

HP Data Protector

Software Version: 8.10

Concepts Guide

Document Release Date: November 2016

Software Release Date: November 2016



Legal Notices

Warranty

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.

The information contained herein is subject to change without notice.

Restricted Rights Legend

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.

Copyright Notice

© Copyright 2016 Hewlett-Packard Development Company, L.P.

Trademark Notices

Adobe® is a trademark of Adobe Systems Incorporated.

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

UNIX® is a registered trademark of The Open Group.

Documentation Updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

To check for recent updates or to verify that you are using the most recent edition of a document, go to: <http://h20230.www2.hp.com/selfsolve/manuals>

This site requires that you register for an HP Passport and sign in. To register for an HP Passport ID, go to: <http://h20229.www2.hp.com/passport-registration.html>

Or click the **New users - please register** link on the HP Passport login page.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HP sales representative for details.

Support

Visit the HP Software Support Online web site at: <http://www.hp.com/go/hpsoftwaresupport>

This web site provides contact information and details about the products, services, and support that HP Software offers.

HP Software online support provides customer self-solve capabilities. It provides a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the support web site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Download software patches
- Manage support contracts
- Look up HP support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HP Passport user and sign in. Many also require a support contract. To register for an HP Passport ID, go to:

<http://h20229.www2.hp.com/passport-registration.html>

To find more information about access levels, go to:

http://h20230.www2.hp.com/new_access_levels.jsp

HP Software Solutions Now accesses the HPSW Solution and Integration Portal Web site. This site enables you to explore HP Product Solutions to meet your business needs, includes a full list of Integrations between HP Products, as well as a listing of ITIL Processes. The URL for this Web site is <http://h20230.www2.hp.com/sc/solutions/index.jsp>

Contents

Contents	3
About this guide	13
Intended audience	13
Chapter 1: About backup and Data Protector	14
About Data Protector	14
Introducing backups and restores	16
What is a backup?	16
What is a restore?	17
Backing up a network environment	17
Data Protector architecture	18
Operations in the cell	19
Backup sessions	20
Restore sessions	21
Enterprise environments	21
Splitting an environment into multiple cells	22
MoM	23
Media management	24
Backup devices	25
User interfaces	26
Data Protector GUI	26
Overview of tasks to set up	27
Chapter 2: Planning your backup strategy	30
Backup strategy planning	30
Defining the requirements of a backup strategy	30
Factors influencing your backup strategy	32
Preparing a backup strategy plan	32
Planning cells	34
One cell or multiple cells?	34
Installing and maintaining client systems	35

Creating cells in the UNIX environment	36
Creating cells in the Windows environment	36
Windows domains	36
Windows workgroups	37
Creating cells in a mixed environment	37
Geographically remote cells	37
Understanding and planning performance	38
The infrastructure	38
Network versus local backups	38
Devices	38
High performance hardware other than devices	39
Advanced high performance configuration	39
Using hardware in parallel	39
Configuring backups and restores	40
Software compression	40
Hardware compression	40
Full and incremental backups	40
Disk image versus filesystem backups	40
Object distribution to media	41
Disk performance	41
SAN performance	42
Online database application performance	42
Planning security	42
Cells	43
Data Protector user accounts	43
Data Protector user groups	44
Data Protector user rights	44
Visibility of backed up data	44
What is backup ownership?	45
Data encryption	45
How Data Protector AES 256-bit encryption works	45
How Data Protector drive-based encryption works	46

Restore from encrypted backups	47
Encrypted control communication	47
How Data Protector encrypted control communication works	48
Data encryption and encrypted control communication	49
Clustering	50
Cluster concepts	50
Cluster support	52
Example cluster environments	53
Cell Manager installed outside a cluster	53
Cell Manager installed outside a cluster, devices connected to the cluster nodes	54
Cell Manager installed in a cluster, devices connected to the cluster nodes	55
Full, incremental, and synthetic backups	57
Full backups	58
Synthetic backup	58
Incremental backups	58
Conventional incremental backup	58
Enhanced incremental backup	58
Incremental backup using Change Log Provider	59
Types of incremental backups	59
Backup generations	61
Synthetic backup	62
Overview	62
Synthetic backup benefits	62
How Data Protector synthetic backup works	62
Synthetic backup and media space consumption	64
Restore and synthetic backup	64
How data protection periods affect restore from synthetic backup	66
Considering restore	66
Keeping backed up data and information about the data	68
Data protection	68
Catalog protection	68
Logging level	69

Browsing files for restore	69
Enabling the browsing of files and quick restore	69
Enabling the restore of files, but not browsing	69
Overwriting backed up files with new data	69
Exporting media from a cell	70
Importing WORM media to a cell	70
Backing up data	70
Creating a backup specification	71
Selecting backup objects	72
Backup sessions	73
Object mirrors	73
Media sets	74
Backup types and scheduled backups	74
Scheduling, backup configurations, and sessions	74
Scheduling tips and tricks	75
When to schedule backups	75
Staggering full backups	75
Optimizing for restore	75
Automated or unattended operation	78
Considerations for unattended backups	78
Duplicating backed up data	80
Copying objects	81
Selection of source devices	82
Selection of destination devices	83
Why use object copy?	83
Replication	87
Why use replication?	87
Object mirroring	88
Copying media	90
Automated media copying	90
Verifying backup media and backup objects	91
What is media verification?	91

What does media verification do for you?	91
What is object verification?	92
What does object verification do for you?	92
Restoring data	93
Restore duration	93
Selection of the media set	94
Selection of devices	94
Operators are allowed to restore	95
End users are allowed to restore	96
Disaster recovery	96
Disaster recovery methods	97
Alternative disaster recovery methods	98
Recovery methods supported by operating system vendors	98
Recovery using third-party tools (for Windows systems)	98
Chapter 3: Devices and media management	100
Devices	100
Device lists and load balancing	101
How load balancing works	102
Device streaming and concurrency	102
Device Filtering	103
Segment size	103
Block size	104
Number of Disk Agent buffers	105
Device locking and lock names	105
Standalone devices	106
Small magazine devices	106
Large libraries	107
Handling of media	107
Size of a library	107
Sharing a library with other applications	108
Enter/eject mail slots	108
Barcode support	108

Cleaning tape support	109
Sharing a library with multiple systems	109
Disk backup	114
Disk backup benefits	115
Data Protector disk-based devices	115
Data Protector and Storage Area Networks	117
Storage Area Networks	117
Fibre Channel	118
Point-to-point topology	119
Loop topology	119
Switched topology	120
Device sharing in SAN	120
Configuring multiple paths to physical devices	120
Device locking	122
Indirect and direct library access	122
Indirect library access	122
Direct library access	123
Device sharing in clusters	124
Static drives	124
Floating drives	124
Media management	125
Media life cycle	126
Media pools	126
Free pools	128
Media pool usage examples	131
Implementing a media rotation policy	133
Media rotation and Data Protector	134
Media needed for rotation	134
Media management before backups begin	135
Initializing or formatting media	135
Labeling Data Protector media	135
Location field	136

Media management during backup sessions	136
Selecting media for backups	137
Adding data to media during backup sessions	137
Writing data to several media sets during backup	139
Calculating media condition	139
Media management after backup sessions	140
Vaulting	140
Restoring from media in a vault	141
Chapter 4: Users and user groups	143
Increased security for users	143
Access to backed up data	143
Users and user groups	143
Data Protector user rights	144
Chapter 5: The Internal Database	145
About the IDB	145
IDB location and internal encoding used	145
The IDB in the Manager-of-Managers environment	146
IDB architecture	146
Media Management Database (MMDB)	147
Catalog Database (CDB)	147
Detail Catalog Binary Files (DCBF)	148
Session Messages Binary Files (SMBF)	149
Encryption keystore and catalog files	149
IDB operation	149
During backup	149
During restore	150
During object copying or object consolidation	150
During object verification	150
Exporting media	151
Removing the Detail Catalog	151
Overview of IDB management	151
IDB growth and performance	152

Key IDB growth and performance factors	152
IDB growth and performance: key tunable parameters	153
Logging level as an IDB key tunable parameter	153
Catalog protection as an IDB key tunable parameter	154
Recommended usage of logging level and catalog protection	154
Chapter 6: Service management	157
Data Protector and service management	157
Data Protector functionality	157
SNMP traps	157
The Data Protector monitor	158
Reporting and notification	158
Event logging and notification	159
Data Protector log files	159
Windows application log	160
Java-based online reporting	160
Data Protector checking and maintenance mechanism	160
Central management, distributed environment	160
Using the data provided by Data Protector	160
Chapter 7: How Data Protector operates	162
Data Protector processes or services	162
Backup sessions	163
Scheduled and interactive backup sessions	163
Backup session data flow and processes	163
Pre-exec and post-exec commands	165
Start-up and location of the commands	166
Windows systems	166
UNIX systems	166
Queuing of backup sessions	167
Mount requests in backup sessions	167
Backing up with disk discovery	168
Resuming backup sessions	168
Restore sessions	168

Restore session data flow and processes	168
Queuing of restore sessions	169
Mount requests in a restore session	170
Parallel restores	170
Fast multiple single file restore	171
Resuming restore sessions	171
Object copy sessions	172
Automated and interactive object copy sessions	172
Object copy session data flow and processes	172
Queuing of object copy sessions	174
Mount requests in an object copy session	174
Replication sessions	174
Automated and interactive replication sessions	175
Replication session data flow and processes	175
Queuing of replication sessions	176
Object consolidation sessions	176
Automated and interactive object consolidation sessions	177
Object consolidation session data flow and processes	177
Queuing of object consolidation sessions	178
Mount requests in an object consolidation session	178
Object verification sessions	178
Automated and interactive object verification sessions	179
Object verification session data flow and processes	179
Media management sessions	180
Media management session data flow	180
Chapter 8: Integration with applications	182
Integration with database applications	182
Overview of database operation	182
Filesystem backup of databases and applications	184
Online backup of databases and applications	184
Integration with virtualization environments	186
Offline filesystem backup of virtual machines	186

Online backup of virtual machines	186
Microsoft Volume Shadow Copy Service	187
Overview	187
Data Protector Volume Shadow Copy integration	190
VSS filesystem and disk image backup and restore	190
Chapter 9: Zero downtime backup and instant recovery	192
Zero downtime backup	192
Creating replicas	193
ZDB types	193
Instant recovery and restore of data	194
Instant recovery	194
Other restore methods	194
Glossary	195
Index	236
We appreciate your feedback!	255

About this guide

This guide describes Data Protector concepts. Read this guide to fully understand the fundamentals and the model of Data Protector.

Intended audience

This guide 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 guide together with the *HP Data Protector Help*.

Chapter 1: About backup and Data Protector

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.

About Data Protector

HP 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 several 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 (CLI) 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.

- **Data security**

To enhance the security of your data, Data Protector lets you encrypt your backups so that they become protected from others. Data Protector offers two data encryption techniques: software-based and drive-based.

- **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.

The Data Protector user interface can access the entire Data Protector functionality on all supported platforms. For a list of supported platforms, see the *HP Data Protector Product Announcements, Software Notes, and References*.

- **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, see the *HP 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 the *HP 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 HP P4000 SAN Solutions, HP P9000 XP Disk Array Family, or EMC Symmetrix.
- 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.

- **Backup object operations**

To provide flexibility in the choice of backup and archive strategy, advanced techniques are available for performing operations on individual backup objects. These include copying of objects from one medium to another, useful for disk staging and archiving purposes, and consolidation of multiple object versions from incremental backups into a single full-backup version. To support such functionality, there is also the ability to verify both original and copied or consolidated backup objects.

- **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.

- **Monitoring, reporting, and notifications**

Various 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 CLI, 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.

In addition, the Data Protector auditing functionality enables you to collect a subset of backup session information and provides an overview of backup operations.

- **Integration with online applications**

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

- **Integration with other products**

Additionally, Data Protector integrates with EMC Symmetrix, Microsoft Cluster Server, HP 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 Data Protector home page at <http://support.openview.hp.com/selfsolve/manuals>.

Introducing backups and restores

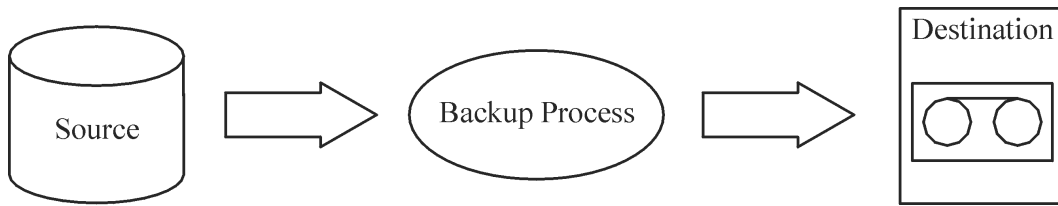
This section explains basic backup and restore concepts.

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 "[Backup process](#) " on the next page.

Figure 1: Backup process



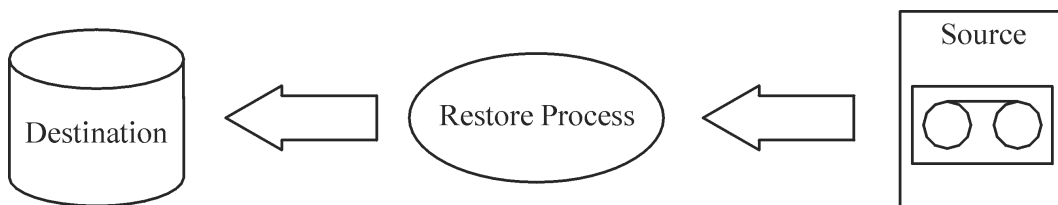
In most cases, the **source** is data on a disk, such as files, directories, databases, and applications. If the backup is expected to be used for disaster recovery, it needs to be consistent.

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 2: Restore process

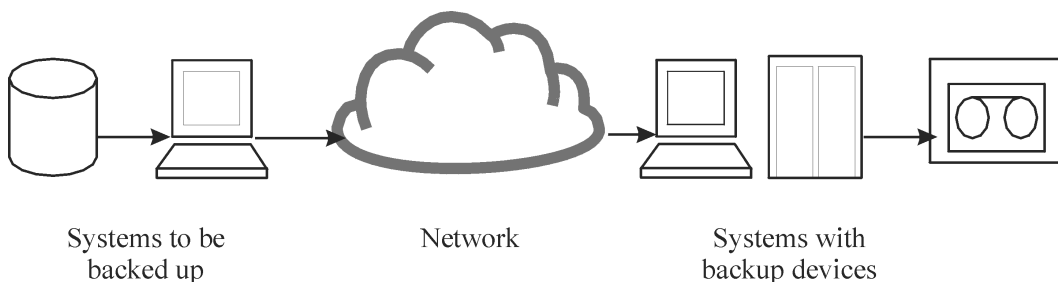


The **source** is a backup copy. A restore application is software that actually writes data to a destination. The **destination** is usually a disk to which the original data is written.

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 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

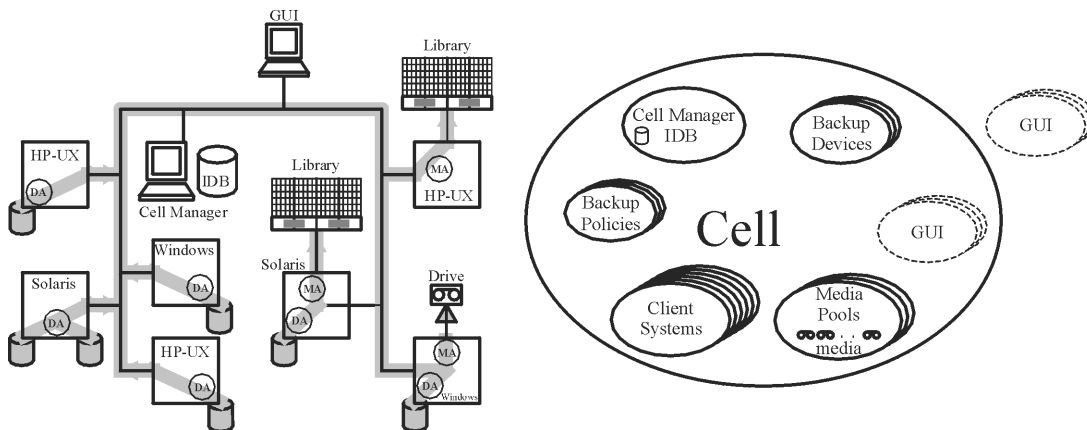
Data Protector architecture

The Data Protector **cell**, shown in "[The Data Protector cell \(physical view and logical view\)](#)" below, 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 4: The Data Protector cell (physical view and logical view)



Note: Data Protector Cell Managers and the Data Protector graphical user interface systems do not have to run the same operating system. For a list of supported operating systems for a particular Data Protector component, see the *HP 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 session IDs

- 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 guide, 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.

While the Cell Manager installation itself provides means of backing up and restoring the Internal Database and the related configuration data, to be able to back up and restore non-Data Protector data residing on the Cell Manager, Disk Agent should also be installed there.

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.

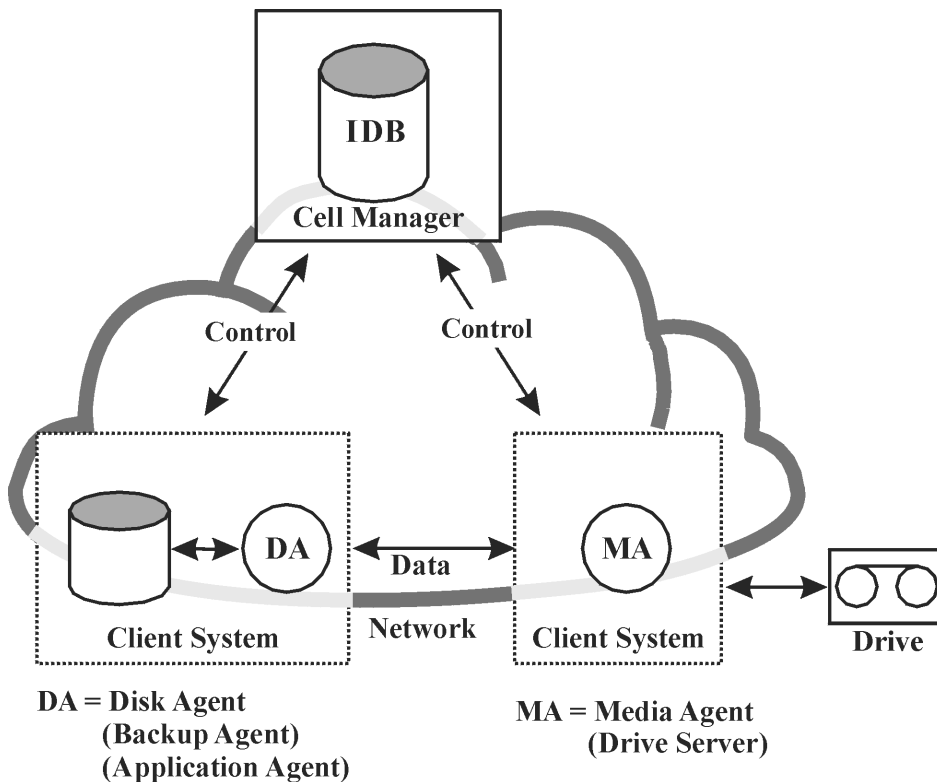
Installation Server

The **Installation Server** holds a repository of the Data Protector installation 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 "[Backup or restore operation](#)" on the next page.

Figure 5: Backup or restore operation



Backup sessions

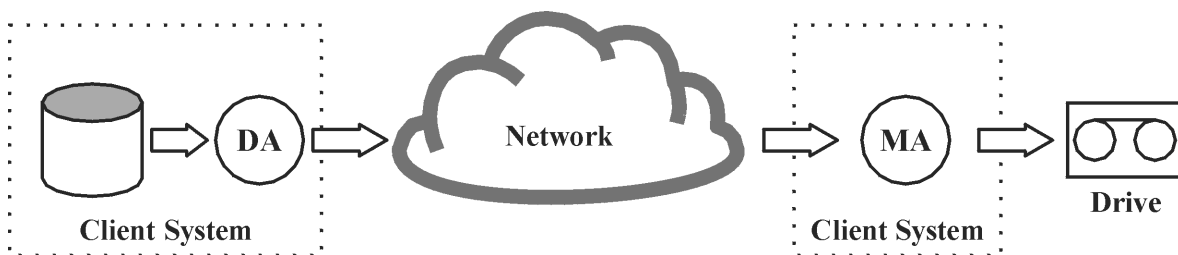
What is a backup session?

A backup session, shown in ["Backup session" below](#), 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 6: Backup session



A typical backup session is more complex than the one shown in ["Backup session" above](#). 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, see ["How Data Protector operates" on page 162](#).

Restore sessions

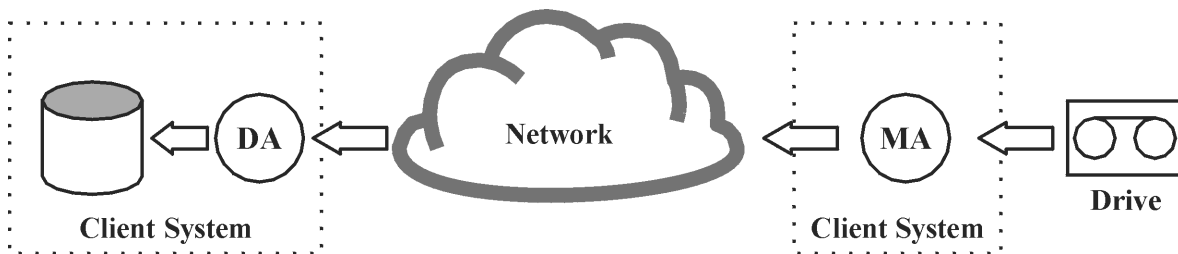
What is a restore session?

A restore session, shown in ["Restore session" below](#), 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 7: Restore session



A restore session may be more complex than the one shown in ["Restore session" above](#). For more information on restore sessions, see ["How Data Protector operates" on page 162](#).

Enterprise environments

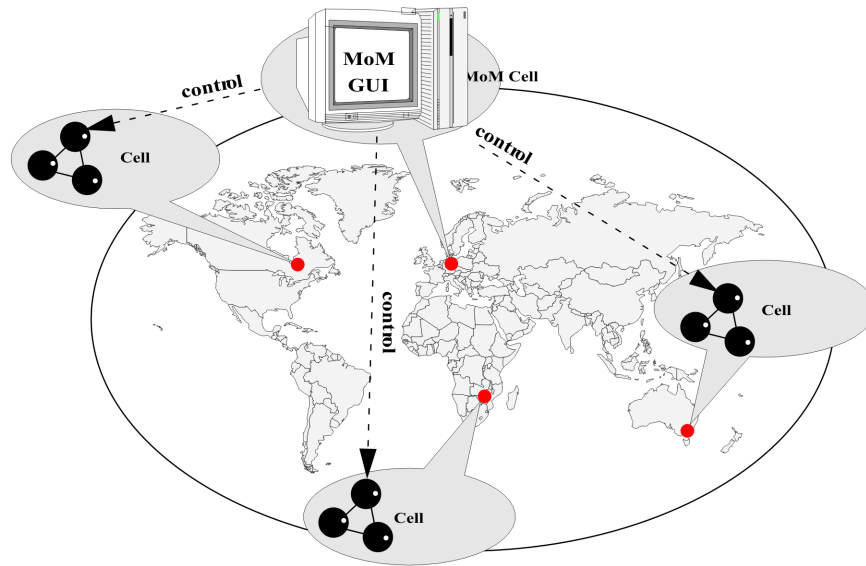
What is an enterprise environment?

A typical enterprise network environment, shown in ["Large Data Protector enterprise environment" on the next page](#), 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 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), see ["MoM" on the next page](#).

Splitting an environment into multiple cells

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

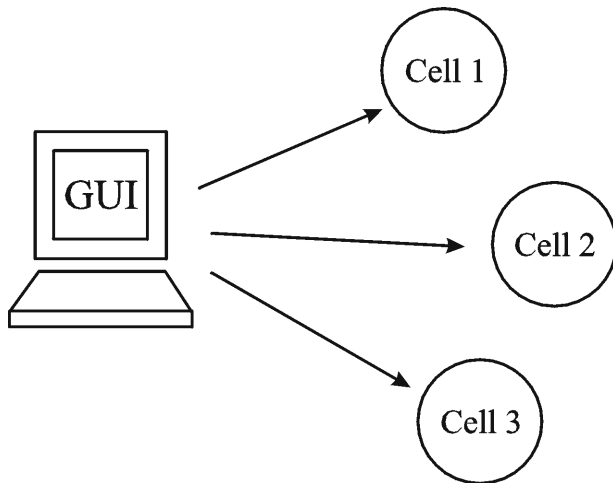
Why split large environments into multiple cells?

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

For a list of considerations in planning your environment, see ["Planning your backup strategy" on page 30](#).

Data Protector allows you to manage multiple cells from a single point.

Figure 9: Single-point management of multiple cells

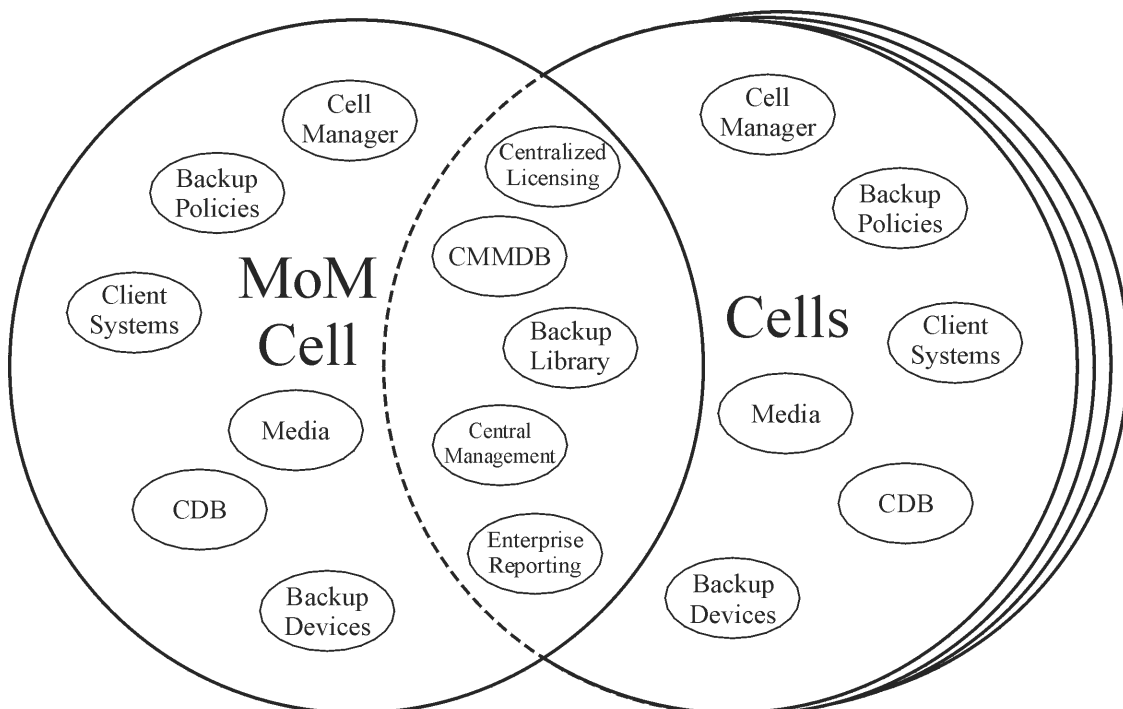


MoM

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 "[Single-point management of multiple cells](#)" above. 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 the controls are sent over the long distance connections, however the backups are performed locally within each Data Protector cell. Nevertheless, this is based on the assumption that each cell has its own Media Management Database.

Figure 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 on the Manager of Managers. 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.

- **Sharing libraries**

With the CMMDB, you can share high-end devices between cells in the multicell 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?

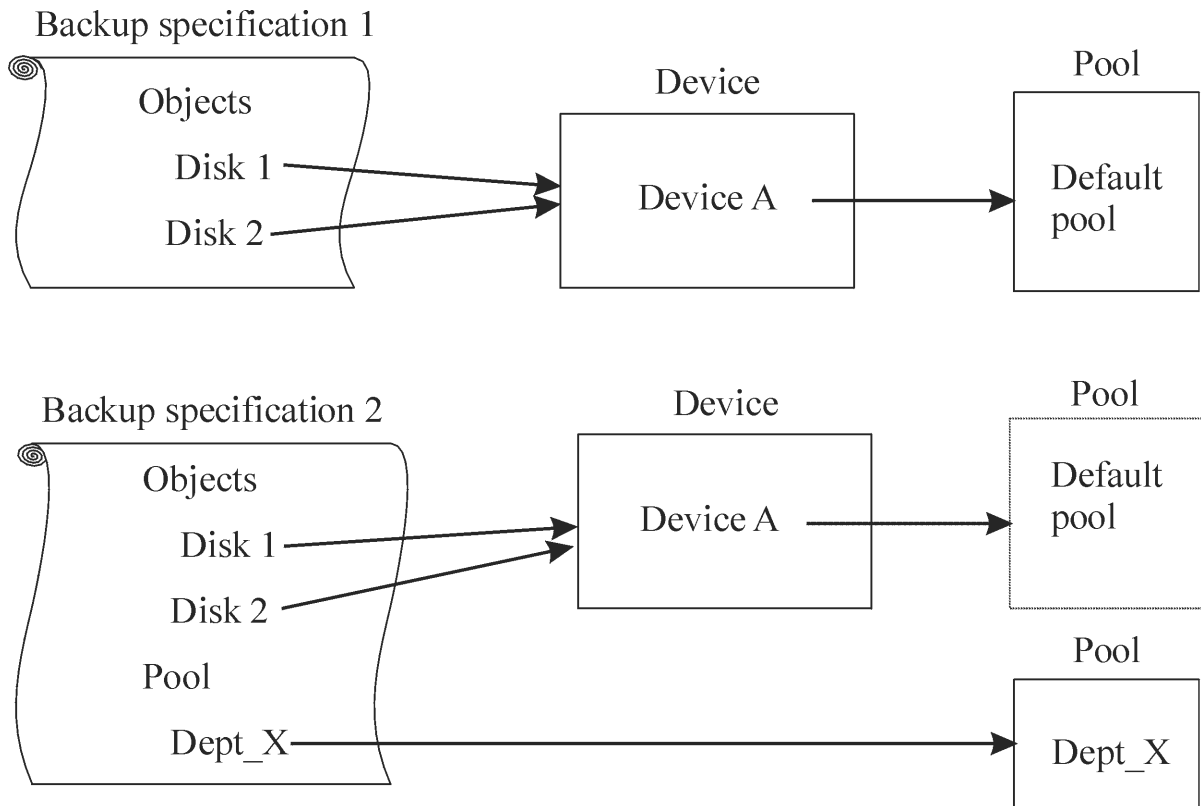
Data Protector uses media pools to manage large numbers of media. A media pool is a logical collection of media of the **same** physical type with common usage policies (properties). Usage is based on the data on the media. The structure and quantity of the pools, as well as which pool contains what type of data on its media, depend 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 11: How backup specifications, devices, and media pools are related



"How backup specifications, devices, and media pools are related" above 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.

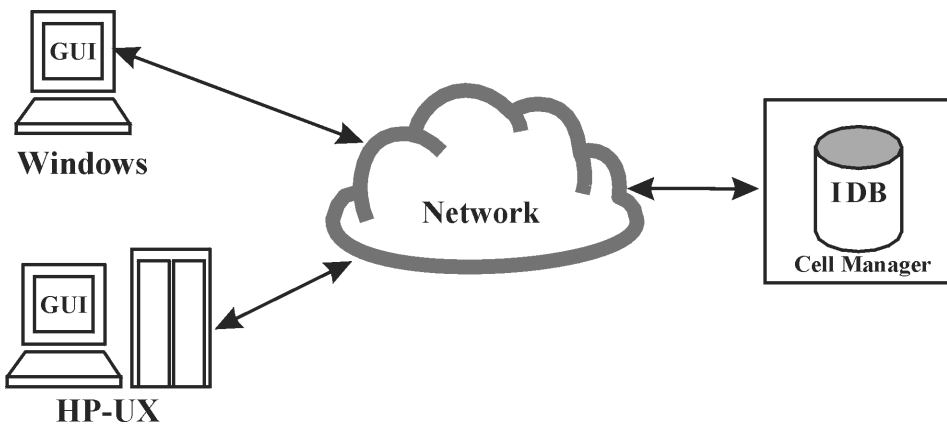
Data Protector supports various devices. For more information, see the *HP Data Protector Product Announcements, Software Notes, and References*.

User interfaces

Data Protector provides easy access to all configuration and administration tasks using the Data Protector GUI on Windows platforms. Additionally, a command-line interface (CLI) is available on Windows and UNIX 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 "[Using the Data Protector user interface](#)" below, the user interface also allows you to transparently manage Data Protector cells with Cell Managers on all supported platforms.

