWorks DLT 4115w Library, and HP StorageWorks DAT24 autoloaders. All actions excluding the actual moving of media to the vault are done by the software solution including queries done internally in the database to prevent the administrator from having to find media that require ejection.

Track the location of media that are moved to a vault. This is important when you want to restore from backups on media that were moved to the vault. Data Protector allows you to perform the following vaulting tasks:

 - Generate reports showing media stored at a specific location with data protection expiring in a specified time

- Generate reports showing media used for backup within a specified time frame
- Display a list of backup specifications that have used specified media during the backup.
- Display a list of media needed for restore and the physical locations where the media are stored.
- Filter media from the media view based on specific criteria, such as media with expired protection.
- Restore
 - Restore by Query

Requests for restores by query will be sent to the administrator. If the files were last backed up less than 3 weeks before the request was placed, then the administrator can use the Restore by Query restore task to select the files and directories to be restored using a specified criteria. The administrator then selects the Overwrite option to replace files and directories on the disk with the versions on the media.
 - Complete Filesystem Restore

Requests for the restore of whole filesystems will be sent to the administrator. If the files were last backed up less than 3 weeks before the request is placed, then the administrator can select the objects for restore and use the Restore Into option.

With the Restore Into option selected, the object is restored with the exact directory structure to a selected directory. Use a Windows or UNIX utility to compare the restored object with the backed up object.
 - Restore from a Vault

To restore data from a vault, which is for instance 3 years old, send a request to the administrator who then:

 1. Identifies the media needed for restore.
 2. Brings the media from a vault, enters the media in the HP StorageWorks DLT 4228w Library, the or HP StorageWorks DLT 4115w Library or other device and then scans the media.

3. Selects the specific object to be restored using the List From Media option, if the media are not in the Data Protector Catalog Database.
4. Performs the restore.



B Further Information

In This Appendix

This appendix provides additional information about some of the aspects of Data Protector concepts, including backup generations, examples of automated media copying, and internationalization.

Backup Generations

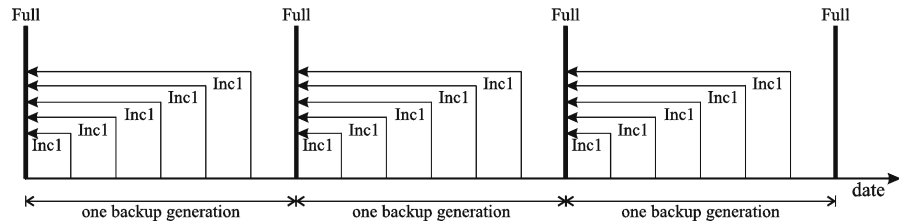
Data Protector provides a time/date related protection model. It is easy to map a generation-based backup model to the time-based model, assuming regular backups are done.

What Is a Backup Generation?

A backup generation, shown in Figure B-1 on page B-3, represents a full backup and all incremental backups based on this full backup. When the next full backup is done, a new backup generation is created.

Backup generations help you to know how many full versions of backed up data you have. For a successful point-in-time restore, you need at least one backup generation (a full backup and all incrementals to that point-in-time). Keep more than one backup generation, three for example, depending on your company policies for data protection.

Figure B-1 Backup Generations



You configure Data Protector to automatically maintain the desired number of backup generations by selecting the appropriate data and catalog protection durations, and scheduling for unattended backups, both full and incremental.

For example, to keep three backup generations while you have weekly full backups and daily leveled incremental backups, specify data protection to $7 \times 3 + 6 = 27$ days. A backup generation represents a full backup and all incremental backups until the next full backup: therefore, the six in the formula represents incremental backups before the next, fourth, backup generation belonging to the third backup generation.

Further Information
Backup Generations

You can set automatic media rotation (for the media with expired protection time) through an appropriate pool usage concept. Refer to “Implementing a Media Rotation Policy” on page 121 for more information.

Examples of Automated Media Copying

After a backup finishes, you can use the automated media copy functionality to copy the media, and then move either the originals or the copies to an off-site vault. You can use either post-backup or scheduled media copying, depending on the availability of devices.

The considerations that must be taken into account are the following:

- It is recommended to perform all backups first and then copy the media.
- During media copying, the media that are being copied are unavailable for restore.
- You can only copy the entire medium, and not specific objects.
- After the copying, the source media that are copied and the copies are marked as non-appendable, which means that you cannot append new backups to these media.
- With scheduled media copying, the necessary devices and media must be available at the scheduled time, otherwise the copy operation will be aborted.

Example 1: Automated Media Copying of Filesystem Backups

Your company has a MoM environment with two cells, each containing 150 computer systems (servers and workstations). On average, each system has 10 GB of data, which means that you have 3000 GB of data that you want to back up.

You want to have daily Incr1 backups of the data, weekly full backups, and monthly full backups for archiving purposes. The backups must be performed outside the company's working hours, which means that they can start after 5 PM and must finish before 8 AM on the next day; they can also run during weekends.

You decide to make copies of the backup media, which will remain on site for restore purposes, and to move the originals to an off-site vault for safety reasons. The media should be copied after the backups finish. To do this, you will use automated media copying.

Examples of Automated Media Copying

You use an HP StorageWorks 6/60 Tape Library with 6 LTO drives, and LTO Ultrium 1 media. Based on previous experience, you assume that the data transfer rate is about 80 GB per hour, and the average capacity of a medium is 153 GB.

After the media copy operation, the source and the target media become non-appendable. Considering this, you may want to minimize the number of media required for the backup. It is recommended to start with empty media and use their maximum capacity. You can achieve this by creating backup specifications with only one device assigned. This ensures that a new medium will be used only after the current medium is full. However, this will increase backup time compared to writing to several media in parallel.

You decide to create 4 backup specifications. To save media space, the data is divided between the backup specifications in such a way that the minimum number of media possible is used. Only one device is used for each backup.

Automated media copying is performed after the backup is completed. You can use all the available devices for the operation. This means that 3 devices will be used for source media, and 3 devices for target media.

It is assumed that the media copying will take approximately the same amount of time as the backup.

Incr1 Backup

Configuring Backups

You schedule Incr1 backups each day from Monday to Thursday at 6 PM. The data protection is set to 4 weeks. Supposing that 30% of the data changes daily, you have 900 GB of data to back up. The data is divided among backup specifications in the following way:

- BackupSpec1 (Drive 1) - 300 GB
- BackupSpec2 (Drive 2) - 300 GB
- BackupSpec3 (Drive 3) - 150 GB
- BackupSpec4 (Drive 4) - 150 GB

BackupSpec1 and BackupSpec2 require 2 media each and the backup takes approximately 4 hours. BackupSpec3 and BackupSpec4 require 1 medium each and the backup takes approximately 2 hours.

Configuring Automated Media Copying

Automated media copying of each backup starts after the backup is completed. You have 6 media to copy, and you can use all the drives in the library for the operation, as soon as the drives are available.

You can use post-backup media copying to copy the media used with BackupSpec1 and BackupSpec2, since two drives (Drive 5 and Drive 6) are free and therefore you do not need to worry about availability of the devices.

You configure post-backup media copying for BackupSpec1 and select Drive 1 as the source device and Drive 6 as the target device. You set the same data protection as original and specify the location of the media (for example, Shelf 1).

You also configure post-backup media copying for BackupSpec2 and select Drive 2 as the source device and Drive 5 as the target device. You set the same data protection as original and specify the location of the media.

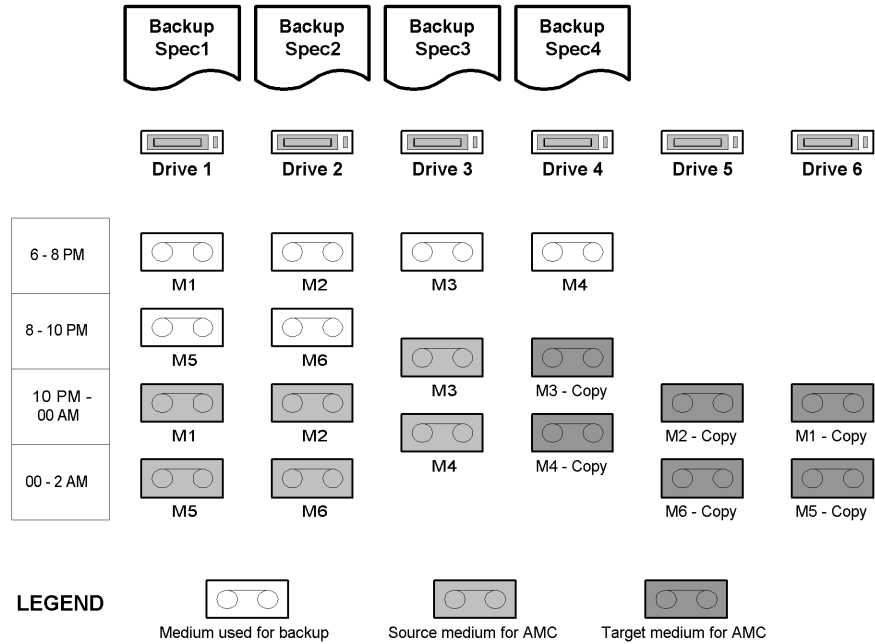
You use scheduled media copying to copy media used in BackupSpec3 and BackupSpec4, because you will be using Drive 3 and Drive 4 for the copy operation, and you have to wait until both backups finish. Note that if the devices are not available at the time the media copying is scheduled, the operation will fail. For this reason, it is recommended to add some time to the estimated backup time when scheduling an automated media copy operation that will use the same devices.

You schedule the media copy operation an hour after the backup is estimated to finish, select both BackupSpec3 and BackupSpec4 to be copied, and select Drive 3 as the source device and Drive 4 as the target device. You set the same data protection as original and specify the location of the media.

For a graphic representation of the Incr1 backup and automated media copying, refer to Figure B-2 on page B-8.

Figure B-2

Incr1 Backup and Automated Media Copying



Full Backup

Configuring Backups

You schedule your weekly full backup on Friday at 6 PM. The data protection is set to 8 weeks. You have 3000 GB of data to back up. The data is divided among backup specifications in the following way:

- BackupSpec1 (Drive 1) - 1000 GB
- BackupSpec2 (Drive 2) - 1000 GB
- BackupSpec3 (Drive 3) - 500 GB
- BackupSpec4 (Drive 4) - 500 GB

BackupSpec1 and BackupSpec2 require 7 media each, BackupSpec3 and BackupSpec4 require 4 media each. The backup is completed in approximately 14 hours.

**Configuring
Automated Media
Copying**

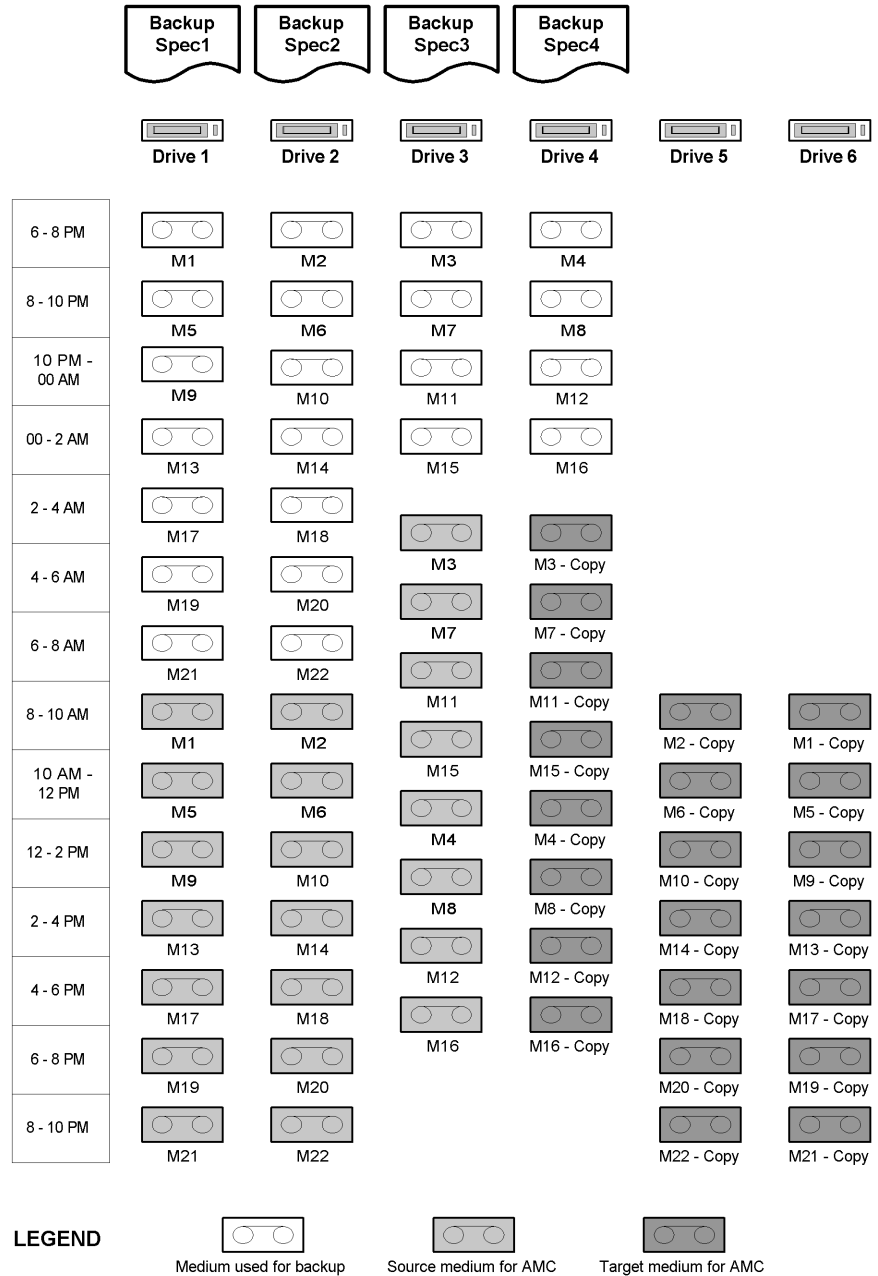
Automated media copying of each backup starts after the backup is completed. You have 22 media to copy, and all the devices are used as soon as they are available.

Again, you use post-backup media copying to copy the media used with BackupSpec1 and BackupSpec2, and scheduled media copying to copy media used in BackupSpec3 and BackupSpec4.

The devices and the data protection settings are the same as those used for the copying of the Incr1 backup. The scheduled media copying starts an hour after the backup is estimated to finish.

For a graphic representation of the full backup and automated media copying, refer to Figure B-3 on page B-10.

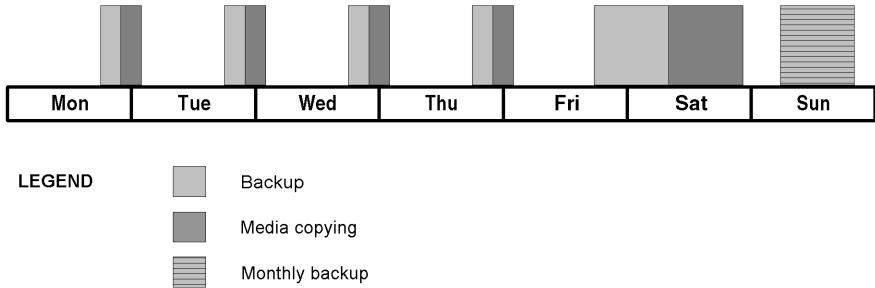
Figure B-3 Full Backup and Automated Media Copying



You schedule your monthly full backup on Sunday at 6 AM. This backup is intended for archiving purposes, so it is normally not copied.

Figure B-4 on page B-11 presents an overview of the time when the devices are busy. Note that this is a rough overview, so the graph ignores the partial overlap of some of the backup and copy sessions.

Figure B-4 Overview of Backup and Automated Media Copy Sessions



Example 2: Automated Media Copying of Oracle Database Backups

Your company has an Oracle database of the size of 500 GB. You want to perform a full backup of the database daily. The backup must be performed outside the company's working hours, which means that it can start after 5 PM and must finish before 8 AM on the next day; it can also run during weekends.

You use automated media copying to make copies of the backup media, which will remain on site for restore purposes. The originals will be moved to an off-site vault for safety reasons. The media should be copied after the backup finishes. To do this, you will use post-backup media copying.

You use an HP StorageWorks 10/700 Tape Library with 10 LTO drives, and LTO Ultrium 1 media. Based on previous experience, you assume that the data transfer rate is about 80 GB per hour, and the average capacity of a medium is 153 GB.

[Further Information](#)

Examples of Automated Media Copying

The media used for backup and media copying become non-appendable after the media copy operation, so you may want to use as much tape space as possible. On the other hand, you want the backup to finish as soon as possible. You use 4 devices for the backup. It is recommended to start with empty media and use their maximum capacity.

Automated media copying starts after the backup is completed. You have 4 media to copy, so you use 8 devices for the operation. This means that 4 devices will be used for source media, and 4 devices for target media.

It is assumed that the media copying will take approximately the same amount of time as the backup.