Figure 12: Using the Data Protector user interface



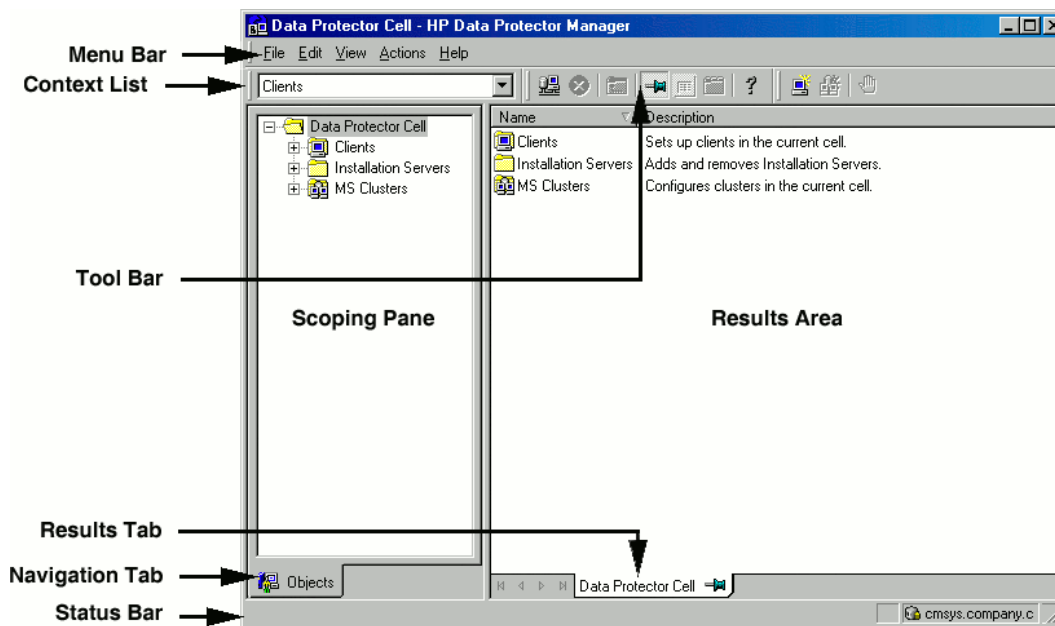
Tip: In a typical mixed environment, install the Data Protector user interface on several systems in the environment, thus providing access to Data Protector from several systems.

Data Protector GUI

The Data Protector GUI, depicted in "[Data Protector GUI](#)" on the next page, 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 Server or those that run in the UNIX environments, such as SAP R/3, Oracle Server, and Informix Server.
- A comprehensive Help system that includes Help topics and context-sensitive Help.

Figure 13: Data Protector GUI



Overview of tasks to set up

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. For more information, see the respective *HP Data Protector Integration Guide*.
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. For details, see the *HP Data Protector Installation and Licensing Guide*.

5. Consider security aspects:
 - Analyze security considerations. See the *HP Data Protector Installation and Licensing Guide*.
 - Consider which user groups you need to configure.
 - Enhance security by writing data to media in an encrypted format.
 - Help preventing unauthorized access by enabling encrypted control communication.
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

Chapter 2: Planning your backup strategy

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.

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. For more information, see ["Disaster recovery" on page 96](#).

- 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. For more information, see ["Full, incremental, and synthetic backups" on page 57](#). Data Protector also backs up some popular online database applications. For more information, see the *HP Data Protector Integration Guide*.

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 ["Synthetic backup" on page 62](#) and ["Disk staging" on page 85](#).

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.

- 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.

- 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 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. If the recommended size is exceeded in a particular cell, the cell becomes less manageable. For information on how to organize Data Protector clients into cells so that they can be efficiently managed, see the *HP Data Protector Product Announcements, Software Notes, and References*.

- Network considerations

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

- 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 (install remotely) 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 removal 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. For details, see the *HP Data Protector Installation and Licensing Guide*.

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 guide, 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 ["Creating cells in the UNIX environment" on the previous page](#). 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. For more information, see ["Splitting an environment into multiple cells" on page 22](#).

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 reestablished 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 disaster. 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.

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 dismount 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 disk read 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, see the *HP 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.
- 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 102](#).

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 the other way round.

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 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, see ["Device streaming and concurrency" on page 102](#).

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.

Disk Agent performance on Windows systems

Disk Agent performance of Windows filesystem backup can be improved by enabling asynchronous reading. Asynchronous reading improves performance of the Disk Agent when backing up data on disk arrays, especially if large files are backed up. It is recommended to perform test backups to establish if asynchronous reading will improve performance in your specific environment and determine the optimum asynchronous reading settings.

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.

For more information on how Data Protector integrates with various applications and for tips on how to improve backup performance, see the respective *HP Data Protector Integration Guide*.

Also see 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
- Data encryption
- Encrypted control communication

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 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.

For more information, see ["Users and user groups" on page 143](#).

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.

For more information, see ["Users and user groups" on page 143](#).

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.

For more information, see ["Users and user groups" on page 143](#).

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 configuring a backup, you can decide whether during a restore the data is visible to everyone (public) or only to the owner of the backup (private). For more information about backup owners, see ["What is backup ownership?" below](#).

What is backup ownership?

Who owns a backup session?

Each backup session and all the data backed up within it is assigned an owner. The owner can be the user who starts an interactive backup, the account under which the CRS process is running, or the user specified as the owner in the backup specification options. For instructions on how to specify a backup owner, see the *HP Data Protector Help* index: "ownership".

Backup ownership and restore

Backup ownership affects the ability of users to see and restore data. Unless the object is marked as Public, only the owner of the media set or an administrator can see the data saved in the media set. The right to see and restore private objects can be granted to groups other than *admin* as well. For instructions on who can see and restore a private object and how this can be applied, see the *HP Data Protector Help* index: "ownership".

Data encryption

Open systems and public networking make data security in large enterprises essential. Data Protector lets you encrypt backed-up data so that it becomes protected from others. Data Protector offers two data encryption techniques: software-based and drive-based.

Data Protector software encryption, referred to as **AES 256-bit encryption**, is based on the AES-CTR (Advanced Encryption Standard in Counter Mode) encryption algorithm that uses random keys of 256-bit length. The same key is used for both encryption and decryption. With AES 256-bit encryption, data is encrypted before it is transferred over a network and before it is written to media.

Data Protector **drive-based encryption** uses the encryption functionality of the drive. The actual implementation and encryption strength depend on the drive's firmware. Data Protector only turns on the feature and manages encryption keys.

The key management functionality is provided by the **Key Management Server (KMS)**, which is located on the Cell Manager. All encryption keys are stored centrally in the keystore file on the Cell Manager and administered by the KMS.

You can encrypt all or selected objects in a backup specification and also combine encrypted and unencrypted sessions on the same medium.

In addition to the encryption functionality, Data Protector also offers the encoding functionality that uses a keyless, built-in algorithm for this purpose.

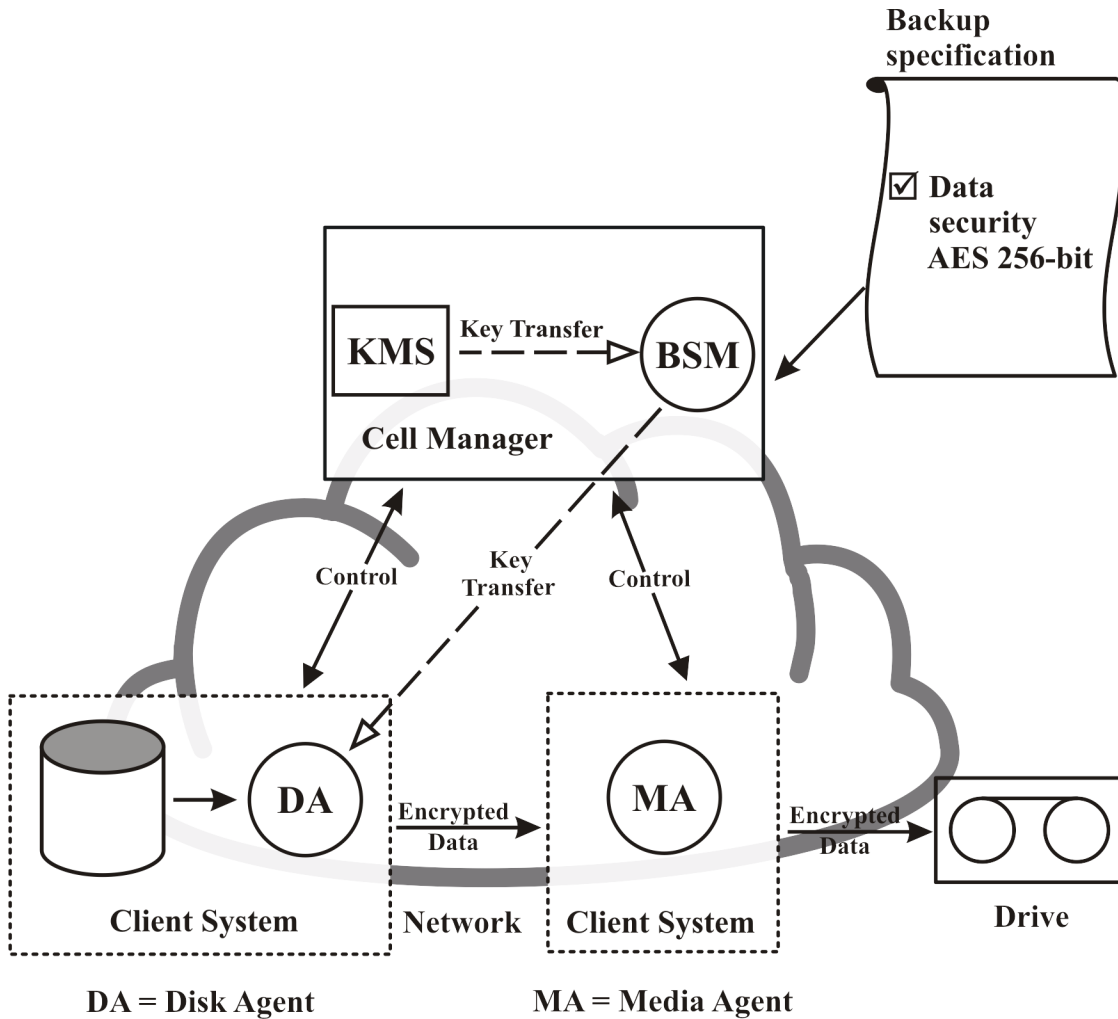
How Data Protector AES 256-bit encryption works

The Backup Session Manager (BSM) reads the backup specification in which the **AES 256-bit** encryption option is selected and requests an active encryption key from the Key Management Server (KMS). The

key is transferred to the Disk Agent (DA), which encrypts the data. Thus the backed up data is encrypted before it is transferred over the network and written to media.

" [Backup session with AES 256-bit encryption](#)" below shows a basic interaction during an encrypted backup session with the **AES 256-bit** encryption option selected.

Figure 14: Backup session with AES 256-bit encryption



How Data Protector drive-based encryption works

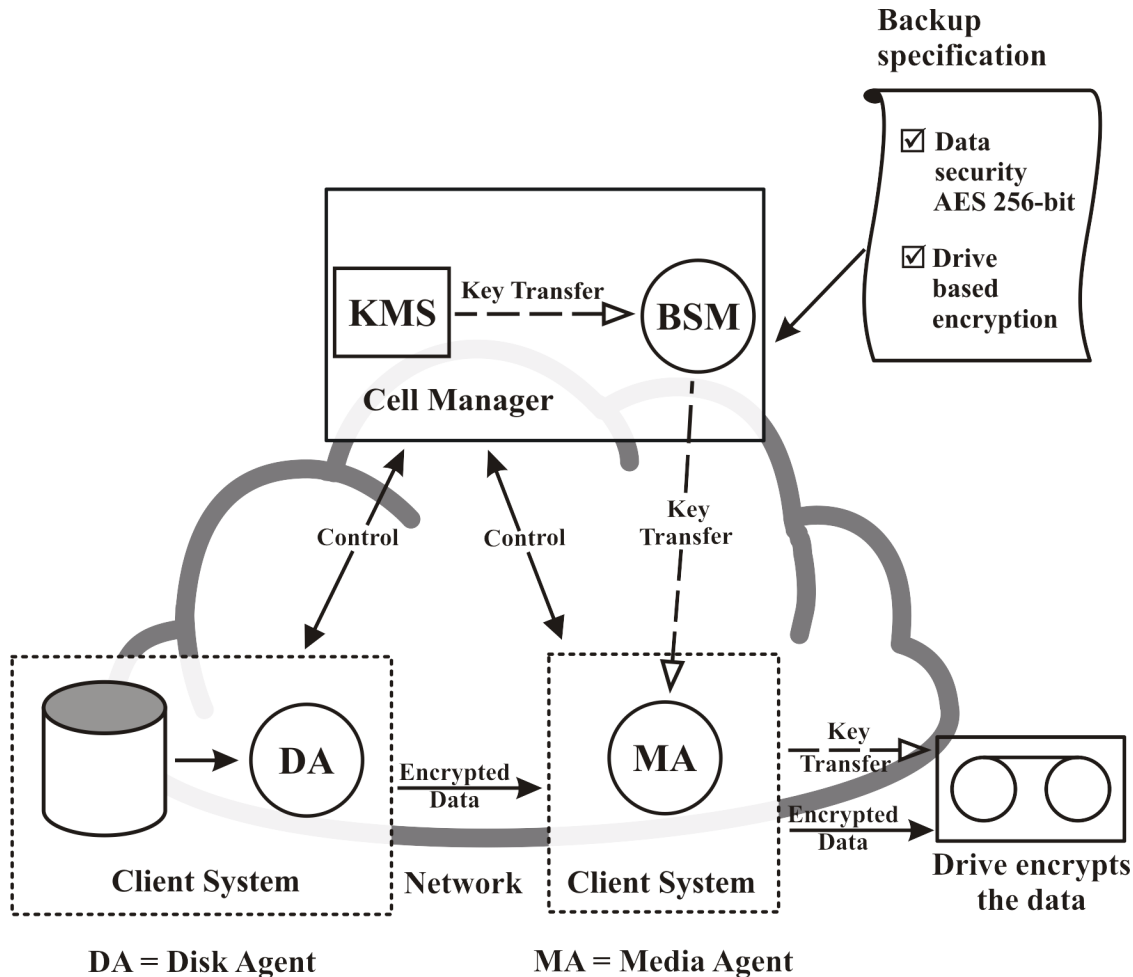
The BSM reads the backup specification in which the **Drive-based encryption** option is selected and requests an active encryption key from the KMS. The key is transferred to the Media Agent (MA), which configures the drive for encryption and sets the encryption key into the drive. The drive encrypts both the data and the metadata that is written to the medium.

In an object copy or object consolidation operation from an encrypted backup, the data is decrypted by the source drives, transferred over the network and encrypted by the destination drives.

If a source medium involved in an automatic media copy session stores encrypted as well as non-encrypted data, all data written to the corresponding target medium will be either encrypted or non-encrypted, depending on current settings for drive-based encryption.

"Backup session with AES 256-bit encryption and drive-based encryption" below shows a basic interaction during an encrypted backup session with the **AES 256-bit encryption** and the **Drive-based encryption** options selected.

Figure 15: Backup session with AES 256-bit encryption and drive-based encryption



Restore from encrypted backups

No additional encryption related preparations are needed for restore of encrypted backups, as Data Protector automatically obtains the appropriate decryption keys.

Encrypted control communication

Data Protector encrypted control communication that helps preventing unauthorized access is based on Secure Socket Layer (SSL), a cryptographic protocol, which provides security for communications over the network. SSL encrypts the segments of network connections and encapsulates the existing Data Protector communication protocol.

Control communication in a Data Protector cell is all communication between Data Protector processes, except the data transfer from Disk Agent (and Integrations) to Media Agent, and the other way round. Data Protector uses export grade SSLv3 algorithms, with up to 512-bit keys for asymmetric, and up to 64-bit keys for symmetric encryption, to encrypt control communication. Since the SSL requires certificates

to establish encrypted communication, Data Protector provides default certificates during the installation or upgrade.

How Data Protector encrypted control communication works

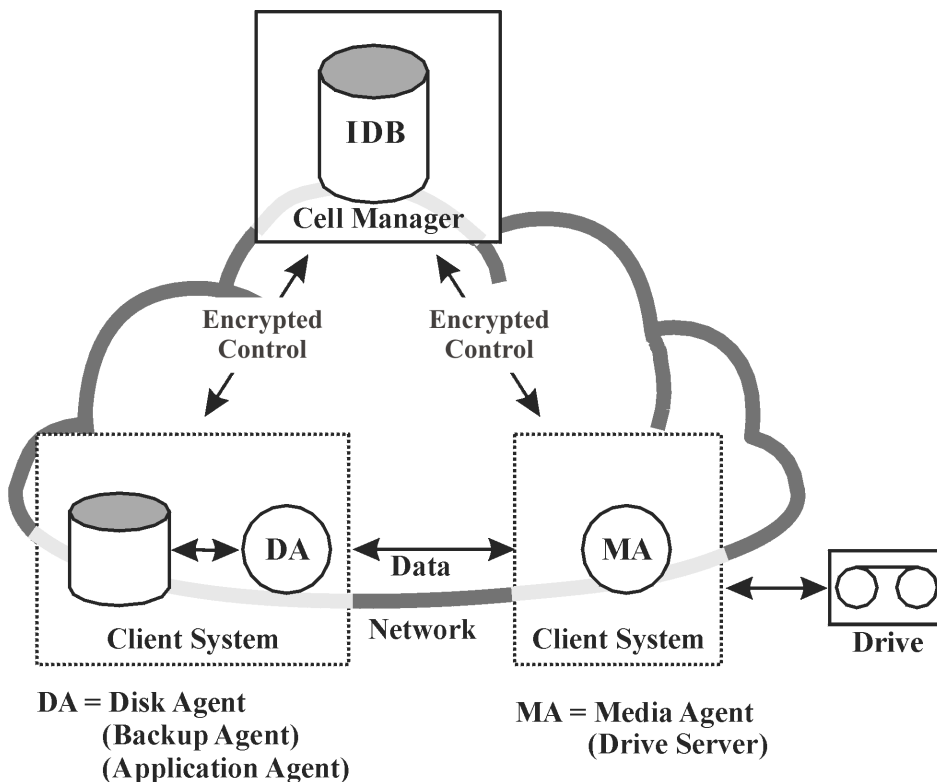
The Data Protector Cell Manager controls backup and restore sessions, which perform all the required actions for a backup or restore, respectively. Encryption is enabled on a per-client basis, which means that encryption is either enabled or disabled for all control communication with the selected client.

You must first enable encryption on a Cell Manager and then on the clients in the cell. Clients that are not supposed to communicate confidentially can be placed in a Cell Manager exception list, which allows those clients to communicate in non-encrypted mode. Before accepting a connection, Data Protector processes check the local configuration (that is whether encryption is enabled, which certificate to use, and so on). The SSL connection is established with local configuration encryption parameters.

During the backup session, the Data Protector processes read encryption parameters from the local configuration. The Backup Session Manager (BSM) establishes an SSL connection with the Media Agent (MA) and the Disk Agent (DA), and the regular Data Protector communication protocol follows. The DA then establishes a data connection to the MA. Data backup follows.

"[Encrypted control communication](#)" below shows basic communication interactions within the Data Protector cell if the encrypted control communication is enabled.

Figure 16: Encrypted control communication



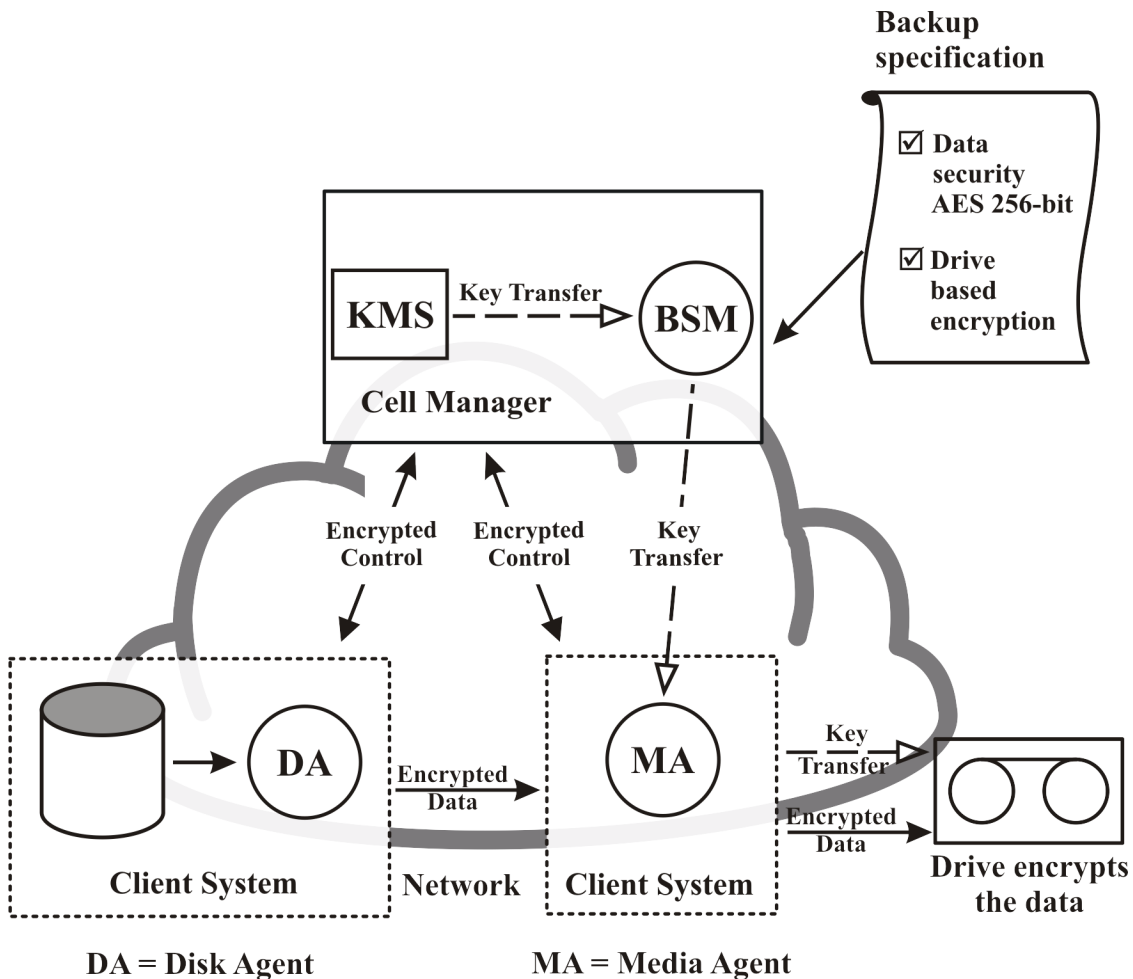
Data encryption and encrypted control communication

By combining data encryption with encrypted control communication, you can with ease maximize the security of your system:

- Software (AES 256-bit) encryption encrypts data before it is transferred over the network and written to media
- Hardware (drive-based) encryption of your backups prevents unauthorized access to your data during media storage and transportation
- Encrypted control communication provides secure communication between the clients in the cell

" [Encrypted control communication and data encryption](#) " below shows basic interactions within the Data Protector cell during an encrypted backup session with the **AES 256-bit encryption** and the **Drive-based encryption** options selected, and encrypted control communication enabled.

Figure 17: Encrypted control communication and data encryption



Clustering

Cluster concepts

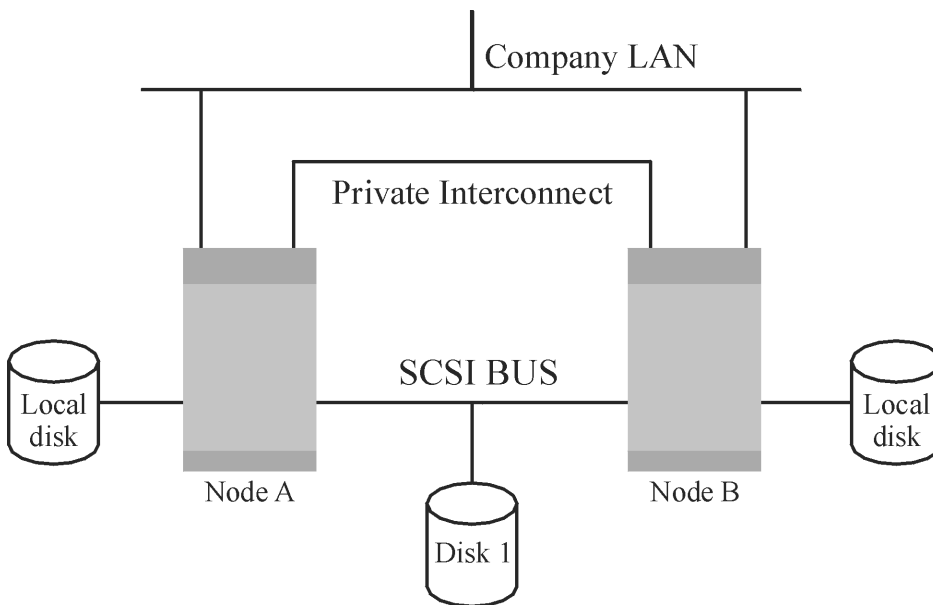
What is a cluster?

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 and MC/Service Guard for HP-UX. For a list of all supported clusters, see the *HP Data Protector Product Announcements, Software Notes, and References*.

Figure 18: 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) 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)
- 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 so that their corresponding sessions are 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. For more information on how to define the criteria, see the *HP Data Protector Help* index: "cluster, managing backups".

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 automatically *restart* operations in the cell 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.

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

For more information on backup session behavior options on failover of the Data Protector Cell Manager, see the *HP Data Protector Help* index: "cluster, backup specification options".

Example cluster environments

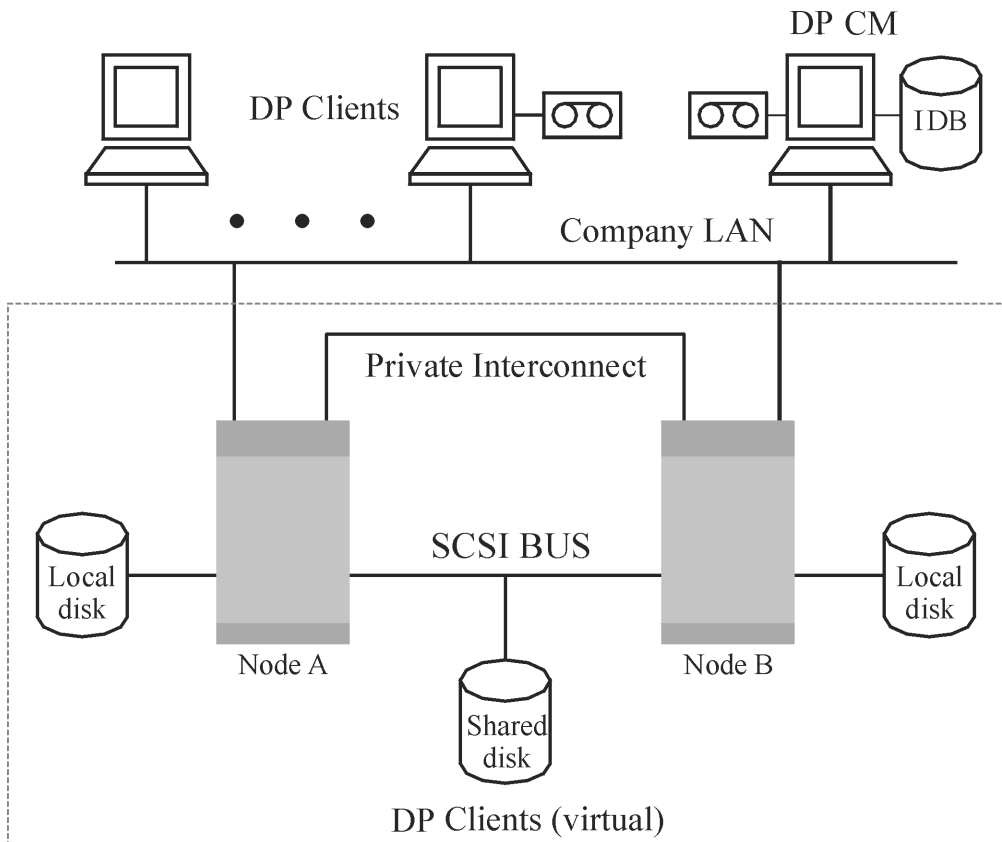
This section gives three example cluster configurations.

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 19: 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.

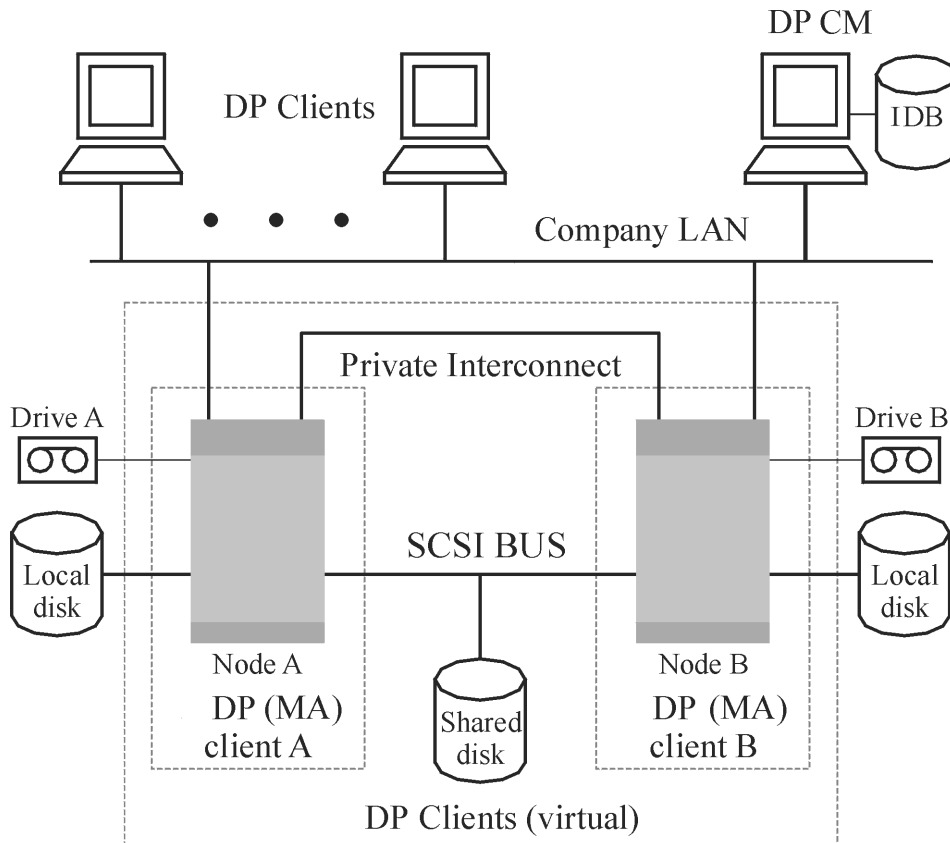
For more information on how to define these options, see the *HP Data Protector Help* index: "cluster, backup specification options".

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 20: 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.

Cell Manager installed in a cluster, devices connected to the cluster nodes

In the environment depicted below:

- The Cell Manager installed in a cluster.

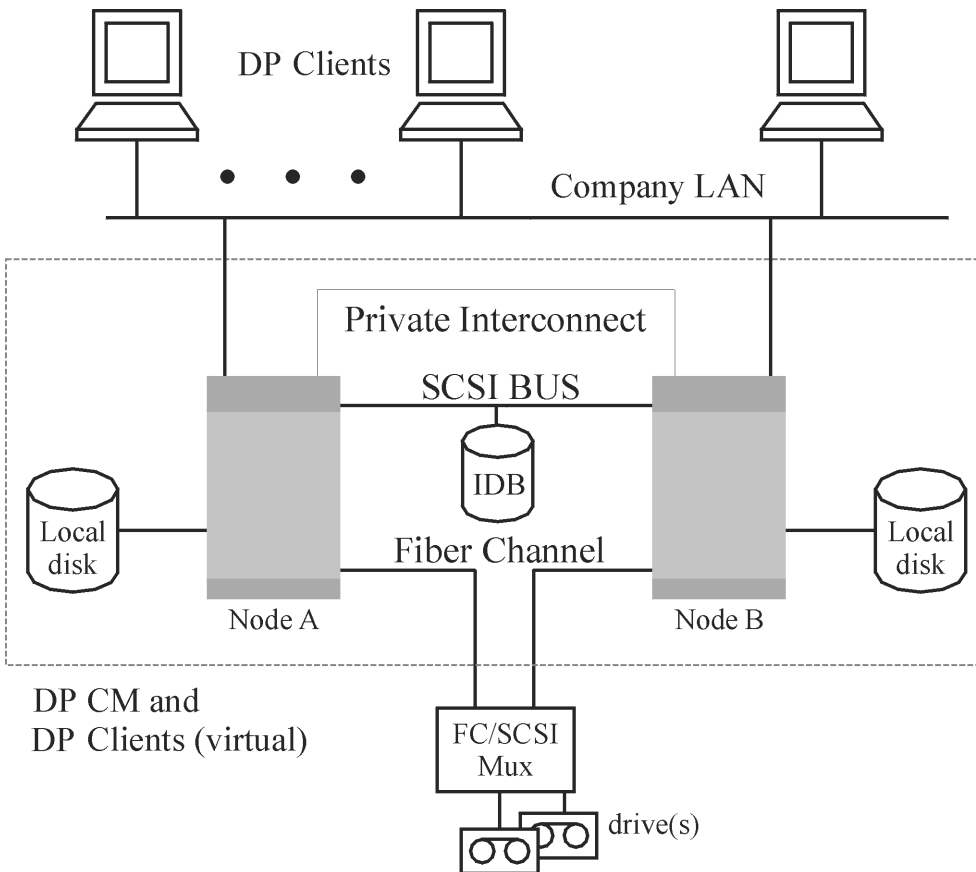
With regard to the Data Protector application integrations, there are two possible ways of configuring Data Protector and an application in such a configuration:

- 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 (HP 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 (HP 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 21: 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.

Additionally, the Data Protector cluster Cell Manager/client can be integrated with the EMC Symmetrix or HP P9000 XP Disk Array Family environment, producing a very highly-available backup environment. For more information, see the *HP Data Protector Zero Downtime Backup Administrator's Guide*.

Full, incremental, and synthetic 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 1: 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 split mirror backup, snapshot backup, and data mover backup) are available with Data Protector application integrations. For more information, see the respective *HP 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 ["Synthetic backup" on page 62](#).

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.

Incremental backups depend on the last full backup. If you specify an incremental backup and there is no protected full backup, a full backup is performed instead.

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 main 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.

Incremental backup using Change Log Provider

You can perform enhanced incremental or conventional incremental backup using the Windows NTFS Change Log Provider. Change Log Provider queries the Windows Change Journal for a list of changed files instead of performing a time-consuming file tree walk. As the Change Journal detects and records all changes made to the files and directories on an NTFS volume, Data Protector can use it as a tracking mechanism to generate a list of files modified since the last full backup. This improves the incremental backup speed, especially in environments containing millions of files only a few of which have changed, and allows to eliminate unnecessary full backups.

Types of incremental backups

Data Protector provides incremental backups of different types:

Incr	A simple incremental backup, shown in "Incremental backups" below , is based on the last backup that is still protected, which can be a full backup or an incremental backup.
Incr1-9	A leveled incremental backup , shown in "Leveled incremental backups" on the next page , depends on the last backup of the next lower level that is still protected. For example, an Incr1 backup saves all changes since the last full backup, while an Incr5 backup saves all changes since the last Incr4 backup. An Incr1-9 backup <i>never</i> references an existing Incr backup.

Figure 22: Incremental backups

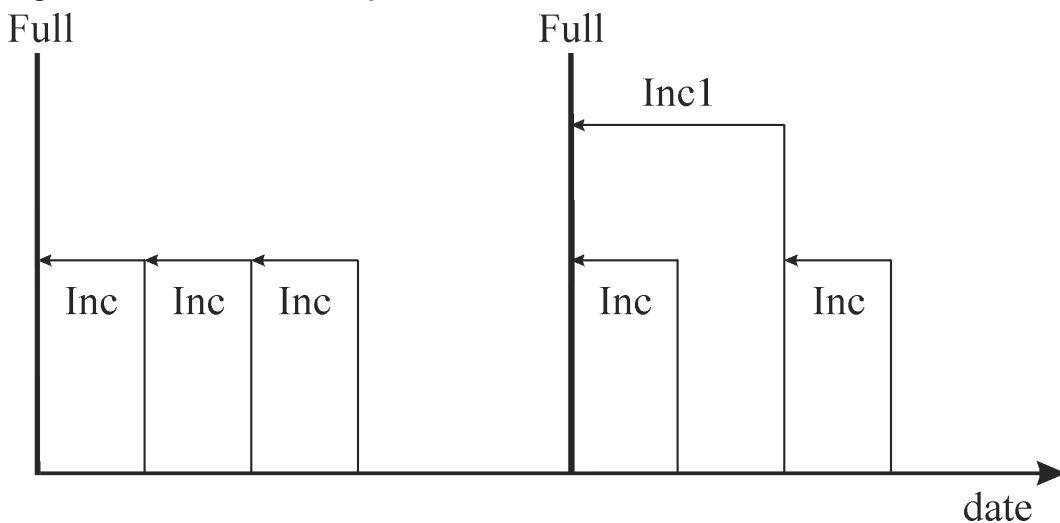
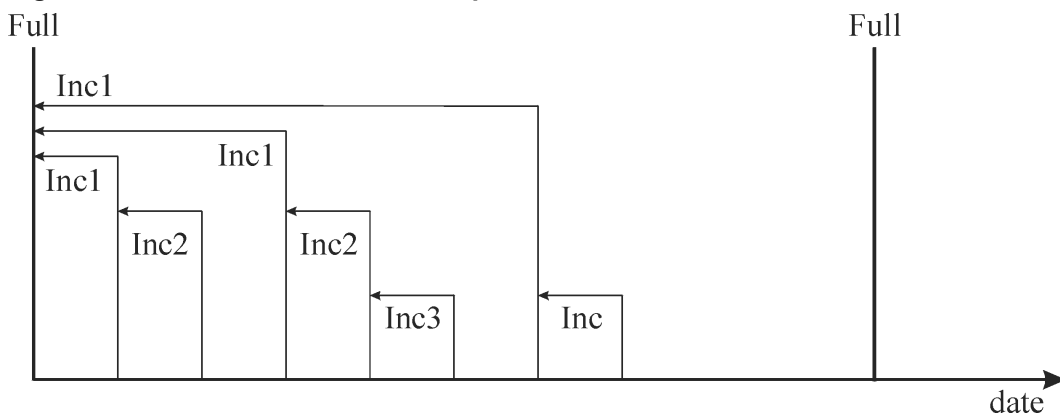


Figure 23: Levelled incremental backups



"Relative referencing of backup runs" below shows the relative referencing of backup runs with various backup types. See the text following the table for a full explanation.

Table 2: Relative referencing of backup runs

1	Full	<---	Incr1				
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 "Relative referencing of backup runs" above

- The rows in "Relative referencing of backup runs" above 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.

Backup generations

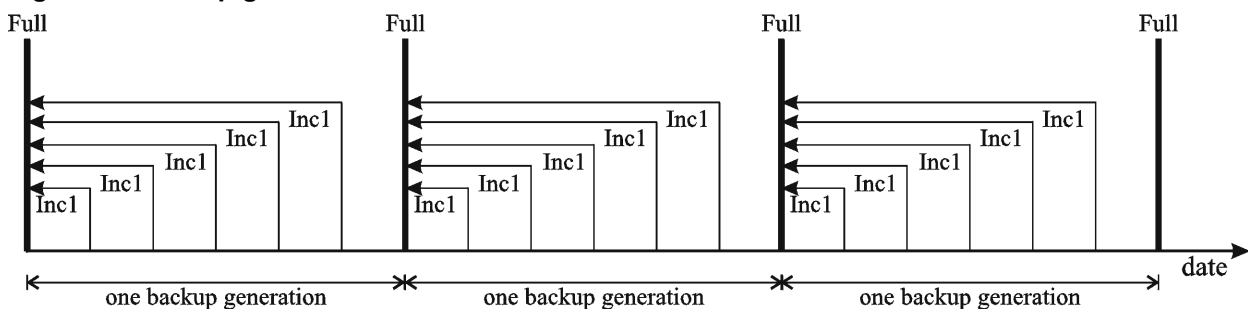
Data Protector provides a time/date related protection model. It is easy to map a generation-based backup model to the time-based model, assuming regular backups are done.

What is a backup generation?

A backup generation, shown in "[Backup generations](#)" below, 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 24: 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.

You can set automatic media rotation (for the media with expired protection time) through an appropriate pool usage concept. For more information, see "[Implementing a media rotation policy](#)" on page 133.

Synthetic backup

This section introduces the concept of synthetic backup and explains the synthetic backup solution provided by Data Protector.

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.

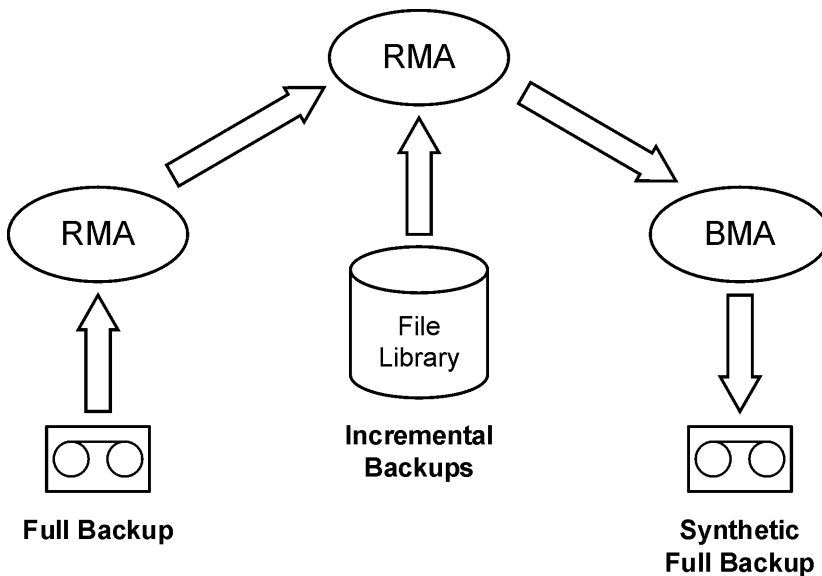
To enable synthetic backup, the use of enhanced incremental backup is required. Enhanced incremental backup must be turned on before the full backup and the incremental backups are performed.

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 that uses distributed file media format, 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.

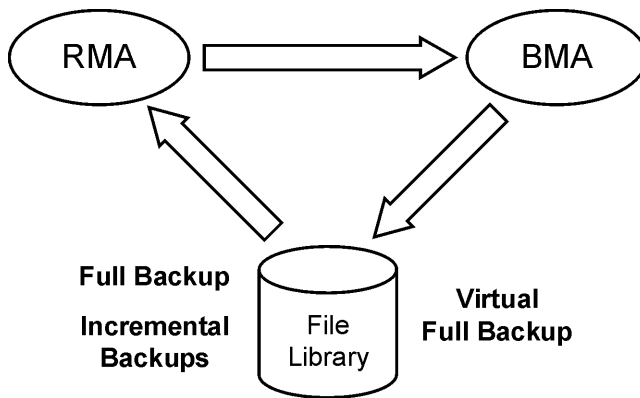
Figure 25: Synthetic backup



"Synthetic backup" above 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 26: Virtual full backup



"Virtual full backup" above shows how a virtual full backup is created. With this type of backup, all the backups reside in a single file library that uses distributed file media format. 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.

Synthetic backup and media space consumption

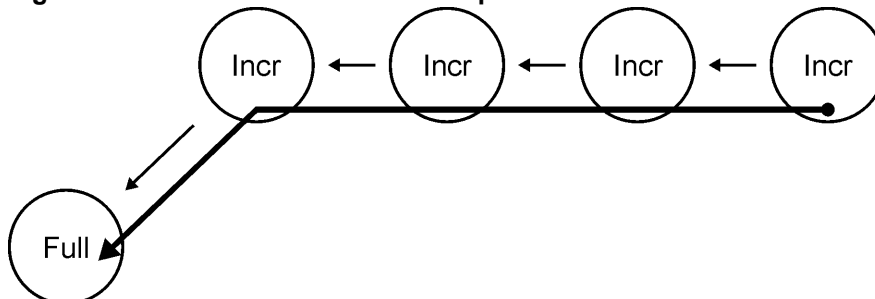
If synthetic backups are performed frequently, and the sources are kept, this typically means significant space consumption on the backup media. However, if virtual full backup is performed, the backup media space consumption is minimized.

With virtual full backup, the space consumption largely depends on the size of the backed up files. If the files are significantly larger than the block size used, virtual full backup achieves maximum savings of the space compared to normal synthetic backup. On the other hand, if the files are smaller than the block size, the savings are rather small.

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.

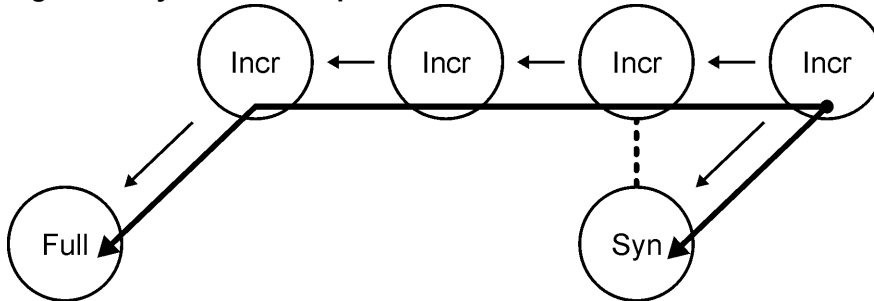
Figure 27: Full and incremental backups



In ["Full and incremental backups"](#) on the previous page, 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.

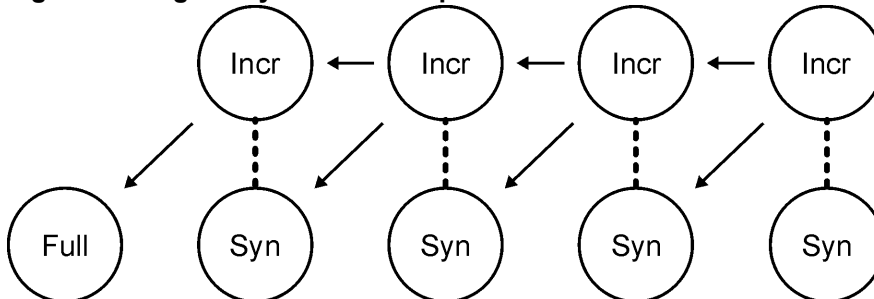
Such a restore can take a considerable amount of time, as each incremental backup must be read. If tape devices are used, time is spent for loading and unloading of several media and seeking for object versions to restore.

Figure 28: Synthetic backup



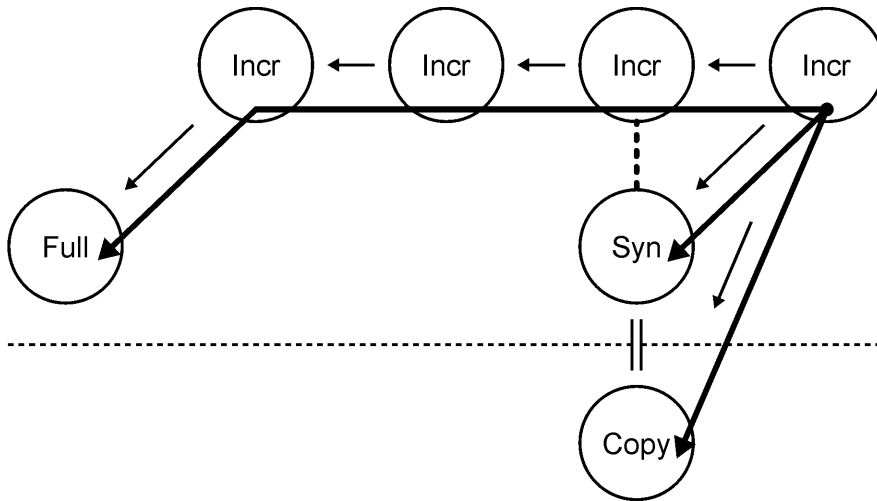
In ["Synthetic backup"](#) above, a synthetic full backup exists, which is used for restore by default. The restore chain consists of only two elements, namely the synthetic full backup and the subsequent incremental backup. The restore is significantly simpler and quicker than that without the synthetic full backup. In the figure, both possible restore chains are shown.

Figure 29: Regular synthetic backup



["Regular synthetic backup"](#) above 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 30: Synthetic backup and object copy



In "[Synthetic backup and object copy](#)" on the previous page, 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.

How data protection periods affect restore from synthetic backup

Data protection of a conventional full backup and all incremental backups that precede synthetic full backup does not compromise a successful restore.

By default, the last synthetic full backup in the backup chain is used for restore, irrespective of whether the preceding backups are still valid or their protection has already expired and the objects are removed from the IDB.

For additional safety, set data protection to permanent so that data on the media is not overwritten unintentionally.

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.

Using simple and leveled incremental backups, as indicated in "[Media needed to restore from simple and leveled incremental backups](#)" on the next page, 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 "[Media needed to restore from leveled incremental backups](#)" on the next page. 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 Incr1 media set for this restore, unless the desired restore point in time would be moved.

Figure 31: Media needed to restore from simple and leveled incremental backups

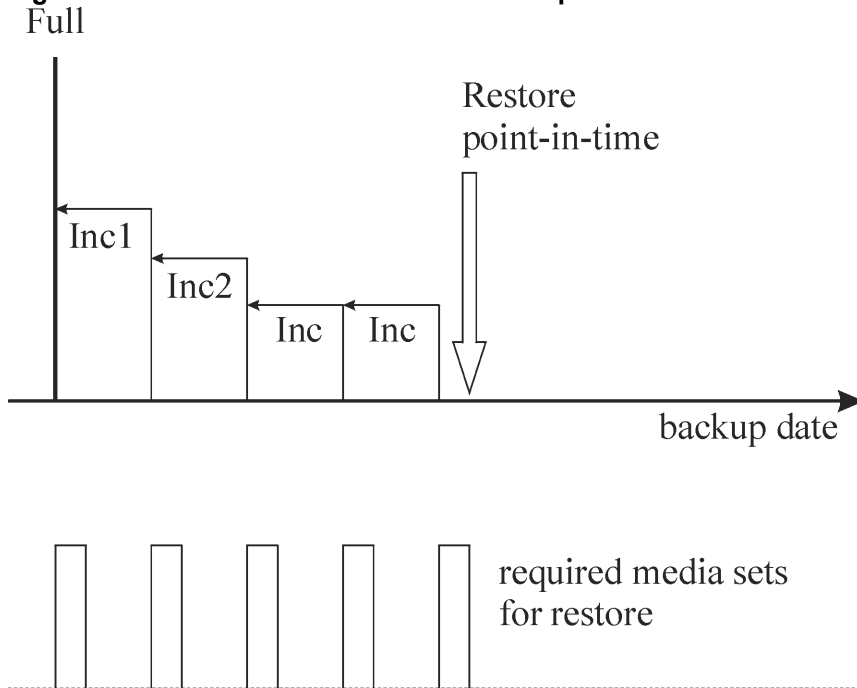
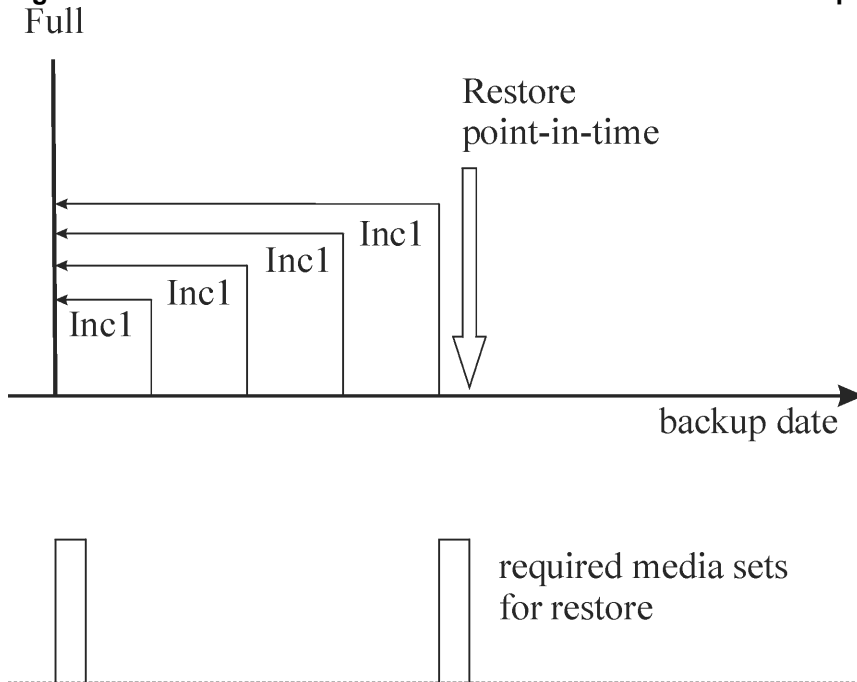


Figure 32: Media needed to restore from leveled incremental backups



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.

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, see the *HP Data Protector Help* index: "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.

For more information on how catalog protection settings influence the IDB growth and performance, see the *HP Data Protector Help*.

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, see ["Logging level as an IDB key tunable parameter" on page 153](#).

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.

During export of media, encryption information relevant to the media is also exported and placed in an export directory as a .csv file. This file is required in order to be able to restore any encrypted backups after re-importing or importing to another cell.

Importing WORM media to a cell

Importing media to a Data Protector cell adds all the information about backed up data on the media and the media themselves to the IDB. You can browse, select, or restore files from imported media using the Data Protector user interface. When importing media on which the data protection has already expired to a Data Protector cell, data on the media is overwritten in one of the subsequent backup sessions.

In case of WORM media, overwriting data on the media is not allowed and the media becomes non-appendable. To allow Data Protector to append data to WORM media, set the data protection beyond the date on which the new data will be written to the media to prevent overwriting of existing data.

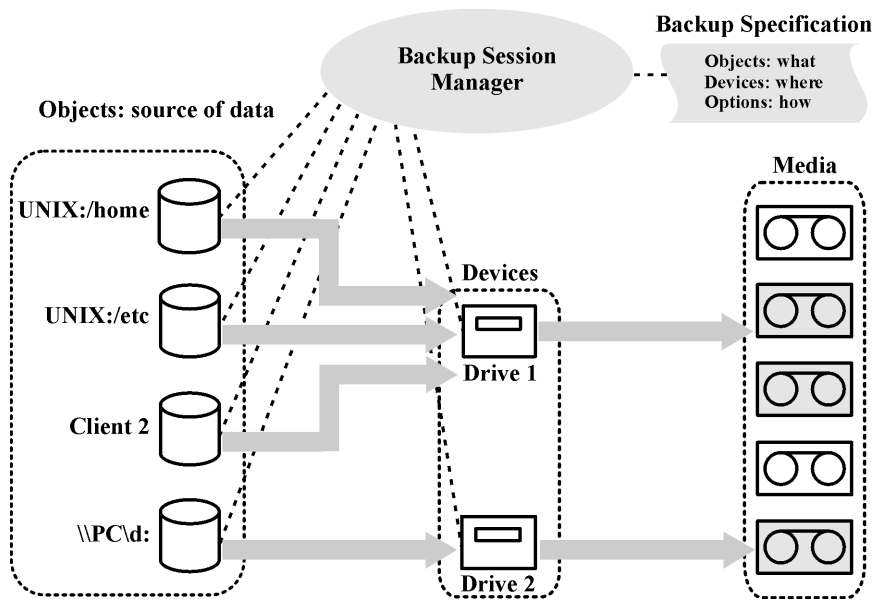
Backing up data

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 33: 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. See ["Backing up with disk discovery" on page 168](#).

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.

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. For more information on logging levels, see ["Logging level as an IDB key tunable parameter" on page 153](#).

- Automatic load balancing

Dynamic device allocation from a specified list. For more information, see ["How load balancing works" on page 102](#).

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, see ["Pre-exec and post-exec commands" on page 165](#).

- Data security

Level of security to be applied to the data.

Data Protector provides three levels of security for backed up data:

- None
- AES 256-bit
- Encode

For more information on encryption, see ["Data encryption" on page 45](#).

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.

For advanced information on backup sessions, and how to control sessions, see ["How Data Protector operates" on page 162](#).

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 88](#).

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 ["Comparison of full and incremental backup" on page 57](#).

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. See ["Staggering full backups" on the next page](#).

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 ZDB to disk, 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, incremental, and synthetic backups" on page 57](#) and ["Keeping backed up data and information about the data" on page 68](#) 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" below](#).

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 3: The staggered approach

	Mon	Tue	Wed	...
system_grp_a	Full	Incr1	Incr1	...
system_grp_b	Incr1	Full	Incr1	...
system_grp_c	Incr1	Incr1	Full	...

Optimizing for restore

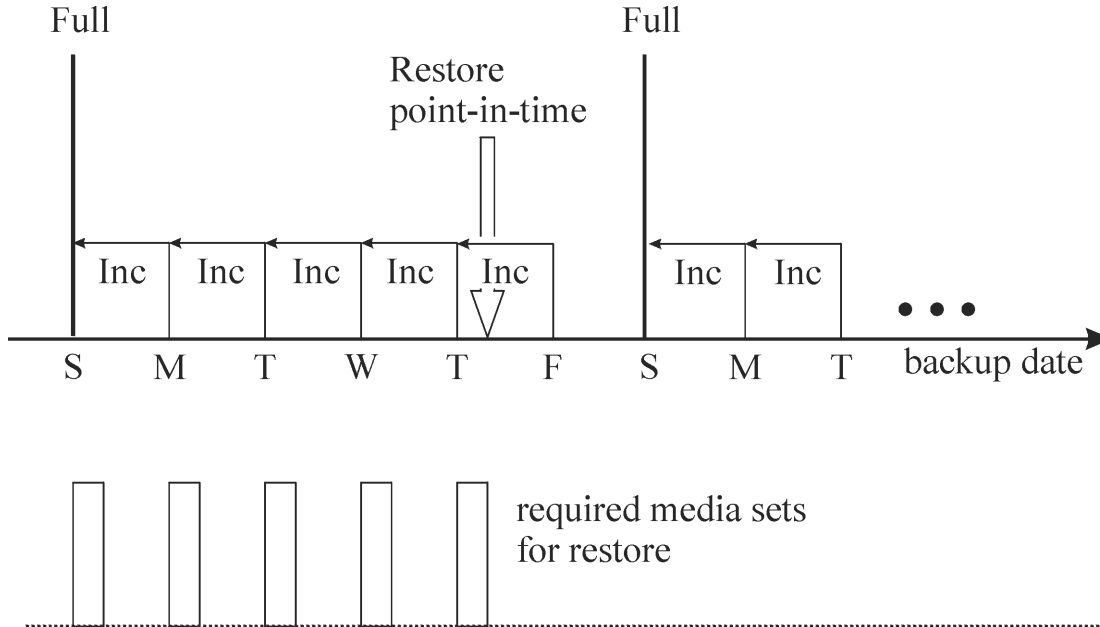
The combination of your scheduling policy with full and incremental backups highly influences the time needed to restore your data. This is illustrated in three examples in this section.

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. For more information on how Data Protector selects media for backups, see ["Selecting media for backups" on page 137](#).

Example 1

"Full backup with daily simple incremental backups" below depicts a scheduling policy based on a full backup plus simple incremental backups.

Figure 34: Full backup with daily simple incremental backups

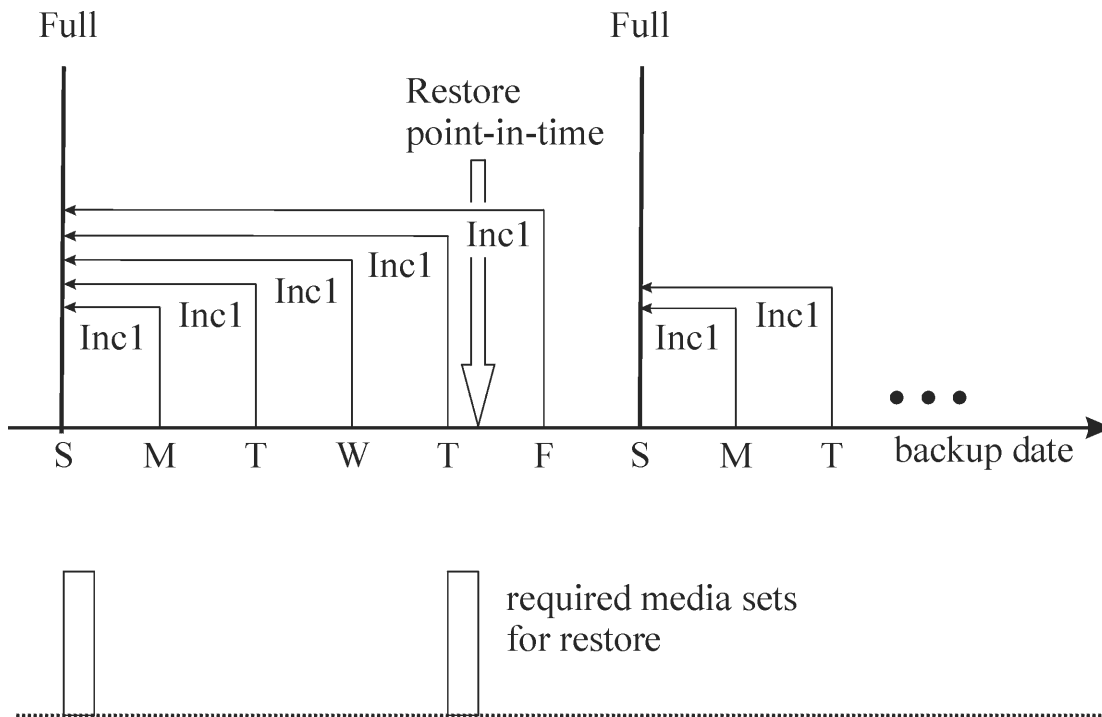


This policy reduces the media space and time needed for backing up, because you only back up changes from the previous day. However, to restore files from a Thursday backup, you need to provide the media for the full and each of the incremental backups until Thursday, that is five media sets. This complicates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

Example 2

"Full backup with daily level 1 incremental backups" on the next page depicts a scheduling policy based on a full backup plus level one incremental backups.

Figure 35: Full backup with daily level 1 incremental backups

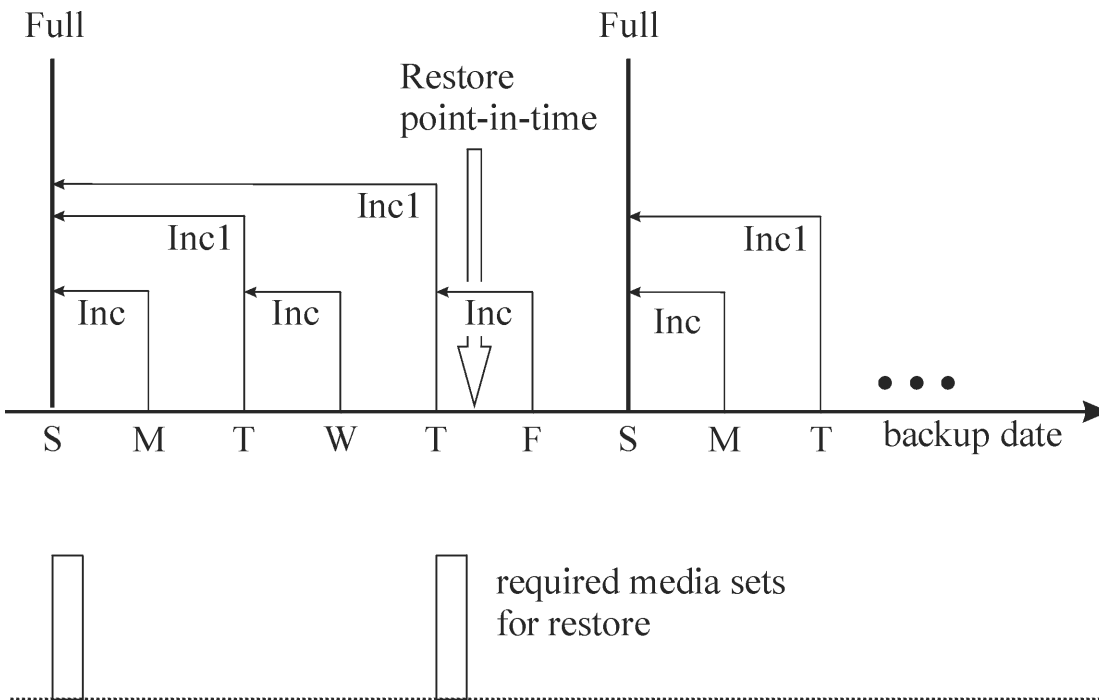


This policy requires slightly more time for backups and also requires a little more media since you back up all the changes from the last full backup every day. To restore files from Thursday's backup, you need to provide media for the full and for Thursday's incremental backup, that is, two media sets 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 36: 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 simple incremental backups and Incr1 (differential) backups to optimize backup performance. To restore files from Thursday's backup, you need to provide media from the full backup and the second Incr1 backup, 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.

- How many media are required?