Full Backup

Configuring Backups

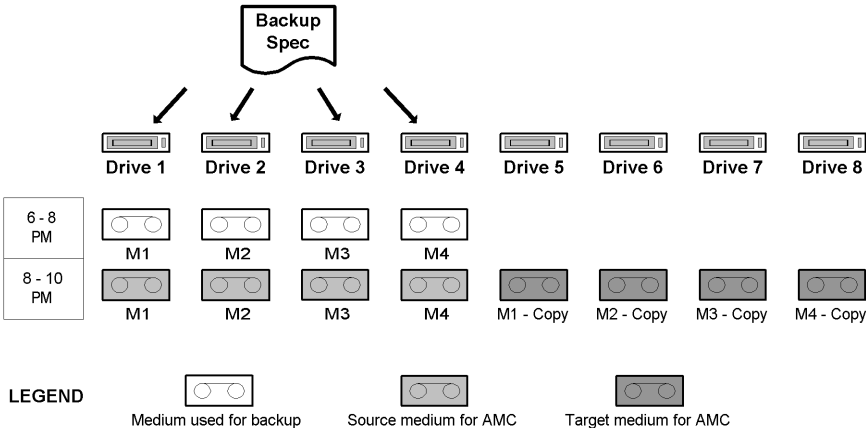
You schedule your daily full backup each day from Monday to Friday at 6 PM. The data protection is set to 4 weeks. You have 500 GB of data to back up. You use Drive 1, Drive 2, Drive 3, and Drive 4. The backup uses 4 media and is completed in approximately 2 hours.

Configuring Automated Media Copying

You use post-backup media copying because you have enough devices available. You specify Drive 1, Drive 2, Drive 3, and Drive 4 as the source devices, and Drive 5, Drive 6, Drive 7, and Drive 8 as the target devices. You set the same data protection as original and specify the location of the media.

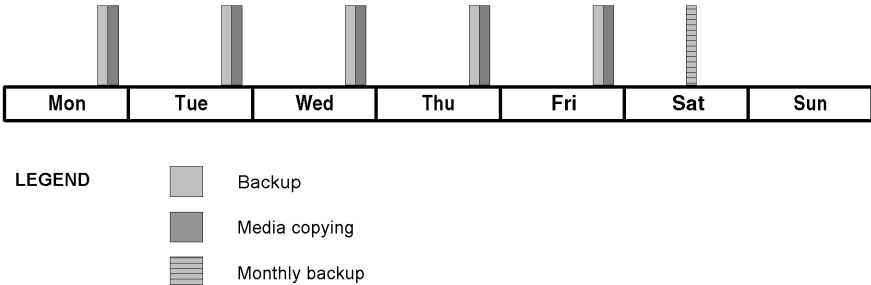
For a graphic representation of the full database backup and automated media copying, refer to Figure B-5 on page B-13.

Figure B-5 Full Database Backup and Automated Media Copying



You schedule your monthly full backup on Saturday at 12 PM. This backup is intended for archiving purposes, so it is normally not copied. Figure B-6 on page B-13 presents an overview of the time when the devices are busy.

Figure B-6 Overview of Backup and Automated Media Copy Sessions



Internationalization

Internationalization is a way to design and implement a software product so that the product interacts with the user's native language and according to the user's locale settings (currency, time, date, number, and other formats). It enables the user to enter their local language text data and correctly display it. Internationalization, as a software development methodology, enables one to implement a single-source, single-binary software that can be localized to several languages by translating the actual texts, which are kept separate from the binaries.

Internationalization is thus a localization-enabling process. Data Protector is an internationalized product that provides several native languages for the user interface.

Localization

Localization is the process of adapting a product or service to a particular language and culture. It relates to the ability to provide localized screens, online Help, error messages, manuals, etc.

Instead of sending actual message strings, Data Protector sends string IDs from agents to the Cell Manager. The Cell Manager then forwards the strings to the GUI, which then displays the messages in the correct language format. Note that file names and directory names are not indexed. They are transmitted as text strings and presented in the GUI as such. The implications of this approach are discussed in the section, “File Name Handling” on page B-14.

Data Protector is localized to various languages. Refer to the *HP OpenView Storage Data Protector Product Announcements, Software Notes, and References*, your supplier, or the local HP sales office for more information on available languages.

File Name Handling

Handling file names in a heterogeneous environment (different operating systems with different local settings, all in one cell) is a significant challenge. Data Protector handles file names under various local settings (such as language, territory, and character sets) that were in effect on the system when the file names have been created. File

names that have been backed up using some locale settings and then viewed or restored using different locale settings, require a specific setup to be displayed correctly.

Background

Different platform vendors have chosen to support different sets of languages using a variety of character set representations or character encoding standards, such as ISO 8859-1, Shift-JIS, EUC, Code Page 932, and Unicode. These encodings conflict with one another - two encodings can use the same value for two *different* characters, or use different values for the *same* character. After the creation of a file name, there is no indication which code set was used. File names passed between systems using different encodings may not display properly in the GUI.

Passing data between different platforms is not problematic if all platforms use the same character set or if they use an implementation of Unicode (UTF-16 on Windows and UTF-xx on other platforms), which accommodates all characters.

Unfortunately, the UTF-xx implementation of Unicode is not yet a standard on UNIX systems. The components of the application can be distributed on several systems and several platforms, like Windows XP Professional, Windows 2000, HP-UX, Solaris, and AIX. Data on all these platforms has to be backed up and restored. Data Protector cannot compensate for the lack of a common industry-wide representation of languages and character sets, but minimizes the impact to the user.

Example

Under certain configurations in heterogeneous environments, the file names can appear corrupted in the GUI. For example, when using Data Protector, it is possible to back up files on HP-UX where the Disk Agent is running and to view those files using the Data Protector GUI running on Windows. Unless identical code sets are used on both platforms, file names may not display properly. This is because the same character value can have a different meaning and appearance under a different coded character sets.

UNIX Incompatibility Example

Three users working on a Solaris system without Data Protector installed, each using a different character set, create files on the same filesystem outside the ASCII character range. If the users then use the `ls` command to display the files they created as well as those created by the other users, the following happens:

- each user views their own file names correctly

- each user views the file names of the other users as corrupted. The corrupted file names may even look different on the different systems.

The corrupted file names were created using a different code set than the one used to perform the `ls` command. They do not have a "tag" indicating the code set which was used for their creation. This happens on systems using native filesystem viewers, for example, `ls` in the terminal window.

File Name Handling During Backup

Data Protector reads file names using the Disk Agent (running on the respective client to be backed up) and saves an original copy to a medium. The file names are also converted to an "internal" code set and logged to the IDB, if the `log filename` option is selected for the backup.

Browsing File Names

The Data Protector GUI can be used to select the files for restore. This is done by viewing the file names in the IDB on the system where the GUI is running. Data Protector offers multiple encodings to view all file names that appear in its GUI. When a specific character encoding is selected, Data Protector uses it to display characters in filenames.

To correctly display filenames, select the same character encoding that was in effect on the system, on which the files were created. Otherwise, file names appear corrupted in the Data Protector GUI.

The correct file names can be restored to the same platform that backup was made on.

Refer to the online Help index keyword `internationalization` for a list of configurations indicating the file name browsing restrictions.

File Name Handling During Restore

Files are typically restored to the same platform as was used for backup. The process is as follows:

- the files to be restored are selected in the GUI
- Data Protector searches the tape for the specified data and restores it
- the original file names (original copies from the tape) are restored

access rights

See user rights.

ACSLs (*StorageTek specific term*)

The Automated Cartridge System Library Server (ACSLs) software that manages the Automated Cartridge System (ACS).

Active Directory (*Windows specific term*)

The directory service in a Windows network. It contains information about resources on the network and makes them accessible to users and applications. The directory services provide a consistent way to name, describe, locate, access and manage resources regardless of the physical system they reside on.

AML (*EMASS/GRAU specific term*)

Automated Mixed-Media library.

application agent

A component needed on a client to back up or restore online database integrations.

See also Disk Agent.

application system (*ZDB specific term*)

A system the application or database runs on. The application or database data is located on source volumes.

See also backup system and source volume.

archived redo log (*Oracle specific term*)

Also called offline redo log. If the Oracle database operates in the ARCHIVELOG mode, as each online redo log is filled, it is copied to one (or more) archived log destination(s). This copy is the Archived Redo Log. The presence or absence of an Archived Redo Log is determined by the mode that the database is using:

- ARCHIVELOG - The filled online redo log files are archived before they are reused. The database can be recovered from an instance and disk failure. The “hot” backup can be performed only when the database is running in this mode.
- NOARCHIVELOG - The filled online redo log files are not archived.

See also online redo log.

archive logging (*Lotus Domino Server specific term*)

Lotus Domino Server database mode where transaction log files are overwritten only after they have been backed up.

ASR Set

A collection of files stored on several diskettes, required for proper reconfiguration of the replacement disk

Glossary

(disk partitioning and logical volume configuration) and automatic recovery of the original system configuration and user data that was backed up during the full client backup.

These files are stored as an ASR archive file on the Cell Manager (in `<Data_Protector_home>\Config\Server\dr\asr` on a Windows Cell Manager or in `/etc/opt/omni/server/dr/asr/` on a UNIX Cell Manager) as well as on the backup medium. The ASR archive file is extracted to three diskettes for 32-bit Windows systems or four diskettes for 64-bit Windows systems after a disaster occurs. You need these diskettes to perform ASR.

autochanger

See **library**

autoloader

See **library**

BACKINT (*SAP R/3 specific term*)

SAP R/3 backup programs can call the Data Protector `backint` interface program via an open interface, which enables them to communicate with Data Protector software. For backup and restore, SAP R/3 programs issue orders for the Data Protector `backint` interface.

backup API

The Oracle interface between the Oracle `backup/restore` utility and the `backup/restore` media management layer. The

interface defines a set of routines to allow the reading and writing of data to the backup media, the creation, searching and removing the backup files.

backup chain

This relates to a situation where full and incremental backups are performed. Based on the level of the incremental backups used (`Incr`, `Incr 1`, `Incr 2`, and so on), simple or rather complex dependencies of incrementals to previous incrementals can exist. The backup chain are all backups, starting from the full backup plus all the dependent incrementals up to the desired point in time.

backup device

A device configured for use with Data Protector, which can write data to and read data from storage media. This can be, for example, a standalone `DDS/DAT` drive or a library.

backup generation

One backup generation includes one full backup and all incremental backups until the next full backup.

backup ID

An identifier of an integration object that equals the session ID of the backup of this object. The backup ID is preserved when an object is copied, exported, or imported.

Glossary

backup object

A backup unit that contains all items backed up from one disk volume (logical disk or mount point). The backed up items can be any number of files, directories, or the entire disk or mount point. Additionally, a backup object can be a database entity or a disk image (rawdisk).

A backup object is defined by:

- **Client name:** hostname of the Data Protector client where the backup object resides.
- **Mount point:** the access point in a directory structure (drive on Windows and mount point on UNIX) on the client where the backup object is located.
- **Description:** uniquely defines backup objects with identical client name and mount point.
- **Type:** backup object type (for example filesystem or Oracle).

backup owner

Each backup object in the IDB has an owner. The default owner of a backup is the user who starts the backup session.

backup session

A process that creates a copy of data on storage media. The activities are

specified in a backup specification or an interactive session. All clients configured in one backup specification are backed up together in one backup session using the same backup type (full or incremental). The result of a backup session is a set of media, which was written to, also called the backup or media set.

*See also **incremental backup** and **full backup**.*

backup set

A complete set of integration objects associated with a backup.

backup set (*Oracle specific term*)

A logical grouping of backed up files that are created using the RMAN backup command. A backup set is a complete set of files associated with a backup. The files can be multiplexed to increase performance. A backup set contains either datafiles or archived logs, but not both together.

backup specification

A list of objects to be backed up, together with a set of devices or drives to be used, backup options for all objects in the specification, days and time that you want backups to be performed. The objects are entire disks/volumes or parts of them such as files, directories, or even the Windows

Glossary

Registry for example. File selection lists such as include-lists and exclude-lists can be specified.

backup system (*ZDB specific term*)

A system connected to target volumes of one or multiple application systems. The backup system is typically connected to a backup device to perform the backup of the data in a replica.

See also **application system, target volume, and replica.**

backup types

See **incremental backup, differential backup, transaction backup, full backup and delta backup.**

backup view

Data Protector provides different views for backup specifications: By Type - according to the type of data available for backups/templates. Default view. By Group - according to the group to which backup specifications/templates belong. By Name - according to the name of backup specifications/templates. By Manager - if you are running MoM, you can also set the Backup view according to the Cell Manager to which backup specifications/templates belong.

BC (*EMC Symmetrix specific term*)

Business Continuance are processes that allow customers to access and manage

instant copies of EMC Symmetrix standard devices.

See also **BCV.**

BC (*HP StorageWorks Disk Array XP specific term*)

The Business Copy XP allows to maintain internal copies of HP StorageWorks Disk Array XP LDEVs for purposes such as data backup or data duplication. The copies (secondary volumes or S-VOLs) can be separated from the primary volumes (P-VOLs) and connected to a different system for various purposes, such as backup and development. For backup purposes, P-VOLs should be connected to the application system, and one of the S-VOL mirror sets to the backup system. *See also* **HP StorageWorks Disk Array XP LDEV, CA, Main Control Unit, application system, and backup system.**

BC Process (*EMC Symmetrix specific term*)

A protected storage environment solution that has defined specially configured EMC Symmetrix devices as mirrors or Business Continuance Volumes to protect data on EMC Symmetrix standard devices.

See also **BCV.**

BC VA (*HP StorageWorks Virtual Array specific term*)

Business Copy VA allows you to

Glossary

maintain internal copies of HP StorageWorks Virtual Array LUNs for data backup or data duplication within the same virtual array. The copies (child or Business Copy LUNs) can be used for various purposes, such as backup, data analysis or development. When used for backup purposes, the original (parent) LUNs are connected to the application system and the Business Copy (child) LUNs are connected to the backup system.

See also **HP StorageWorks Virtual Array LUN, application system, and backup system.**

BCV (*EMC Symmetrix specific term*) Business Continuance Volumes, or BCV devices, are dedicated SLDs that are pre-configured in the ICDA on which the business continuation operation runs. BCV devices are assigned separate SCSI addresses, differing from the addresses used by the SLDs they mirror. The BCV devices are used as splittable mirrors of the primary EMC Symmetrix SLDs that need to be protected.

See also **BC** and **BC Process.**

Boolean operators

The Boolean operators for the full text search functionality of the online Help system are AND, OR, NOT, and NEAR. Used when searching, they enable you to define your query precisely by creating a relationship between search terms. If no operator is specified in a

multi-word search, AND is used by default. For example, the query manual disaster recovery is equivalent to manual AND disaster AND recovery.

boot volume/disk/partition

A volume/disk/partition with files required for the initial step of the boot process. Microsoft terminology defines the boot volume/disk/partition as a volume/disk/partition containing the operating system files.

BRARCHIVE (*SAP R/3 specific term*)

An SAP R/3 backup tool that allows you to archive redo log files. BRARCHIVE also saves all the logs and profiles of the archiving process.

See also **SAPDBA, BRBACKUP** and **BRRESTORE.**

BRBACKUP (*SAP R/3 specific term*)

An SAP R/3 backup tool that allows an online or offline backup of the control file, of individual data files, or of all tablespaces and, if necessary, of the online redo log files.

See also **SAPDBA, BRARCHIVE** and **BRRESTORE.**

BRRESTORE (*SAP R/3 specific term*)

An SAP R/3 tool that can be used to restore files of the following type:

- Database data files, control files, and online redo log files saved with BRBACKUP

Glossary

- Redo log files archived with BRARCHIVE
- Non-database files saved with BRBACKUP

You can specify files, tablespaces, complete backups, log sequence numbers of redo log files, or the session ID of the backup.

See also **SAPDBA**, **BRBACKUP** and **BRARCHIVE**.

BSM

The Data Protector Backup Session Manager controls the backup session. This process always runs on the Cell Manager system.

CA (*HP StorageWorks Disk Array XP specific term*)

Continuous Access XP allows you to create and maintain remote copies of HP StorageWorks Disk Array XP LDEVs for purposes such as data duplication, backup, and disaster recovery. CA operations involve the main (primary) disk arrays and the remote (secondary) disk arrays. The main disk arrays contain the CA primary volumes (P-VOLs), which contain the original data and are connected to the application system. The remote disk arrays contain the CA secondary volumes (S-VOLs) connected to the backup system.

See also **BC** (*HP StorageWorks Disk*

Array XP specific term), **Main Control Unit** and **HP StorageWorks Disk Array XP LDEV**.

CAP (*StorageTek specific term*)

Cartridge Access Port is a port built into the door panel of a library. The purpose is to enter or eject media.

catalog protection

Defines how long information about backed up data (such as file names and file versions) is kept in the IDB.

See also **data protection**.

CDB

The Catalog Database is a part of the IDB that contains information about backups, object copies, restores, media management sessions, and backed up data. Depending on the selected logging level, it also contains file names and file versions. This part of the database is always local to the cell.

See also **MMDB**.

CDF file (*UNIX specific term*)

A Context Dependent File is a file consisting of several files grouped under the same pathname. The system ordinarily selects one of the files using the context of the process. This mechanism allows machine dependent executables, system data, and device files to work correctly from all hosts in a cluster while using the same pathname.

Glossary

cell

A set of systems that are under the control of a Cell Manager. The cell typically represents the systems on a site or an organizational entity, which are connected to the same LAN. Central control is available to administer the backup and restore policies and tasks.

Cell Manager

The main system in the cell where the essential Data Protector software is installed and from which all backup and restore activities are managed. The GUI used for management tasks can be located on a different system. Each cell has one Cell Manager system.

centralized licensing

Data Protector allows you to configure centralized licensing for the whole enterprise environment consisting of several cells. All Data Protector licenses are installed and kept on the Enterprise Cell Manager system. You can then allocate licenses to specific cells to suit your needs.