Define a media rotation policy. See ["Implementing a media rotation policy" on page 133](#). 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. or more information, see ["Large libraries" on page 107](#).

- 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 files open, so running a backup would produce an inconsistent backup. This can be avoided by using 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 "[Data Protector data duplication methods](#)" below for an overview of the main characteristics of these methods.

Table 4: Data Protector data duplication methods

	Object copy	Replication	Object mirror	Media copy
What is duplicated	Any combination of object versions from one or several backup, object copy, or object consolidation sessions	A set of objects from a backup session, object copy session, or object consolidation session	A set of objects from a backup session	An entire medium
Time of duplication	Any time after the completion of a backup	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	Data can be replicated only to B2D devices of the same type	Can be different	Must be the same
Size of source and target media	Can be different	The target device must have enough space for the deduplicated data	Can be different	Must be the same
Appendability of target media	Yes	N/A	Yes	No ¹
Result of the operation	Media containing the selected object versions	An identical copy stored on the target B2D device	Media containing the selected object versions	Media identical to the source media

¹You 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

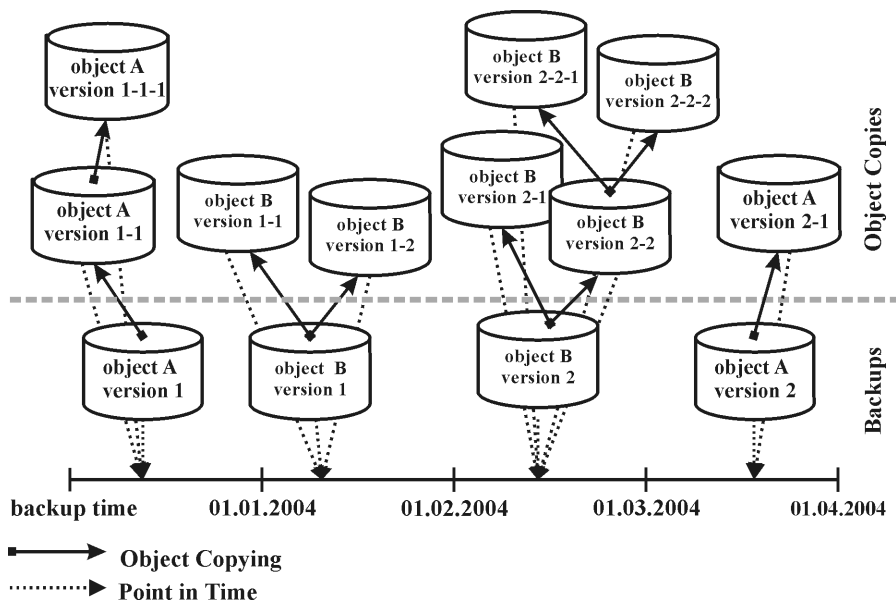
What is object copy?

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, object copy, 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.

"Object copy concept" below 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 37: 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 as well as post-copy and post-consolidation object copying, which are subsets of post-backup object copying, take place after the completion of a session that is specified in the automated

object copy specification. They copy 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.

Selection of source devices

By default, Data Protector automatically selects the source devices for an object copy according to device policies set within the device configuration. This ensures optimum usage of the available resources. You can disable the automatic device selection, if you want to use the original device, or select a specific device:

- Automatic device selection (default):

Data Protector will automatically use available source device. This device is selected for an object copy and is from the same library and of the same media type (for example, LTO) as the replaced original one.

Data Protector attempts to use the device that was used for writing the object (the original device) first. If the original device is not selected for an object copy, then a global option is considered. To use alternative device first or to prevent the use of the original device all together, modify the global option `AutomaticDeviceSelectionOrder`.

You can group devices into device groups for different purposes by specifying a device tag. Devices with the same tag are considered compatible and can substitute each other. The unavailable original devices can be replaced with the alternative devices which have the same device tag and are from the same library. By default, no device tags are defined.

Note that if the original device was deleted, a device from the same library and of the same media type replaces it. It is not examined whether this device is selected for an object copy and has the same device tag as the original had.

The object copy can be started with fewer devices than were used during backup.

- Original device selection:

Data Protector will use the original device as a source device for an object copy, and will wait in case the device is unavailable.

Selection of destination 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 94](#).

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, copied, or consolidated data are created for multiple purposes:

- Vaulting

You can make copies of backed up, copied, or consolidated 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

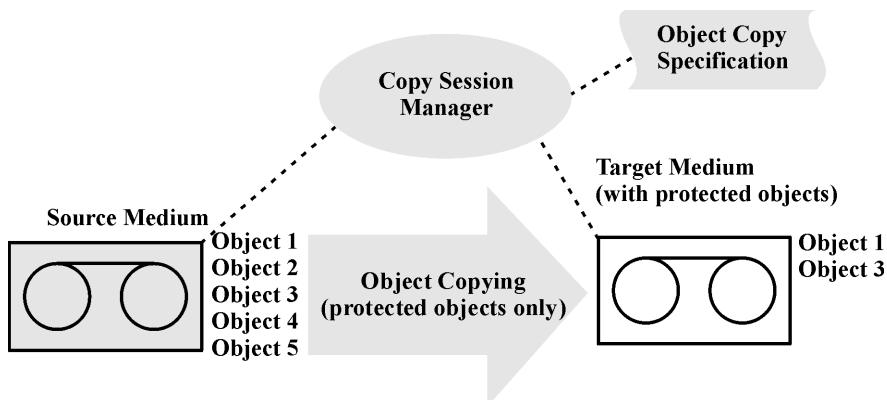
Vaulting is a process of storing media in a safe place, often called a vault, where they are kept for a specific period of time. For details, see ["Vaulting" on page 140](#).

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 ["Freeing media" below](#).

Figure 38: Freeing media

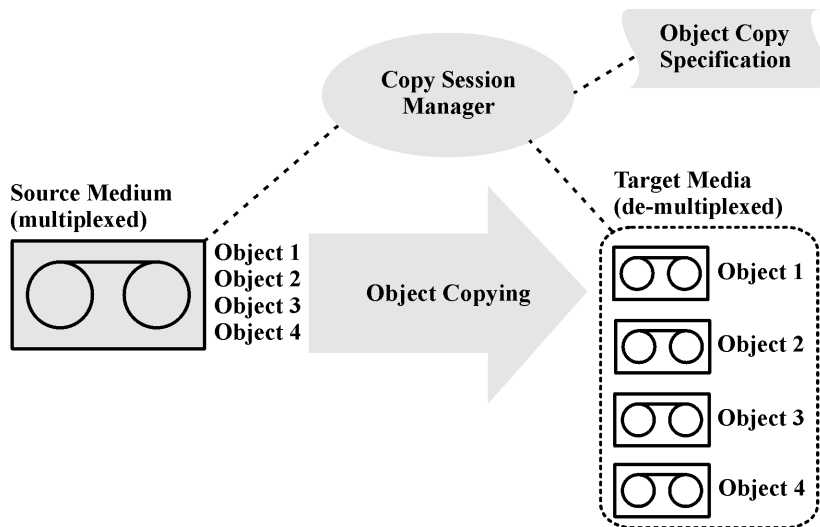


Demultiplexing of media

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 ["Demultiplexing a medium" on the next page](#).

Figure 39: 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

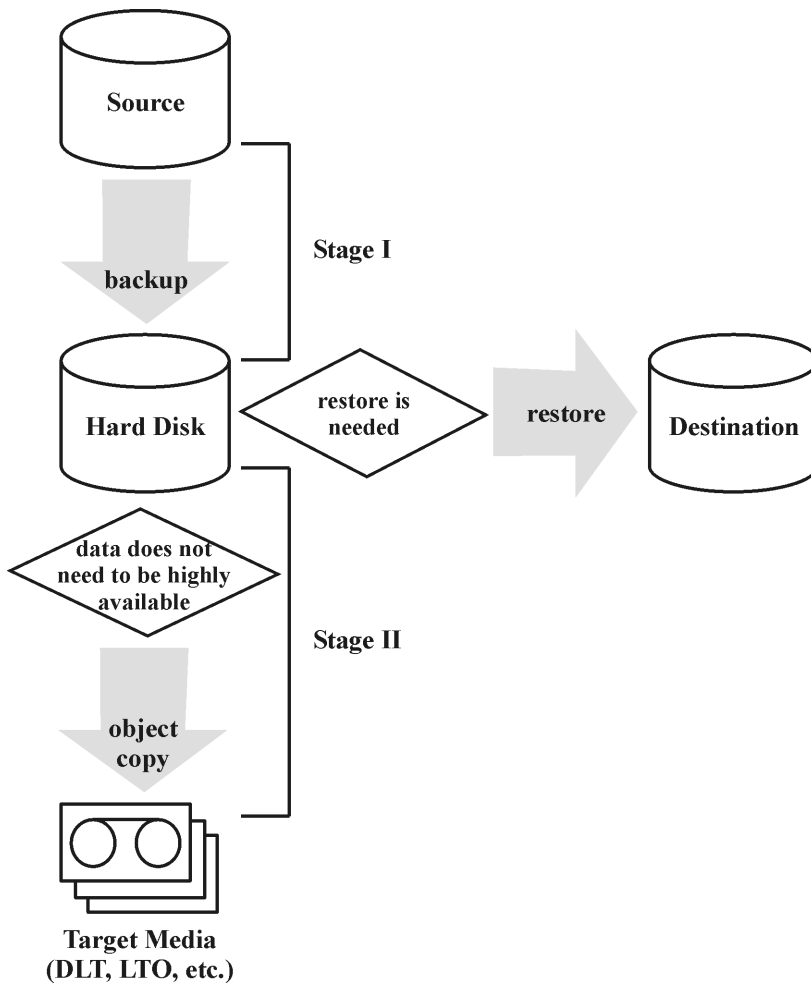
You can migrate backed up data to another media type. For example, you can copy objects from file devices to LTO devices or from DLT devices to LTO devices.

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 ["Disk staging concept" on the next page](#).

Figure 40: Disk staging concept



This process can be performed as an automated operation.

Consider the following example, which briefly describes an approach simple to implement as a standard operation, while providing extra data security. It uses options for setting source and target protection independently. The requirement is for fast restore capability from disk for the first 15 days and then standard restore from tape for a further 30 days.

- The initial backup is performed to disk using a file library, with the data and catalog protection set to the overall requirement of 45 days.
- A post-backup copy operation is then performed, in which the backup objects are copied to tape, leaving the initial backup on the file library. If the copy to tape is successful, the data and catalog protection for it are set to 45 days.
- A successful copy having been created, the protection time for the disk backup can be reduced to 15 days, the period for which fast restore is required. After this time, it can be deleted, leaving the tape copy for longer term security. Until then, the tape copy provides extra security in case the disk copy is damaged.

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.

Replication

What is replication?

The Data Protector replication functionality enables you to replicate objects between two Backup to Disk (B2D) devices capable of replication, without transferring data through Media Agents.

You can select one or several backup sessions, object copy sessions, or object consolidation sessions. During the replication session, Data Protector reads the object from one backup session and initiates the replication from the source B2D device to the target device. The result of a replication session is a copy of all objects from the session you specified.

You can enable replication when you create an object copy specification by choosing devices capable of replication and selecting appropriate options in the object copy specification.

Start of replication session

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

Post-backup replication

Post-backup as well as post-copy and post-consolidation replication, which are subsets of post-backup replication, take place after the completion of a session that is specified in the automated replication specification. They replicate objects selected according to the automated replication specification that were written in that particular session.

Scheduled replication

Scheduled replication takes place at a user-defined time. Different sessions can be replicated in a single scheduled replication session.

Selection of devices

You need separate devices to be used as the source device and the target device. Only B2D devices can be used for replication.

Why use replication?

Replication can be used for many of the purposes that object copying is used for (such as vaulting and so on), except for operations involving media. See also ["Why use object copy?" on page 83](#).

Additionally, compared to object copying, replication between B2D devices offers the following advantages:

- Data is replicated directly between the B2D devices. This reduces the load on the Media Agent clients.
- Only unique (deduplicated) data is replicated. This reduces the network load.

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:

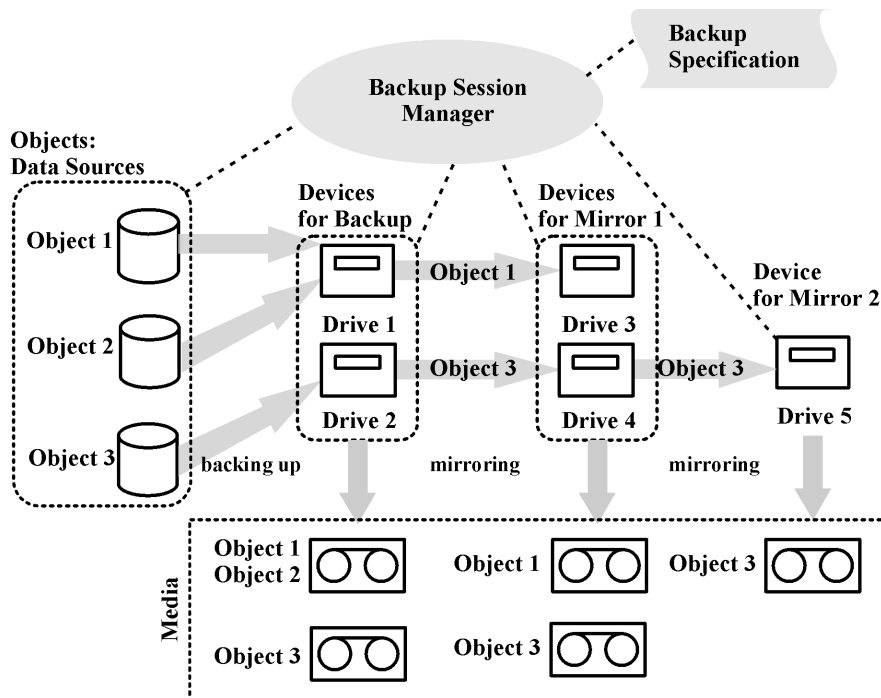
- It increases the availability of backed up data due to the existence of multiple copies.
- It enables easy multi-site vaulting, as the backed up data can be mirrored to remote sites.
- It improves the fault tolerance of backups, as the same data is written to several media. A media failure on one medium 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 ["Object mirroring" on the next page](#).

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 41: 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, see ["Automated media copying" below](#).

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, see the *HP Data Protector Help*.

Verifying backup media and backup objects

As a backup administrator, it is not sufficient to merely backup your important data regularly. It is just as important to have confidence that you will be able to restore the backed-up data successfully in the event of problems, particularly with some of the more sophisticated backup techniques now available. With Data Protector backup media and backup object verification, you have the ability to check restore capability to various levels of confidence.

What is media verification?

Data Protector media verification allows you to check whether the data format of any medium is valid and update the information about the medium in the IDB. You can use this to interactively check any complete, single, Data Protector resident medium. Examples of times when you might want to use media verification are:

- You have copied a medium for archive purposes and you want to check the validity of the copy before placing it in a vault.
- A backup medium has become full and you would like to check all the objects on it before sending for long-term storage.

What does media verification do for you?

When you run media verification, Data Protector:

- Checks the media identification, description, and location information in the Data Protector headers
- Reads all blocks on the medium and verifies block format
- If a cyclic redundancy check (CRC) was performed during backup, recalculates the CRC and compares it with the one stored on the medium

The first two checks, if successful, confirm that the hardware status of the tape is good and that all data could be read from it successfully, providing a medium level of confidence in restore capability from that medium.

The third check, if successful, confirms that the backup data itself is consistent within each block, giving a high level of confidence in restore capability from that medium.

What is object verification?

Data Protector object verification allows you to check the validity of backup objects, as opposed to backup media. You can use it to check:

- single or multiple objects
- on single or multiple media
- interactively, or in scheduled or post-operation sessions

You might want to use object verification:

- after an object copy to a different medium
- after performing object consolidation on the restore chain of an object backed up incrementally
- to check all backup objects produced within a specified time-frame after a backup device change

What does object verification do for you?

When you run object verification, Data Protector provides the same levels of data verification as with media verification. However, whereas with media verification it can only check complete single media, with object verification it can check, for instance:

- a single backup object, without having to check the complete medium, potentially saving a lot of time with large backup media
- large objects that span more than one medium
- several objects on several media
- a specific object version (interactive only)

In addition, you can perform the verification on:

- the Media Agent host, avoiding any network traffic
- another host, factoring in network effects

Information on object verification specifications and sessions can be viewed in various Session Specifications and Session in Timeframe reports.

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. For more information, see ["Full, incremental, and synthetic backups" on page 57](#).
- Media and devices used for backup. For more information, see ["Devices and media management" on page 100](#).
- Speed of networks and systems. For more information, see ["Understanding and planning performance" on page 38](#).
- The application you are recovering, for example, Oracle database files. For more information, see the appropriate *HP Data Protector Integration Guide*.
- The use of parallel restore. Several objects can be restored with a single read operation, depending on how the data was backed up. See ["Parallel restores" on page 170](#).
- 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. See ["Logging level as an IDB key tunable parameter" on page 153](#).

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.

Selection of devices

By default, Data Protector automatically selects the devices for a restore according to device policies set within the device configuration. This ensures optimum usage of the available resources. You can disable the automatic device selection, if you want to use the original device, or select a specific device:

- Automatic device selection (default):

Data Protector will automatically use available device. This device is selected for a restore and is from the same library and of the same media type (for example, LTO) as the replaced original one.

Data Protector attempts to use the device that was used for writing the object (the original device) first. If the original device is not selected for a restore, then a global option is considered. To use alternative device first or to prevent the use of the original device all together, modify the global option `AutomaticDeviceSelectionOrder`.

You can group devices into device groups for different purposes by specifying a device tag. Devices with the same tag are considered compatible and can substitute each other. The unavailable original

devices can be replaced with the alternative devices which have the same device tag and are from the same library. By default, no device tags are defined.

Note that if the original device was deleted, a device from the same library and of the same media type replaces it. It is not examined whether this device is selected for a restore and has the same device tag as the original had.

The restore can be started with fewer devices than were used during backup.

- Original device selection:

Data Protector will use the original device for a restore, and will wait in case the device is unavailable. This is the preferred option for Data Protector SAP MaxDB and DB2 UDB integrations. Such databases are usually backed up with interdependent data streams and, consequently, restore must be started with the same number of devices as used during backup.

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 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, and so on. 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 affected 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 affected 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 latest support matrices at <http://support.openview.hp.com/selfsolve/manuals>.

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. Reinstall the operating system from scratch.
2. Reinstall 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 systems)

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.

Chapter 3: Devices and media management

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

Devices

Data Protector supports a number of devices available on the market. For an up-to-date list of supported devices, see <http://support.openview.hp.com/selfsolve/manuals>.

Device types

The devices can be grouped in the following categories:

- Tape devices:
 - Standalone devices. See ["Standalone devices" on page 106](#).
 - Small magazine devices. See ["Small magazine devices" on page 106](#).
 - Large libraries. See ["Large libraries" on page 107](#).
- Disk-based devices. See ["Disk backup" on page 114](#).

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, see <http://support.openview.hp.com/selfsolve/manuals>.

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.

For related information, see ["Backup sessions" on page 163](#).

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.

Device Filtering

What is device filtering?

Device filtering is a mechanism to select device targets for backup based on the active node of the cluster setup.

Device filters can be used to assign backup targets to clients located in close proximity to reduce network traffic. For example, an Oracle RAC metro cluster could span across two data centers. Device filtering is a mechanism to identify the node from which data originates and assign it to the destination target local to that node.

How device filtering works?

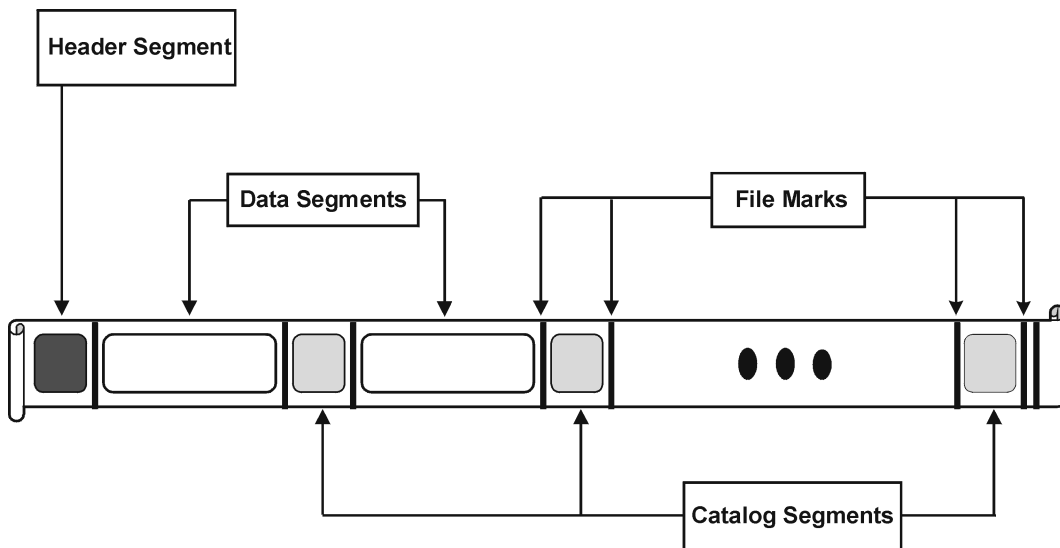
In a cluster setup, each node is identified by a unique name called the host tag. The same name is assigned to devices that could be potential targets. When selecting destination devices in the backup specification, select devices that match all the node host tags. While backing up based on the node in which the data is streamed, the matching device will be used.

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 ["Data format" on the next page](#).

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

Figure 42: 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 most device types is 256 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.

Caution: Before increasing the block size for a device controlled by the Data Protector Media Agent running on a particular operating system, make sure the desired block size does not exceed the default maximum block size supported by the operating system. If the limitation is exceeded, Data Protector cannot restore data from such a device. For information if and how the maximum supported block size can be adjusted, see the operating system documentation.

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:

$$\text{DAConcurrency} * \text{NumberOfBuffers} * \text{BlockSize}$$

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

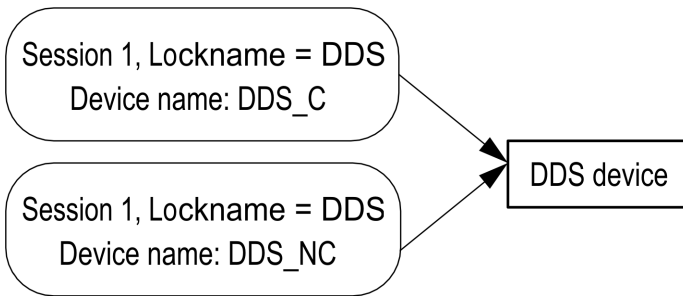
When specifying a device used for backup, you may specify one device name in one backup specification and another device name of the same physical device in a different backup specification. Depending on the backup schedule, this may result in Data Protector trying to use the same physical device in 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 ["Device locking and device names" on the next page](#). Specify the same lockname, DDS, for both devices.

Figure 43: 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, see <http://support.openview.hp.com/selfsolve/manuals>.

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 Libraries, StorageTek/ACSLs and ADIC/GRAU AML. For a list of supported devices, see <http://support.openview.hp.com/selfsolve/manuals>.

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. See "[Implementing a media rotation policy](#)" on page 133.

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.

For more information, see the documentation about your device.

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.

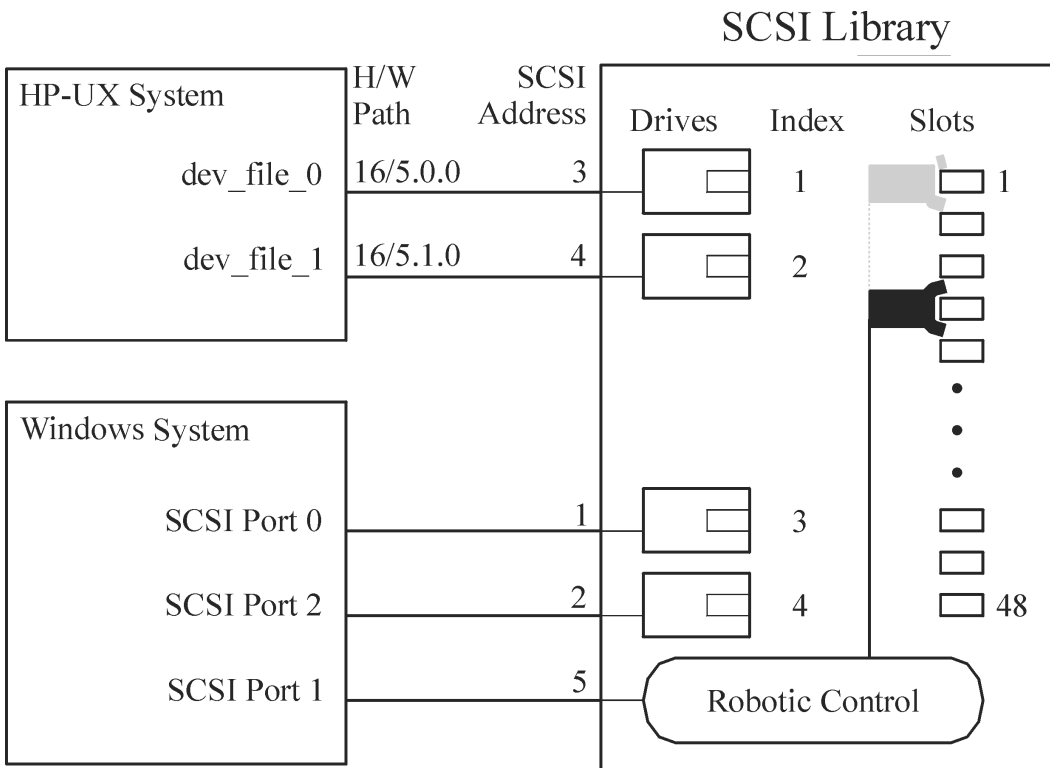
You cannot continue your backup without cleaning the drive, since the backup may fail because data may not be correctly written and stored on the media.

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 44: 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.

"[Required Data Protector Media Agent for drive control](#)" below 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 5: Required Data Protector Media Agent for drive control

	Drive control protocol	
	NDMP	SCSI
Robotic control protocol (ADIC/GRAU, StorageTek ACS, SCSI, 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.

"[Required Data Protector Media Agent for robotic control](#)" below 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.

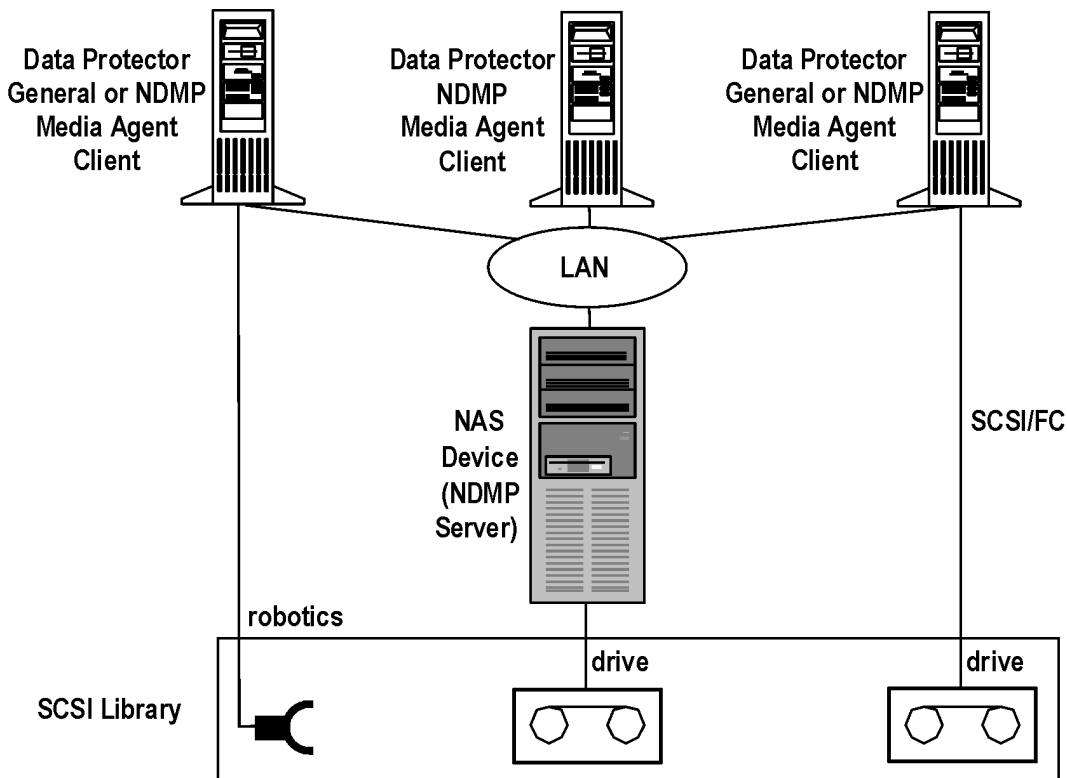
Table 6: Required Data Protector Media Agent for robotic control

	Robotic control protocol			
	ADIC/GRAU	StorageTek ACS	SCSI	NDMP
Drive control protocol (NDMP, 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

"[Sharing a SCSI library \(robotics attached to a Data Protector Client System\)](#)" on the next page through "[Sharing an ADIC/GRAU or StorageTek ACS library](#)" on page 114 show exemplary configurations of shared drives in libraries and Data Protector Media Agents distributions in such configurations.

Figure 45: Sharing a SCSI library (robotics attached to a Data Protector Client System)

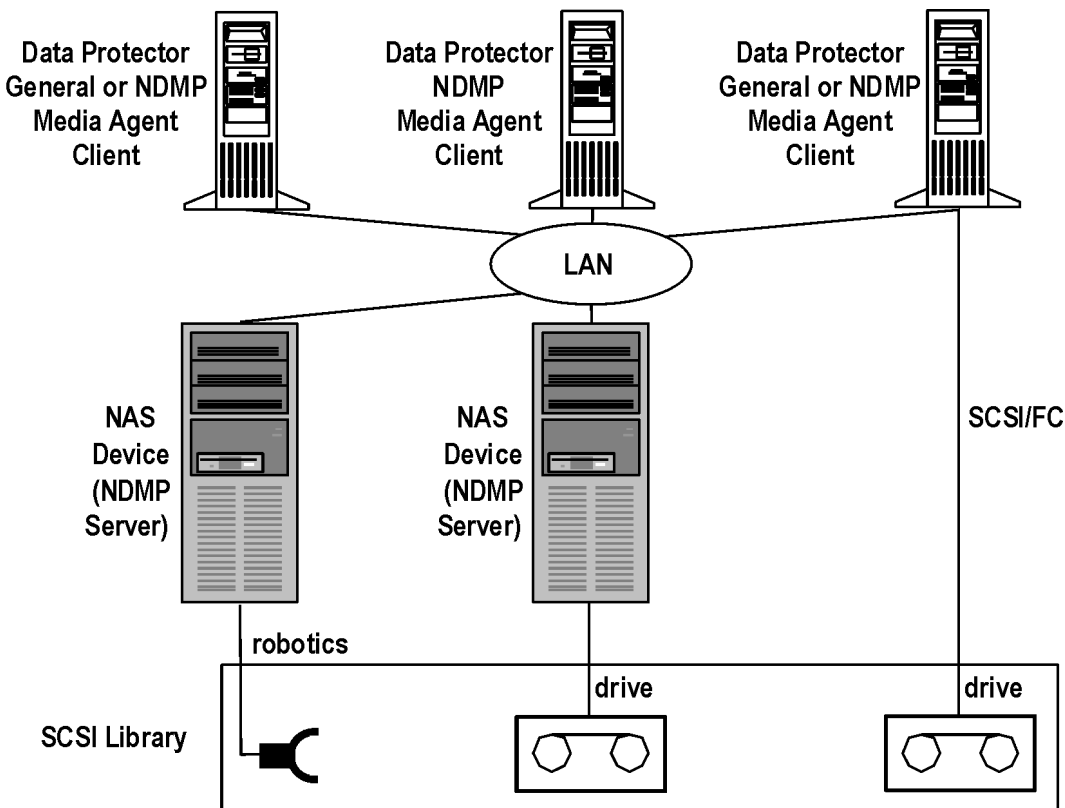


"Sharing a SCSI library (robotics attached to a Data Protector Client System)" above shows a SCSI library, with its robotics attached to and configured on the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI robotic control protocol is used 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 46: Sharing a SCSI library (robotics attached to an NDMP Server)



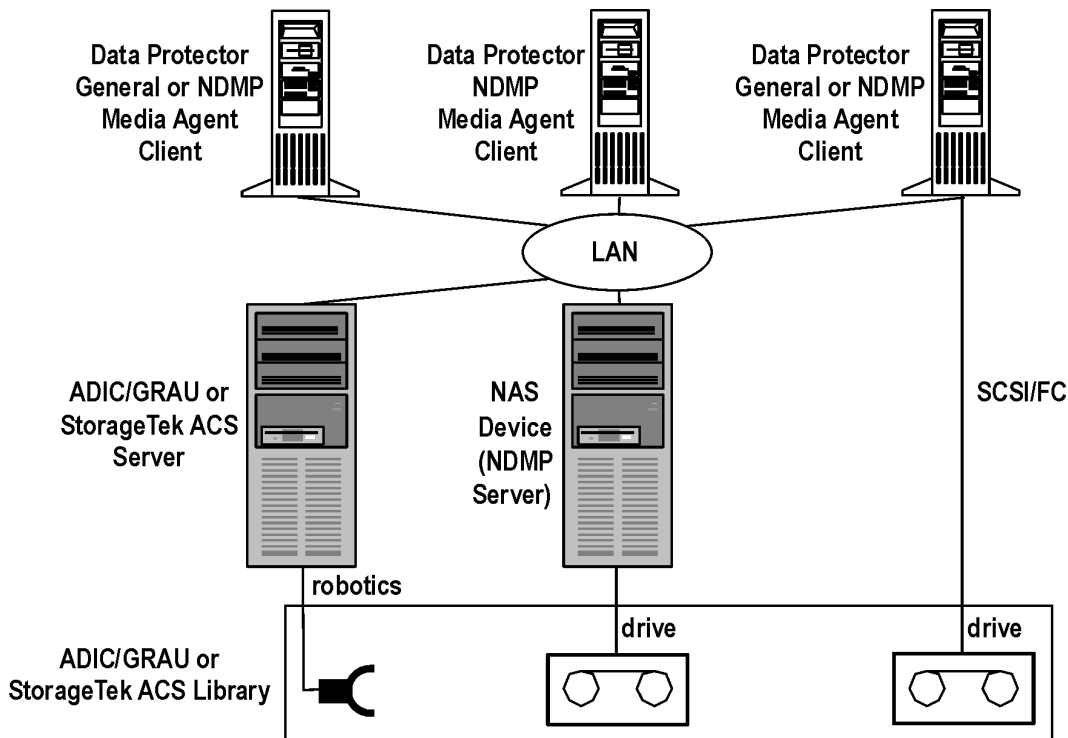
"Sharing a SCSI library (robotics attached to an NDMP Server)" above shows a SCSI library, with its robotics attached to an NDMP Server and configured on the Data Protector client system with either the General Media Agent or the NDMP Media Agent installed. The SCSI robotic control protocol is used by 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 has 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 47: Sharing an ADIC/GRAU or StorageTek ACS library



"Sharing an ADIC/GRAU or StorageTek ACS library" above 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.

Disk backup

This section 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.

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 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 dismount 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 * 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

HP 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, see the *HP Data Protector Help* index: “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.

With file libraries used for **virtual full backup**, distributed file media format must be used. Select this format in the device’s properties.

Data Protector and Storage Area Networks

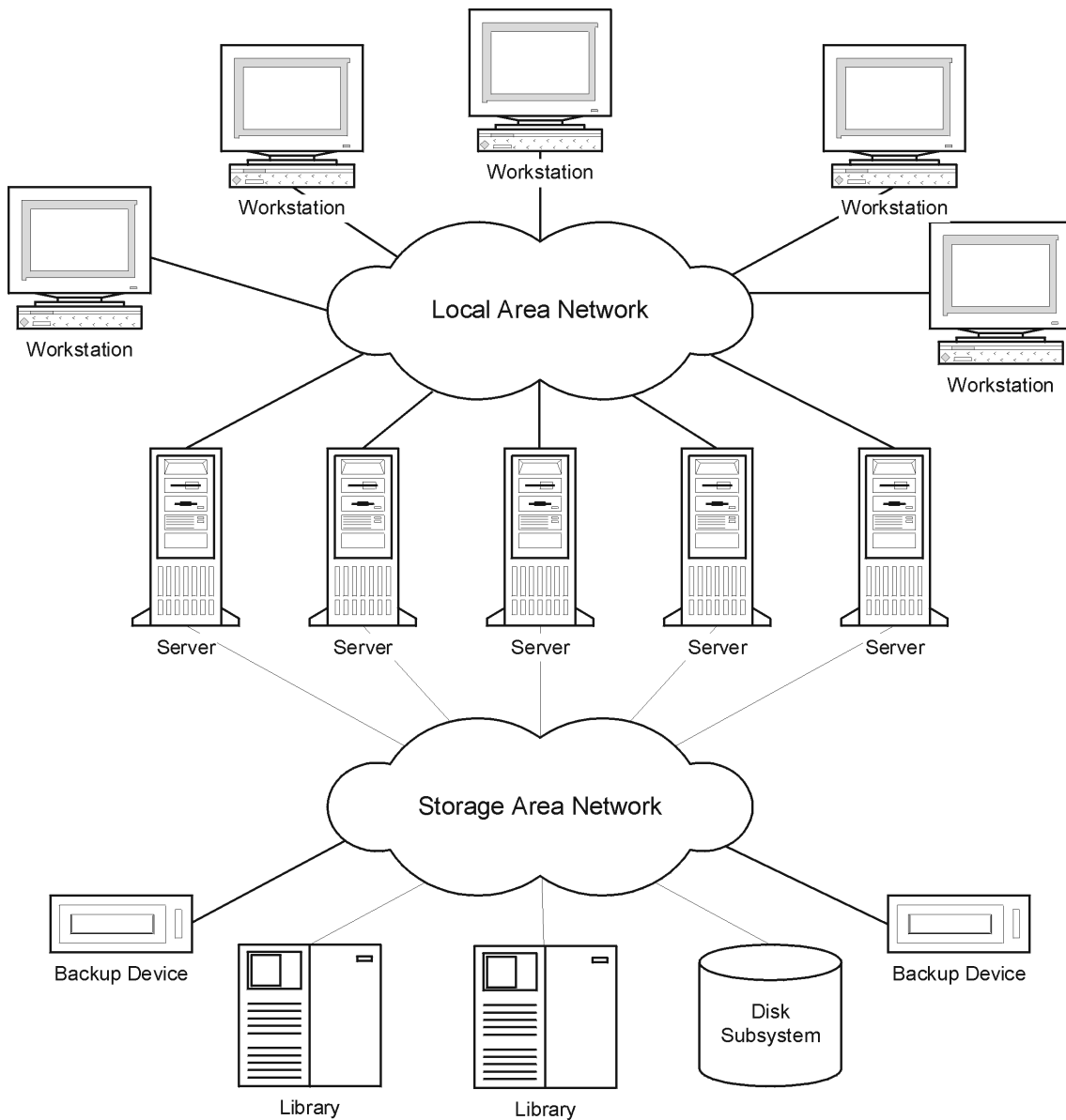
Where and how you store data in your enterprise may have a serious impact on your business. Information is becoming increasingly mission-critical to most companies. Today, terabytes of data must be accessible to users across the network. The Data Protector implementation of SAN-based Fibre Channel technology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

Storage Area Networks

A Storage Area Network (SAN) 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.

Figure 48: 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. 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.

For a list of supported configurations, see the <http://support.openview.hp.com/selfsolve/manuals>.

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

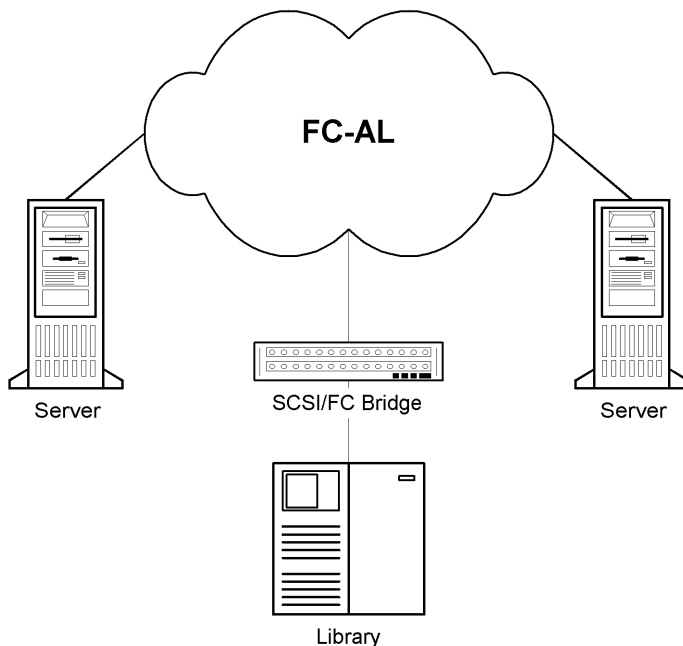
The loop topology is based on the Fibre Channel Arbitrated Loop (FC-AL) standard. Nodes include servers, backup devices, hubs, and switches. Any node in a loop can communicate with any other node in the loop, and all nodes share the same bandwidth. An FC-AL loop is typically implemented using an FC-AL hub with automatic port by-pass. Automatic port by-pass allows the hot-plug of nodes into the loop.

LIP

A Loop Initialization Primitive (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). See "[Loop initialization protocol](#)" below.

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 49: 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.

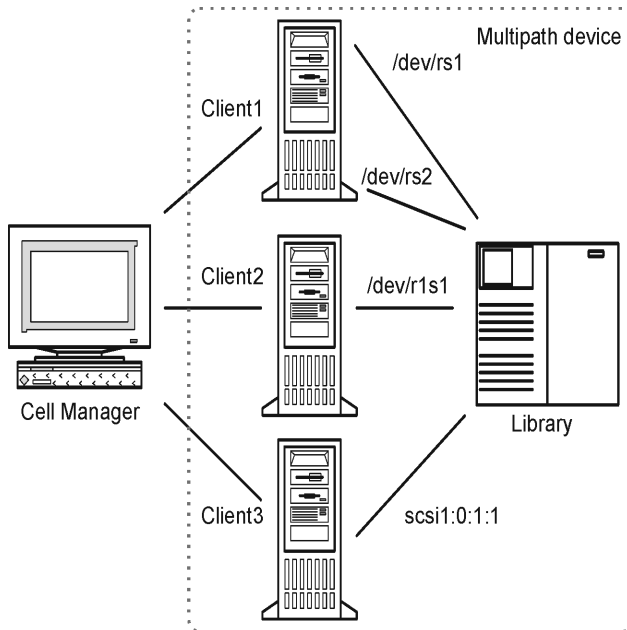
SAN technology provides an excellent way to manage the robotics of a library from multiple systems. This allows the option to manage the robotics from one system (classic) or allow each system that uses the library to access the robotics directly, provided the requests to the robotics are synchronized 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 50: 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.

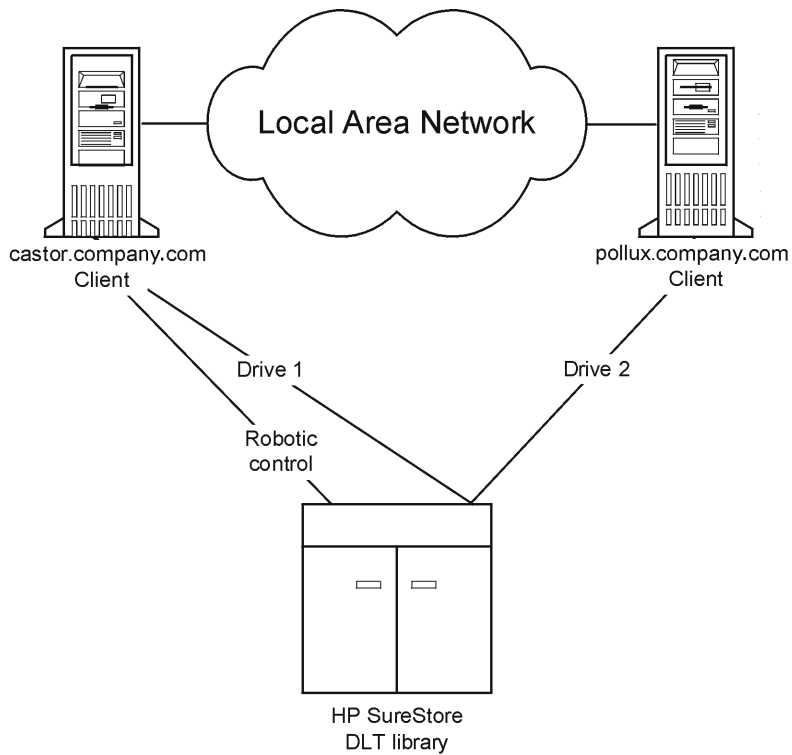
Indirect and direct library access

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 ["Indirect library access " on the next page](#), two client systems are attached to an HP 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 51: Indirect library access



Note that you cannot use a shared library if the client system that controls the robotics, *castor* in our example, fails.

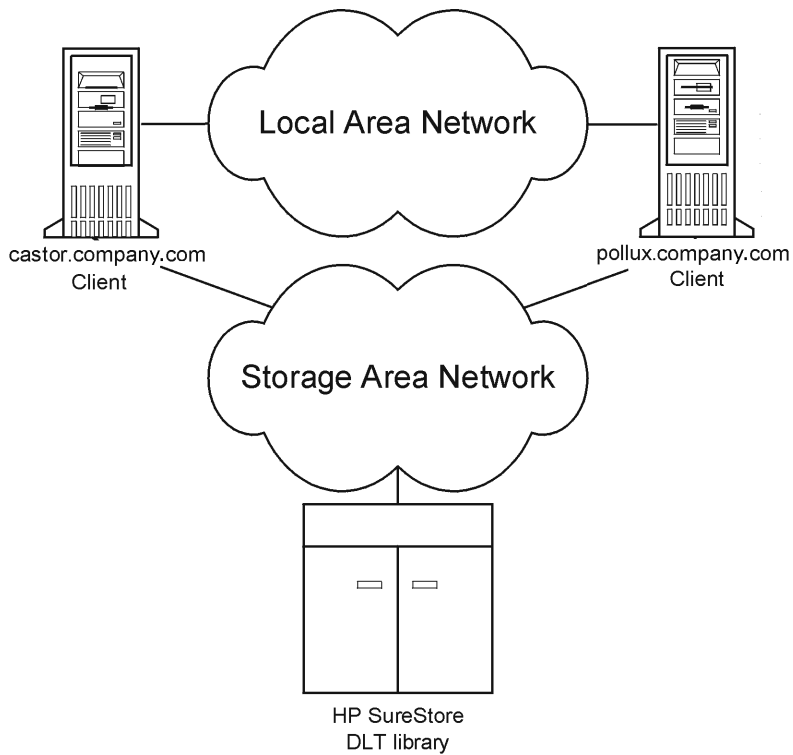
Direct library access

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.

"[Direct library access](#)" on the next page shows an HP 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 52: 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. For details, see ["Device locking" on page 122](#).

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.

Media management

Serious challenges can arise when administering 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.
- The capability to transfer all media-related catalog data from one Data Protector Cell Manager to another one without physically accessing the media.
- 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.
- 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.

For more information, see ["Media management before backups begin" on page 135](#).

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.

For more information, see ["Media management during backup sessions" on page 136](#).

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.

For more information on vaulting, see ["Media management after backup sessions" on page 140](#).

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.

See ["Calculating media condition" on page 139](#).

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, see the *HP Data Protector Help* index: “media pools, properties of”.

Media pools and DC directories

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

For information on the DCBF part of the IDB and DC directories, see ["IDB architecture" on page 146](#).

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 Vista systems, and one for Windows 7 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.

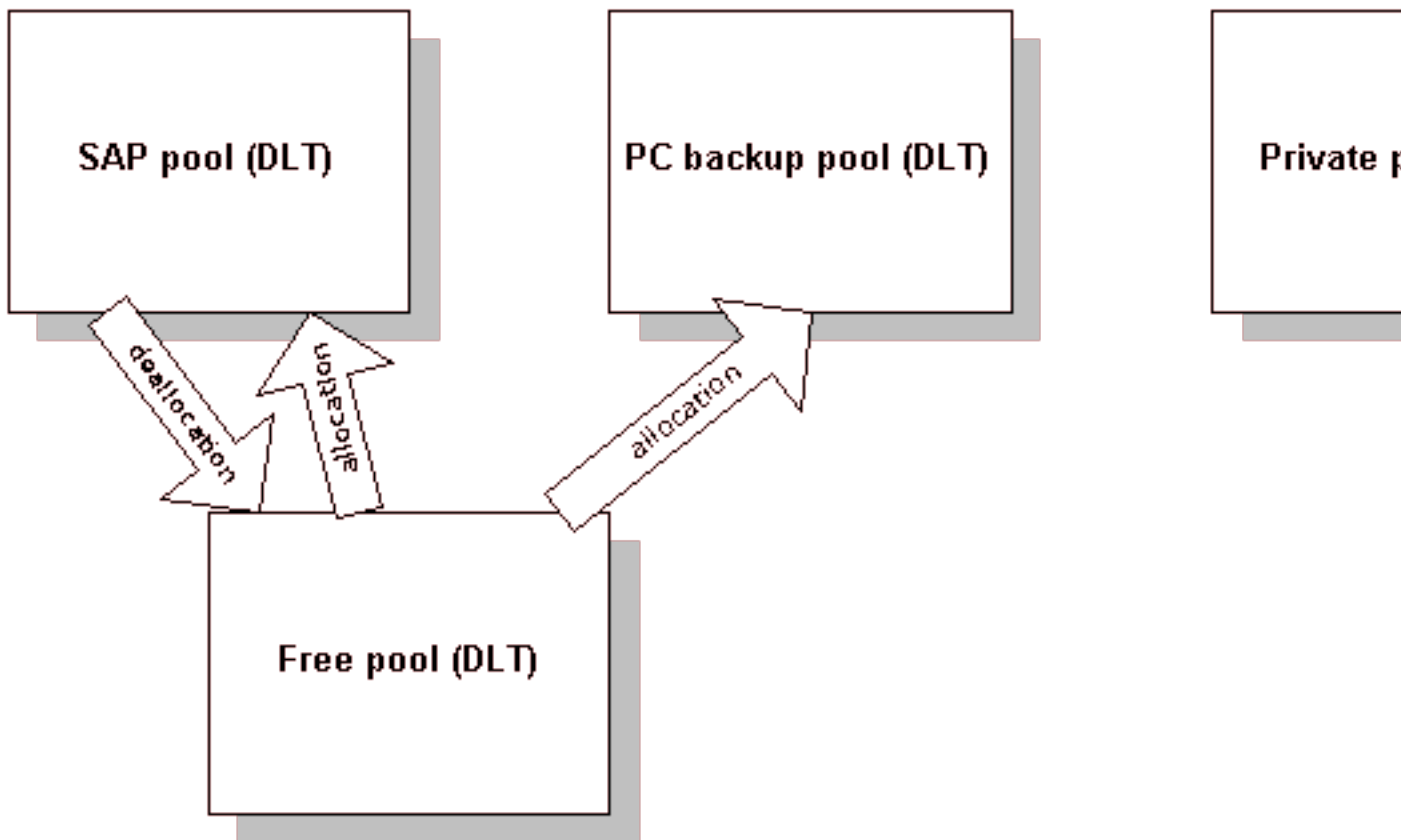
Free pools

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, and so on. 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 53: Free pools



When is a free pool used?

Media are moved between regular and free pools on two events ("[Free pools](#)" above):

- 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 "[Free pools](#)" above, 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 "[Free pools](#)" above, 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, see the *HP Data Protector Help* index: “free pools, characteristics”.

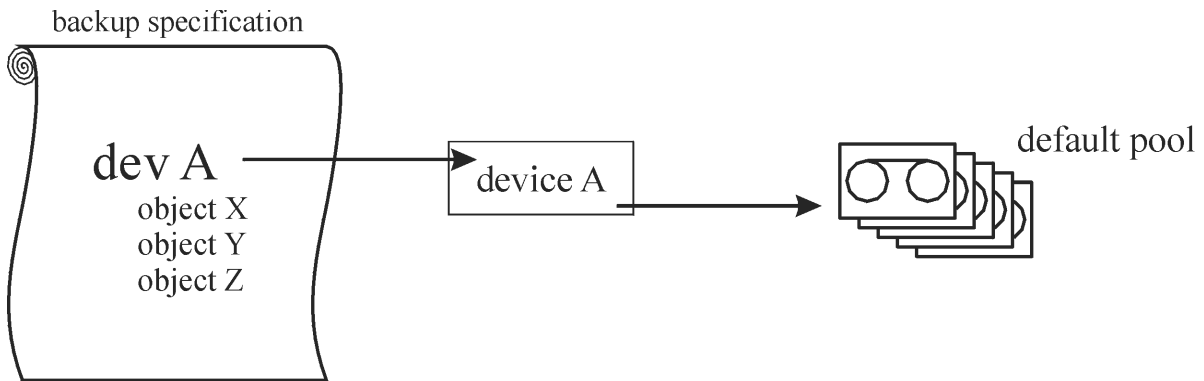
Media pool usage examples

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 ["A simple one device/one media pool relation"](#) below, 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 54: A simple one device/one media pool relation

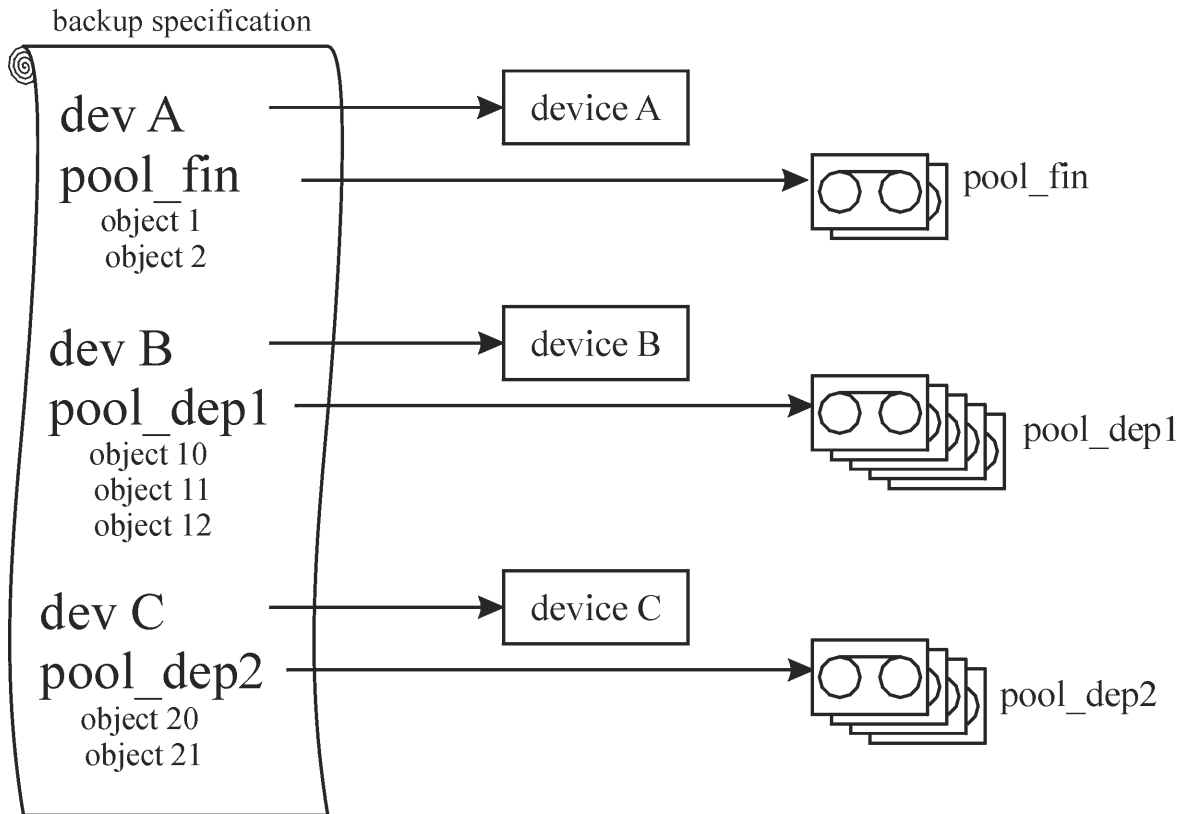


Example 2

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 ["Configuration of media pools for large libraries"](#) on the next page, 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.

For details about the relation between media pools and large library devices, see ["Large libraries"](#) on page 107.

Figure 55: Configuration of media pools for large libraries

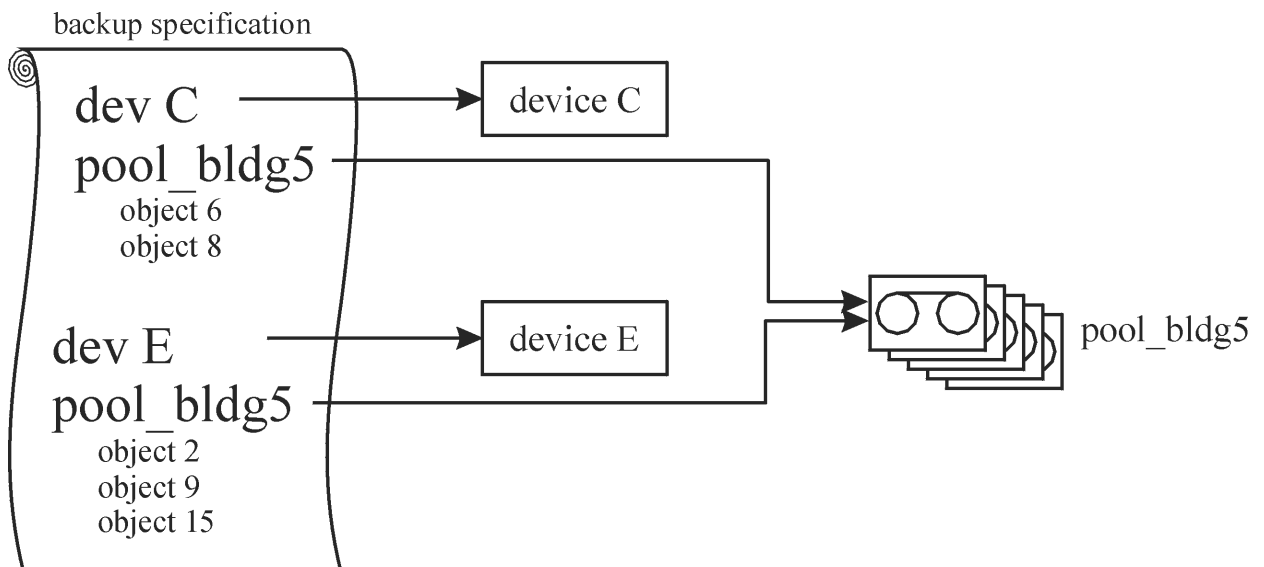


Example 3

"Multiple devices, single media pool " below 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.

For more information, see "Device lists and load balancing" on page 101.

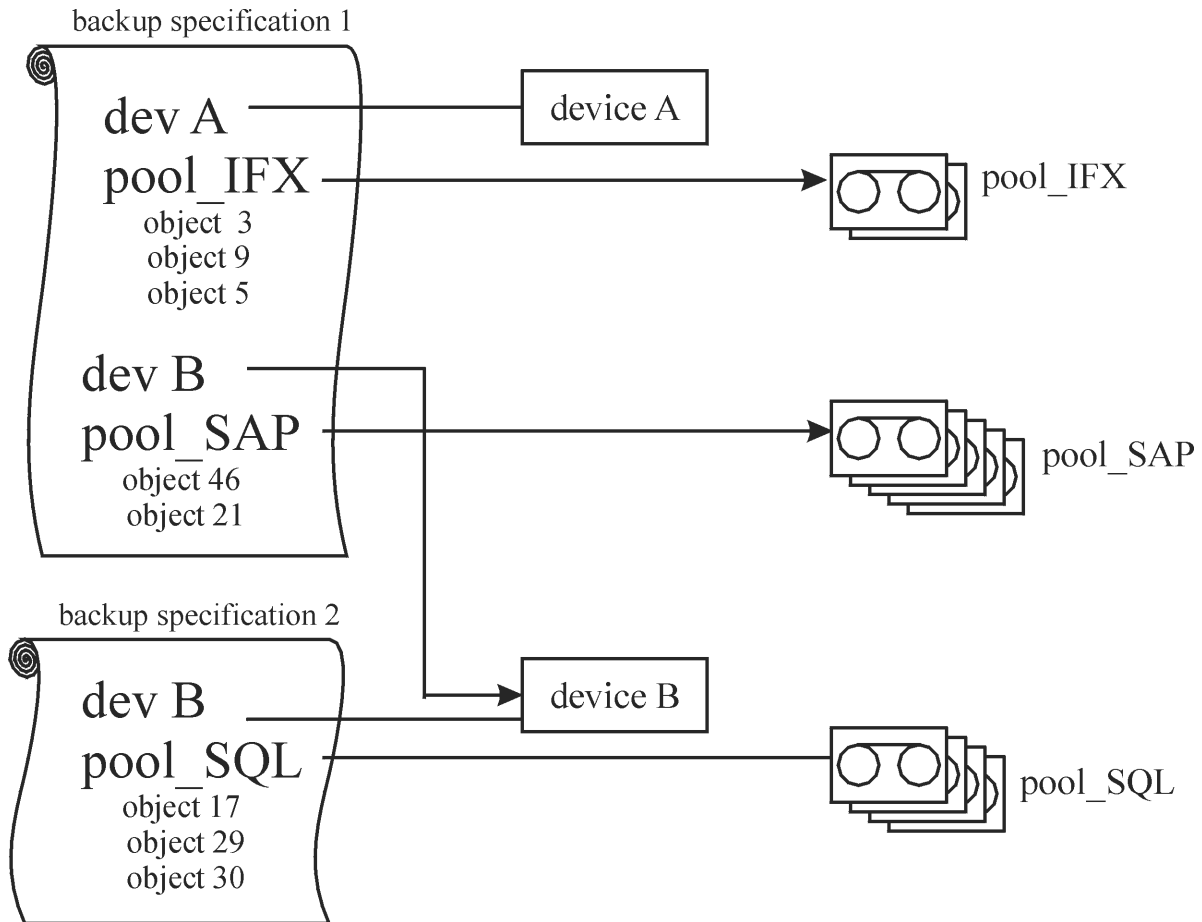
Figure 56: Multiple devices, single media pool



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 57: 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

Automatic media rotation and media handling

Data Protector automates media rotation and media handling as follows:

- 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. See ["Selecting media for backups" on page 137](#).

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. For more information, see ["Selecting media for backups" on page 137](#).

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 94](#).

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

For related information, see the following sections:

- ["Full and incremental backups" on page 40](#)
- ["Media pools" on page 126](#)

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. For more information, see ["Calculating media condition" on page 139](#).