See also MoM.

Centralized Media Management Database (CMMDB)

See CMMDB.

channel (*Oracle specific term*)

An Oracle Recovery Manager resource allocation. Every allocated channel starts a new Oracle process, which

performs backup, restore, and recovery actions. The type of channel allocated determines the type of media used:

- type “disk”
- type ‘SBT_TAPE’

If the specified channel is type ‘SBT_TAPE’ and Oracle is integrated with Data Protector, the server process will attempt to read backups from or write data files to Data Protector.

circular logging (*Microsoft Exchange Server and Lotus Domino Server specific term*)

Circular logging is a Microsoft Exchange Server database and Lotus Domino Server database mode, in which the transaction log file contents are periodically overwritten after the corresponding data is committed to the database. Circular logging reduces disk storage space requirements.

client backup

A backup of all filesystems mounted on a client. Filesystems mounted on the client after the backup specification was created are not automatically detected.

client backup with disk discovery

A backup of all filesystems mounted on a client. When the backup starts, Data Protector discovers the disks on the clients. Client backup with disk

Glossary

discovery simplifies backup configuration and improves backup coverage of systems that often mount or dismount disks.

client or client system

Any system configured with any Data Protector functionality and configured in a cell.

cluster-aware application

It is an application that supports the cluster Application Programming Interface. Each cluster-aware application declares its own critical resources (disk volumes (on Microsoft Cluster Server), volume groups (on MC/ServiceGuard), application services, IP names and addresses ...).

CMD Script for OnLine Server

(Informix specific term)

Windows CMD script that is created in INFORMIXDIR when Informix OnLine Server is configured. The CMD script is a set of system commands that export environment variables for OnLine Server.

CMMDB

The Data Protector Centralized Media Management Database (CMMDB) is the result of merging MMDBs from several cells in the MoM environment. It allows you to share high-end devices and media across multiple cells in a MoM environment. One cell can control the

robotics, serving the devices that are connected to systems in other cells. The CMMDB must reside on the MoM Manager. A reliable network connection between the MoM cell and the other Data Protector cells is highly recommended
See also MoM.

COM+ Registration Database

(Windows specific term)

The COM+ Registration Database and the Windows Registry store COM+ application attributes, class attributes, and computer-level attributes. This guarantees consistency among these attributes and provides common operation on top of these attributes.

command-line interface

A set of DOS and UNIX like commands that you can use in shell scripts to perform Data Protector configuration, backup, restore, and management tasks.

Command View (CV) EVA

(HP StorageWorks EVA specific term)

The user interface that allows you to configure, manage, and monitor your HP StorageWorks EVA storage system. It is used to perform various storage management tasks, for example, creating virtual disk families, managing storage system hardware, creating snapclones and snapshots of virtual disks. The Command View EVA software runs on the HP OpenView

Glossary

Storage Management Appliance, and is accessed by a Web browser.

See also **HP StorageWorks EVA Agent (legacy)** and **HP StorageWorks EVA SMI-S Agent**.

concurrency

See **Disk Agent concurrency**.

control file (*Oracle and SAP R/3 specific term*)

An Oracle data file that contains entries specifying the physical structure of the database. It provides database consistency information used for recovery.

CRS

The Cell Request Server process (service) runs on the Data Protector Cell Manager. It starts and controls the backup and restore sessions. The service is started as soon as Data Protector is installed on the Cell Manager. CRS runs under the account root on UNIX systems, and under any Windows account. By default, it runs under the account of the user, specified at installation time.

CSM

The Data Protector Copy Session Manager process controls the object copy session and runs on the Cell Manager system.

data file (*Oracle and SAP R/3 specific term*)

A physical file created by Oracle that contains data structures such as tables and indexes. A data file can only belong to one Oracle database.

data protection

Defines how long the backed up data on media remains protected, that is, Data Protector will not overwrite it. When the protection expires, Data Protector will be able to reuse the media in one of the next backup sessions.

See also **catalog protection**.

Data Protector Event Log

A central repository of all Data Protector related notifications. By default, all notifications are sent to the Event Log. The Event Log is accessible only to Data Protector users in the Admin group and to Data Protector users who are granted the Reporting and notifications user rights. You can view or delete all events in the Event Log.

Data Protector user account

You can use Data Protector only if you have a Data Protector user account, which restricts unauthorized access to Data Protector and to backed up data. Data Protector administrators create this account specifying a user logon name, the systems from which the user can log on, and a Data Protector user group

Glossary

membership. This is checked whenever the user starts the Data Protector user interface or performs specific tasks.

data stream

Sequence of data transferred over the communication channel.

database library

A Data Protector set of routines that enables data transfer between Data Protector and a server of an online database integration, for example, the Oracle Server.

database parallelism

More than one database is backed up at a time if the number of available devices allows you to perform backups in parallel.

database server

A computer with a large database stored on it, such as the SAP R/3 or Microsoft SQL database. A server has a database that can be accessed by clients.

Dboject (*Informix specific term*)

An Informix physical database object. It can be a blobspace, dbspace, or logical-log file.

DC directory

The Detail Catalog (DC) directory consists of DC binary files, which store information about file versions. It represents the DCBF part of the IDB,

which occupies approximately 80% of the IDB. The default DC directory is called the dcbf directory and is located in the `<Data_Protector_home>\db40` directory on a Windows Cell Manager and in the `/var/opt/omni/server/db40` directory on a UNIX Cell Manager. You can create more DC directories and locate them as appropriate to you. Up to 10 DC directories are supported per cell. The default maximum size of a DC directory is 4 GB.

DCBF

The Detail Catalog Binary Files (DCBF) part of the IDB stores information about file versions and attributes. It occupies approximately 80% of the IDB. One DC binary file is created for each Data Protector medium used for backup.

delta backup

A delta backup is a backup containing all the changes made to the database from the last backup of any type.

See also **backup types**

device

A physical unit which contains either just a drive or a more complex unit such as a library.

device chain

A device chain consists of several standalone devices configured for sequential use. When a medium in one

Glossary

device gets full, the backup automatically continues on a medium in the next device in the device chain.

device group (*EMC Symmetrix specific term*)

A logical unit representing several EMC Symmetrix devices. A device cannot belong to more than a single device group. All devices in a device group must be on the same EMC Symmetrix unit. You can use a device group to identify and work with a subset of the available EMC Symmetrix devices.

device streaming

A device is streaming if it can feed enough data to the medium to keep it moving forward continuously. Otherwise, the tape has to be stopped, the device waits for more data, reverses the tape a little and resumes to write to the tape, and so on. In other words, if the data rate written to the tape is less or equal the data rate which can be delivered to the device by the computer system, then the device is streaming. Streaming significantly improves the performance of the device and use of space.

DHCP server

A system running the Dynamic Host Configuration Protocol (DHCP) providing dynamic configuration of IP addresses and related information.

differential backup

An incremental backup (incr) based on any previous Data Protector backup (full or any incremental), which must still be protected.

See **incremental backup**.

differential backup (*MS SQL specific term*)

A database backup that records only the data changes made to the database after the last full database backup.

See also **backup types**.

differential database backup

A differential database backup records only those data changes made to the database after the last full database backup.

direct backup

A SAN-based backup solution in which data movement directly from disk to tape (or to other secondary storage) is facilitated by the SCSI Extended Copy (Xcopy) command. Direct backup lessens the backup I/O load on systems in a SAN environment. The data movement is facilitated directly from disk to tape (or to other secondary storage) by the SCSI Extended Copy (XCOPY) command. The command is provided by any element of the infrastructure including bridges, switches, tape libraries, and disk subsystems.

See also **XCOPY engine**.

Glossary

directory junction (*Windows specific term*)

Directory junctions use the reparse point concept of Windows. An NTFS 5 directory junction allows you to redirect a directory/file request to another location.

disaster recovery

A process to restore a client's main system disk to a state close to the time when a (full) backup was performed.

Disk Agent

A component needed on a client to back it up and restore it. The Disk Agent controls reading from and writing to a disk. During a backup session, the Disk Agent reads data from a disk and sends it to the Media Agent, which then moves it to the device. During a restore session the Disk Agent receives data from the Media Agent and writes it to the disk.

Disk Agent concurrency

The number of Disk Agents that are allowed to send data to one Media Agent concurrently.

disk discovery

The detection of disks during client backup with disk discovery. During this backup, Data Protector discovers (detects) the disks that are present on the client — even though they might not have been present on the system when the backup was configured — and backs

them up. This is particularly useful in dynamic environments, where configurations change rapidly. After the disks are expanded, each inherits all options from its master client object. Even if pre- and post-exec commands are specified once, they are started many times, once per each object.

disk group (*Veritas Volume Manager specific term*)

The basic unit of data storage in VxVM system. A disk group can consist of one or more physical volumes. There can be more than one disk group on the system.

disk image (rawdisk) backup

A high-speed backup where Data Protector backs up files as bitmap images. A disk image (rawdisk) backup does not track the files and directory structure stored on the disk, but stores a disk image structure on byte level. You can perform a disk image backup of either specific disk sections or a complete disk.

disk quota

A concept to manage disk space consumption for all or a subset of users on a computer system. This concept is used by several operating system platforms.

disk staging

The process of backing up data in several phases to improve the

Glossary

performance of backups and restores, reduce costs of storing the backed up data, and increase the data availability and accessibility for restore. The backup stages consist of backing up data to one media type first (for example disk) and later copying it to a different media type (for example tape).

Distributed File System (DFS)

A service that connects file shares into a single namespace. The file shares can reside on the same or on different computers. DFS provides client access to the resources in a location-transparent manner.

DMZ

The Demilitarized Zone (DMZ) is a network inserted as a "neutral zone" between a company's private network (intranet) and the outside public network (Internet). It prevents outside users from getting direct access to company servers in the intranet.

DNS server

In the DNS client-server model, this is the server containing information about a portion of the DNS database that makes computer names available to client resolvers querying for name resolution across the Internet.

domain controller

A server in a network that is responsible for user security and verifying passwords within a group of other servers.

DR image

Data required for temporary disaster recovery operating system (DR OS) installation and configuration.

DR OS

A disaster recovery operating system is an operating system environment in which disaster recovery runs. It provides Data Protector a basic runtime environment (disk, network, tape, and filesystem access). The OS has to be installed and configured before the Data Protector disaster recovery can be performed. DR OS not only hosts the Data Protector disaster recovery process but is also a part of the restored system because it replaces its own configuration data with the original configuration data.

drive

A physical unit that receives data from a computer system and can write it onto a magnetic medium (typically a tape drive). It can also read the data from the medium and send it to the computer system.

Glossary

drive index

A number that identifies the mechanical position of a drive inside a library device. This number is used by the robotic control to access a drive.

dynamic client

See **client backup with disk discovery**.

EMC Symmetrix Agent (SYMA)

(EMC Symmetrix specific term)

See **Symmetrix Agent (SYMA)**

emergency boot file *(Informix specific term)*

An Informix configuration file that resides in the <INFORMIXDIR>\etc directory (on HP-UX) or <INFORMIXDIR>/etc directory (on Windows) and is called ixbar.<server_id>, where <INFORMIXDIR> is the OnLine Server home directory and <server_id> is the value of the SERVERNUM configuration parameter. Each line of the emergency boot file corresponds to one backup object.

Enterprise Backup Environment

Several cells can be grouped together and managed from a central cell. The enterprise backup environment includes all clients located in several Data Protector cells which are managed and

administered from a central cell using the Manager-of-Managers concept. See also **MoM**.

Event Logs

Files in which Windows logs all events, such as the starting or stopping of services and the logging on and off of users. Data Protector can back up Windows Event Logs as part of the Windows configuration backup.

exchanger

Also referred to as SCSI Exchanger. See also **library**.

exporting media

A process that removes all data about backup sessions, such as systems, objects, and file names, which reside on the media from the IDB. Information about the media and their relation to a pool is also removed from the IDB. The data on the media remains unchanged. See also **importing media**.

Extensible Storage Engine (ESE)

(Microsoft Exchange Server specific term)

A database technology used as a storage system for information exchange in Microsoft Exchange Server.

failover

Transferring of the most important cluster data, called group (on Windows)

Glossary

or package (on Unix) from one cluster node to another. A failover can occur mostly because of software or hardware failures or maintenance on the primary node.

FC bridge

See **Fibre Channel bridge**

Fibre Channel

An ANSI standard for high-speed computer interconnection. Using either optical or copper cables, it allows the high speed bidirectional transmission of large data files and can be deployed between sites kilometers apart. Fibre Channel connects nodes using three physical topologies: point-to-point, loop, and switched.

Fibre Channel bridge

A Fibre Channel bridge or multiplexer provides the ability to migrate existing parallel SCSI devices, like RAID arrays, solid state disks (SSD), and tape libraries to a Fibre Channel environment. On one side of the bridge or multiplexer there is a Fibre Channel interface while on the other side there are parallel SCSI ports. The bridge or multiplexer enables SCSI packets to be moved between the Fibre Channel and parallel SCSI devices.

file depot

A file containing the data from a backup to a file library device.

file jukebox device

A device residing on disk consisting of multiple slots used to store file media.

file library device

A device which resides on a disk emulating a library with several media, hence containing multiple files, referred to as file depots.

File Replication Service (FRS)

A Windows service that replicates the domain controller store logon scripts and group policies. FRS also enables replication of Distributed File System (DFS) shares between systems and allows any server to perform replication activity.

file version

The same file can be backed up multiple times in case of full backups and incremental backups (if the file changed). If the log level ALL is selected for backup, Data Protector retains one entry in the IDB for the filename itself and one for each version (date/time) of the file.

filesystem

The organization of files on a hard disk. A filesystem is backed up so that the file attributes and the file contents are stored on the backup media.

Glossary

first level mirror (*HP StorageWorks Disk Array XP specific term*)

HP StorageWorks Disk Array XP allows up to three mirror copies of a Primary Volume and each of these copies can have additional two copies. The three mirror copies are called first level mirrors.

See also **Primary Volume**, and **MU numbers**.

fnames.dat

The fnames.dat files of the IDB contain information on the names of the backed up files. Typically, these files occupy about 20% of the IDB, if filenames are stored.

formatting

A process that erases any data contained on a medium and prepares it for use with Data Protector. Information about media (media ID, description, and location) is saved in the IDB as well as on the respective media (media header). Data Protector media with protected data are not formatted until the protection expires or the media are unprotected/ recycled.

free pool

An auxiliary source of media for use by media pools when they run out of media. The media pools must be configured to use free pools.

full backup

A backup in which all selected objects are backed up, whether or not they have been recently modified.

See also **backup types**.

full database backup

A backup of all data in a database, not only the data that has been changed after the last (full or incremental) backup of the database. A full database backup does not depend on any other backup.

full mailbox backup

A full mailbox backup is a backup of the entire mailbox content.

full ZDB

A ZDB backup in which all selected objects are backed up, even if there are no changes from the previous backup.

See also **incremental ZDB**.

global options file

A file that allows you to customize Data Protector. It explains the global options, which cover various aspects of Data Protector, typically time-outs and limits, and affect the entire Data Protector cell. The file is located in the /etc/opt/omni/server/options directory on HP-UX and Solaris systems and in the <Data_Protector_home>\Config\Server\Options directory on Windows systems.

Glossary

group (*Microsoft Cluster Server specific term*)

A collection of resources (for example disk volumes, application services, IP names and addresses) that are needed to run a specific cluster-aware applications.

GUI

A cross-platform (HP-UX, Solaris, and Windows) graphical user interface, provided by Data Protector for easy access to all configuration, administration, and operation tasks.

hard recovery (*Microsoft Exchange Server specific term*)

A Microsoft Exchange Server database recovery that is performed after a restore by the database engine, using transaction log files.

heartbeat

A cluster data set with a time stamp carrying information about the operational status of a particular cluster node. This data set or packet is distributed among all cluster nodes.

Hierarchical Storage Management (HSM)

A method for optimizing the use of expensive hard disk storage by migrating less frequently used data to less expensive optical platters. When needed, the data is migrated back to

hard disk storage. This balances the need for fast retrieval from hard disk with the lower cost of optical platters.

Holidays file

A file that contains information about holidays. You can set different holidays by editing the Holidays file: `/etc/opt/omni/server/Holidays` on the UNIX Cell Manager and `<Data_Protector_home>\Config\Server\holidays` on the Windows Cell Manager.

host backup

See **client backup with disk discovery**.

hosting system

A working Data Protector client used for Disk Delivery Disaster Recovery with a Data Protector Disk Agent installed.

HP ITO

See **OVO**.

HP OpC

See **OVO**.

HP OpenView SMART Plug-In (SPI)

A fully integrated, out-of-the-box solution which "plugs into" HP OpenView Operations, extending the managed domain. Through the Data Protector integration, which is implemented as an HP OpenView SMART Plug-In, a user can have an

Glossary

arbitrary number of Data Protector Cell Managers monitored as an extension to HP OpenView Operations (OVO).

HP OVO

See **OVO**.

HP StorageWorks Disk Array XP LDEV

A logical partition of a physical disk within an HP StorageWorks Disk Array XP. LDEVs are entities that can be replicated in the Continuous Access XP (CA) and Business Copy XP (BC) configurations, or can be used as standalone entities.

See also **BC** (*HP StorageWorks Disk Array XP specific term*), **CA** (*HP StorageWorks Disk Array XP specific term*), and **replica**.

HP StorageWorks EVA Agent (legacy)

A Data Protector software module that executes all tasks required for the HP StorageWorks Enterprise Virtual Array integration operating on HP StorageWorks EVA with Command View (CV) EVA software v3.1 or lower, and the EVA VCS firmware v3.01x or lower.

See also **Command View (CV) EVA** and **HP StorageWorks EVA SMI-S Agent**.

HP StorageWorks EVA SMI-S Agent

A Data Protector software module that executes all tasks required for the HP StorageWorks Enterprise Virtual Array integration operating on HP StorageWorks EVA with Command View (CV) EVA software starting with v3.2. With the EVA SMI-S Agent, the control over the array is established through HP StorageWorks SMI-S EVA provider, which directs communication between incoming requests and CV EVA.

See also **Command View (CV) EVA**, **HP StorageWorks SMI-S EVA provider**, and **HP StorageWorks EVA Agent (legacy)**.

HP StorageWorks SMI-S EVA provider

An interface used for controlling HP StorageWorks Enterprise Virtual Array. SMI-S EVA provider runs as a separate service on the HP OpenView Storage Management Appliance system and acts as a gateway between incoming requests and Command View EVA. With the Data Protector HP StorageWorks EVA integration, SMI-S EVA provider accepts standardized requests from the EVA SMI-S Agent, communicates with Command View EVA for information or method invocation, and returns standardized responses.

See also **HP StorageWorks EVA SMI-**

Glossary

S Agent and Command View (CV) EVA.

HP StorageWorks Virtual Array LUN

A logical partition of a physical disk within an HP StorageWorks Virtual Array. LUNs are entities that can be replicated in the HP StorageWorks Business Copy VA configuration, or can be used as standalone entities.
See also **BC VA** and **replica**.

HP VPO
See **OVO**.

ICDA (*EMC Symmetrix specific term*)
EMC's Symmetrix Integrated Cached Disk Arrays (ICDA) is a disk array device that combines a set of physical disks, a number of FWD SCSI channels, an internal cache memory, and control and diagnostic software commonly referred to as the microcode.

IDB

The Data Protector Internal Database is an embedded database located on the Cell Manager that keeps information regarding which data is backed up, on which media it is backed up, how backup and restore sessions are run, and which devices and libraries are configured.

importing media

A process that re-reads all data about backup sessions which are on the medium back into the IDB. This then allows for fast and convenient access to the data on the media.
See also **exporting media**.

incremental backup

A backup that selects only files that have changed since a previous backup. Several levels of incremental backup are available, allowing selective backup of only files that have changed since the last incremental backup.
See also **backup types**.

incremental backup (*Microsoft Exchange Server specific term*)

A backup of the Microsoft Exchange Server data that has changed since the last full or incremental backup. With the incremental backup, only the transaction log files are backed up.
See also **backup types**.

incremental mailbox backup

An incremental mailbox backup backs up all the changes made to the mailbox after the last backup of any type.

incremental1 mailbox backup

An incremental1 mailbox backup backs up all the changes made to the mailbox after the last full backup.

Glossary

incremental (re)-establish (*EMC Symmetrix specific term*)

A BCV or SRDF control operation. In BCV control operations, an incremental establish causes the BCV device to be synchronized incrementally and to function as an EMC Symmetrix mirrored medium. The EMC Symmetrix devices must have been previously paired.

In SRDF control operations, an incremental establish causes the target (R2) device to be synchronized incrementally and to function as an EMC Symmetrix mirrored medium. The EMC Symmetrix devices must have been previously paired.

incremental restore (*EMC Symmetrix specific term*)

A BCV or SRDF control operation. In BCV control operations, an incremental restore reassigns a BCV device as the next available mirror of the standard device in the pair. However, the standard devices are updated with only the data that was written to the BCV device during the time of the original pair split, and the data that was written to the standard device during the split is overwritten with data from the BCV mirror. In SRDF control operations, an incremental restore reassigns a target (R2) device as the next available mirror of the source (R1) device in the pair. However, the source (R1) devices are updated with only the data that was

written to the target (R2) device during the time of the original pair split, and the data that was written to the source (R1) device during the split is overwritten with data from the target (R2) mirror.

incremental ZDB

A ZDB to tape or ZDB to disk+tape session in which only changes from the last full or incremental protected backup are streamed to tape.

See also **full ZDB**.

Inet

A process that runs on each UNIX system or service that runs on each Windows system in the Data Protector cell. It is responsible for communication between systems in the cell and for starting other processes needed for backup and restore. The Inet service is started as soon as Data Protector is installed on a system. The Inet process is started by the inetd daemon.

Information Store (*Microsoft Exchange Server specific term*)

The Microsoft Exchange Server service that is responsible for storage management. Information Store in Microsoft Exchange Server manages two kinds of stores: mailbox stores and public folder stores. A mailbox store consists of mailboxes that belong to individual users. A public folder store contains public folders and messages

Glossary

that are shared among several users.
See also **Key Management Service** and **Site Replication Service**.

initializing
See **formatting**.

Installation Server

A computer system that holds a repository of the Data Protector software packages for a specific architecture. The Installation Server is used for remote installation of Data Protector clients. In mixed environments at least two Installation Servers are needed: one for UNIX systems and one for Windows systems.

instant recovery (*ZDB specific term*)
A process in which a replica, produced by a ZDB-to-disk or a ZDB-to-disk+tape sessions, is used to restore the contents of the source volumes to their states at the time at which the replica was created, avoiding the need to perform a restore from tape. Depending on the application/database concerned, this may be all that is required, or other steps, such as the application of transaction log files, may be required for full recovery.
See also **replica, zero downtime backup (ZDB), ZDB to disk, and ZDB to disk+tape**.

integrated security (*MS SQL specific term*)
Integrated security allows the Microsoft SQL Server to use Windows authentication mechanisms to validate Microsoft SQL Server logins for all connections. Using integrated security means that users have one password for both Windows and Microsoft SQL Server. Integrated security should be used in environments where all clients support trusted connections. Connections validated by Windows Server and accepted by Microsoft SQL Server are referred to as trusted connections. Only trusted connections are allowed.

integration object
A backup object of a Data Protector integration, such as Oracle or SAP DB.

Internet Information Server (IIS)

(*Windows specific term*)
Microsoft Internet Information Server is a network file and application server that supports multiple protocols. Primarily, IIS transmits information in Hypertext Markup Language (HTML) pages by using the Hypertext Transport Protocol (HTTP).

IP address
Internet Protocol address is a numeric address of a system used to uniquely identify the system on the network. The

Glossary

IP address consists of four groups of numbers separated by periods (full stops).

ISQL (*Sybase specific term*)

A Sybase utility used to perform system administration tasks on Sybase SQL Server.

ITO

See **OVO**.

jukebox

See **library**.

jukebox device

A device consisting of multiple slots used to store either optical or file media. When being used to store file media, the jukebox device is known as the “file jukebox device”.

Key Management Service (*Microsoft Exchange Server specific term*)

The Microsoft Exchange Server service that provides encryption functionality for enhanced security.

See also **Information Store** and **Site Replication Service**.

LBO (*EMC Symmetrix specific term*)

A Logical Backup Object (LBO) is an object of data storage/retrieval in the EMC Symmetrix environment. It is stored/retrieved by EMC Symmetrix as one entity and can only be restored as a whole.

library

Also called autochanger, jukebox, autoloader, or exchanger. A library contains media in repository slots. Each slot holds one medium (for example, DDS/DAT). Media are moved between slots and drives by a robotic mechanism, allowing random access to media. The library can contain multiple drives.

lights-out operation or **unattended operation**

A backup or restore operation that takes place outside of normal business hours without an operator. This implies that no operator personnel is present to work with the backup application or service mount requests, for example.

LISTENER.ORA (*Oracle specific term*)

An Oracle configuration file that describes one or more Transparent Network Substrate (TNS) listeners on a server.

load balancing

By default, Data Protector automatically balances the usage of devices selected for backup, so that they are used evenly. Load balancing optimizes the device usage by balancing the number of objects written to each device. Since load balancing is done automatically during backup time, you do not need to manage how the data is actually backed up. You just specify the devices to be

Glossary

used. If you do not want to use load balancing, you can select which device will be used for each object in the backup specification. Data Protector will access the devices in the specified order.

local and remote recovery

Remote recovery is performed if all Media Agent hosts specified in the SRD file are accessible. If any of them fails, the disaster recovery process fails over to the local mode. This means that the target system is searched for locally attached devices. If only one device is found, it is automatically used. Otherwise, Data Protector prompts you to select the device, which will be used for restore.

lock name

You can configure the same physical device several times with different characteristics, by using different device names.

The lock name is a user specified string that is used for locking all such device configurations to prevent collision if several such devices (device names) are used concurrently. Use an identical lock name for all device definitions which use the same physical device.

log_full shell script (*Informix UNIX specific term*)

A script provided by ON-Bar that you

can use to start backing up logical-log files when OnLine Server issues a log-full event alarm. The Informix ALARMPROGRAM configuration parameter defaults to the `<INFORMIXDIR>/etc/log_full.sh`, where `<INFORMIXDIR>` is the OnLine Server home directory. If you do not want logical logs to be backed up continuously, set the ALARMPROGRAM configuration parameter to `<INFORMIXDIR>/etc/no_log.sh`.

logging level

The logging level determines the amount of details on files and directories written to the IDB during backup or object copying. You can always restore your data, regardless of the logging level used during backup. Data Protector provides four logging levels: Log All, Log Directories, Log Files, and No Log. The different logging level settings influence the IDB growth, backup speed, and the convenience of browsing data for restore.

logical-log files

This applies to online database backup. Logical-log files are files in which modified data is first stored before being flushed to disk. In the event of a failure, these logical-log files are used to roll forward all transactions that have been

Glossary

committed as well as roll back any transactions that have not been committed.

login ID (*MS SQL Server specific term*)

The name a user uses to log on to Microsoft SQL Server. A login ID is valid if Microsoft SQL Server has an entry for that user in the system table syslogin.

login information to the Oracle Target Database (*Oracle and SAP R/3 specific term*)

The format of the login information is <user_name>/<password>@<service>, where:

- <user_name> is the name by which a user is known to Oracle Server and to other users. Every user name is associated with a password and both have to be entered to connect to an Oracle Target Database. This user must have been granted Oracle SYSDBA or SYSOPER rights.
- <password> is a string used for data security and known only to its owner. Passwords are entered to connect to an operating system or software application. The password has to be the same as the password specified in the Oracle password file (orapwd), which is used for authentication of users performing database administration.

- <service> is the name used to identify an SQL*Net server process for the target database.

login information to the Recovery Catalog Database (*Oracle specific term*)

The format of the login information to the Recovery (Oracle) Catalog Database is <user_name>/<password>@<service>, where the description of the user name, password, and service name is the same as in the Oracle SQL*Net V2 login information to the Oracle target database. In this case, <service> is the name of the service to the Recovery Catalog Database, not the Oracle target database.

Note that the Oracle user specified here has to be the owner of the Oracle Recovery (Oracle) Catalog.

Lotus C API (*Lotus Domino Server specific term*)

An interface for the exchange of backup and recovery information between Lotus Domino Server and a backup solution, like Data Protector.

LVM

A Logical Volume Manager is a subsystem for structuring and mapping physical disk space to logical volumes on UNIX systems. An LVM system

Glossary

consists of several volume groups, where each volume group has several volumes.

Magic Packet

See **Wake ONLAN**.

mailbox (*Microsoft Exchange Server specific term*)

The location to which e-mail is delivered, which is set up by the administrator for each user. If a set of personal folders is designated as the e-mail delivery location, e-mail is routed from the mailbox to this location.

Mailbox Store (*Microsoft Exchange Server specific term*)

A part of the Information Store that maintains information about user mailboxes. A mailbox store consists of a binary rich-text .edb file and a streaming native internet content .stm file.

Main Control Unit (MCU) (*HP StorageWorks Disk Array XP specific term*)

An HP StorageWorks XP disk array that contains the primary volumes for the Continuous Access configuration and acts as a master device.

See also **BC** (*HP StorageWorks Disk Array XP specific term*), **CA** (*HP StorageWorks Disk Array XP specific term*), and **HP StorageWorks Disk Array XP LDEV**.

Manager-of-Managers (MoM)

See **Enterprise Cell Manager**.

Media Agent

A process that controls reading from and writing to a device, which reads from or writes to a medium (typically a tape).

During a backup session, a Media Agent receives data from the Disk Agent and sends it to the device for writing it to the medium. During a restore session, a Media Agent locates data on the backup medium and sends it to the Disk Agent. The Disk Agent then writes the data to the disk. A Media Agent also manages the robotics control of a library.

MAPI (*Microsoft Exchange specific term*)

The MAPI (Messaging Application Programming Interface) is the programming interface that lets applications and messaging clients interact with messaging and information systems.

media allocation policy

Determines in which sequence media are used for backup. The Strict allocation policy directs Data Protector to prompt for a specific medium. The Loose policy directs Data Protector to prompt for any suitable medium. The Formatted First policy directs Data Protector to give preference to unknown media, even if unprotected media are available in the library.

Glossary

media condition

The quality of a medium as derived from the media condition factors. Heavy usage and age result in an increased number of read and write errors with tape media. Media need to be replaced when they are marked as POOR.

media condition factors

The user-assigned age threshold and overwrite threshold used to determine the state of a medium.

media ID

A unique identifier assigned to a medium by Data Protector.

media label

A user-defined identifier used to describe a medium.

media location

A user-defined physical location of a medium, such as "building 4" or "off-site storage".

media management session

A session performing some action on a medium, such as initializing, scanning the content, verifying data on a medium, or copying a medium.

media pool

A set of media of the same type (such as DDS) used and tracked as a group. Media are formatted and assigned to a media pool.

media set

The result of a backup session is data backed up on a group of media called media set. Depending on the media usage policy, several sessions can share the same media.

media type

The physical type of media, such as DDS or DLT.

media usage policy

The media usage policy controls how new backups are added to the already used media. It can be Appendable, Non-Appendable, or Appendable for incrementals only.

merging

This defines one mode to resolve file conflicts during restore. If the file to be restored already exists at the destination, the one with the more recent modification date is kept. Files not present on the disk are always restored. *See also* **overwrite**.

MFS

The Migrating File System enables a standard JFS filesystem with migration capabilities (on HP-UX 11.00). The MFS is accessed via a standard filesystem interface (DMAPI), it is mounted to a directory the same way as any HP-UX filesystem. In an MFS, only the superblock, the inode and the 'extended attribute' information remain

Glossary

permanently on the hard disk and are never migrated.

See also **VBFS**.

Microsoft Exchange Server

A “client-server” messaging and a workgroup system that offers a transparent connection to many different communication systems. It provides users with an electronic mail system, individual and group scheduling, online forms, and workflow automation tools. It provides a developer with a platform on which to build custom information-sharing and messaging-service applications.

Microsoft Management Console (MMC) *(Windows specific term)*

An administration model for Windows-based environments. It provides a simple, consistent, and integrated administration user interface allowing management of many applications through the same GUI, provided that the applications adhere to the MMC model.

Microsoft SQL Server

A database management system designed to meet the requirements of distributed “client-server” computing.

Microsoft Volume Shadow Copy service (VSS)

A software service that provides a unified communication interface to coordinate backup and restore of a VSS-

aware application regardless of its specific features. This service collaborates with the backup application, writers, shadow copy providers, and the operating system kernel to implement the management of volume shadow copies and shadow copy sets.

See also **shadow copy, shadow copy provider, writer**.

mirror *(EMC Symmetrix and HP StorageWorks Disk Array XP specific term)*

See **target volume**.

mirror rotation *(HP StorageWorks Disk Array XP specific term)*

See **replica set rotation**.

MMD

The Media Management Daemon process (service) runs on the Data Protector Cell Manager and controls media management and device operations. The process is started when Data Protector is installed on the Cell Manager.

MMDB

The Media Management Database (MMDB) is a part of the IDB that contains information about media, media pools, devices, libraries, library drives, and slots configured in the cell, as well as the Data Protector media used for backup. In an enterprise backup

Glossary

environment, this part of the database can be common to all cells.
See also **CMMDB, CDB.**

MoM

Several cells can be grouped together and managed from a central cell. The management system of the central cell is the Manager-of-Managers (MoM). The MoM allows you to configure and manage multiple cells from a central point.

mount request

A screen prompt that tells you to insert a specific medium into a device. Once you respond to the mount request by providing the required medium and confirm the mount request, the session continues.

mount point

The access point in a directory structure for a disk or logical volume, for example /opt or d:. On UNIX, the mount points are displayed using the bdf or df command.

MSM

The Data Protector Media Session Manager, which runs on the Cell Manager and controls media sessions, such as copying media.

MU number (*HP StorageWorks Disk Array XP specific term*)

A Mirror Unit number is an integer

number (0, 1 or 2), used to indicate a first level mirror.
See also **first level mirror.**

multi-drive server

A license that allows you to run an unlimited number of Media Agents on a single system. This license, which is bound to the IP address of the Cell Manager, is no longer available.

obdrindex.dat

An IDB file with information about IDB backups, media, and devices used for the backup. This information can significantly simplify IDB recovery. It is recommended to relocate the file, together with IDB transaction logs, on a separate physical disk from other IDB directories, and, additionally, to make a copy of the file and locate it where you want.