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

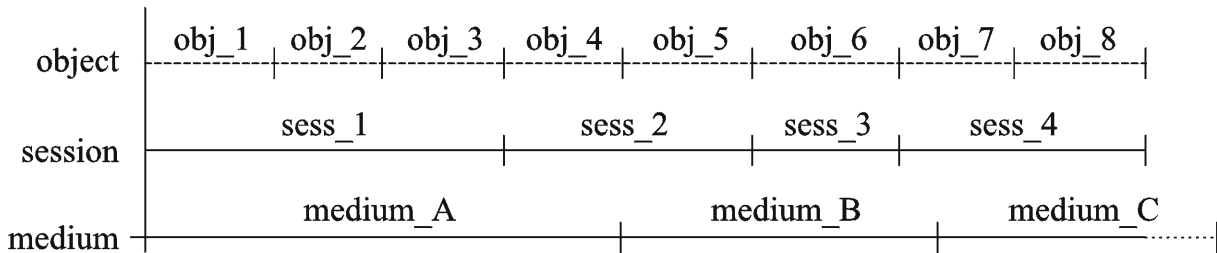
The available media usage policies are listed below:

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

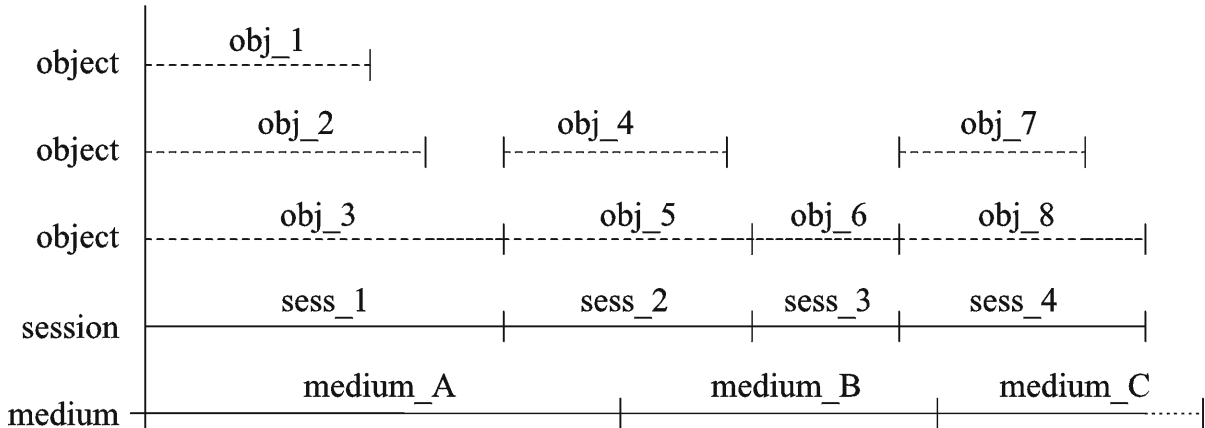
The following figures show some examples of how objects can be distributed over media:

Figure 58: Multiple objects and sessions per medium, sequential writes



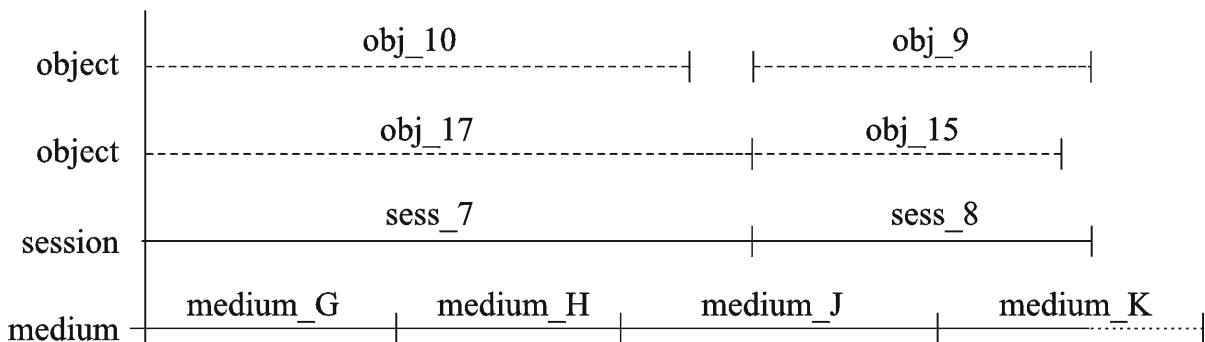
"Multiple objects and sessions per medium, sequential writes" above shows an example of eight sequential writes over four sessions, using the appendable media usage policy. The data was written in four sessions, one object at a time. The three media belong to the same media pool. *Medium_A* and *medium_B*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

Figure 59: Multiple objects and sessions per medium, concurrent writes



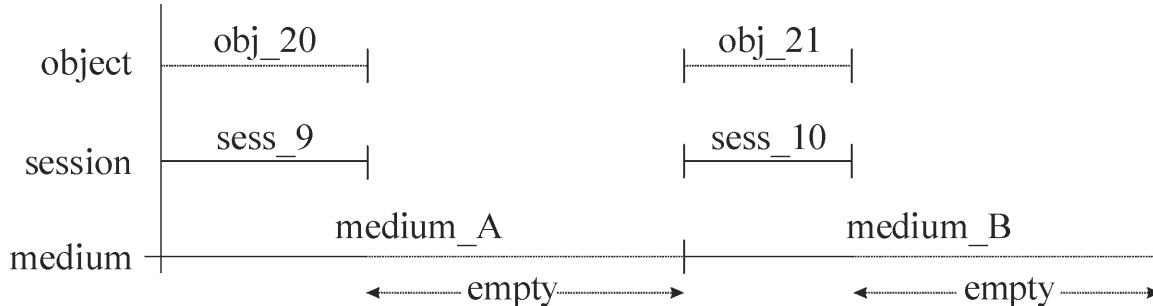
"Multiple objects and sessions per medium, concurrent writes" above shows an example of eight objects that have been written during four sessions with the concurrency settings that allow for simultaneous writes. In this case, *obj_1*, *obj_2*, and *obj_3* have been backed up concurrently in *sess_1*; *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 60: Multiple media per session, multiple media per object



"Multiple media per session, multiple media per object" on the previous page 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 61: Each object written on a separate medium



"Each object written on a separate medium" above 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, see ["Full and incremental backups" on page 40](#).

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 88](#).

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, 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. For more information on data and catalog protection, see ["Keeping backed up data and information about the data" on page 68](#).

- 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.

See ["Duplicating backed up data" on page 80](#).

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, see ["Keeping backed up data and information about the data" on page 68](#).

Chapter 4: Users and user groups

This chapter discusses Data Protector security, users, user groups, and user rights.

Increased security for 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.

Visibility of backed up data

- 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. For instructions, see the *HP Data Protector Help*.

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: admin, operator, and end-user. For example, only members of the admin user group can access all Data Protector functionality. Operators can, by default, start and monitor backups. For more information, see the *HP Data Protector Help* index: “user groups”.

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.

Customizing user groups

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, see the *HP Data Protector Help* index: “user groups, admin”.

Data Protector user rights

Data Protector users have the Data Protector user rights of the user group they belong to.

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 “*”.

Additionally, you can complement the user security layer provided by Data Protector user groups with restrictions of user actions to certain systems of the cell.

For a detailed description of the Data Protector user rights for each user group, see the *HP Data Protector Help*.

Chapter 5: The Internal Database

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.

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 backup media it resides, the result of backup, restore, object copy, object consolidation, object verification, and media management sessions, and what backup devices and libraries are configured.

Why is the IDB used?

The information stored in the IDB enables the following:

- Fast and convenient restore: The information stored in the IDB enables you to quickly find the backup 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 to allocate media during backup, object copy, and object consolidation sessions, track media attributes, group media in different media pools, and track media locations in tape libraries.
- Encryption/decryption management: The information stored in the IDB enables Data Protector to allocate encryption keys for encrypted backup or object copy sessions, and to supply the decryption key required for the restore of encrypted backup objects.

IDB size and growth consideration

The IDB may grow very large and its size may have 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**. See also "[IDB growth and performance](#)" on page 152.

IDB location and internal encoding used

IDB location

The IDB is located on the Cell Manager in the following directory:

Windows systems: `Data_Protector_program_data\server\db80`

UNIX systems: `/var/opt/omni/server/db80`

Internal text encoding in the IDB

The IDB stores all text information in either the Unicode double-byte format or the UTF-8 format. These formats allow for full support of filenames and messages localized to other languages.

The IDB in the Manager-of-Managers environment

In the Manager-of-Managers (MoM) environment, you can use the Centralized Media Management Database (CMMDB) instead of local Media Management Databases (MMDBs). This allows you to share devices and media across several cells. For more information on the MoM functionality, see ["Enterprise environments" on page 21](#).

IDB architecture

The IDB consists of the following parts:

- MMDB (Media Management Database)
- CDB (Catalog Database)
- DCBF (Detail Catalog Binary Files)
- SMBF (Session Messages Binary Files)
- Encryption Keystore

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. See ["IDB parts" on the next page](#).

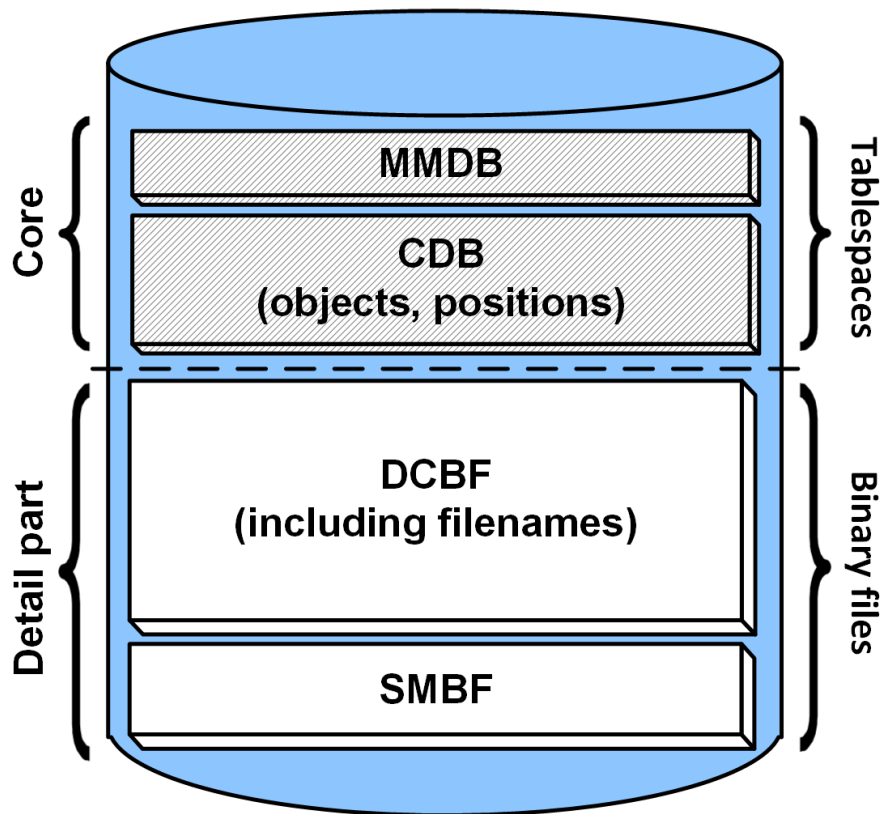
For robustness considerations, see the *HP Data Protector Help* index: "robustness of IDB".

Underlying technology

The MMDB and CDB parts are implemented using an embedded database consisting of tablespaces. This database is controlled by the `hdp-idb`, `hdp-idb-cp`, and `hdp-as` processes. The CDB and the MMDB parts represent the core part of the IDB.

The DCBF and SMBF parts of the IDB consist of binary files. Updates are direct, without using transactions.

Figure 62: 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.

Catalog Database (CDB)

CDB records

The Catalog Database stores information about the following:

- Backup, restore, object copy, object consolidation, object verification, 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. In the case of encrypted object versions, key identifiers (KeyID-StoreID) are also stored.
- 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.

Size and growth of the CDB

The CDB records occupy a minor share of space in the IDB. For more details, see ["IDB growth and performance" on page 152](#).

Detail Catalog Binary Files (DCBF)

DCBF information

The Detail Catalog Binary Files part of the IDB stores:

- Pathnames of backed up files (filenames) together with client system names. The filenames created between backups are added to the DCBF.
- File metadata. This is information about backed up file versions, their file sizes, modification times, attributes/protection, and positions of the backup copies on the backup media.

One DC (Detail Catalog) binary file is created for each Data Protector medium used for backup. When the medium is overwritten, its DC binary file is removed and a new one is created.

Size and growth of the filenames and file attributes part

The biggest and fastest growing part of the DCBF is the filenames part. The growth of the filenames part is proportional to the growth and dynamics of the backup environment as well as to the number of backups.

A file or directory occupies approximately 100 bytes in the IDB.

Size and growth of the rest of DCBF

In an environment where filesystem backups using the **Log All** option are typical, the DCBF occupies the largest part of the IDB. See ["IDB growth and performance: key tunable parameters" on page 153](#).

By default, five DC directories, dcbf0 through dcbf4, are configured for the DC binary files. You can create more DC directories and have them on different disks on the Cell Manager, thus extending IDB size. For more information, see the *HP Data Protector Help* index: "DC directories". For default and maximum values related to limitations concerning DC directories and DC binary files, see the *HP Data Protector Product Announcements, Software Notes, and References*, chapter *Limitations and recommendations*, section *Internal Database scalability*.

Session Messages Binary Files (SMBF)

SMBF records

The Session Messages Binary Files stores session messages generated during any Data Protector 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. 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. For more details, see the *HP Data Protector Help*.

Encryption keystore and catalog files

All the keys created, either manually or automatically, during encrypted backups are stored in a keystore. The keys can also be used for object copy, object verification, and restore sessions. In the case of hardware encryption, they can also be used for object consolidation sessions.

In the case of software encryption, the key identifiers (each consisting of a KeyID and a StoreID) are mapped to the object versions encrypted. This mapping is stored in the Catalog Database. Different objects in a medium can have different (software) encryption keys.

For hardware encryption, the key identifiers are mapped to medium ID and these mappings are stored in a catalog file. This file contains the information required to allow an encrypted medium to be exported to another cell. For more details, see the *HP Data Protector Help* index: "encryption".

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 IDB and contain information about what data was backed up, when it was backed up, and where it was backed up to.

If software encryption has been requested for the backup, the active encryption keys for the entities involved (hosts) are obtained from the keystore, used for the backup, and the key identifiers (KeyID-StoreID) are linked to the object versions and included in the CDB records. The mappings of the hosts to the KeyID-StoreIDs are also stored in a catalog in the keystore.

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. If the media involved are in drives for which hardware encryption has been requested, first the active encryption keys for the entities (media) are

obtained from the keystore. The mappings of the media to the KeyID-StoreIDs are recorded in a catalog in the keystore and also written to the media.

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 created.

All session messages generated during backups are stored in session messages binary files (the SMBF part).

Depending on the configuration of your Internal Database backup specification, the IDB backup process can remove the backed up archived log files and starts creating new ones that are necessary for 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 the hierarchy (file systems, application objects) 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 may 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. If the media involved have been hardware encrypted, the Media Agent first detects the key identifiers (KeyID-StoreID) and requests the key which is retrieved from the keystore by the Key Management Server (KMS).

If software encryption has been used for the backups concerned, when the Disk Agents receive the encrypted data, they submit the detected KeyID-StoreIDs to the KMS and request the relevant decryption keys, which are retrieved from the keystore.

During object copying or object consolidation

During an object copy or object consolidation session, the same 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 the previous page](#) and ["During restore" above](#). This does not apply for object consolidation with software encryption, since this is not supported.

During object verification

During an object verification session, the same database processes run as during a restore session. Basically, data is read from the source media, as if it were being restored, and is sent to the host Disk Agent(s) where the verification is performed. An object verification session has the same effect on the IDB operation as a restore session. All session messages generated during verification sessions are stored in session messages binary files. For details, see ["During restore" above](#).

Exporting media

When a medium is exported, if it contains encrypted information, the relevant keys are exported from the keystore to a `filename.csv` file on the Cell Manager. This file is required for successful import of the medium in another cell.

Removed items

In addition, several items are 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. 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 for a specific medium is removed, 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. You can therefore run a restore of an entire backup object, but not a restore of a single 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: HP strongly recommends 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.

Caution: Always back up the Internal Database after any modification in the IDB configuration, for example, after changing the password of the Internal Database Service and Application Server user account. Failing to do so may result in inability to successfully perform online IDB restore as well as offline IDB recovery.

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.

For detailed information, see the *HP Data Protector Help* index: "IDB, recovery".

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. A more detailed logging level has a stronger influence on the IDB storage space consumption. For details, see ["IDB growth and performance: key tunable parameters" on the next page](#).

- 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 the next page](#).

- 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 40](#).

- 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 may have a significant impact on the growth of the filenames part of the IDB. 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.

- 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.

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.

Logging level as an IDB key tunable parameter

What is logging level?

Data Protector logging level defines the amount of detail on the backed up files and directories that is written to the IDB during backup. You can always restore your data, regardless of the logging level used.

Four logging levels are available:

Log All	Logs all detailed information about backed up files and directories (names, versions, and attributes).
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 and the convenience of browsing for data to be restored.

Logging level and backup speed

The backup speed is approximately the same regardless of the logging level chosen.

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 effective logging level, 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 corresponding backup session was run with either Log All, Log Directories, or Log Files logging level.

If the backup session was run with the No Log logging level, the restore speed may reduce when restoring single files. In this case, 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 entire backup object is read anyway, so the logging level does 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 have impact on performance of browsing the backup objects.

Catalog protection and restore

When catalog protection expires, data is restored as if it were backed up using the No Log option. See ["Logging level as an IDB key tunable parameter" on the previous page](#).

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 size increase is 13 times larger if the catalog protection is set to one year.

It is recommended that catalog protection includes at least the last full backup. For example, you can set a catalog protection of eight 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 numerous user workstations. Since the dynamics of these systems are rather different, it is difficult to prescribe one setting that suits them all. Therefore, it is recommended to create several backup specifications with the following logging level settings:

Mail servers: Use the **Log Directories** option.

Database servers: No logging is necessary as they have their own restore policies. Therefore, use the **No Log** option.

Workstations and file servers: 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, see the *HP Data Protector Help* index: "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 relatively small and will remain such, 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 extremely large, or there are that many files generated on a daily basis, and you use the **Log All** option, then 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, and setting the `Log Directories` option can reduce its size even further. The actual gain, 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 to it.

Chapter 6: Service management

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 service management products.

Data Protector and service management

Data Protector provides service management supports and can be integrated with service management applications.

Data Protector functionality

The functionality described in the following sections comes with Data Protector “out of the box”.

Key functions

- 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 systems 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.
- Data Protector backup session auditing stores information about all backup tasks that were performed over extended periods for the whole Data Protector cell, and provides this information on demand in an integral and printable fashion for auditing and administrative purposes.
- The Data Protector capability to send major and critical events into the Windows Event Log opens up a variety of interesting integration possibilities.
- 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.

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, see the *HP Data Protector Help* index: “SNMP, reports send methods”.

The Data Protector 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 multicell 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 *Clients not Configured for Data Protector* report, which contains information about unprotected systems, the *Session Specification Schedule* report, which lists all scheduled backups, object copy, and object consolidation as well as the *List of Pools* 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 *Configured Devices not Used by Data Protector* report, which lists devices that are currently not used for backup, object copy, or object consolidation and are consequently available.
- Problem Reports such as the *Session Statistics* report, which consists of information about failed backup, copy, and consolidation sessions. 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 numerous pre-configured reports (including, but not limited to, reports such as sessions in a specific time frame, IDB report, and device usage report)
- Specify your own parameters for those reports (such as time frames, backup, copy, and consolidation 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 systems 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, copy, and consolidation 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. Events that are logged in the Data Protector Event Log are either process-triggered or user-triggered. 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 HP software management applications. In addition to reports, log entries feed HP software management applications via the Data Protector SPI (SMART Plug-In) so that they can trigger preventive or corrective actions.

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.

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 Data Protector **Event Log** using the Event Log Viewer.

Data Protector log files

Some service management applications 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..

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, see the *HP Data Protector Troubleshooting Guide* and the *HP Data Protector Help*.

Windows application log

Some service management applications monitor the Windows Application Log.

For more information on how to enable automatic forwarding of all Data Protector messages and messages about the Data Protector services (if they are stopped) to Windows Application Log, see the *HP 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, see the *HP Data Protector Help* index: "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, see the *HP Data Protector Help* index: "MoM environment".

Using the data provided by Data Protector

What can I do with the data?

Here are some examples of what you can do with the data that Data Protector provides:

- Regular e-mail reports to back up operators, end users, and management (Data Protector built-in reporting with the capability to send e-mails).

- 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 (Data Protector built-in notification engine sending SNMP traps).

Chapter 7: How Data Protector operates

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.

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.

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).
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.
KMS	The KMS (Key Management Server) process (service) runs on the Cell Manager and provides key management for the Data Protector encryption functionality. The process is started when Data Protector is installed on the Cell Manager.
hdpd-idb	The Data Protector Internal Database Service (hdpd-idb) is the service under which the IDB runs. The service is accessed locally on the Cell Manager by processes that need information from the internal database. This service is accessed remotely only for media management information about transfer from the IDB on the Cell Manager to the IDB on the Manager-of-Manager (MoM).
hdpd-idb-cp	The Data Protector Internal Database Connection Pooler (hdpd-idb-cp) service offers a pool of open connections to the hdpd-idb which can be used on request instead of opening a new connection for every request, thus ensuring hdpd-idb connection scalability. The service runs on the Cell Manager and is accessed only by local processes.
hdpd-as	The Data Protector Application Server (hdpd-as) service is used for connecting the GUI to the IDB through a HTTPS connection (web services). It runs on the Cell Manager and has a local connection to the hdpd-idb-cp service.

For instructions on how to manually start or stop the Data Protector processes and services, see the *HP Data Protector 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 ["Backup session information flow \(1\)" on the next page](#). 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 split mirror backup, see 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.

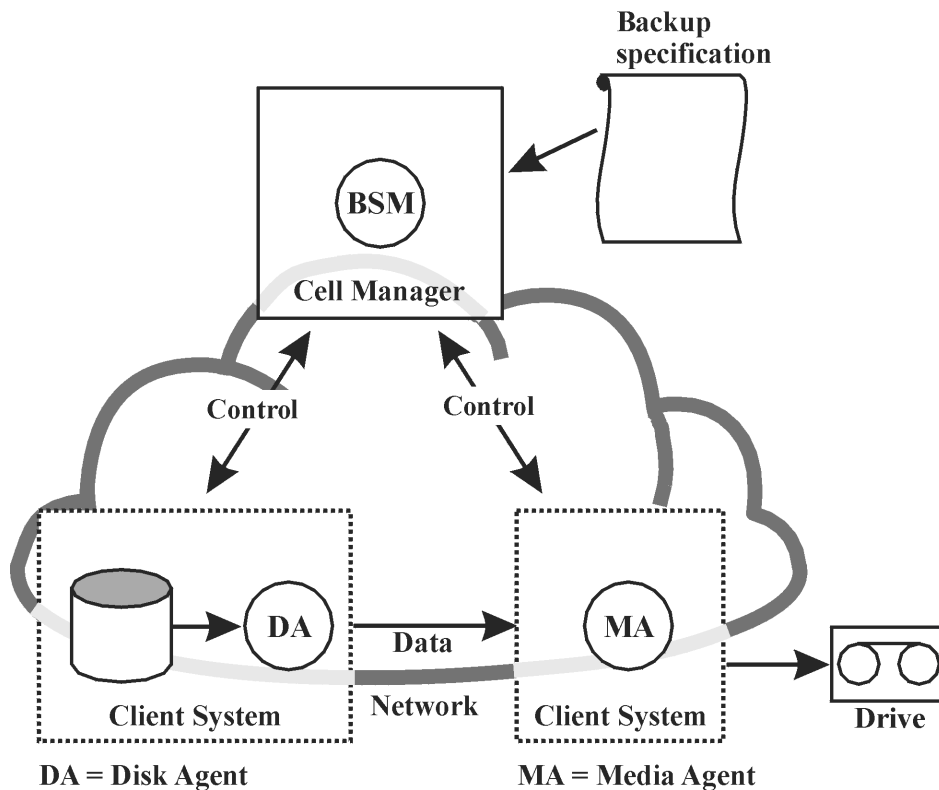
5. Disk Agents read data from disks and send it to the Media Agents that write data to media.

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 63: Backup session information flow (1)

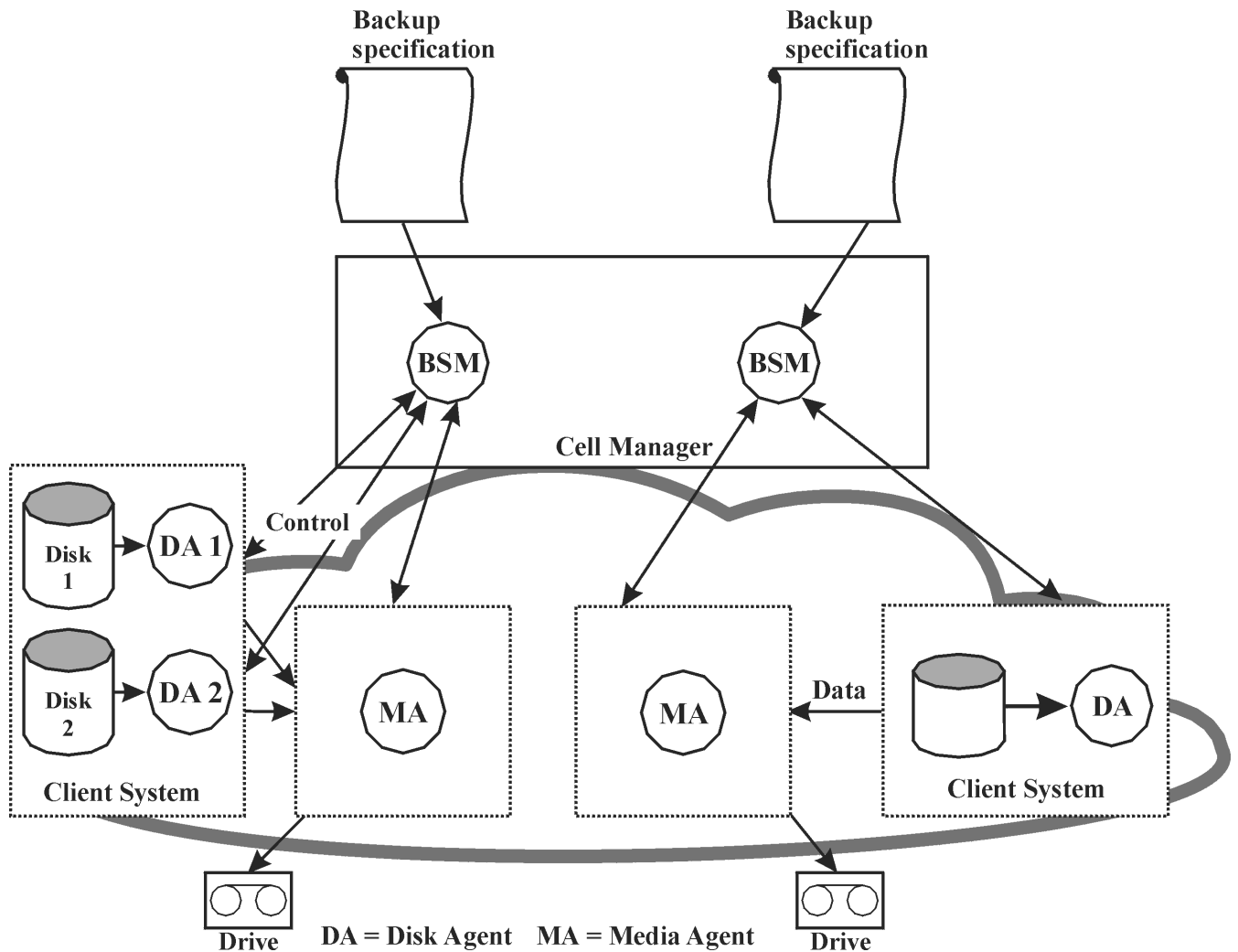


How many sessions can run concurrently?

A number of backup sessions can run in the cell at the same time. This number is limited by resources in the cell, such as the availability of devices and the configuration of the Cell Manager, for instance, processor speed, main memory size, and similar. To prevent Data Protector processes from exceeding system capabilities, the maximum number of concurrent backup sessions is limited. The limit is configurable.

"Backup session information flow - multiple sessions" below shows multiple sessions running concurrently.

Figure 64: 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.

Start-up and location of the commands

Pre- and post-exec commands for a backup session are started before and after the backup session, respectively. They are executed on the Cell Manager by default, but you can choose another system.

Windows systems

Pre- and post-exec scripts are started by the Data Protector CRS when executed on the Cell Manager; and under the Data Protector Inet Service account (by default, the local system account) when executed remotely.

On the Cell Manager, the scripts can be located in any directory. On the systems other than the Cell Manager, the scripts must be located in the *Data_Protector_home*\bin directory.

For the scripts located in the *Data_Protector_home*\bin directory, specify only the filename, otherwise, specify the full pathname of the script.

Only .bat, .exe, and .cmd are supported extensions for pre- and post-exec commands. To run a script with unsupported extension (for example, .vbs), create a batch file that starts the script. Then configure Data Protector to run the batch file as a pre- or post-exec command, which then starts the script with the unsupported extension.

If you use quotes (") to specify a pathname, do not use the combination of backslash and quotes (\"). If you need to use a trailing backslash at the end of the pathname, use the double backslash (\\).

Note: The direct usage of perl.exe is prohibited.

UNIX systems

Pre- and post-exec scripts are started by the backup session owner, unless the backup session owner has Backup as root permission; the commands are then started under root.

On the Cell Manager, the commands for backup specifications can reside in any directory.

On a remote UNIX client, the exec commands for backup specifications must be located as follows:

HP-UX, Solaris, and Linux systems: /opt/omni/sbin

Other UNIX systems: /usr/omni/bin

For the commands located in the /opt/omni/sbin or in the /usr/omni/bin directory, specify only the filename, otherwise, specify the full pathname.

For more information, see *Pre- and Post-Exec Commands for a Backup Specification* in *HP Data Protector Help*.

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.

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 using the Data Protector global options. 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.

For more information, see ["Adding data to media during backup sessions" on page 137](#) and ["Selecting media for backups" on page 137](#).

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. For more information, see ["Reporting and notification" on page 158](#).

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.

Resuming backup sessions

Backup sessions that did not complete successfully (for example, due to some network problems) or were aborted, can be resumed using the Data Protector resume session functionality. When you resume a failed backup session, Data Protector continues with backup, starting right from where the failed session left off.

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

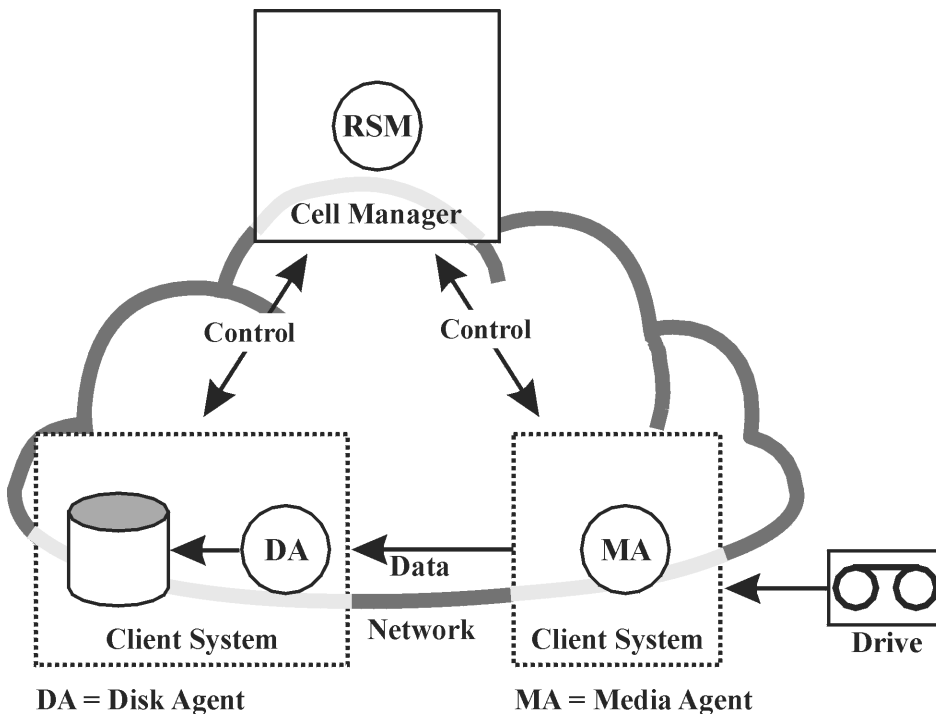
What happens in a restore session?

When a restore session, as shown in ["Restore session information flow" on the next page](#), 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. For more information, see ["Parallel restores" on the next page](#).
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 65: 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

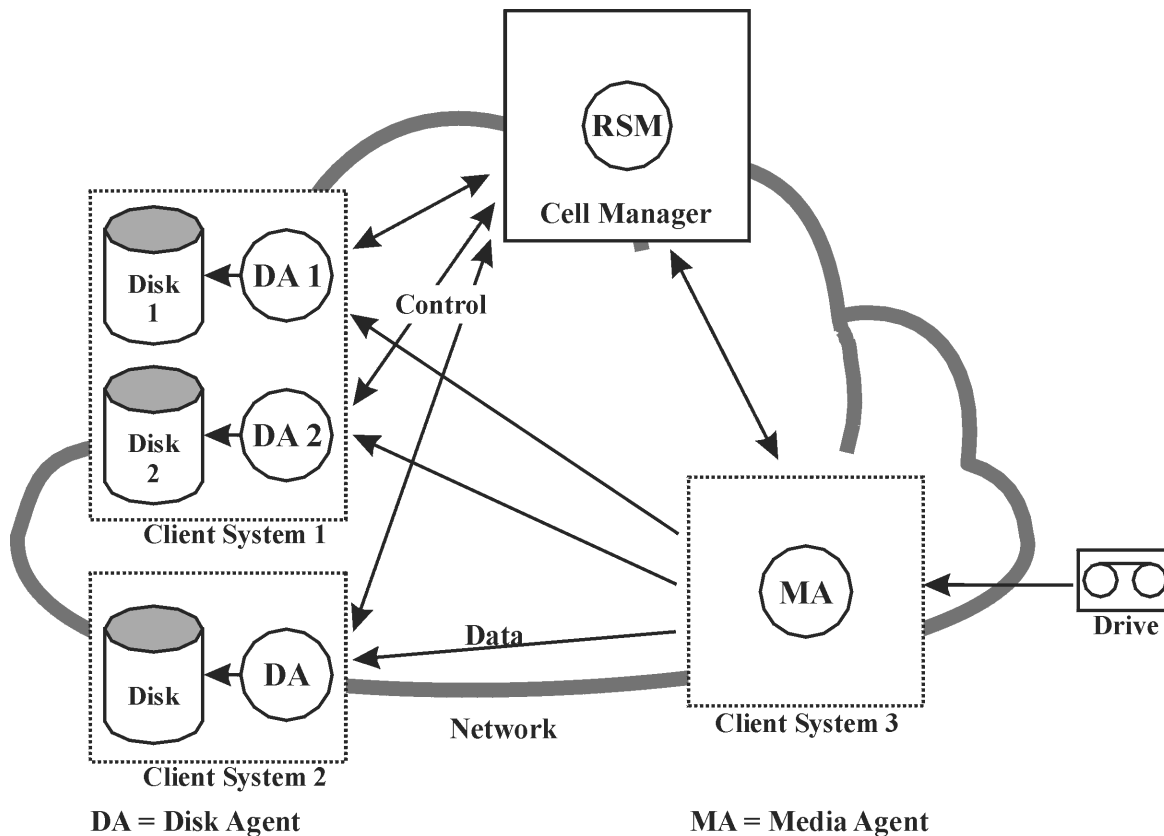
What is a parallel restore?