OBDR capable device

A device that can emulate a CD-ROM drive loaded with a bootable disk and can thus be used as a backup or boot device for disaster recovery purposes.

object

See **backup object**

object copy

A copy of a specific object version that is created during an object copy session or a backup session with object mirroring.

Glossary

object copy session

A process that creates an additional copy of the backed up data on a different media set. During an object copy session, the selected backed up objects are copied from the source to the target media.

object copying

The process of copying selected object versions to a specific media set. You can select object versions from one or several backup sessions to be copied.

Object ID (*Windows specific term*)

The object IDs (OIDs) enable access to NTFS 5 files no matter where in the system the files reside. Data Protector treats the OIDs as alternate streams of the files.

object mirror

A copy of a backup object created using object mirroring. Object mirrors are often referred to as object copies.

object mirroring

The process of writing the same data to several media sets during a backup session. Data Protector enables you to mirror all or some backup objects to one or more media sets.

offline backup

A backup during which an application database cannot be used by the application.

- For simple backup methods (non ZDB), the database is generally put into a quiescent state that allows use by the backup system, but not the application, for the whole backup period (~minutes/hours). For instance, for backup to tape, until streaming of data to the tape is finished.
- For ZDB methods, the database is also put into the quiescent state, but for the period of the data replication process only (~seconds). Normal database operation can then be resumed for the rest of the backup process.

*See also **zero downtime backup (ZDB)** and **online backup**.*

offline recovery

Offline recovery is performed if the Cell Manager is not accessible, for example, due to network problems. Only standalone and SCSI library devices can be used for offline recovery. Recovery of the Cell Manager is always offline.

offline redo log

*See **archived redo log***

OmniStorage

Software providing transparent migration of less frequently used data to the optical library while keeping more

Glossary

frequently used data on the hard disk. HP OmniStorage runs on HP-UX systems.

On-Bar (*Informix specific term*)

A backup and restore system for OnLine Server. ON-Bar enables you to create a copy of your OnLine Server data and later restore the data. The ON-Bar backup and restore system involves the following components:

- onbar utility
- Data Protector, as the backup solution
- XBSA interface
- ON-Bar catalog tables, which are used to back up dobjects and track instances of dobjects through multiple backups.

onbar utility (*Informix specific term*)

The Informix utility that communicates backup and restore requests to OnLine Server. The utility uses XBSA to exchange control data and back up and restore data with Data Protector.

ONCONFIG (*Informix specific term*)

An environment variable that specifies the name of the active ONCONFIG configuration file. If the ONCONFIG environment variable is not present, OnLine uses the configuration values

from the file `<INFORMIXDIR>/etc/onconfig` (on HP-UX) or `<INFORMIXDIR>\etc\onconfig` (on Windows).

online backup

A backup performed while a database application remains available for use. The database is placed into a special backup mode of operation for the time period that the backup application requires access to the original data objects. During this period, the database is fully operational, but there may be a small performance impact and log files may grow very quickly.

- For simple backup methods (non ZDB), backup mode is required for the whole backup period (~minutes/ hours). For instance, for backup to tape, until streaming of data to tape is finished.
- For ZDB methods, backup mode is required for the short period of the data replication process only (~seconds). Normal database operation can then be resumed for the rest of the backup process.

In some cases, transaction logs may also have to be backed up to allow a consistent database to be restored. *See also* **zero downtime backup (ZDB)** and **offline backup**.

Glossary

online redo log (*Oracle specific term*)

Redo logs that have not been archived, but are either available to the instance for recording database activity or are filled and waiting to be archived or reused.

See also **archived redo log**.

OnLine Server (*Informix specific term*)

Refers to INFORMIX-OnLine Dynamic Server.

OpC

See **OVO**.

Oracle instance (*Oracle specific term*)

Each installation of an Oracle database on one or more systems. One computer system can have several instances of a database running.

ORACLE_SID (*Oracle specific term*)

A unique name for an Oracle Server instance. To switch among Oracle Servers, specify the desired `<ORACLE_SID>`. The `<ORACLE_SID>` is included in the `CONNECT DATA` parts of the connect descriptor in a `TNSNAMES.ORA` file and in the definition of the TNS listener in the `LISTENER.ORA` file.

original system

The system configuration backed up by Data Protector before a computer disaster hits the system.

overwrite

An option that defines one mode to resolve file conflicts during restore. All files are restored from a backup even if they are older than existing files.

See also **merging**.

OVO

HP OpenView Operations for Unix provides powerful capabilities for operations management of a large number of systems and applications on in a network. Data Protector provides an integration into this management product. This integration is implemented as a SMART Plug-In for OVO management servers on HP-UX and Solaris. Earlier versions of OVO were called IT/Operation, Operations Center and Vantage Point Operations.

See also **merging**.

ownership

The ownership of a backup determines who can restore from the backup. The user who starts an interactive backup is the session owner. If a user starts an existing backup specification without modifying it, the session is not considered interactive. In that case, if the backup owner has been defined in the backup specification, they remain the session owner. Otherwise, the session owner becomes the user who started the backup in question. For the scheduled backups, by default, the session owner is for the UNIX Cell

Glossary

Manager: root.sys@<Cell Manager>, and for the Windows Cell Manager, the user that was specified during the installation of the Cell Manager. It is possible to modify the ownership, so that the specific user becomes the session owner.

P1S file

P1S file contains information on how to format and partition all disks installed in the system during Enhanced Automated Disaster Recovery (EADR). It is created during full backup and is saved on backup medium and on Cell Manager into

<*Data_Protector_home*>\Config\Server\dr\p1s directory on a Windows Cell Manager or in /etc/opt/omni/server/dr/p1s directory on a UNIX Cell Manager with the filename recovery.p1s.

package (*MC/ServiceGuard and Veritas Cluster specific term*)

A collection of resources (for example volume groups, application services, IP names and addresses) that are needed to run a specific cluster-aware application.

pair status (*HP StorageWorks Disk Array XP specific term*)

A mirrored pair of disks can have various status values depending on the action performed on it. The three most important status values are:

- **COPY** - The mirrored pair is currently resynchronizing. Data is transferred from one disk to the other. The disks do not contain the same data.
- **PAIR** - The mirrored pair is completely synchronized and both disks (the primary volume and the mirrored volume) contain identical data.
- **SUSPENDED** - The link between the mirrored disks is suspended. That means that both disks are accessed and updated independently. However, the mirror relationship is still maintained and the pair can be resynchronized without transferring the complete disk.

parallel restore

Restoring backed up data to multiple disks at the same time (that is, in parallel) by running multiple Disk Agents, that receive data from one Media Agent. For the parallel restore to work, select data that is located on different disks or logical volumes and during backup, the data from the different objects must have been sent to the same device using a concurrency of 2 or more. During a parallel restore, the data for multiple objects selected for restore is read from media at the same time, thereby improving performance.

Glossary

parallelism

The concept of reading multiple data streams from an online database.

physical device

A physical unit that contains either a drive or a more complex unit such as a library.

post-exec

A backup option that executes a command or script after the backup of an object or after the entire session completes. Post-exec commands are not supplied by Data Protector. You need to create your own. They can be written as executables or batch files on Windows and as shell scripts on UNIX.

See also **pre-exec**.

pre- and post-exec commands

Pre- and post-exec commands are used to perform additional action before and after a backup or restore session. They are not supplied by Data Protector. You need to create your own commands. They can be written as executables or batch files on Windows and as shell scripts on UNIX.

prealloc list

A subset of media in a media pool that specifies the order in which media are used for backup.

pre-exec

A backup option that executes a command or script before the backup of an object or before the entire session is started. Pre-exec commands are not supplied by Data Protector. You need to create your own. They can be written as executables or batch files on Windows and as shell scripts on UNIX.

See also **post-exec**.

Primary Volume (P-VOL) (*HP*

StorageWorks Disk Array XP specific term)

Standard HP StorageWorks Disk Array XP LDEVs that act as a primary volume for the CA and BC configurations. The P-VOL is located in the MCU.

See also **Secondary Volume (S-VOL)**.

protection

See **data protection** and also **catalog protection**.

public folder store (*Microsoft Exchange Server specific term*)

The part of the Information Store that maintains information in public folders. A public folder store consists of a binary rich-text .edb file and a streaming native internet content .stm file.

public/private backed up data

When configuring a backup, you can select whether the backed up data will be:

Glossary

- public, that is visible (and accessible for restore) to all Data Protector users
- private, that is, visible (and accessible for restore) only to the owner of the backup and administrators

RAID

Redundant Array of Inexpensive Disks.

RAID Manager Library (*HP StorageWorks Disk Array XP specific term*)

The RAID Manager Library is used internally by Data Protector on Solaris systems to allow access to HP StorageWorks Disk Array XP configuration, status, and performance data and to key HP StorageWorks Disk Array XP features through the use of function calls translated into a sequence of low level SCSI commands.

RAID Manager XP (*HP StorageWorks Disk Array XP specific term*)

The RAID Manager XP application provides an extensive list of commands to report and control the status of the CA and BC applications. The commands communicate through a RAID Manager instance with the HP StorageWorks Disk Array XP Disk Control Unit. This instance translates the commands into a sequence of low level SCSI commands.

rawdisk backup

See disk image backup.

RCU (*HP StorageWorks specific term*)

The Remote Control Unit acts as a slave of an MCU in a CA configuration. In bidirectional configurations, the RCU can act as an MCU.

RDBMS

Relational Database Management System.

RDF1/RDF2 (*EMC Symmetrix specific term*)

A type of SRDF device group. Only RDF devices can be assigned to an RDF group. An RDF1 group type contains source (R1) devices and an RDF2 group type contains target (R2) devices.

RDS

The Raima Database Server process (service) runs on the Data Protector Cell Manager and manages the IDB. The process is started when Data Protector is installed on the Cell Manager.

Recovery Catalog (*Oracle specific term*)

A set of Oracle tables and views that are used by Recovery Manager to store information about Oracle databases. This information is used by Recovery Manager to manage the backup, restore,

Glossary

and recovery of Oracle databases. The recovery catalog contains information about:

- The physical schema of the Oracle target database
- Data file and archived log backup sets
- Data file copies
- Archived Redo Logs
- Stored scripts.

Recovery Catalog Database (*Oracle specific term*)

An Oracle database that contains a recovery catalog schema. You should not store the recovery catalog in your target database.

RecoveryInfo

When backing up Windows configuration files, Data Protector collects the information about the current system configuration (information on disk layout, volume, and network configuration). This information is needed for disaster recovery.

Recovery Manager (RMAN) (*Oracle specific term*)

An Oracle command-line interface that directs an Oracle Server process to back

up, restore, or recover the database it is connected to. RMAN uses either the recovery catalog or the control file to store information about backups. This information can be used later in restore sessions.

recycle

A process that removes the data protection from all backed up data on a medium, allowing Data Protector to overwrite it during one of the next backups. Data that belongs to the same session(s) but resides on other media is also unprotected. Recycling does not actually alter the data on the medium.

redo log (*Oracle specific term*)

Every Oracle database has a set of two or more redo log files. The set of redo log files for the database is known as the database's redo log. Oracle uses the redo log to record all changes made to data.

Remote Control Unit (RCU)

(HP StorageWorks Disk Array XP specific term)

The Remote Control Unit (RCU) acts as a slave of an MCU in a CA configuration. In bidirectional configurations, the RCU can act as an MCU.

Removable Storage Management Database (*Windows specific term*)

A Windows service used for managing removable media (such as tapes and

Glossary

disks) and storage devices (libraries). Removable Storage allows applications to access and share the same media resources.

reparse point (*Windows specific term*)

A system-controlled attribute that can be associated with any directory or file. The value of a reparse attribute can have user-controlled data. The format of the data is understood by the application that stored the data and a filesystem filter that was installed to interpret the data and process such files. Whenever the filesystem encounters a file with a reparse point, it attempts to find the filesystem filter associated with the data format.

replica (*ZDB specific term*)

An image, at a particular point in time, of the data in source volumes that contain user-specified backup objects. Depending on the hardware/software with which it is created, the image may be an independent exact duplicate (clone) of the storage blocks at the physical disk level (for example, a split mirror), or a virtual copy (for example, a snapshot). From a host's perspective, on a basic UNIX or Windows system, the complete physical disk containing a backup object is replicated. However, if a volume manager is used on UNIX, the whole volume/disk group containing a

backup object is replicated.

See also **snapshot**, **snapshot creation**, **split mirror**, and **split mirror creation**.

replica set (*ZDB specific term*)

A group of replicas, all created using the same backup specification.

See also **replica** and **replica set rotation**.

replica set rotation (*ZDB specific term*)

The use of a replica set for regular backup production: Each time the same backup specification requiring the use of a replica set is run, a new replica is created and added to the set, until the maximum number for the set is reached. After that, the oldest replica in the set is replaced, maintaining the maximum number of replicas in the set.

See also **replica** and **replica set**.

restore session

A process that copies data from backup media to a client.

RMAN (*Oracle specific term*)

See **Recovery Manager**.

RSM

The Data Protector Restore Session Manager controls the restore session. This process always runs on the Cell Manager system.

Glossary

RSM (*Windows specific term*)

Removable Storage Manager (RSM) includes a media management service that facilitates communication among applications, robotic changers, and media libraries. It enables multiple applications to share local robotic media libraries and tape or disk drives and to manage removable media.

SAPDBA (*SAP R/3 specific term*)

An SAP R/3 user interface that integrates the BRBACKUP, BRARCHIVE, and BRRESTORE tools.

scan

A function that identifies the media in a device. This synchronizes the MMDB with the media that are actually present at the selected locations (for example, slots in a library).

scanning

A function which identifies the media in a device. This synchronizes the MMDB with the media that are actually present at the selected locations (for example, slots in a library). It is useful to perform a scan and check the actual media in the device if someone has manually manipulated media without using Data Protector to eject or enter, for example.

Scheduler

A function that controls when and how often automatic backups occur. By setting up a schedule, you automate the start of backups.

Secondary Volume (S-VOL) (*HP StorageWorks Disk Array XP specific term*)

Secondary Volumes, or S-VOLs, are XP LDEVs that act as a secondary CA or BC mirror of another LDEV (a P-VOL). In the case of CA, S-VOLs can be used as failover devices in a MetroCluster configuration. The S-VOLs are assigned separate SCSI addresses, different from the addresses used by the P-VOLs. *See also* **Primary Volume (P-VOL)**.

session

See **backup session**, **media management session**, and **restore session**.

session ID

An identifier of a backup, restore, object copy, or media management session, consisting of the date when the session ran and a unique number.

session key

This environment variable for the Pre- and Post-exec script is a Data Protector unique identification of any session, including preview sessions. The session key is not recorded in the database, and

Glossary

it is used for specifying options for the omnimnt, omnistat and omniabort CLI commands.

shadow copy (*MS VSS specific term*)

A volume that represents a duplicate of the original volume at a certain point in time. The data is then backed up from the shadow copy and not from the original volume. The original volume continues to change as the backup process continues, but the shadow copy of the volume remains constant. |

See also **Microsoft Volume Shadow Copy service**.

shadow copy provider (*MS VSS specific term*)

An entity that performs the work on creating and representing the volume shadow copies. Providers own the shadow copy data and expose the shadow copies. Providers can be software (for example, system providers) or hardware (local disks, disk arrays).

See also **shadow copy**.

shadow copy set (*MS VSS specific term*)

A collection of shadow copies created at the same point in time.

See also **shadow copy**.

shared disks

A Windows disk on another system that has been made available to other users

on the network. Systems with shared disks can be backed up without a Data Protector Disk Agent installed.

SIBF

The Serverless Integrations Binary Files (SIBF) is a part of the IDB that stores raw NDMP meta data. This data is necessary to perform restore of NDMP objects.

Site Replication Service (*Microsoft Exchange Server specific term*)

The Microsoft Exchange Server 2000/2003 service that permits compatibility with Microsoft Exchange Server 5.5 by emulating the Exchange Server 5.5 directory service.

See also **Information Store** and **Key Management Service**.

slot

A mechanical position in a library. Each slot can hold a medium, such as a DLT tape. Data Protector references each slot with a number. To read a medium, a robotic mechanism moves the medium from a slot into the drive.

SMB

See **split mirror backup**.

SMBF

The Session Messages Binary Files (SMBF) part of the IDB stores session messages generated during backup, object copy, restore, and media

Glossary

management sessions. One binary file is created per session. The files are grouped by year and month.

snapshot (*HP StorageWorks VA and HP StorageWorks EVA specific term*)

A form of replica produced using snapshot creation techniques. A range of snapshot types is available, with different characteristics, depending on the arrays/techniques used. Such replicas are dynamic and may be either virtual copies, still reliant upon the contents of the source volumes, or independent exact duplicates (clones), depending on the snapshot type and the time since creation.

See also **replica** and **snapshot creation**.

snapshot backup (*HP StorageWorks VA and HP StorageWorks EVA specific term*)

See **ZDB to tape**, **ZDB to disk**, and **ZDB to disk+tape**.

snapshot creation (*HP StorageWorks VA and HP StorageWorks EVA specific term*)

A replica creation technique, in which copies of source volumes are created using storage virtualization techniques. The replicas are considered to be created at one particular point-in-time, without pre-configuration, and are immediately available for use. However background

copying processes normally continue after creation.

See also **snapshot**.

source (R1) device (*EMC Symmetrix specific term*)

An EMC Symmetrix device that participates in SRDF operations with a target (R2) device. All writes to this device are mirrored to a target (R2) device in a remote EMC Symmetrix unit. An R1 device must be assigned to an RDF1 group type.

See also **target (R2) device**.

source volume (*ZDB specific term*)

A storage volume containing data to be replicated.

sparse file A file that contains data with portions of empty blocks. Examples are: -A matrix in which some or much of the data contains zeros -files from image applications -high-speed databases If sparse file processing is not enabled during restore, it might be impossible to restore this file.

split mirror (*EMC Symmetrix and HP StorageWorks Disk Array XP specific term*)

A replica created using split mirror techniques. Such a replica provides an independent, exact duplicate, or clone, of the contents of the source volumes.

See also **replica** and **split mirror creation**.

Glossary

split mirror backup (*EMC Symmetrix specific term*)

See **ZDB to tape**.

split mirror backup (*HP StorageWorks Disk Array XP specific term*)

See **ZDB to tape**, **ZDB to disk**, and **ZDB to disk+tape**.

split mirror creation (*EMC Symmetrix and HP StorageWorks Disk Array XP specific term*)

A replica creation technique, in which a pre-configured set of target volumes (a mirror) is kept synchronized with a set of source volumes until the time at which a replica of the contents of the source volumes is required. Then, the synchronization is stopped (the mirror is split) and a split mirror replica of the source volumes at the time of the split remains in the target volumes.

See also **split mirror**.

split mirror restore (*EMC Symmetrix and HP StorageWorks Disk Array XP specific term*)

A process in which data backed up in a ZDB-to-tape or a ZDB-to-disk+tape session is restored from tape media to a split mirror replica, which is then synchronized to the source volumes. Individual backup objects or complete sessions can be restored using this method.

See also **ZDB to tape**, **ZDB to disk+tape**, and **replica**.

sqlhosts file (*Informix specific term*)

An Informix connectivity-information file that contains the names of each of the database servers and any aliases to which the clients on a host computer can connect.

SRD file

The Data Protector System Recovery Data (SRD) file contains system information required for installing and configuring the operating system in case of a disaster. The SRD file is an ASCII file, generated when a CONFIGURATION backup is performed on a Windows client and stored on the Cell Manager.

SRDF (*EMC Symmetrix specific term*)

The EMC Symmetrix Remote Data Facility is a business continuation process that enables effective, real-time data replication of SLDs between dislocated processing environments. These environments could be situated within the same root computer environment or separated by long distances.

SSE Agent (*HP StorageWorks Disk Array XP specific term*)

A Data Protector software module that executes all tasks required for a split mirror backup integration. It communicates with the HP StorageWorks Disk Array XP storing system using the RAID Manager XP

Glossary

utility (HP-UX and Windows systems) or RAID Manager Library (Solaris systems).

sst.conf file

The file `/usr/kernel/drv/sst.conf` is required on each Data Protector Sun Solaris client to which a multi-drive library device is connected. It must contain an entry for the SCSI address of the robotic mechanism of each library device connected to the client.

st.conf file

The file `/kernel/drv/st.conf` is required on each Data Protector Solaris client with a backup device connected. It must contain device information and a SCSI address for each backup drive connected to that client. A single SCSI entry is required for a single-drive device and multiple SCSI entries are required for a multi-drive library device.

stackers

Devices with multiple slots for media storage usually with only one drive. A stacker selects media from the stack sequentially. In contrast, a library can randomly select media from its repository.

standalone file device

A file device is a file in a specified directory to which you back up data.

standard security (*MS SQL specific term*)

Standard security uses the login validation process of the Microsoft SQL Server for all connections. Standard security is useful in network environments with a variety of clients, some of which may not support trusted connections. It also provides backward compatibility for older versions of the Microsoft SQL Server.

See also **integrated security**.

Storage Group

(*Microsoft Exchange Server specific term*)

A collection of databases (stores) that share a common set of transaction log files. Exchange manages each storage group with a separate server process.

StorageTek ACS library

(*StorageTek specific term*)

Automated Cartridge System is a library system (also known as Silo) consisting of one Library Management Unit (LMU) and one to 24 Library Storage Modules (LSM) connected to the unit.

storage volume (*ZDB specific term*)

A storage volume represents an object that may be presented to an operating system or some other entity (for example, a virtualization mechanism) upon which volume management systems, file systems, or other objects may exist. The volume management

Glossary

systems, file systems are built on this storage. Typically, these can be created or exist within a storage system such as a disk array.

switchover

See failover

Sybase Backup Server API (*Sybase specific term*)

An industry-standard interface developed for the exchange of backup and recovery information between a Sybase SQL Server and a backup solution like Data Protector.

Sybase SQL Server (*Sybase specific term*)

The server in the Sybase “client-server” architecture. Sybase SQL Server manages multiple databases and multiple users, keeps track of the actual location of data on disks, maintains mapping of logical data description to physical data storage, and maintains data and procedure caches in memory.

Symmetrix Agent (SYMA) (*EMC Symmetrix specific term*)

The Data Protector software module that prepares the EMC Symmetrix environment for backup and restore operations.

System Backup to Tape (*Oracle specific term*)

An Oracle interface that handles the

actions required to load, label, and unload correct backup devices when Oracle issues a backup or restore request.

system databases (*Sybase specific term*)

The four system databases on a newly installed Sybase SQL Server are the:

- master database (master)
- temporary database (tempdb)
- system procedure database (sybsystemprocs)
- model database (model).

system disk

A system disk is a disk containing operating system files. Microsoft terminology defines the system disk as a disk containing the files required for initial step of boot process.

system partition

A system partition is a partition containing operating system files. Microsoft terminology defines a system partition as a partition containing the files required for initial step of boot process.

System State (*Windows specific term*)

The System State data comprises the Registry, COM+ Class Registration

Glossary

database, system startup files, and the Certificate Services database (if the server is a certificate server). If the server is a domain controller, Active Directory directory services and the Sysvol directory are also contained in the System State data. If the server is running the Cluster service, the System State data also includes resource registry checkpoints and the quorum resource recovery log, which contains the most recent cluster database information.

system volume/disk/partition

A volume/disk/partition containing operating system files. Microsoft terminology defines the system volume/disk/partition as the volume/disk/partition containing files required for the initial step of the boot process.

SysVol (*Windows specific term*)

A shared directory that stores the server copy of the domain's public files, which are replicated among all domain controllers in the domain.

tablespace

A part of a database structure. Each database is logically divided into one or more tablespaces. Each tablespace has data files or raw volumes exclusively associated with it.

tapeless backup (*ZDB specific term*)

See ZDB to disk.

target database (*Oracle specific term*)

In RMAN, the target database is the database that you are backing up or restoring.

target (R2) device (*EMC Symmetrix specific term*)

An EMC Symmetrix device that participates in SRDF operations with a source (R1) device. It resides in the remote EMC Symmetrix unit. It is paired with a source (R1) device in the local EMC Symmetrix unit and receives all write data from its mirrored pair. This device is not accessed by user applications during normal I/O operations. An R2 device must be assigned to an RDF2 group type. *See also source (R1) device*

target system (*Disaster Recovery specific term*)

A system after a computer disaster has occurred. The target system is typically in a non-bootable state and the goal of disaster recovery is to restore this system to the original system configuration. The difference between a crashed system and a target system is that a target system has all faulty hardware replaced.

target volume (*ZDB specific term*)

A storage volume to which data is replicated.

Glossary

Terminal Services (*Windows specific term*)

Windows Terminal Services provide a multi-session environment that allows clients to access a virtual Windows desktop session and Windows-based programs running on the server.

thread (*MS SQL Server specific term*)

An executable entity that belongs to only one process. It comprises a program counter, a user-mode stack, a kernel-mode stack, and a set of register values. Several threads can run at a time within one process.

TimeFinder (*EMC Symmetrix specific term*)

A business continuation process that creates an instant copy of single or multiple Symmetrix Logical Devices (SLDs). The instant copy is created on specially preconfigured SLDs called BCVs and is accessible via a separate device address to the system(s).

TLU

Tape Library Unit.

TNSNAMES.ORA (*Oracle and SAP R/3 specific term*)

A network configuration file that contains connect descriptors mapped to service names. The file may be maintained centrally or locally, for use by all or individual clients.

transaction

A mechanism for ensuring that a set of actions is treated as a single unit of work. Databases use transactions to keep track of database changes.

transaction backup

Transaction backups generally use fewer resources than database backups, so they can be created more frequently than database backups. By applying transaction backups, you can recover the database to a specific point in time prior to when a problem occurred.

transaction backup (*Sybase and SQL specific term*)

A backup of the transaction log providing a record of changes made since the last full or transaction backup.

transaction log backup

Transaction log backups generally use fewer resources than database backups so they can be created more frequently than database backups. By applying transaction log backups, you can recover the database to a specific point in time.

transaction log files

Files that record transactions of the database modifications, and provide fault tolerance in case of a database disaster.

Glossary

transaction logs (*Data Protector specific term*)

Keeps track of IDB changes. The archiving of transaction logs should be enabled to prevent you from losing the transaction log files that are created after the last IDB backup and are necessary for IDB recovery.

transaction log table (*Sybase specific term*)

A system table in which all changes to the database are automatically recorded.

transportable snapshot (*MS VSS specific term*)

A shadow copy that is created on the application system and can be presented to the backup system which performs the backup.

See also **Microsoft Volume Shadow Copy service (VSS)**.

TSANDS.CFG file (*Novell NetWare specific term*)

A file that allows you to specify the names of containers where you want backups to begin. It is text file located in the SYS:SYSTEM\TSA directory on the server where TSANDS.NLM is loaded.

unattended operation

See **lights-out operation**.

user account

You can use Data Protector only if you have a Data Protector user account,

which restricts unauthorized access to Data Protector and to backed up data. Data Protector administrators create this account specifying a user logon name, the systems from which the user can log on, and a Data Protector user group membership. This is checked whenever the user starts the Data Protector user interface or performs specific tasks.

user disk quotas

NTFS quota-management support enables an enhanced tracking mechanism and control over disk space usage on shared storage volumes. Data Protector backs up user disk quotas on the whole system and for all configured users at a time.

user group

Each Data Protector user is member of a User Group. Each User Group has a set of user rights that are given to every user in that User Group. The number of User Groups with their associated user rights can be defined as desired. Data Protector provides three default user groups: admin, operator, and user.

user profile (*Windows specific term*)

Configuration information retained on a user basis. This information includes desktop settings, screen colors, network connections, and so on. When the user logs on, the user profile is loaded and the Windows environment is set accordingly.

Glossary

user rights

User rights or access rights are the permissions needed to perform specific Data Protector tasks. Configuring a backup, starting a backup session, or starting a restore session are typical user rights. Users have the access rights of the user group to which they belong.

vaulting media

The process of storing media to a safe and remote place. The media are brought back to the data center when they are needed for restore or are ready for reuse in subsequent backups. The vaulting procedure depends on your company's backup strategy and policies for data protection/reliability.

VBFS (*OmniStorage specific term*)

A Very Big File System is an extension of the standard HP-UX file system on HP-UX 9.x. It is mounted to a directory the same way as any HP-UX file system. In a VBFS, only the superblock, the inode and the 'extended attribute' information remain permanently on the hard disk and are never migrated.
See also MFS.

verify

A function that lets you check whether the Data Protector data on a specified medium is readable. Additionally, consistency within each block can be

checked if the backup was performed with the cyclic redundancy check (CRC) option ON.

Virtual Controller Software (VCS)

(HP StorageWorks EVA specific term)

The firmware that manages all aspects of storage system operation, including communication with Command View EVA through the HSV controllers.

See also Command View (CV) EVA.

Virtual Device Interface (*MS SQL Server specific term*)

This is a SQL Server programming interface that allows fast backup and restore of large databases.

virtual disk (*HP StorageWorks EVA specific term*)

A unit of storage allocated from an HP StorageWorks Enterprise Virtual Array storage pool. Virtual disks are the entities that are replicated using the HP StorageWorks Enterprise Virtual Array snapshot functionality.

See also source volume and target volume.

virtual server

A virtual machine in a cluster environment defined in a domain by a network IP name and address. Its address is cached by the cluster software and mapped to the cluster node that is currently running the virtual server

Glossary

resources. This way all requests for a particular virtual server are cached by a specific cluster node.

volser (*ADIC and STK specific term*)

A VOLume SERial number is a label on the medium to identify the physical tape used in very large libraries. A volser is a naming convention specific to ADIC/ GRAU and StorageTek devices.

volume group

A unit of data storage in an LVM system. A volume group can consist of one or more physical volumes. There can be more than one volume group on the system.

volume mountpoint (*Windows specific term*)

An empty directory on a volume that can be used to mount another volume. The volume mount point acts as a gateway to the target volume. Provided that the volume is mounted, users and applications can refer to the data on the mounted volume by the full (merged) filesystem path as if both volumes are one integral part.

Volume Shadow Copy service

See **Microsoft Volume Shadow Copy service**.

VPO

See **OVO**.

VSS

See **Microsoft Volume Shadow Copy service**.

VxFS

Veritas Journal Filesystem.

VxVM (Veritas Volume Manager)

A Veritas Volume Manager is a system for managing disk space on Solaris platforms. A VxVM system consists of an arbitrary group of one or more physical volumes organized into logical disk groups.

Wake ONLAN

Remote power-up support for systems running in power-save mode from some other system on the same LAN.

Web reporting

The Data Protector functionality that allows you to view reports on backup status and Data Protector configuration using the Web interface.

wildcard character

A keyboard character that can be used to represent one or many characters. The asterisk (*), for example, typically represents one or more characters, and the question mark (?) typically represents a single character. Wildcard characters are often used in operating systems as a means of specifying more than one file by name.

Glossary

Windows CONFIGURATION backup

Data Protector allows you to back up Windows CONFIGURATION, including Windows Registry, user profiles, Event Logs, and WINS and DHCP server data (if configured on a system) in one step.

Windows Registry

A centralized database used by Windows to store configuration information for the operating system and the installed applications.

WINS server A system running Windows Internet Name Service software that resolves Windows networking computer names to IP addresses. Data Protector can back up WINS server data as part of the Windows configuration.

writer

(MS VSS specific term)

A process that initiates change of data on the original volume. Writers are typically applications or system services that write persistent information on a volume. Writers also participate in the shadow copy synchronization process by assuring data consistency.

XBSA interface *(Informix specific term)*

The onbar utility and Data Protector communicate with each other through

the X/Open Backup Specification Services Programmer's Interface (XBSA).

XCOPY engine *(direct backup specific term)*

A SCSI-3 copy command that allows you to copy data from a storage device having a SCSI source address to a backup device having a SCSI destination address, thus enabling direct backup. The data flows from a source device (either block or streaming, that is, disk or tape) to the destination device (either block or streaming) through XCOPY. This releases the controlling server of reading the data from the storage device into memory and then writing the information to the destination device.

See also **direct backup**.

ZDB

See **zero downtime backup (ZDB)**.

ZDB database *(ZDB specific term)*

A part of the IDB, storing ZDB related information such as source volumes, replicas and security information. The ZDB database is used for ZDB, instant recovery, and split mirror restore.

See also **zero downtime backup (ZDB)**.

ZDB to disk *(ZDB specific term)*

A form of zero downtime backup where the replica produced is kept on the disk

Glossary

array as a backup of the source volumes at a specific point in time. Multiple replicas, produced using the same backup specification at different times, can be kept in a replica set. A replica from ZDB to disk can be restored using the instant recovery process.

See also **zero downtime backup (ZDB), ZDB to tape, ZDB to disk+tape, instant recovery, and replica set rotation.**

ZDB to disk+tape (*ZDB specific term*)

A form of zero downtime backup where the replica produced is kept on the disk array as a backup of the source volumes at a specific point in time, in the same way as with ZDB to disk. However, data in the replica is also streamed to a backup medium, as with ZDB to tape. If this backup method is used, data backed up in the same session can be restored using the instant recovery process, the standard Data Protector restore from tape, or on split mirror arrays, split mirror restore.

See also **zero downtime backup (ZDB), ZDB to disk, ZDB to tape, instant recovery, replica, and replica set rotation.**

ZDB to tape (*ZDB specific term*)

A form of zero downtime backup where data in the replica produced is streamed to a backup medium, typically tape. Instant recovery is not possible from such a backup, so the replica need not be

retained on the disk array after backup completion. The backed-up data can be restored using standard Data Protector restore from tape. On split mirror arrays, split mirror restore can also be used.

See also **zero downtime backup (ZDB), ZDB to disk, instant recovery, ZDB to disk+tape, and replica.**

zero downtime backup (ZDB)

A backup approach in which data replication techniques, provided by a disk array, are used to minimize the impact of backup operations on an application system. A replica of the data to be backed up is created first. All subsequent backup operations are performed on the replicated data rather than the original data, while the application system can be returned to normal operation.