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. For more information, see ["Parallel restore session flow" on the next page](#).

How does it compare to a standard restore?

Data from multiple Disk Agents is (most of the time) multiplexed and stored on the media. See ["Multiple objects and sessions per medium, sequential writes" on page 138](#). 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 66: 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.

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.

Resuming restore sessions

Restore sessions that did not complete successfully (for example, due to some network problems) can be resumed using the Data Protector resume session functionality. When you resume a failed session, Data Protector continues with the restore in a new session, starting right from where the failed session left off.

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, copied, or consolidated data on a different media set. During an object copy session, the selected backed up, copied, or consolidated 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, object copy, or object consolidation. A scheduled object copy session is started at the time you have specified using the Data Protector Scheduler. A post-backup or a post-copy or a post-consolidation 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 ["Object copy session information flow" on page 174](#). 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.
3. The CSM locks the devices. 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. The Media Agents load the 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. For all objects successfully copied, the CSM updates the IDB protection entries according to the options specified for the copy session.

The protection of any failed source objects is also updated to allow recycling if the recycle option was specified for the session.

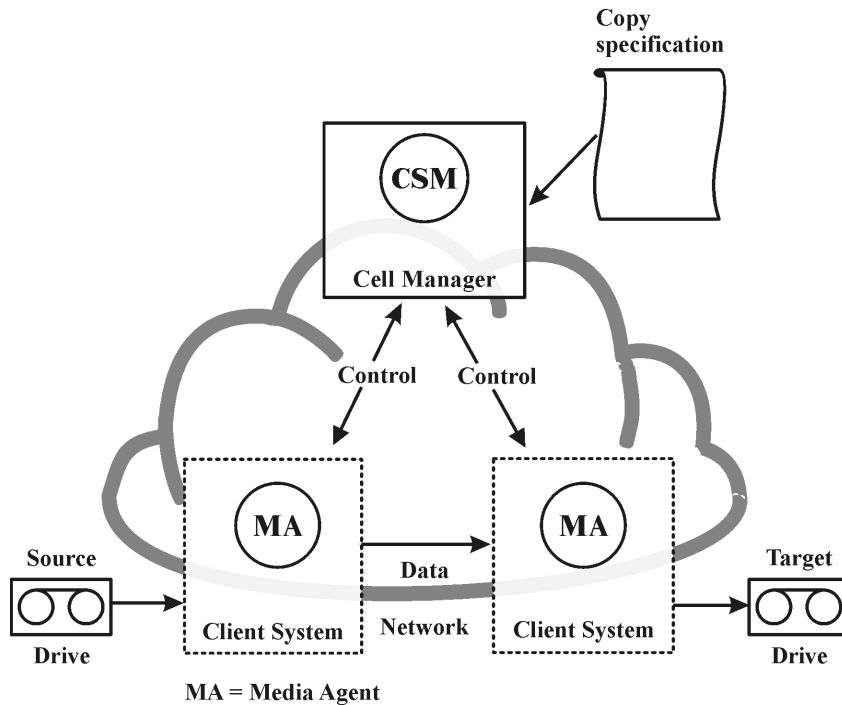
8. When the object copy session is completed, the CSM closes the session.

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.

However, it is not possible to run two or more object copy sessions from the same object copy specification in parallel.

Figure 67: 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.

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.

Replication sessions

This section describes how a replication session is started, what happens during the session, and the processes and services involved.

What is a replication session?

A replication session is a process that creates an additional copy of the backed up, copied, or consolidated data on a different Backup to Disk (B2D) device capable of replication. During a replication session, the selected backed up, copied, or consolidated objects are replicated from the source to the target device, directly from one device to another, without transferring data through Media Agent clients. Additionally, the network load is reduced as only unique (deduplicated) data is transferred over the network.

Automated and interactive replication sessions

Automated replication session

An automated replication session can either be scheduled or started immediately after a backup, object copy, or object consolidation. A scheduled replication session is started at the time you have specified using the Data Protector Scheduler. A post-backup or a post-copy or a post-consolidation replication session is started after the specified session finishes. You can view the progress of the automated replication session in the Data Protector monitor.

Interactive replication session

An interactive replication 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 replication session. You may want to stop monitoring by disconnecting the user interface from the session. The session will then continue in the background.

Replication session data flow and processes

What happens in a replication session?

The information flow of a replication session is shown in "[Object copy session information flow](#)" on the [previous page](#). When a replication 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 copy specification (with enabled replication options) for information on what to replicate and which options and devices to use. It also controls the replication session.
2. The CSM opens the IDB, reads the information about the devices needed for replicating, and writes the information about the replication session, such as generated messages, to the IDB.
3. The CSM locks the devices. If the resources are still unavailable after the timeout, the session is aborted.
4. The CSM initiates the replication process between the devices configured for replication.
5. For all objects successfully replicated, the CSM updates the IDB protection entries according to the options specified for the replication session.

The protection of any failed source objects is also updated to allow recycling if the recycle option was

specified for the session.

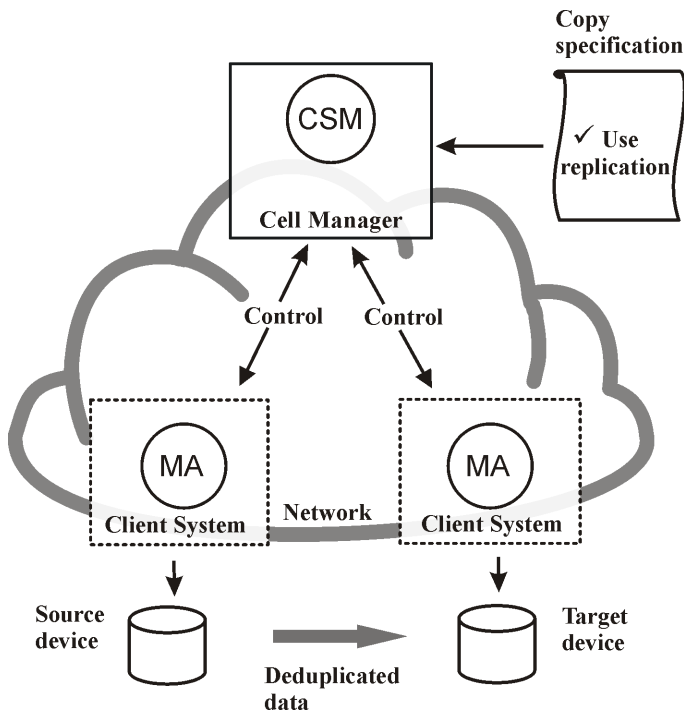
6. When the replication session is completed, the CSM closes the session.

How many sessions can run concurrently?

A number of replication 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.

However, it is not possible to run two or more replication sessions from the same replication specification in parallel. You also cannot run two or more interactive replication sessions in parallel.

Figure 68: Replication session information flow



Queuing of replication sessions

Timeout

When a replication 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.

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 ["Synthetic backup" on page 62](#).

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 devices. 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. The Media Agents load the source and target media allocated according to the backup policies.

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
5. One Media Agent reads the full object version. It sends the data to another Media Agent that reads incremental object versions. The latter Media Agent does the actual consolidation and sends the data to the Media Agent that writes the data to the target media.
- If the full backup and the incremental backups reside in the same file library, the same Media Agent reads all the backups and consolidates them.
- If the block size of the source device is smaller than that of the destination device, blocks are repackaged.
6. 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. Object consolidations sessions are treated like backup sessions and their number is limited by the same factors.

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.

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.

Object verification sessions

This section describes how an object verification session is started, what happens during the session, and the processes and services involved.

What is an object verification session?

An object verification session is a process that verifies the media segments allocated to a specified object or specified objects, checking the information in the header segments and reading the data blocks in the data segments to verify their format. If a cyclic redundancy check (CRC) was performed during the original backup, it also recalculates the CRC and compares it with the original.

Data Protector can perform the verification on the host that was the source of the backup, effectively verifying the Data Protector components in the restore path, on another host, verifying restore capability to a different location, or directly on the host with the Media Agent involved, verifying the data only.

Automated and interactive object verification sessions

Automated object verification session

You can specify an automatic object verification session to run at a specified time, using the Data Protector Scheduler, or to run as a post-backup object verification session immediately after completion of a specified backup, object copy, or object consolidation session. You can view the progress of such sessions in the Data Protector Monitor.

Interactive object verification session

You can start an interactive object verification session directly from the Data Protector user interface. The Data Protector monitor starts immediately and you can view the progress of the session. Multiple users can monitor the same object verification session. You can perform other operations with the user interface and let the session continue in the background, if required.

Object verification session data flow and processes

What happens in an object verification session?

When an object verification session is started, the basic process flow is as follows:

1. The Restore Session Manager (RSM) process is started on the Cell Manager system, triggered either by:
 - the Data Protector Scheduler, for a scheduled session
 - the End of Session event, for post-backup sessions
 - the user from the GUI or the CLI, for interactive sessions

This process controls the verification session.

2. The RSM opens the IDB, reads the information about the objects to be verified, and writes information about the verification session, such as generated messages, to the IDB.
3. The RSM starts the Media Agents (MA) on the source systems involved in the verification. For each drive used in parallel, a new Media Agent is started.
4. Verification of the data is performed by the Disk Agents (DA) on the destination hosts, so the RSM

starts a Disk Agent for each destination disk in parallel. The actual number of Disk Agents started depends on the objects you selected for verification. The process is similar to that for restore. For more information, see ["Parallel restores" on page 170](#).

5. The Media Agents read the object data from the media and send it to the Disk Agents that perform the verification. The RSM monitors the progress of the session and starts new Disk Agents and new Media Agents as necessary.
6. When the object verification session is completed, the RSM closes the session.

Variations in process flow with object verification

The object verification process emulates the restore process from the point at which data is requested for restore to the point at which the data reaches the destination host. Beyond that point, the verification process does not write any data and, for application integration objects, there is no communication with the application integration.

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.

Chapter 8: Integration with applications

This chapter gives a brief description of the Data Protector integration with database applications.

Integration with database applications

This section gives a brief description of the Data Protector integration with database applications. For a detailed list of supported applications, see the latest support matrices at

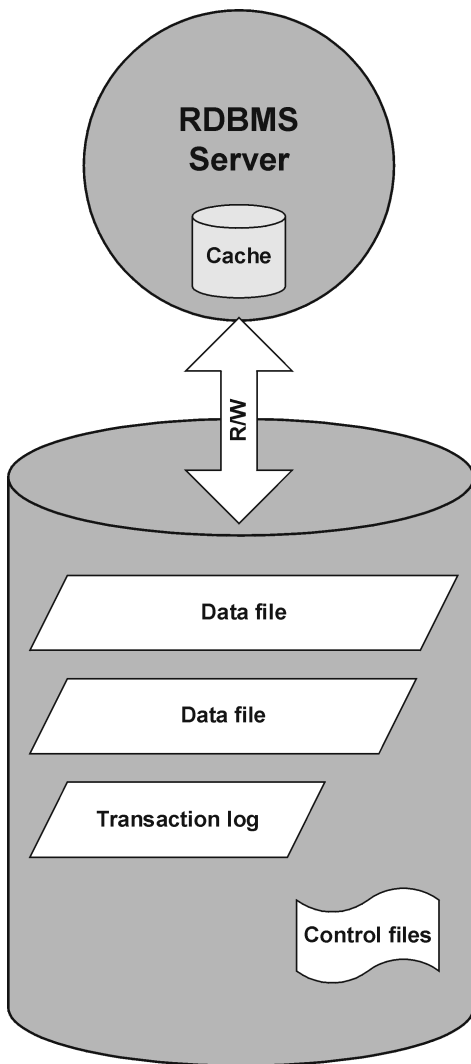
<http://support.openview.hp.com/selfsolve/manuals>.

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 have various names, such as **tablespaces** (for example, in Oracle), **dbspaces** (for example, in Informix Server), and **segments** (for example, in Sybase).

Figure 69: Relational database



"Relational database" above 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.
- Choose two commands or prepare two scripts or applications 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 command, script, or application as a **pre-exec command** and the start-up command, script, or application as a **post-exec command**.

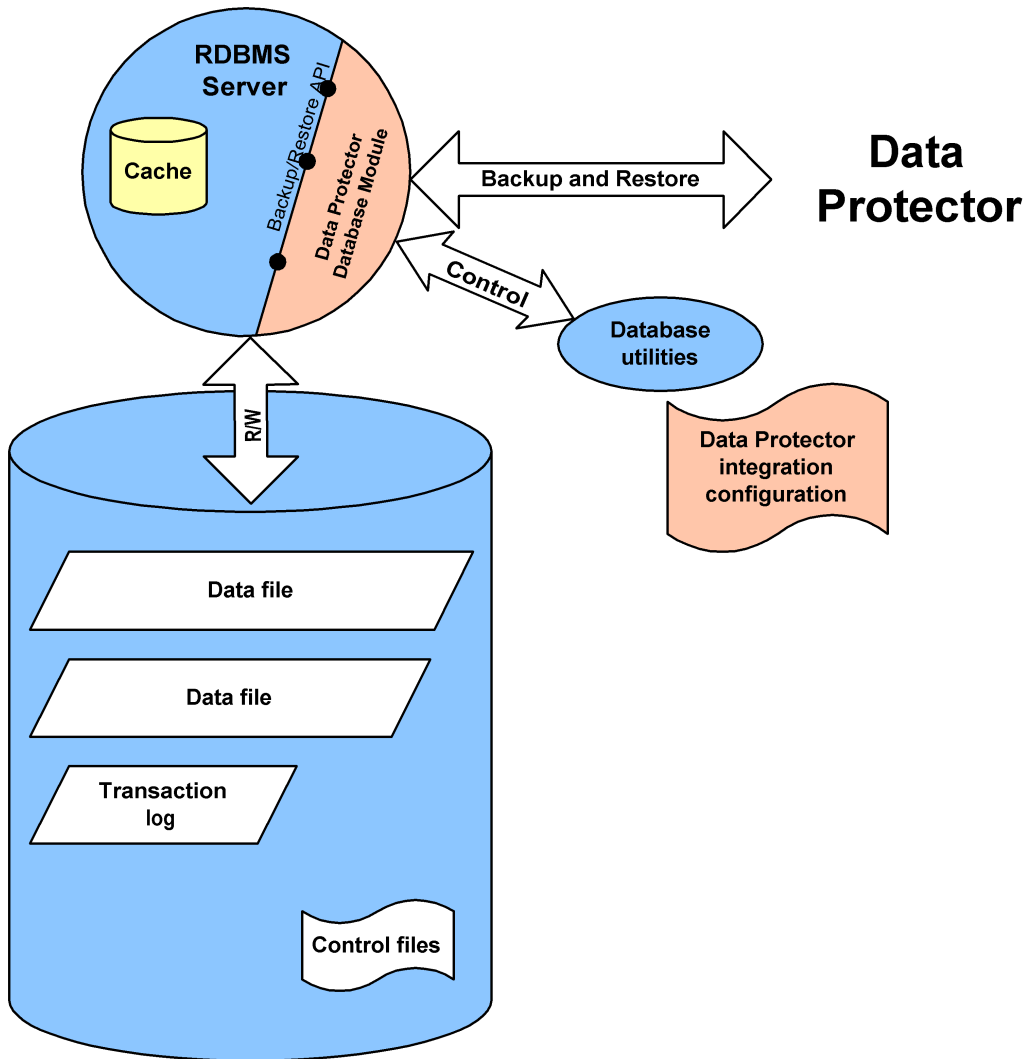
This method is relatively straightforward and easy to configure but has a key disadvantage of the database not being accessible during the backup, which is unacceptable for most business environments.

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, see the *HP 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 70: Data Protector integration with databases



"Data Protector integration with databases" above 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.

Integration with virtualization environments

This section gives a brief description of Data Protector integrations with virtualization environments. For a detailed list of supported environments, see the latest support matrices at <http://support.openview.hp.com/selfsolve/manuals>.

For more information, see the *HP Data Protector Integration Guide for Virtualization Environments*.

Offline filesystem backup of virtual machines

As virtual machines are constantly changing while they are online, you must put the virtual machines in a special mode or even shut them down before you start a filesystem backup.

Files on the disk that belong to a virtual machine have to be in a consistent state, otherwise the backup image created for that virtual machine is of no use.

To configure a filesystem backup of a virtual machine, you should:

1. Identify all the virtual machine files.
2. Choose two commands or prepare two scripts or applications that are able to shut down and start up the virtual machine.
3. Create a filesystem backup specification with all the virtual machine files included, and specify the shut-down command, script, or application for the as the pre-exec command and the start-up command, script, or application as a post-exec command.

This method is relatively straightforward, but has one key disadvantage: the virtual machine cannot be actively used during the backup.

Online backup of virtual machines

Data Protector can use specific interfaces provided by the virtualization environments to perform a backup of virtual machines while they are running (online backup). Depending on the virtualization environment, applications inside the virtual machines can also be put into a consistent state before the backup starts.

Besides the key advantage of the virtual machine 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.
- Virtualization environments are aware of the backup operation and keep track of which parts have been backed up.
- Several modes of backup are possible.
- Several modes of restore are possible.

Microsoft Volume Shadow Copy Service

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. An effective way to reduce the number of application specific implementation is the introduction of 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 Microsoft Windows operating systems. This service collaborates with the backup application, applications to be backed up, shadow copy providers, and the operating system to implement the management of volume shadow copies and shadow copy sets.

VSS provides a unified communication interface that can coordinate backup and restore of any application regardless of their specific features. A backup application does not need to handle each application to be backed up specifically, as long as the application confirms to the VSS specification.

What is a shadow copy?

A **shadow copy** refers to a volume that represents a copy of the original volume at a particular moment 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 Microsoft 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 has been a part of Windows operating systems starting with Windows Server 2003.

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.

["Actors of the traditional backup model" below](#) and ["Actors of the VSS backup model" on the next page](#) show differences between the traditional backup model and the model with the VSS coordinator.

Figure 71: Actors of the traditional backup model

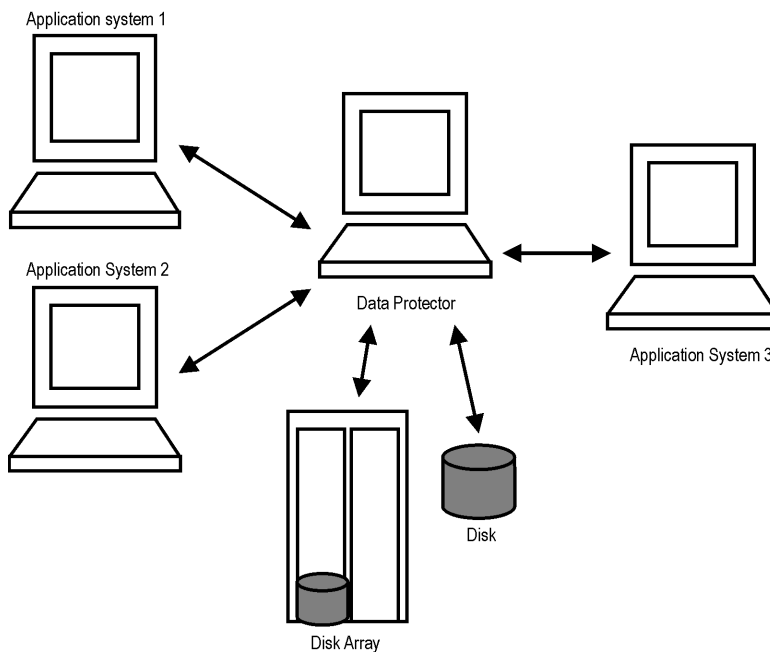
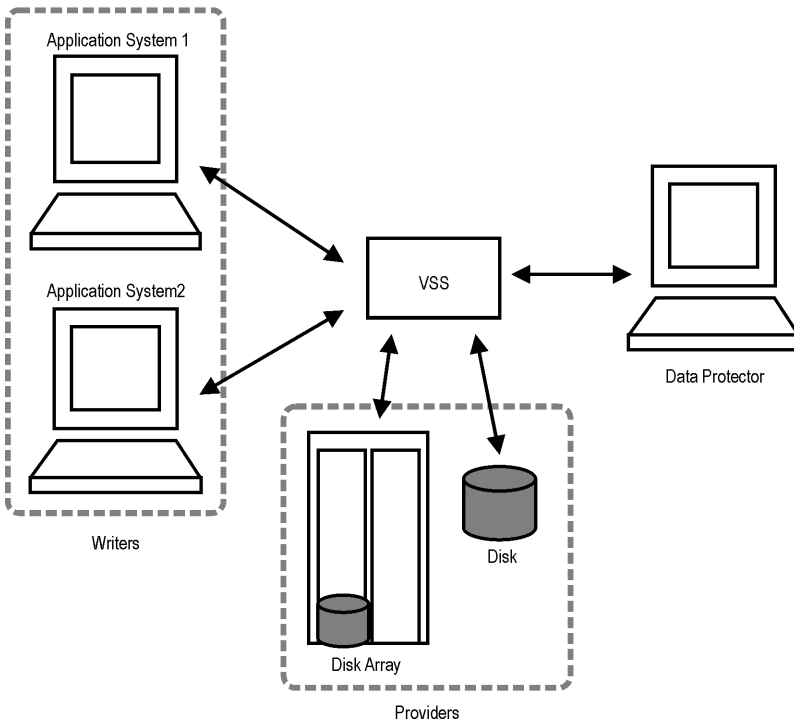


Figure 72: Actors of the VSS backup model



VSS benefits

The advantages of using Volume Shadow Copy Service are as follows

- A unified backup interface for all writers.
- A unified backup interface for all shadow copy providers.
- Writers provide data integrity at application level. Intervention from the backup application is unnecessary.

Data Protector supports the Microsoft Volume Shadow Copy Service at two levels:

- Within the Microsoft Volume Shadow Copy Service integration, Data Protector provides a shadow copy backup and restore of VSS-aware writers, including ZDB and instant recovery functionality.
- Within the Disk Agent functionality, Data Protector provides VSS filesystem backup.

The Data Protector VSS integration supports a consistent shadow copy backup only for VSS-aware writers. Consistency in this case is provided by the writer. Whenever applications are not VSS-aware the consistency of the shadow copy data is not guaranteed at application level, however, it is improved in comparison to a non-VSS filesystem backup.

The table below outlines the differences between using Data Protector VSS integration backup, VSS filesystem backup, and non-VSS filesystem backup:

Table 7: 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.
Locked files	No locked files.	No locked files.	If files are locked, backup skips them.
Data integrity	Provided by the writer.	Crash-consistent state (in the event of a power failure, for example).	None (inherent).

Data Protector Volume Shadow Copy integration

The Data Protector integration with Microsoft Volume Shadow Copy service provides full support for VSS-aware writers. This includes automatic detection of VSS-aware writers, and backup and restore functionality. The integration is primarily intended to back up application data.

For detailed information on the integration, see the *HP Data Protector Integration Guide for Microsoft Volume Shadow Copy Service*.

VSS filesystem and disk image backup and restore

Some applications are not aware of the Volume Shadow Copy Service. Such applications cannot guarantee consistency of data during the creation of a shadow copy. The VSS mechanism cannot coordinate the activities of these applications in order to perform a consistent backup. However, it can still assure a higher level of data consistency compared to a file system backup. Microsoft calls this state of data consistency “crash-consistent state”. This means that the VSS mechanism commits all pending I/O operations and holds incoming writing requests during the 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 a 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 with applications that are not aware of the VSS mechanism.

Important: When applications that are not aware of the VSS mechanism are being backed up, data consistency from the applications’ point of view cannot be guaranteed. Data consistency is the same as in the event of a power failure. Data Protector cannot guarantee any data consistency when applications are not actively participating in the creation of a shadow copy.

The consistency of data in a VSS filesystem and disk image backup is improved in comparison to a non-VSS backup. VSS allows you to create shadow copy backups of volumes and 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 or disk image backup. In this way, files that have changed during the backup procedure are copied correctly.

The advantages of VSS filesystem and disk image backup are as follows:

- 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 creation of the shadow copy.
- Backups can be performed at any time without locking out users.
- There is little or no impact on the performance of the application system during the backup process.

Backup and restore

VSS filesystem and VSS disk image backups are implemented as additional backup options on Windows Server 2003 and later operating systems. To enable VSS filesystem backup, you should specify it as a WinFS option. During the disk image backup, VSS writers are used by default. The level of data integrity is slightly improved in comparison to a traditional backup of active volumes. For detailed information on Windows filesystem and disk image backup and restore, see the *HP Data Protector Help*.

During a VSS filesystem and VSS disk image backup, applications cannot effectively contribute to data consistency because they are not aware of the VSS mechanism. However, Data Protector and a provider can still cooperate in creating volume shadow copies. VSS backup offers the option of backing up data as it appears at a certain point-in-time, regardless of system I/O activity during the backup.

When Data Protector requests a 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 it uses the newly created shadow copy instead for the source volume during the backup. If shadow copy creation fails, Data Protector will proceed with a traditional backup if a fallback 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 a backup. 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.

On Windows Vista, Windows 7, Windows 8, Windows Server 2008, and Windows Server 2012 systems, when preparing for EADR or OBDR, you can use VSS disk image backup functionality to back up volumes as disk images. Dismounted volumes, volumes mounted to NTFS folders, and the CONFIGURATION object cannot be backed up as disk images. Thus, such objects should be backed up as file system objects.

Note: To customize VSS disk image backup, use the omnirc options.

Chapter 9: Zero downtime backup and instant recovery

This chapter provides a basic overview of zero downtime backup and instant recovery concepts. These concepts are discussed in more detail in the *HP Data Protector Zero Downtime Backup Concepts Guide*.

Conventional methods of backing up data are not well suited to applications operating on large volumes of data; for example database applications. Either the application has to be taken offline or, if the application supports it, put into “hot-backup” mode while data in it is streamed to tape. The first can cause major disruption to the application’s operation. The second can produce large transaction log files, putting extra load on the application system.

Zero downtime backup (ZDB) and instant recovery (IR) have two great advantages over conventional backup and restore techniques:

- Minimal downtime or impact on the application system during the session
- Shorter restore times

Disk arrays and storage virtualization techniques

Large disk arrays that use RAID technology can hold large application databases, containing vast amounts of data. By using storage virtualization, disk arrays are usually divided into many virtual disks. These can easily be copied within a disk array, perhaps many times dependent on disk array technology and the available storage space. This makes it possible to perform operations on copies of data without exposing original data to risk. In particular, it enables effective backup solutions for applications in high-availability and mission-critical areas.

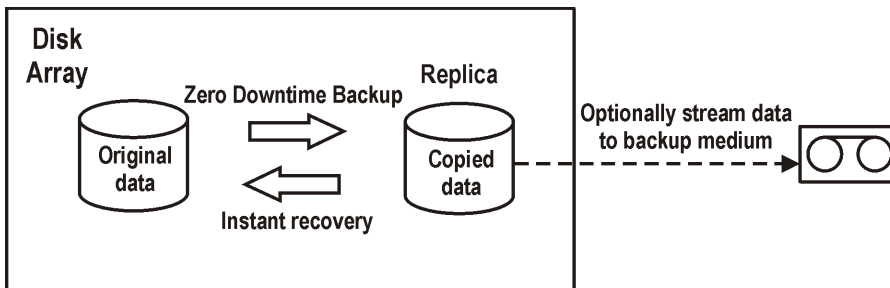
Zero downtime backup

Zero downtime backup uses disk array technology to minimize the disruption caused by the backup process. In very general terms, a copy or **replica** of the data is created or maintained on a disk array. This is very fast and has little impact on the application’s performance. The replica can itself form the backup, or it can be streamed to backup media without further interruption to the application’s use of the source database.

Depending on the hardware and software with which it is created, a replica may be an exact duplicate or a virtual copy of the data being backed up.

In ZDB, **replication** (the process of creating or maintaining a replica) is the critical factor in minimizing interruption to the application.

Figure 73: Zero downtime backup and instant recovery concept



Creating replicas

Replication creates a replica of application data or filesystem data at a particular moment.

The volumes containing the original data objects to be replicated are referred to as **source volumes**. These are replicated to an equivalent number of **target volumes**. When the replication process is complete, the data in the target volumes constitutes the replica.

Data Protector utilizes two basic replication techniques:

- **Split mirror**

A mirror is a dynamic duplicate of the source data, synchronized with it. Any changes to the source are automatically applied to the mirror as well.

To create a replica, the mirror is temporarily split from the source. Data is backed up from the mirror and the mirror is then resynchronized with the source.

- **Snapshot**

A snapshot replica is created by making a copy of data at a particular moment. The snapshot can be a full copy, thus independent of the source volume, or a virtual copy that still depends on the source volume.

ZDB types

After a replica has been created, it can be backed up to backup media. It is mounted to a **backup system** connected to the disk array on which the replica was created. To take full advantage of ZDB, this should be a separate system. There are then three forms of ZDB:

- **ZDB to tape**

- a. Data in the replica is streamed to tape according to the backup type you have selected (Full, Incr, Incr1-9).
- b. The replica can then be discarded.

Data can be restored from the tape using *standard* Data Protector restore techniques.

- **ZDB to disk**

The replica is kept on the disk array and used as the backup.

Data can be restored using *instant recovery*, which recovers the complete replica. See "[Instant recovery and restore of data](#)" below.

- **ZDB to disk+tape**
 - a. Data in the replica is streamed to tape according to the tape backup type you have selected (Full, Incr, Incr1-9).
 - b. The replica is kept on the disk array.

This provides extra flexibility because data can be restored in two ways:

- Using standard Data Protector restore from backup media (allowing restore of individual backup objects)
- Directly from the replica using instant recovery of the complete replica

Instant recovery and restore of data

Instant recovery

Instant recovery requires a replica to exist on the same disk array to which the data is to be restored. The application and backup systems are disabled and the contents of the replica are either restored directly to their original locations or presented to the system in place of contents of the source volumes. Because the instant recovery is performed internally within the disk array, it usually runs at very high speed.

Once the instant recovery is completed, the sections of the database or filesystem concerned are returned to their states at the time the replica was created and the application system is then re-enabled.

Depending on the application/database concerned, this may be all that is required. In some cases, additional action is required for full recovery, such as applying archived transaction log files that have been backed up separately.

Other restore methods

Data backed up to backup media can be restored using the **standard Data Protector restore** procedure.

However, with specific disk array families, it is possible to first restore data from backup media to update a replica and then restore the replica contents to their original locations. This is known as **split mirror restore**. Restoring the replica contents to their original locations is a similar process to instant recovery. It is only necessary to suspend application operation during this stage, minimizing the impact on the application.

Glossary

A

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.

AES 256-bit encryption

The Data Protector software encryption, based on the AES-CTR (Advanced Encryption Standard in Counter Mode) encryption algorithm that uses random keys of 256-bit length. The same key is used for both encryption and decryption. With AES 256-bit encryption, data is encrypted before it is transferred over a network and before it is written to media.

AML (ADIC/GRAU specific term)

Automated Mixed-Media library.

AMU (ADIC/GRAU specific term)

Archive Management Unit.

application agent

A component needed on a client system 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.

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.

archived log files (Data Protector specific term)

Files that keep track of changes made to the Data Protector Internal Database (IDB). They are used for online or offline IDB restore and recovery where the IDB needs to be recreated either in its latest possible state, beyond the time of the most recent IDB backup session, or in a state between the times of two consecutive IDB backup sessions.

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 an archived log destination. This copy is the archived redo log. The presence or absence of an archived redo log is determined by the mode the database is using: ARCHIVELOG - The filled online redo log files are archived before they are reused. The database can be recovered if an instance or a disk fails. A "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.

ASR set

A collection of files stored on several diskettes, required for proper reconfiguration of the replacement disk (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 the directory <Data_Protector_program_data>\Config\Server\dr\asr (Windows systems) or /etc/opt/omni/server/dr/asr (UNIX systems), as well as on the backup medium. After a disaster occurs, the ASR archive file is extracted to diskettes which you need to perform ASR.

audit logs

Data files to which auditing information is stored.

audit report

User-readable output of auditing information created from data stored in audit log files.

auditing information

Data about every backup session that was performed over an extended, user-defined period for the whole Data Protector cell.

autochanger

See library.

autoloader

See library.