See also **ZDB to disk, ZDB to tape, ZDB to disk+tape, and instant recovery.**

Glossary

A

- adding data to media during backups, 127
- ADIC (EMASS/GRAU) AML, 143
- admin user group, 167
- alarms, 195
- alternative disaster recovery methods, 107
 - operating system vendors, 107
 - third-party tools, 107
- ANSI X3.27 labels, 124
- any-to-any connectivity, 152
- Application Agents, 11
- application client
 - snapshot backup, 287
 - split mirror backup, 274
- Application Response Measurement, 194, 195
 - real time alerts, 195
 - response time, 195
 - transactions, 195
- architecture
 - backup devices, 10
 - Cell Managers, 10
 - cells, 10
- archive log backup
 - snapshot backup, 287
 - split mirror backup, 274
- ARM 2.0, 195
- autoloaders, 143
 - See also* libraries
- automated media copying, 39
 - examples, B-5
- automated operation, 5, 87

B

- backed up data
 - hiding from other users, 50
 - visibility, 50
- backing up data, 77–86
 - procedure, 77
- backup
 - IDB operation, 178
 - to disk, 253
- Backup Agents, 11
- backup client
 - snapshot backup, 287
 - split mirror backup, 274
- backup client as failover server
 - snapshot backup, 296
 - split mirror backup, 276
- backup concurrency, 137, A-18, A-40
- backup configuration, 81

- backup devices, 20, 42
 - overview, 134
- backup duration
 - example calculations, A-15, A-35
- backup environment growth
 - database growth and performance key factors, 182
- backup environments, A-5, A-22
- backup generations, 123, A-14, A-34, B-3
- backup interfaces, 237
- backup object, 78
- backup options, A-18, A-39
- backup overview, 7
- backup ownership, 51
- backup performance, 137
- backup policies, 132
 - enterprise environment
 - backup policies, 15
- backup process
 - destination, 7
 - source, 7
- backup scenarios (company ABC), A-22–A-42
- backup scenarios (company XYZ), A-5–A-21
- backup session (definition), 80, 212
- Backup Session Manager, 212
- backup session ownership, 51
- backup sessions, 12, 51, 77, 81, 212–217
 - backup configuration, 81
 - interactive, 212
 - mount requests, 216
 - scheduled, 212
 - timeout, 216
- backup specifications, 20, 78, A-17, A-37
- backup strategy, 27
- backup strategy factors, 31
- backup strategy planning, 27–107
 - backup policies, 32
 - catalog protection, 33
 - data protection, 33
 - data types, 32
 - defining requirements, 29
 - definition, 29
 - device configuration, 33
 - media management, 33
 - scheduling backups, 33
 - system availability, 32
- backup strategy requirements, A-8, A-24
- backup types, 82
 - differential, 68
 - full, 44, 66, 67
 - incremental, 44, 66, 67

Index

- planning performance, 44
- backup with disk discovery, 217
- backups
 - adding data to media, 127
 - automated, 87
 - backup objects, 78
 - backup specifications, 78
 - configuring, 44
 - devices, 134
 - direct, 41
 - disk discovery vs. standard backup, 217
 - disk image, 45
 - filesystem, 45
 - lights-out, 87
 - local, 41
 - network, 41
 - scheduled, 80
 - scheduling policies, 80
 - sessions, 81
 - staggering, 82
 - standard backup vs. disk discovery, 217
 - unattended, 87
- barcode support, 144
- barcodes, 144
- benefits
 - disk backup, 256
 - synthetic backup, 264
 - Volume Shadow Copy service, 304
- benefits of online integrations, 239
- block size
 - backup devices, 139
 - default, 139
 - devices, 139
 - performance, 139
- broadcasts, 194
- browsing files, 75
- BSM, 212

- C**
- cache memory, 46, 235
- Campus Cluster with LVM Mirroring, 295, 296
- Catalog Database, 175
 - do not log any details, 74
 - filename size and growth, 175
 - location, 175
 - log all detailed information, 74
 - log directory names only, 74
 - log level of information, 79
 - records, 175
 - size and growth for CDB Records other than filenames, 175
- Catalog Database growth factors
 - catalog protection, 74
 - level of details, 74
- catalog protection, 74, A-18
 - as an IDB key tunable parameter, 185
 - backup generations, B-3
 - browsing files, 75
 - IDB size and growth, 171
 - catalog protection as an IDB key tunable parameter, 185
 - expired, 185
 - impact on backup performance, 185
 - restoring data when catalog protection expires, 185
- CDB location
 - Catalog Database, 175
- CDB records
 - Catalog Database, 175
- CDB. *See* Catalog Database
- Cell Managers, 36
 - high availability, 55
 - optimizing the load, 216
- Cell Request Server, 211
- cells
 - backup operation, 12
 - Cell Managers, 11
 - logical view, 10
 - mixed environment, 39
 - multiple, 16, 35
 - physical view, 10
 - planning, 35
 - planning security, 48
 - remote, 40
 - restore operation, 12
 - single-point management, 16
 - splitting, 16
 - UNIX environment, 37
 - Windows 2000 environment, 38
 - Windows domains, 38
 - Windows environment, 38
 - Windows workgroups, 39
- centralized licensing, 17
- Centralized Media Management Database, 17, 172, A-28
- character encoding standards, B-15
- checkpoints, 236
- cleaning tape detection, 144
- cleaning tape support, 145

- magazine devices, 142
 - magazines, 142
 - client systems, 11
 - clients, 11
 - installing, 36
 - maintaining, 36
 - cluster (definition), 52
 - cluster heartbeat, 53
 - cluster integrations
 - overview, 55
 - cluster node, 53
 - clustering, 52–65
 - automatic restart, 55
 - Cell Manager availability, 55
 - device sharing, 161
 - failover, 54
 - floating drives, 162
 - group, 54
 - heartbeat, 53
 - load balancing, 55
 - MC/Service Guard, 52
 - Microsoft Cluster Server, 52
 - nodes, 53
 - package, 54
 - primary node, 54
 - secondary node, 54
 - shared disks, 53
 - Veritas Cluster, 52
 - virtual cluster node backup, 57, 59, 62
 - virtual server, 54
 - CMMDB, 17, A-28
 - CMMDB. *See* Centralized Media Management Database
 - code sets, B-15
 - collision, 140
 - commands
 - omniclus command, 63
 - post-exec, 215, 236
 - pre-exec, 215, 236
 - company backup policies, 132
 - comparison
 - disk based devices, 258
 - complete filesystem restore, A-20, A-41
 - compression
 - hardware, 42, 44
 - software, 44
 - concepts
 - snapshot backup, 286
 - split mirror backup, 273
 - concurrency, 136, 137
 - configuring backup specifications, 78
 - configuring cells, A-11, A-28
 - configuring devices, 134
 - large libraries, 143
 - magazines, 142
 - standalone devices, 141
 - consolidating a restore chain, 94
 - control files, 235
 - conventions, 1xiii
 - copying backed up data, 89
 - copying media, 98
 - automated, 99
 - copying objects, 90
 - for vaulting purposes, 93
 - to consolidate a restore chain, 94
 - to demultiplex a medium, 93
 - to free a medium, 93
 - to implement disk staging, 94
 - to migrate to another media type, 94
 - creating backup specifications, 78
 - creating cells
 - mixed environment, 39
 - UNIX environment, 37
 - Windows 2000 environment, 38
 - Windows domains, 38
 - Windows environment, 38
 - Windows workgroups, 39
 - CRS, 211
- ## D
- daily maintenance
 - IDB operation, 179
 - data
 - hiding from other users, 50
 - visibility, 50
 - data encoding, 50
 - data encryption, 50
 - data files, 234
 - data protection, 73, A-18
 - Data Protector architecture
 - cell, 10
 - Cell Managers, 10
 - client systems, 10
 - devices, 10
 - logical view, 10
 - physical view, 10
 - Data Protector concepts
 - Cell Managers, 10
 - cells, 10
 - clients, 10

Index

- devices, 10
- Data Protector features, 3
- Data Protector functionality, 3
- Data Protector GUI, 23
- Data Protector Inet, 211
- Data Protector operation, 209–229
- Data Protector processes, 211–229
 - Cell Request Server, 211
 - Data Protector Inet, 211
 - Media Management Daemon, 211
 - Raima Database Server, 211
- Data Protector security (definition), 48
- Data Protector services, 211–229
 - Cell Request Server, 211
 - Data Protector Inet, 211
 - Media Management Daemon, 211
 - Raima Database Server, 211
- Data Protector setup, 24
- Data Protector user accounts, 49
- Data Protector user groups, 49
- Data Protector user interfaces, 11, 22
- Data Protector user rights (definition), 50
- Data Source Integration, 196
- database
 - advantages, 171
 - architecture, 173
 - Catalog Database, 175
 - catalog protection, 171
 - Detail Catalog Binary Files, 176
 - growth and performance, 181
 - IDB management, 180
 - in the Manager-of-Managers environment, 172
 - Media Management Database, 174
 - on the HP-UX and Solaris Cell Managers, 172
 - on the Windows Cell Manager, 172
 - operation, 178
 - Serverless Integrations Binary Files, 177
 - Session Messages Binary Files, 176
 - size and growth, 171
- database architecture, 173
- database growth and performance key factors, 181
 - backup environment growth, 182
 - filesystem dynamics, 181
- database growth and performance key tunable parameters, 182
- catalog protection, 185
- influence of logging level and catalog protection on IDB growth scheme, 183
- logging level, 183
 - usage of logging level and catalog protection, 185
- database in the Manager-of-Managers environment, 172
 - Centralized Media Management Database, 172
- Database Library, 238
- database on the HP-UX and Solaris Cell Managers
 - IDB format, 172
 - IDB location, 172
- database on the Windows Cell Manager, 172
 - IDB format, 172
 - IDB location, 172
- database operation, 233
- database size estimation, 187
- databases, 233
 - backup interfaces, 237
 - cache memory, 235
 - Centralized Media Management Database, 17
 - checkpoints, 236
 - control files, 235
 - data files, 234
 - dbspaces, 233
 - files, 233
 - online backups, 237
 - segments, 233
 - tables, 233
 - tablespaces, 233
 - transaction logs, 234
- dbspaces, 233
- DC binary file
 - Detail Catalog Binary Files, 176
 - IDB operation, 178
- DC directory
 - Detail Catalog Binary Files, 176
- DCBF information
 - Detail Catalog Binary Files, 176
- DCBF location
 - Detail Catalog Binary Files, 176
- DCBF size and growth
 - Detail Catalog Binary Files, 176
- DCBF. See Detail Catalog Binary Files
- default block size, 139
- default media pools, 114
- demultiplexing media, 93
- Detail Catalog Binary Files, 176
 - DC binary file, 176
 - DC directory, 176

- DCBF size and growth, 176
 - information, 176
 - location, 176
 - device chaining, 136
 - device chains, 141
 - device collision, 140
 - device configuration, 134
 - device lists, 135
 - device locking, 140
 - device sharing in clusters, 161
 - device sharing in SAN, 156
 - drives, 158
 - robotics, 158
 - device streaming (definition), 136
 - devices, 20, 42, 134–162
 - ADIC (EMASS/GRAU) AML, 143
 - autoloaders, 143
 - cleaning tape support, 145
 - concurrency, 136
 - configuring, 134
 - device chaining, 136
 - device lists, 135
 - device locking, 140
 - device streaming, 136
 - disk based, 258
 - exchangers, 143
 - GRAU/EMASS, 143
 - HP StorageWorks DAT Autoloaders, A-34
 - HP StorageWorks DAT24 Autoloaders, A-14
 - HP StorageWorks DLT 4115w Libraries, A-14
 - HP StorageWorks DLT 4228w Libraries, A-33
 - jukeboxes, 143
 - library management console, support, 134
 - load balancing, 135
 - lock names, 140
 - multiple devices, 135
 - number of buffers, 139
 - overview, 134
 - physical device collision, 140
 - planning performance, 42
 - SCSI libraries, 143
 - segment size, 137
 - standalone, 141
 - StorageTek/ACSLs, 143
 - TapeAlert support, 135
 - differential backups, 68
 - Direct Backup, 241
 - overview, 243
 - supported configurations, 251
 - Direct backup
 - requirements, 250
 - Direct Library Access, 160
 - dirty drive detection, 145
 - disaster, 104
 - Disaster Recovery
 - concepts, 104
 - overview, 104
 - Phase 0, 104
 - Phase 2, 105
 - Phase 3, 105
 - disaster recovery, 105
 - alternative methods, 107
 - disaster recovery methods
 - alternative, 107
 - disk
 - backup to, 253
 - Disk Agent concurrency, 137, A-18, A-40
 - Disk Agents, 11
 - disk backup
 - benefits, 256
 - disk based devices
 - comparison, 258
 - overview, 255, 263
 - disk discovery (definition), 217
 - disk discovery vs. standard backup, 217
 - disk fragmentation, 46
 - disk image backups, 45, 46
 - disk image vs. filesystem backups, 45
 - disk performance, 45, 46
 - cache memory, 46
 - compression, 46
 - disk image backups, 46
 - disk staging, 94
 - do not log any details, Catalog Database, 74
 - Drive Servers, 11
 - drives, 158
 - connecting to multiple systems, 146
 - floating, 162
 - static, 161
 - duplicating backed up data, 89
- ## E
- e-mail, 194
 - EMC Symmetrix, 273
 - encoding, 50
 - encryption, 50
 - end-user user group, 167
 - enterprise environment, 15
 - enterprise reporting, 18

Index

- environment
 - enterprise, 15
 - Manager-of-Managers, 15
 - mixed, 39
 - network, 8
 - UNIX, 37
 - Windows, 38
- examples
 - backup scenarios, A-2
 - media pool usage, 118
 - reporting and notification, 199
 - scheduling policies, 84
 - using data provided by Data Protector, 201
 - vaulting usage, 132
- examples of media usage policies, 127
- exchangers, 143
 - See also* libraries
- expired catalog protection, 185
- exporting media, 76
 - IDB operation, 179
 - removed objects, 179

F

- factors affecting restore duration, 100
- factors influencing backup strategies, 31
- failover, 54, 55
- FC-AL, 154
- features of Data Protector, 3
- fibre channel
 - planning performance, 46
- Fibre Channel (definition), 153
- Fibre Channel Arbitrated Loop, 154
- Fibre Channel topologies, 154
 - loop topology, 154
 - point-to-point, 154
 - switched topology, 155
- file jukebox device, 258
- file library device, 258
- file versions purge, 179
- Filename Handling, B-14
- filename size and growth
 - Catalog Database, 175
 - fnames.dat file, 175
- filenames purge
 - IDB operation, 179
- filesystem backup, 45
 - Volume Shadow Copy service, 304, 308
- filesystem dynamics
 - database growth and performance key factors, 181

- filesystem vs. disk image backups, 45
- floating drives, 162
- fnames.dat file
 - filename size and growth, 175
- formatting media, 113
- fragmentation, 46
- freeing media, 93
- full and incremental backups, 66–81
- full backups, 44
 - advantages, 67
 - disadvantages, 67
 - staggering, 82
- functionality of Data Protector, 3
- further information, B-1

G

- General Media Agent, 146
- geographically remote cells, 40
- GRAU/EMASS, 143
- group, 54

H

- hardware compression, 42, 44
- heartbeat, 53
- high availability, 4, 55
 - snapshot backup, 285
 - split mirror backup, 274
- HP OpenView Operations, 195, 196
- HP OpenView Performance Agent, 192, 195
- HP OpenView Performance Agent
 - integration, 196
- HP StorageWorks DAT24 Autoloaders, A-14, A-34
- HP StorageWorks Disk Array XP, 273
- HP StorageWorks DLT 4115w Libraries, A-14
- HP StorageWorks DLT 4228w Libraries, A-33
- HP StorageWorks Enterprise Virtual Array, 285
- HP StorageWorks Virtual Array, 285
- HTML, 194

I

- IDB, 169, 171
 - advantages, 171
 - architecture, 173
 - Catalog Database, 175
 - Detail Catalog Binary Files, 176
 - in the Manager-of-Managers environment, 172
 - management, 180

- Media Management Database, 174
 - on the HP-UX and Solaris Cell Managers, 172
 - on the Windows Cell Manager, 172
 - operation, 178
 - Serverless Integrations Binary Files, 177
 - Session Messages Binary Files, 176
 - size and growth, 171
 - IDB advantages, 171
 - IDB architecture, 173
 - Catalog Database, 175
 - Detail Catalog Binary Files, 176
 - IDB parts, 173
 - IDB parts scheme, 174
 - Media Management Database, 174
 - Serverless Integrations Binary Files, 177
 - Session Messages Binary Files, 176
 - IDB configuration
 - creating a backup specification for the IDB backup, 180
 - IDB management, 180
 - IDB format
 - database on the HP-UX and Solaris Cell Managers, 172
 - database on the Windows Cell Manager, 172
 - IDB growth and performance, 181
 - backups as key factors, 181
 - database growth and performance key tunable parameters, 182
 - database size estimation, 187
 - key factors, 181
 - IDB in the Managers-of-Managers environment
 - Centralized Media Management Database, 172
 - IDB location
 - database on the HP-UX and Solaris Cell Managers, 172
 - database on the Windows Cell Manager, 172
 - IDB maintenance
 - IDB management, 180
 - IDB management
 - IDB configuration, 180
 - IDB maintenance, 180
 - IDB recovery, 180
 - overview, 180
 - setting up the Data Protector backup environment, 180
 - IDB operation, 178
 - backup, 178
 - daily maintenance, 179
 - DC binary file, 178
 - exporting media, 179
 - filenames purge, 179
 - media position record, 178
 - restore, 178
 - session messages binary files, 178
 - IDB parts
 - architecture, 173
 - IDB parts scheme
 - IDB architecture, 174
 - IDB recovery
 - IDB management, 180
 - IDB size and growth, 171
 - catalog protection, 171
 - logging level, 171
 - incremental backup types
 - differential backups, 68
 - leveled incremental backups, 68
 - incremental backups, 44
 - advantages, 67
 - disadvantages, 67
 - types, 68
 - Indirect Library Access, 159
 - influence of logging level and catalog protection on IDB growth scheme, 183
 - initializing media, 113
 - media ID, 124
 - Installation Servers, 11, 37
 - instant recovery
 - snapshot backup, 288
 - split mirror backup, 275
 - integration with database applications, 6, 231–239
 - integrations
 - ManageX, 196
 - OVO, 196
 - Volume Shadow Copy service, 306
 - interactive backup sessions, 212
 - internal database. *See* IDB
 - Internationalization, B-14
 - IT management, 191
- J**
- java reporting, 200
 - java-based online reporting, 200
 - jukeboxes, 143
 - See also* libraries

L

- labeling media, 124
- labels, 124
- LAN-free backups, 156
- large libraries, 143–151
- level 1 incremental backups, A-17, A-37
- leveled incremental backups, 68
- libraries, 18
 - barcode support, 144
 - cleaning tape support, 145
 - connecting to multiple systems, 146
 - drives, 146
 - entering and ejecting mail slots, 144
 - HP StorageWorks DAT Autoloaders, A-34
 - HP StorageWorks DAT24 Autoloaders, A-14
 - HP StorageWorks DLT 4115w Libraries, A-14
 - HP StorageWorks DLT 4228w Libraries, A-33
 - management console, support, 134
 - media handling, 143
 - multiple slots, 144
 - sharing, 144
 - silo, 143
 - size, 143
 - slot range, 143
 - slots, 143
- Library Access
 - Direct, 160
 - Indirect, 159
- library management console, support, 134
- library sharing, 145
- library size, 143
- life cycle, media, 113
- lights-out operation, 5, 87
- LIP, 154
- load balancing, 43, 55, 79, 135
- load balancing (definition), 135
- Localization, B-14
- location fields, 125
- lock names, 140, 158
- log all detailed information, Catalog Database, 74
- log directory names only, Catalog Database, 74
- log level of information, 79
- logging level
 - IDB size and growth, 171
- logging level as an IDB key tunable parameter, 183

- enabling restore regardless of the logging level set, 184
- impact on ability to browse for restore, 184
- impact on IDB speed and backup processes, 184
- impact on restore speed, 184
- Log All, 183
- Log Directories, 184
- Log File, 184
- No Log, 184
- Loop Initialization Primitive (Protocol), 154
- loop topology, 154

M

- magazine devices
 - cleaning, 142
- management console. *See* library management console
- Manager-of-Managers, 16, 17, A-29
 - enterprise reporting, 18
 - remote cells, 40
 - sharing libraries, 18
- ManageX, 192, 194, 196, 197
- MC/Service Guard, 52
- media
 - age, 130
 - barcode support, 144
 - barcodes, 144
 - catalog segments, 137
 - cleaning tape support, 145
 - copying, 98
 - copying, automated, 99
 - data segments, 137
 - device errors, 130
 - ejecting mail slots, 144
 - entering mail slots, 144
 - estimating quantity of needed media, 123
 - exporting, 76
 - file marks, 137
 - formatting, 113
 - header segments, 137
 - initializing, 113, 124
 - labeling, 124, 144
 - location fields, 125
 - mail slots, 144
 - number of overwrites, 130
 - object distribution, 45
 - preparing, 113
 - retiring, 113

- vaulting, 113, 131
 - Media Agents, 11
 - General Media Agent, 146
 - NDMP Media Agent, 146
 - media allocation policies, 114, 122, 126
 - loose, 126
 - strict, 126
 - media condition, 130
 - calculating, 130
 - fair, 127
 - good, 127
 - poor, 127
 - media condition factors, 130
 - media copies, 98
 - media description, 124
 - media handling, 122, 143
 - media life cycle, 113
 - media location, 124
 - media management, 19, 109–133
 - adding data to media, 127
 - copies, 98
 - copying media, 98
 - labeling media, 124
 - media allocation policies, 126
 - media condition, 127
 - media copies, 98
 - media life cycle, 113
 - media pools, 19, 114
 - media rotation policies, 121
 - pre-allocation policies, 126
 - selecting media, 126
 - vaulting, 131
 - media management after backing up, 131
 - media management before backing up, 124
 - media management concepts, 19
 - Media Management Daemon, 211
 - Media Management Database, 174
 - location, 174
 - records, 174
 - size and growth, 174
 - media management during backing up, 126
 - media management functionality, 19, 111
 - media management session (definition), 229
 - media pool properties
 - append incrementals only, 114
 - appendable, 114
 - media allocation policy, 114
 - media pool usage examples, 118
 - large library configuration, 119
 - multiple devices/multiple pools, 121
 - multiple devices/single pool, 120
 - one device/one pool, 118
 - media pools, 19, 20, 114, A-16, A-35
 - default, 114
 - definition, 114
 - properties, 114
 - usage examples, 115, 118
 - media recognition, 144
 - media rotation policies, 121
 - media rotation policy (definition), 121
 - Media Session Managers, 229
 - media set, 71
 - media set (definition), 80
 - media sets, 51, 80, 212
 - media usage, 113
 - media usage policies, 127
 - appendable, 127
 - appendable of incrementals only, 127
 - examples, 127
 - non-appendable, 127
 - media vaulting, 113
 - Microsoft Cluster Server, 52
 - migrating to another media type, 94
 - mirroring objects, 96
 - miscellaneous information, B-1
 - mixed environment, 39
 - MMD, 211
 - MMDB location
 - Media Management Database, 174
 - MMDB records
 - Media Management Database, 174
 - MMDB size and growth
 - Media Management Database, 174
 - MMDB. *See* Media Management Database
 - MoM, 16, 17
 - monitoring, 6, 197
 - mount prompt handling, 88
 - mount requests, 216, 225, 228
 - automating, 216
 - notification, 216
 - responding, 216, 220
 - mount requests (restore sessions), 220
 - MSM, 229
 - multiple cells, 16, 35
 - multiple devices, 135
 - multiple slots, 144
- ## N
- NDMP Media Agent, 146
 - network environment, 8
 - node
 - cluster, 53

Index

- primary, 54
- secondary, 54
- notification, 6
- number of buffers, 139
- number of cells, 35
 - considerations, 35

O

- object consolidation sessions, 226
 - mount requests, 228
 - queuing, 228
- object copy sessions, 223
 - mount requests, 225
 - queuing, 225
- object copy tasks, 92
- object copying, 90
- object distribution to media, 45
- object mirroring, 96
- omniclus command, 63
- online backup of databases, 237
- online database backup
 - archive log backup, snapshot, 287
 - archive log backup, split mirror, 274
 - snapshot backup, 287
 - split mirror backup, 274
- online integrations, 239
- online reporting, 200
- OpenView
 - Operations, 194
- operator user group, 167
- optimizing the load on Cell Managers, 216
- overview
 - backup, 7
 - Direct Backup, 243
 - Disaster Recovery, 104
 - IDB management, 180
 - restore, 7
 - snapshot backup, 285
 - split mirror backup, 273
 - Volume Shadow Copy service, 301
- OVO, 192, 194, 195, 196
- ownership, 51
 - backup sessions, 51
 - restore sessions, 51

P

- package, 54
- parallel restore vs. standard restore, 220
- parallel restores, 220
- parallelism, 43

- Phase 1, 105
- physical device collision, 140
- planning cells, 35–40
 - Cell Managers, 36
 - Installation Servers, 37
 - number of cells, 35
- planning performance, 41–47
 - backup types, 44
 - cache memory, 46
 - compression, 42, 46
 - devices, 42
 - direct backups, 41
 - disk fragmentation, 46
 - disk performance, 45
 - fibre channel, 46
 - hardware compression, 44
 - infrastructure, 41
 - load balancing, 43
 - local backups, 41
 - network backups, 41
 - parallelism, 43
 - software compression, 44
- planning security, 48–65
 - cells, 48
 - data encoding, 50
 - Data Protector user accounts, 49
 - Data Protector user groups, 49
 - visibility of backed up data, 50
- point-to-point topology, 154
- post-backup media copying, 99
- post-backup object copying, 91
- post-exec commands, 215, 236
- post-exec scripts, 79
- predefined user groups, 166, 167
- pre-exec and post-exec scripts, 215
- pre-exec commands, 215, 236
- pre-exec scripts, 79
- preparing a backup strategy plan, 32
- preparing media, 113
- preventing collision, 140
- primary node, 54
- processes, 211
 - backup, 7
 - Backup Session Manager, 212
 - restore, 7
 - Restore Session Managers, 218
- properties of media pools, 114
- protection types
 - catalog, 74
 - data, 73
- purging

file versions, 179
filenames, 179

Q

queuing (restore sessions), 219
queuing of object consolidation sessions, 228
queuing of object copy sessions, 225

R

RAID

snapshot backup, 285
split mirror backup, 277

Raima Database Server, 211

RDS, 211

real time alerts, 195

recovery, 105

disaster recovery, 105

recycling media, 113

remote cells, 40

replica

snapshot backup, 286
split mirror backup, 273

replica set

snapshot backup, 288
split mirror backup, 276

replica set rotation

snapshot backup, 288
split mirror backup, 276

reporting, 6, 197

reporting and notification, A-19, A-40

broadcasts, 194

e-mail, 194

examples, 199

HTML, 194

SNMP, 194

requirements

Direct backup, 250

response time, 195

restore

IDB operation, 178
media location priority, 101
selecting media, 101
Volume Shadow Copy service, 307

restore by query, A-20, A-41

restore chain, 71

restore duration, 100
factors affecting, 100
parallel restore, 101

restore options, A-18

restore overview, 7

restore policies, 100

end users, 103

operators, 102

Restore Session Managers, 218

restore sessions, 14, 51, 218–221

definition, 218

mount requests, 220

queuing, 219

timeout, 219

restores, 100, 218

complete filesystem restore, A-20, A-41

configuring, 44

duration, 100

end users

end-user user group, 103

operators, 102

optimizing, 82

parallel, 220

restore by query, A-20, A-41

vaulting, 133

restoring data, 100–103

restoring from media in a vault, 133

retiring media, 113

robotics, 158

RSM, 218

S

SAN

See Storage Area Networks

scheduled backup sessions, 212

scheduled backups, 80

scheduled media copying, 99

scheduled object copying, 91

scheduling

backup configuration, 81

scheduling policies, 80, 82

scheduling policy examples, 84

scheduling tips and tricks, 81

scripts

post-exec, 79

pre-exec, 79

pre-exec and post-exec, 215

secondary node, 54

security

data encoding, 165

definition, 48

unauthorized access of data, 165

user groups, 165

user-related, 165

visibility of backed up data, 165

Index

- security features, 48
- segment size, 137
- segments, 233
- selecting backup objects, 78
- selecting media for backup
 - media
 - selecting for backup, 126
 - selecting media for restore, 101
- Serverless Integrations Binary Files, 177
 - data, 177
 - location, 177
 - size and growth, 177
- Service Management, 5, 189–202
 - Application Response Measurement, 194
 - monitor, 197
 - notification, 197
 - operative analyses of trends, 191, 192
 - overview, 191
 - reporting, 197
- service management applications
 - HP OpenView Performance Agent, 192
 - ManageX, 192
 - OVO, 192
- Service Management examples, 201
- service monitoring, 197
- services, 211
- Session Messages Binary Files, 176
 - location, 177
 - records, 176
 - size and growth, 177
- sessions
 - backup, 12, 212
 - media management, 229
 - object consolidation, 226
 - object copy, 223
 - restore, 14, 218
- setting catalog protection
 - usage of logging level and catalog protection, 185
- setting up Data Protector (overview), 24
- setting up the Data Protector backup environment
 - IDB management, 180
- shadow copy, 301
- shadow copy provider, 302
- shadow copy set, 301
- shared disks, 53
- sharing devices in SAN, 156
 - drives, 158
 - robotics, 158
- sharing libraries, 18, 143, 144, 145
- SIBF data
 - Serverless Integrations Binary Files, 177
- SIBF location
 - Serverless Integrations Binary Files, 177
- SIBF size and growth
 - Serverless Integrations Binary Files, 177
- silos, 143
- single file restore, 221
- size
 - libraries, 143
- size and growth for CDB Records other than filenames
 - Catalog Database, 175
- slot range, 143
- slots, 143
- SMBF location
 - Session Messages Binary Files, 177
- SMBF records
 - Session Messages Binary Files, 176
- SMBF size and growth
 - Session Messages Binary Files, 177
- SMBF. *See* Session Messages Binary Files
- snapclones, 290
- snapshot backup, 283
 - application client, 287
 - archive log backup, 287
 - backup client, 287
 - backup client as failover server, 296
 - concepts, 286
 - configuration, Campus Cluster with LVM Mirroring, 296
 - configuration, disk arrays - single host, 294
 - configuration, LVM mirroring, 295
 - configuration, multiple application hosts - single backup host, 294
 - configuration, multiple disk arrays - dual host, 293
 - configuration, other, 296
 - configuration, single disk array - dual host, 291
 - configurations, 291
 - high availability, 285
 - instant recovery, 288
 - online database backup, 287
 - overview, 285
 - RAID, 285
 - replica, 286
 - replica set, 288
 - replica set rotation, 288
 - source volume, 286

- target volume, 286
 - ZDB to disk, 288
 - ZDB to disk+tape, 288
 - ZDB to tape, 288
 - snapshot configurations, 291
 - Campus Cluster with LVM Mirroring, 296
 - disk arrays - single host, 294
 - LVM mirroring, 295
 - multiple application hosts - single backup host, 294
 - multiple disk arrays - dual host, 293
 - other, 296
 - single disk array - dual host, 291
 - snapshots
 - types of, 289
 - snapshots with the preallocation of disk space, 289
 - snapshots without the preallocation of disk space, 289
 - SNMP, 194
 - software compression, 44
 - solutions for backup scenarios, A-9, A-26
 - source volume
 - snapshot backup, 286
 - split mirror backup, 273
 - split mirror backup
 - application client, 274
 - archive log backup, 274
 - backup client, 274
 - backup client as failover server, 276
 - concepts, 273
 - configuration, local mirror-dual host, 277
 - configuration, local mirror-single host, 278
 - configuration, local/remote mirror, 280
 - configuration, other, 281
 - configuration, remote mirror, 278
 - configurations, 277
 - high availability, 274
 - instant recovery, 275
 - online database backup, 274
 - overview, 273
 - RAID, 277
 - replica, 273
 - replica set, 276
 - replica set rotation, 276
 - source volume, 273
 - target volume, 273
 - ZDB to disk, 275
 - ZDB to disk+tape, 275
 - ZDB to tape, 275
 - split mirror configurations, 277
 - local mirror-dual host, 277
 - local mirror-single host, 278
 - local/remote mirror, 280
 - other configurations, 281
 - remote mirror, 278
 - stacker devices, 141
 - staggering full backups, 82
 - standalone devices, 141
 - standalone file device, 258
 - standard backup vs. disk discovery, 217
 - standard restore vs. parallel restore, 220
 - static drives, 161
 - Storage Area Networks, 152–162
 - any-to-any connectivity, 152
 - concepts, 152
 - device sharing, 156
 - device sharing in clusters, 161
 - Direct Library Access, 160
 - Fibre Channel, 153
 - Fibre Channel topologies, 154
 - Indirect Library Access, 159
 - LAN-free backups, 156
 - lock names, 158
 - sharing devices, 156
 - storage duration of backed up data, 73–76
 - storage virtualization, 285
 - StorageTek/ACSLs, 143
 - supported configurations for Direct Backup, 251
 - switched topology, 155
 - synthetic backup, 261
 - benefits, 264
 - operation, 265
 - systems to be backed up, 11
 - systems with backup devices, 11
- ## T
- tablespaces, 233
 - TapeAlert support, 135
 - Target System, 104
 - target volume
 - snapshot backup, 286
 - split mirror backup, 273
 - timeout, 216
 - timeout (restore sessions), 219
 - transaction logs, 234
 - transactions, 195
 - types of incremental backups, 68
 - differential backups, 68
 - leveled incremental backups, 68

Index

typographical conventions, 1xiii

U

unattended operation, 5, 87, 141
usage of logging level and catalog protection, 185
 setting catalog protection, 185
 specifics for large cells, 187
 specifics for small cells, 186
 using different logging levels in the same cell, 186
usage of media pools, 115
user groups, 166
 admin, 167
 end-user, 167
 operator, 167
 predefined, 166, 167
user interfaces, 11, 22
 Data Protector GUI, 23
user rights, 166, 167
user-related security, 165
users, 166
users and user groups, 163–167

V

vaulting, 113, 131–133, A-19, A-40
 definition, 131
 restoring, 133
 restoring from a vault, A-20, A-41
vaulting usage example, 132
Veritas Cluster, 52
virtual cluster nodes, 57, 59, 62
virtual server, 54
visibility of backed up data, 50, 165
Volume Shadow Copy service (VSS)
 backup, 306
 backup model, 302
 benefits, 304
 filesystem backup, 304
 filesystem backup and restore, 308
 integration with Data Protector, 304, 306
 overview, 301
 restore, 307
 shadow copy, 301
 shadow copy provider, 302
 shadow copy set, 301
 writer, 302
VSS
 See Volume Shadow Copy service, 301
VSS backup, 306

VSS backup model, 302

W

Windows domains, 38
Windows workgroups, 39
writer, 302
Writer Metadata Document (WMD), 306

Z

ZDB to disk
 snapshot backup, 288
 split mirror backup, 275
ZDB to disk+tape
 snapshot backup, 288
 split mirror backup, 275
ZDB to tape
 snapshot backup, 288
 split mirror backup, 275
Zero Downtime Backup
 snapshot backup, 285
 split mirror backup, 275