Automatic Storage Management (ASM) (Oracle specific term)

A filesystem and volume manager integrated into Oracle which manages Oracle database files. It eliminates complexity associated with data and disk management and optimizes performance by providing striping and mirroring capabilities.

auxiliary disk

A bootable disk that has a minimal operating system with networking and

Data Protector Disk Agent installed. It can be carried around and used to boot the target system in Phase 1 of Disk Delivery Disaster Recovery of UNIX clients.

B

BACKINT (SAP R/3 specific term)

A Data Protector interface program that lets the SAP R/3 backup programs communicate with the Data Protector software via calls to an open interface. For backup and restore, SAP R/3 programs issue commands through the Data Protector backint interface.

backup API (Oracle specific term)

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 reading and writing of data to media and also to allow creation, searching, and removing of backup files.

backup chain

See restore chain.

backup device

A device configured for use with Data Protector that 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.

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/application 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 (for filesystem objects - the access point in a directory structure on the client where the backup object is located (drive on Windows systems and mount point on UNIX systems), for integration objects - backup stream identification, indicating the backed up database/application items), description (for filesystem objects - uniquely defines objects with identical client name and mount point, for integration objects - displays the integration type), and type (for filesystem objects - filesystem type, for integration objects - "Bar").

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. The result of a backup session is a set of media, which was written to, also called the backup or media set. See also backup specification, 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; and 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 Registry (for example). File selection lists, such as include-lists and exclude-lists, can be specified. All clients configured in one backup specification are backed up at the same time in one backup session using the same backup type (full or incremental).

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 of your 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 Continuances are processes that allow customers to access and manage instant copies of EMC Symmetrix standard devices.

BC Process (EMC Symmetrix specific term)

A protected storage environment solution, which uses specially configured EMC Symmetrix devices as mirrors or Business Continuation Volumes to protect data on EMC Symmetrix standard devices. See also BCV.

BCV (EMC Symmetrix specific term)

Business Continuation 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 Data Protector 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 containing 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 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 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; 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 BRARCHIVE and BRBACKUP.

BSM

The Data Protector Backup Session Manager, which controls the backup session. This process always runs on the Cell Manager system.

C

CAP (StorageTek specific term)

The Cartridge Access Port built into the door panel of a library. The purpose is to enter or eject media.

Catalog Database (CDB)

A part of the Data Protector Internal Database (IDB) that contains information about backup, restore, object copy, object consolidation, object verification, and media management sessions. This part of the IDB is always local to the cell. It is stored in the embedded database. See also MMDB.

catalog protection

Defines how long information about backed up data (such as filenames and file attributes) is kept in the IDB. See also data protection.

CDB

See Catalog Database (CDB).

CDF file (UNIX systems 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.

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.

Certificate Server

A Windows certificate server can be installed and configured to provide certificates for clients. It provides customizable services for issuing and managing certificates for the enterprise. These services issue, revoke, and manage certificates employed in public key-based cryptography technologies.

Change Journal (Windows specific term)

A Windows filesystem feature that logs a record of each change as it occurs to the files and directories on a local NTFS volume.

Change Log Provider

A module that can be queried to determine which objects on a filesystem have been created, modified, or deleted.

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' or type 'sbt_tape'. If the specified channel is of 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 volumes (filesystems) mounted on a Data Protector client. What is actually backed up depends on how you select objects in a backup specification. If you select the check box next to the client system name, a single backup object of the Client System type is created. As a result, at the time of the backup, Data Protector first detects all volumes that are mounted on the selected client and then backs them up. On Windows clients, CONFIGURATION is also backed up. If you individually select all volumes that are mounted on the client system, a separate backup object of the Filesystem type is created for each volume. As a result, at the time of the backup, only the selected volumes are backed up. Volumes that have been potentially mounted on the client after the

backup specification was created are not backed up.

client or client system

Any system configured with any Data Protector functionality and configured in a cell.

cluster continuous replication (Microsoft Exchange Server specific term)

Cluster continuous replication (CCR) is a high availability solution that uses cluster management and failover options to create and maintain an exact copy (CCR copy) of a storage group. A storage group is replicated to a separate server. CCR removes any single point of failure in your Exchange back-end servers. You can perform backups using VSS on your passive Exchange Server node where a CCR copy is located and thus reducing the load on the active node. A CCR copy is used for disaster recovery since you can switch to the CCR copy in a few seconds. A replicated storage group is represented as a new instance of Exchange writer called Exchange Replication Service and can be backed up (using VSS) like an ordinary storage group. See also Exchange Replication Service and local continuous replication.

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 HP Serviceguard), application services, IP names and addresses ...).

CMD script for Informix Server (Informix Server specific term)

A Windows CMD script that is created in INFORMIXDIR when an Informix Server database is configured. The CMD script

is a set of system commands that export environment variables for Informix 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 Manager-of-Managers. A reliable network connection between the MoM cell and the other Data Protector cells is highly recommended. See also MoM.

COM+ Class Registration Database (Windows specific term)

The COM+ Class Registration Database and the Windows Registry store COM+ application attributes, class attributes, and computer-level attributes. This guaranties consistency among these attributes and provides common operation on top of these attributes.

command device (HP P9000 XP Disk Array Family specific term)

A dedicated volume in the disk array which acts as the interface between a management application and the disk array's storage system. It cannot be used for data storage and only accepts requests for operations that are then executed by the disk array.

command-line interface (CLI)

A set commands that you can use in shell scripts to perform Data Protector configuration, backup, restore, and management tasks.

concurrency

See Disk Agent concurrency.

container (HP P6000 EVA Disk Array Family specific term)

Space on a disk array, which is pre-allocated for later use as a standard snapshot, vsnap, or snapclone.

control file (Oracle and SAP R/3 specific term)

A data file that contains entries specifying the physical structure of the database. It provides database consistency information used for recovery.

copy set (HP P6000 EVA Disk Array Family specific term)

A pair that consists of the source volumes on a local P6000 EVA and their replica on a remote P6000 EVA. See also source volume, replica, and HP Continuous Access + Business Copy (CA+BC) P6000 EVA.

CRS

The Data Protector Cell Request Server process (service), which runs on the Cell Manager, starts and controls the backup and restore sessions. The service is started as soon as Data Protector is installed on the Cell Manager. The CRS runs under the account root on UNIX systems. On Windows systems it runs under the account of the user, specified at installation time.

CSM

The Data Protector Copy and Consolidation Session Manager process controls the object copy and object consolidation sessions and runs on the Cell Manager system.

D**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 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.

data replication (DR) group (HP P6000 EVA Disk Array Family specific term)

A logical grouping of HP P6000 EVA Disk Array Family virtual disks. It can contain up to eight copy sets provided they have common characteristics and share a common HP CA P6000 EVA log. See also copy set.

data stream

Sequence of data transferred over the communication channel.

Data_Protector_home

A reference to the directory containing Data Protector program files (on Windows Vista, Windows 7, Windows Server 2008, and Windows Server 2012) or the directory containing Data Protector program files and data files (on other Windows operating systems). Its default path is %ProgramFiles%\OmniBack, but the path can be changed in the Data Protector Setup Wizard at installation time. See also Data_Protector_program_data.

Data_Protector_program_data

A reference to the directory containing Data Protector data files on Windows Vista, Windows 7, Windows Server 2008, and Windows Server 2012. Its default path is %ProgramData%\OmniBack, but the path can be changed in the Data Protector Setup Wizard at installation time. See also Data_Protector_home.

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, 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.

Dbobject (Informix Server specific term)

An Informix Server physical database object. It can be a blob space, db space, or logical log file.

DC directory

A directory that contains DC binary files, one for each configured Data Protector backup medium. DC directories constitute the Detail Catalog Binary Files part of the Data Protector Internal Database. See also Detail Catalog Binary Files (DCBF) and Internal Database (IDB).

DCBF

See Detail Catalog Binary Files (DCBF).

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.

Detail Catalog Binary Files (DCBF)

A part of the Data Protector Internal Database that stores names, versions, and metadata of the backed up items. It consists of DC directories with DC binary files. See also DC directory and Internal Database (IDB).

device

See backup device.

device chain

A device chain consists of several standalone devices configured for sequential use. When a medium in one 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. -Get configuration, status, and performance statistics by device group. -Issue control operations that apply to all devices in the device group.

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 writing to the tape, and so on. If data is written to the tape slower than it is delivered to the device then the device is streaming. Streaming significantly improves the use of space and the performance of the device.

DHCP server

A system running the Dynamic Host Configuration Protocol (DHCP) providing dynamic IP address assignment and network configuration for DHCP clients.

differential backup

An incremental backup that backs up changes made since the last full backup. To perform this type of backup, specify the Incr1 backup type. See also incremental backup.

differential backup (Microsoft SQL Server 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.

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.

disaster recovery operating system

See DR OS.

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. During an object verification session the Disk Agent receives data from the Media Agent and performs the verification process, but no data is written to disk.

Disk Agent concurrency

The number of Disk Agents that are allowed to send data to one Media Agent concurrently.

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 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 media format

A media format, available with the file library, which supports a space efficient type of synthetic backup called virtual full backup. Using this format is a prerequisite for virtual full backup. See also virtual full backup.

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.

DR image

Data required for temporary disaster recovery operating system (DR OS) installation and configuration.

DR OS

An operating system environment in which disaster recovery runs. It provides Data Protector with a basic runtime environment (disk, network, tape, and filesystem access). It has to be installed on disk or loaded into memory and configured before the Data Protector disaster recovery can be performed. DR OS can be either temporary or active. A temporary DR OS is used exclusively as a host environment for some other operating system restore along with the target operating system configuration data. It is deleted after the target system is restored to the original system configuration. An active DR OS not only hosts the Data Protector disaster recovery process but can also be 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. It can also read the data from the medium and send it to the computer system.

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.

drive-based encryption

The Data Protector drive-based encryption uses the encryption functionality of the drive. While performing the backup, the drive encrypts both the data and the metadata that is written to the medium.

E

EMC Symmetrix Agent

A Data Protector software module that prepares the EMC Symmetrix environment for backup and restore operations.

emergency boot file (Informix Server specific term)

The Informix Server configuration file `ixbar.<server_id>` that resides in the directory `<INFORMIXDIR>/etc` (on Windows systems) or `<INFORMIXDIR>/etc` (on UNIX systems). `<INFORMIXDIR>` is the Informix 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.

encrypted control communication

Data Protector secure communication between the clients in the Data Protector cell is based on a Secure Socket Layer (SSL) that uses export grade SSLv3 algorithms to encrypt control communication. Control communication in a Data Protector cell is all communication between Data Protector processes, except the data transfer from Disk Agent (and Integrations) to Media Agent, and the other way round.

encryption key

A 256-bit randomly generated number used by the Data Protector encryption algorithm to encode information during backups for which AES 256-bit software encryption or drive-based encryption has been specified. The same key is used for subsequent decryption of the information. Encryption keys for a Data Protector cell are stored in a central keystore on the Cell Manager.

encryption KeyID-StoreID

Combined identifier used by the Data Protector Key Management Server to identify and administer encryption keys used by Data Protector. KeyID identifies the key within the keystore. StoreID identifies the keystore on the Cell Manager. If Data Protector has been upgraded from an earlier version with encryption functionality, there may be several StoreIDs used on the same Cell Manager.

enhanced incremental backup

Conventional incremental backup backs up files that have changed since a previous backup, but has certain limitations in detection of changes. Unlike conventional incremental backup, enhanced incremental backup reliably detects and backs up also renamed and moved files, as well as files with changes in attributes.

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 Log (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.

Event Logs (Windows specific term)

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.

Exchange Replication Service (Microsoft Exchange Server specific term)

The Microsoft Exchange Server service that represents storage groups that were replicated using either local continuous replication (LCR) or cluster continuous replication (CCR) technology. See also cluster continuous replication and local continuous replication.

exchanger

See library.

exporting media

A process that removes all data about backup sessions, such as systems, objects, and filenames, 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.

F**failover**

Transferring of the most important cluster data, called group (on Windows systems) or package (on UNIX systems) from one cluster node to another. A failover can occur mostly because of software or hardware failures or maintenance on the primary node.

**failover (HP P6000 EVA Disk Array
Family specific term)**

An operation that reverses the roles of source and destination in HP Continuous Access + Business Copy (CA+BC) P6000 EVA configurations. See also HP Continuous Access + Business Copy (CA+BC) P6000 EVA.

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 bi-directional 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 tree walk

The process of traversing a filesystem to determine which objects have been created, modified, or deleted.

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.

first-level mirror (HP P9000 XP Disk Array Family specific term)

A mirror of an internal disk (LDEV) of a disk array of the HP P9000 XP Disk Array Family which can be further mirrored itself, producing second-level mirrors. For Data Protector zero downtime backup and instant recovery purposes, only first-level mirrors can be used. See also primary volume and mirror unit (MU) number.

flash recovery area (Oracle specific term)

A directory, filesystem, or Automatic Storage Management (ASM) disk group managed by Oracle that serves as a centralized storage area for files related to backup, restore, and database recovery (recovery files). See also recovery files.

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). Use the Force operation option to reformat Data Protector media with non-protected data. 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-to-tape or ZDB-to-disk+tape session in which all selected objects are streamed to tape, even if there are no changes from the previous backup. See also incremental ZDB.

G

global options

A set of options that define behavior of the entire Data Protector cell. The options are stored in a plain text file on the Cell Manager.

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 graphical user interface provided by Data Protector for easy access to all configuration, administration, and operation tasks. It is available for Microsoft Windows operating systems.

H**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 that resides on the Cell Manager at the following location: <Data_Protector_program_data>\Config\Server\holidays (Windows systems) and /etc/opt/omni/server/Holidays (UNIX systems).

hosting system

A working Data Protector client used for Disk Delivery Disaster Recovery with a Data Protector Disk Agent installed.

HP Business Copy (BC) P6000 EVA (HP P6000 EVA Disk Array Family specific term)

A local replication software solution that enables creation of point-in-time copies (replicas) of the source volumes using the snapshot and clone capabilities of the P6000 EVA firmware. See also replica, source volume, snapshot, and HP Continuous Access + Business Copy (CA+BC) P6000 EVA.

HP Business Copy (BC) P9000 XP (HP P9000 XP Disk Array Family specific term)

An HP P4000 SAN Solutions configuration that enables creation and maintenance of internal copies of LDEVs for various purposes, such as data duplication and backup. The copies (secondary volumes or S-VOLs) can be separated from the primary volumes (P-VOLs) and connected to a different system. For Data Protector zero downtime backup purposes, P-VOLs should be available to the application system, and one of the S-VOL sets should be available to the backup system. See also LDEV, HP Continuous Access (CA) P9000 XP, Main Control Unit (MCU), application system, and backup system.

HP Command View (CV) EVA (HP P6000 EVA Disk Array Family specific term)

The user interface that enables you to configure, manage, and monitor your P6000 EVA storage system. It is used to perform various storage management tasks, for example, creating virtual disk families, managing storage system hardware, and creating snapshots, snapclones, or mirrorclones of virtual disks. The HP Command View EVA software runs on the HP Storage Management Appliance, and is accessed

by a Web browser. See also HP P6000 / HP 3PAR SMI-S Agent and HP SMI-S P6000 EVA Array provider.

HP Continuous Access (CA) P9000 XP (HP P9000 XP Disk Array Family specific term)

An HP P9000 XP Disk Array Family configuration that enables creation and maintenance of remote copies of LDEVs for purposes such as data duplication, backup, and disaster recovery. HP CA P9000 XP operations involve main (primary) disk array units and remote (secondary) disk array units. The main disk array units are connected to the application system and contain primary volumes (P-VOLs), which store original data. The remote disk array units are connected to the backup system and contain secondary volumes (S-VOLs). See also HP BC P9000 XP (HP P9000 XP Disk Array Family specific term), Main Control Unit (MCU), and LDEV.

HP Continuous Access + Business Copy (CA+BC) P6000 EVA (HP P6000 EVA Disk Array Family specific term)

An HP P6000 EVA Disk Array Family configuration that enables creation and maintenance of copies (replicas) of the source volumes on a remote P6000 EVA, and later use of these copies as the source for local replication on this remote array. See also HP BC P6000 EVA, replica, and source volume.

HP P6000 / HP 3PAR SMI-S Agent

A Data Protector software module that executes all tasks required for the HP P6000 EVA Disk Array Family integration. With the P6000 / 3PAR SMI-S Agent, the control over the array is established through an appropriate SMI-S provider, which directs communication between incoming requests and the storage system's native interface. See

also HP Command View (CV) EVA and HP SMI-S P6000 EVA Array provider.

HP P9000 XP Agent

A Data Protector software component that executes all tasks needed by the Data Protector HP P9000 XP Disk Array Family integration. It communicates with the P9000 XP Array storage system via the RAID Manager Library.

HP SMI-S P6000 EVA Array provider

An interface used for controlling HP P6000 EVA Disk Array Family. SMI-S P6000 EVA Array provider runs as a separate service on the HP Storage Management Appliance system and acts as a gateway between incoming requests and HP Command View EVA. With the Data Protector HP P6000 EVA Disk Array Family integration, SMI-S P6000 EVA Array provider accepts standardized requests from the HP P6000 / HP 3PAR SMI-S Agent, communicates with HP Command View EVA for information or method invocation, and returns standardized responses. See also HP P6000 / HP 3PAR SMI-S Agent and HP Command View (CV) EVA.

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

See Internal Database (IDB).

IDB recovery file

A file that maintains information about completed IDB backup sessions and the

backup media and backup devices used in them. If available, it significantly simplifies and speeds up offline recovery of the Internal Database in case of a Cell Manager disaster. Its filename is obdrindex.dat.

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 (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 backup

A backup that selects only files that have changed since a previous backup. Several levels of incremental backup are available, which enables detailed control of restore chain length. 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.

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 filesystem ZDB-to-tape or ZDB-to-disk+tape session in which only changes from the last protected full or incremental backup are streamed to tape. See also full ZDB.

Incremental1 Mailbox Backup

An incremental1 mailbox backup backs up all the changes made to the mailbox after the last full backup.

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 that are shared among several users. See also Key Management Service and Site Replication Service.

Informix Server (Informix Server specific term)

Refers to Informix Dynamic Server.

initializing

See formatting.

Installation Server

A computer system that holds a repository of the Data Protector installation 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.

integration object

A backup object of a Data Protector integration, such as Oracle or SAP MaxDB.

Internal Database (IDB)

An entity in Data Protector that keeps information regarding which data was backed up, to which media it was backed up, how and when backup, restore, and other sessions were run, which devices, libraries, and disk arrays are configured, and so on. It is implemented with an embedded database and a collection of proprietary data files which reside on the Cell Manager. See also DC directory and Detail Catalog Binary Files (DCBF).

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).

ISQL (Sybase specific term)

A Sybase utility used to perform system administration tasks on Sybase SQL Server.

J**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'.

K

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.

keychain

A tool that eliminates the supply of a passphrase manually when decrypting the private key. It needs to be installed and configured on the Installation Server if you perform remote installation using secure shell.

keystore

All encryption keys are centrally stored in the keystore on the Cell Manager and administered by the Key Management Server (KMS).

KMS

Key Management Server (KMS) is a centralized service that runs on the Cell Manager and provides key management for the Data Protector encryption functionality. The service is started as soon as Data Protector is installed on the Cell Manager.

L

LBO (Symmetric specific term)

A 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.

LDEV (HP P9000 XP Disk Array Family specific term)

A logical partition of a physical disk of a disk array of the HP P9000 XP Disk Array Family. An LDEV is the entity that can be replicated using the split-mirror or snapshot functionality of such disk array. See also HP Business Copy (BC) P9000 XP, HP Continuous Access (CA) P9000 XP, and replica.

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 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.

local continuous replication (Microsoft Exchange Server specific term)

Local continuous replication (LCR) is a single-server solution that creates and maintains an exact copy (LCR copy) of a storage group. An LCR copy is located on the same server as the original storage group. When an LCR copy is created, it is kept up to date through change propagation (log replay) technology. The replication feature in LCR guarantees that logs that have not been replicated are not deleted. The implication of this behavior is that running backups in a mode that deletes logs may not actually free space if replication is sufficiently far behind in its log copying. An LCR copy is used for disaster recovery because you can switch to the LCR copy in a few seconds. If an LCR copy is used for backup and if it is located on a different disk than the original data, then the I/O load on a production database is minimal. A replicated storage group is represented as a new instance of Exchange writer called Exchange Replication Service and

can be backed up (using VSS) as a normal storage group. See also cluster continuous replication and Exchange Replication Service.

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 Server UNIX systems specific term)

A script provided by ON-Bar that you can use to start backing up logical log files when Informix Server issues a logfull event alarm. The Informix Server ALARMPROGRAM configuration parameter defaults to the `<INFORMIXDIR>/etc/log_full.sh`, where `<INFORMIXDIR>` is the Informix 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

An option that determines the amount of details on files and directories written to the IDB during backup, object copying, or object consolidation. 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 mainly influence the IDB growth 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 committed as well as roll back any transactions that have not been committed.

login ID (Microsoft SQL Server specific term)

The name a user needs 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 Oracle SYSDBA or SYSOPER rights. <password> must 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 must be the owner of the Oracle Recovery 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 consists of several volume groups, where each volume group has several volumes.

M**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 in 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 P9000 XP Disk Array Family specific term)

An HP P9000 XP Disk Array Family unit that contains primary volumes (P-VOLs) for the HP CA P9000 XP Array or HP CA+BC P9000 XP Array configuration and acts as a master device. See also HP Business Copy (BC) P9000 XP, HP Continuous Access (CA) P9000 XP, and LDEV.

maintenance mode

An operating mode you can initiate on the Cell Manager to prevent changes in the Internal Database. It enables you to perform various maintenance tasks, such as upgrading and patching the installation.

make_net_recovery

make_net_recovery is an Ignite-UX command, which allows you to create a recovery archive over the network onto the Ignite-UX server or any other specified system. The target system can be recovered across subnets after booting either from a bootable tape created by the Ignite-UX make_boot_tape command or the system boots directly from the Ignite-UX server. Booting directly from the Ignite-UX server can be automated with the Ignite-UX bootsys command or interactively specified on the boot console.

make_tape_recovery

make_tape_recovery is a command on Ignite-UX which creates a bootable recovery (installation) tape, customized for your system and enables you unattended disaster recovery by connecting the backup device directly to the target system and booting the target system from the bootable recovery tape. The backup device has to be locally connected to the client during the creation of the archive and recovery of the client.

Manager-of-Managers

See MoM.

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.

MCU

See Main Control Unit (MCU).

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 or object verification session, a Media Agent locates data on the backup medium and sends it to the Disk Agent for processing. For a restore session, the Disk Agent then writes the data to the disk. A Media Agent also manages the robotics control of a library.

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.

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 assert the condition of a medium.

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.

medium ID

A unique identifier assigned to a medium by Data Protector.

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.

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, replica, and writer.

mirror (EMC Symmetrix and HP P9000 XP Disk Array Family specific term)

See target volume.

mirror rotation (HP P9000 XP Disk Array Family specific term)

See replica set rotation.

mirror unit (MU) number (HP P9000 XP Disk Array Family specific term)

A non-negative integer number that determines a secondary volume (S-VOL) of an internal disk (LDEV) located on a disk array of the HP P9000 XP Disk Array Family. See also first-level mirror.

mirrorclone (HP P6000 EVA Disk Array Family specific term)

A dynamic replica of a storage volume, which is kept updated with changes made to the original storage volume via a local replication link. Replication between the original storage volume and its mirrorclone can be suspended. For each storage volume, a single mirrorclone can be created on the disk array.

MMD

The Media Management Daemon process (service) (MMD) 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 environment, this part of the database can be common to all cells. See also CMMDB and Catalog Database (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 cells are called MoM clients. The MoM enables you to configure and manage multiple cells from a central point.

mount point

The access point in a directory structure for a disk or logical volume, for example /opt or d:. On UNIX systems, the mount points are displayed using the bdf or df command.

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.

MSM

The Data Protector Media Session Manager, which runs on the Cell Manager and controls media sessions, such as copying media.

multisnapping (HP P6000 EVA Disk Array Family specific term)

Simultaneous creation of target volumes so that the backup data is consistent not only on each individual target volume, but also across all the volumes that constitute a snapshot. See also snapshot.

O

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.

obdrindex.dat

See IDB recovery file.

object

See backup object.

object consolidation

The process of merging 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. The process is a part of the synthetic backup procedure. The result is a synthetic full backup of the specified backup object.

object consolidation session

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.

object copy

A copy of a specific object version that is created during an object copy session or a backup session with object mirroring.

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.

object verification

The process of verifying the data integrity of backup objects, from the Data Protector point of view, and the ability of Data Protector to deliver them to the required destination. The process can be used to provide a level of confidence in the ability to restore object versions created by backup, object copy, or object consolidation sessions.

object verification session

A process that verifies the data integrity of specified backup objects or object versions and the ability of selected Data Protector network components to deliver them to a specified host. Object verification sessions can be run interactively, or as specified in automated post-backup, or scheduled specifications.

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. Cell Manager can only be recovered offline.

offline redo log

See archived redo log.

ON-Bar (Informix Server specific term)

A backup and restore system for Informix Server. ON-Bar enables you to create a copy of your Informix Server data and later restore the data. The ON-Bar backup and restore system involves the following components: the onbar command, Data Protector as the backup solution, the XBSA interface, and ON-Bar catalog tables, which are used to back up dbobjects and track instances of dbobjects through multiple backups.

ONCONFIG (Informix Server specific term)

An environment variable that specifies the name of the active ONCONFIG configuration file. If the ONCONFIG environment variable is not present, Informix Server uses the configuration values from the onconfig file located in

the directory <INFORMIXDIR>\etc (on Windows systems) or <INFORMIXDIR>/etc/ (on UNIX systems).

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.

online recovery

A type of the Internal Database recovery that you can use when the Cell Manager is accessible. In this case, the Cell Manager runs sessions, the sessions are logged into the IDB, and you can monitor the session progress using the GUI.

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.

OpenSSH

A set of network connectivity tools used to access remote machines securely, by using a variety of authentication and encryption methods. It needs to be installed and configured on the Installation Server and the client if you perform remote installation using secure shell.

Oracle Data Guard (Oracle specific term)

Oracle Data Guard is Oracle's primary disaster recovery solution. Oracle Data Guard is able to maintain up to nine standby databases, each of which is a real-time copy of the production (primary) database, to protect against corruptions, data failures, human errors, and disasters. If a failure occurs on the production (primary) database, then a failover to one of the standby databases which becomes the new primary database is possible. In addition, planned downtime for maintenance can be reduced because the production processing can be moved from the current primary database to a standby database and back quickly.

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.

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 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. When copying or consolidating backup objects, the owner of the resulting objects is the user who started the original backup session.

P

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 the

directory <Data_Protector_program_data>\Config\Server\dr\p1s (Windows systems) or /etc/opt/omni/dr/p1s (UNIX systems) with the filename recovery.p1s.

package (HP 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 P9000 XP Disk Array Family specific term)

The status of a disk pair (secondary volume and its corresponding primary volume) of a disk array of the HP P9000 XP Disk Array Family. Depending on the circumstances, the paired disks can be in various states. The following states are particularly important for the operation of the Data Protector HP P9000 XP Agent:

PAIR - The secondary volume is prepared for zero downtime backup. If it is a mirror, it is completely synchronized, and if it is a volume to be used for snapshot storage, it is empty.

SUSPENDED - The link between the disks is suspended. However, the pair relationship is still maintained, and the secondary disk can be prepared for zero downtime backup again at a later time.

COPY - The disk pair is currently busy and making a transition into the PAIR state. If the secondary volume is a mirror, it is re-synchronizing with the primary volume, and if it is a volume to be used for snapshot storage, its contents are getting cleared.

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.

parallelism

The concept of reading multiple data streams from an online database.

phase 0 of disaster recovery

Preparation for disaster recovery - the prerequisite condition for a successful disaster recovery.

phase 1 of disaster recovery

Installation and configuration of DR OS, establishing previous storage structure.

phase 2 of disaster recovery

Restoration of operating system (with all the configuration information that defines the environment) and Data Protector.

phase 3 of disaster recovery

Restoration of user and application data.

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 systems and as shell scripts on UNIX systems. 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 systems and as shell scripts on UNIX systems.

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 systems and as shell scripts on UNIX systems. See also post-exec.

primary volume (P-VOL) (HP P9000 XP Disk Array Family specific term)

An internal disk (LDEV) of a disk array of the HP P9000 XP Disk Array Family for which a secondary volume (S-VOL), either its mirror or a volume to be used for its snapshot storage, exists. In the HP CA P9000 XP and HP CA+BC P9000 XP configurations, primary volumes are located in the Main Control Unit (MCU). See also secondary volume (S-VOL) and Main Control Unit (MCU).

protection

See data protection and 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: 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

R

RAID

Redundant Array of Independent Disks.

RAID Manager Library (HP P9000 XP Disk Array Family specific term)

A software library that is used for accessing the configuration, status, and performance measurement data of a P9000 XP Array storage system, and for invoking operations on the disk array. It translates function calls into sequences of low-level SCSI commands. See also HP P9000 XP Agent.

RAID Manager P9000 XP (HP P9000 XP Disk Array Family specific term)

A software application that provides a command-line interface to disk arrays of the HP P9000 XP Disk Array Family. It offers an extensive set of commands for reporting and controlling the status of a P9000 XP Array storage system, and for performing various operations on the disk array.

rawdisk backup

See disk image backup.

RCU

See Remote Control Unit (RCU).

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.

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, 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, and 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.

recovery files (Oracle specific term)

Recovery files are Oracle specific files that reside in the flash recovery area: the current control file, online redo logs, archived redo logs, flashback logs, control file autobackups, datafile copies, and backup pieces. See also flash recovery area.

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.

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.

recycle or unprotect

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 P9000 XP Disk Array Family specific term)

An HP P9000 XP Disk Array Family unit that acts as a slave device to the Main Control Unit (MCU) in the HP CA P9000 XP or HP CA+BC P9000 XP configuration. In bidirectional configurations, the RCU can also act as an MCU.

Removable Storage Management Database (Windows specific term)

A Windows service used for managing removable media (such as tapes and 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 or 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 perspective of a basic operating system, the complete physical disk containing a backup object is replicated. However, if a volume manager is used on a UNIX system, the whole volume/disk group containing a backup object is replicated. If partitions are used on a Windows system, the whole physical volume containing the selected partition 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 chain

Backup images that are needed to restore a backed up object to the state it was in at the selected point in time. In general, a restore chain of an object consists of its full backup image and one or more related incremental backup images.

restore session

A process that copies data from backup media to a client.

resync mode (HP P9000 XP Disk Array Family VSS provider specific term)

One of two P9000 XP Array VSS hardware provider operation modes. When the P9000 XP Array provider is in the resync mode, the source volume (P-VOL) and its replica (S-VOL) are in the suspended mirror relationship after a backup. The maximum number of replicas (S-VOLs per a P-VOL) rotated is three provided that MU range is 0-2 or 0, 1, 2. Restore from a backup in such a configuration is possible only by re-synchronization of an S-VOL with its P-VOL. See also VSS compliant mode, source volume, primary volume (P-VOL), replica, secondary volume (S-VOL), MU number, and replica set rotation.

RMAN (Oracle specific term)

See Recovery Manager.

RSM

The Data Protector Restore Session Manager controls restore and object

verification sessions. This process always runs on the Cell Manager system.

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.

S**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 P9000 XP Disk Array Family specific term)

An internal disk (LDEV) of a disk array of the HP P9000 XP Disk Array Family which is paired with another LDEV: a primary volume (P-VOL). It can act as a mirror of the P-VOL or as a volume to be used for the P-VOL's snapshot storage. An S-VOL is assigned a SCSI address different from the one used for the P-VOL. In an HP CA P9000 XP configuration, the S-VOLs acting as mirrors can be used as failover devices in a MetroCluster configuration. See also primary volume (P-VOL) and Main Control Unit (MCU).

session

See backup session, media management session, and restore session.

session ID

An identifier of a backup, restore, object copy, object consolidation, object verification, 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 it is used for specifying options for the omnimnt, omnistat, and omniabort CLI commands.

shadow copy (Microsoft 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 and replica.

shadow copy provider (Microsoft 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 (Microsoft VSS specific term)

A collection of shadow copies created at the same point in time. See also shadow copy and replica set.

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.

Site Replication Service (Microsoft Exchange Server specific term)

The Microsoft Exchange Server 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, restore, object copy, object consolidation, object verification, and media management sessions. One binary file is created per session. The files are grouped by year and month.

SMI-S Agent (SMISA)

See HP P6000 / HP 3PAR SMI-S Agent.

snapshot (HP P4000 SAN Solutions, HP P6000 EVA Disk Array Family, HP P9000 XP Disk Array Family, and HP 3PAR StoreServ Storage specific term)

A type of target volumes created using a specific replication technology.

Depending on the disk array model and the chosen replication technique, a range of snapshot types with different characteristics is available. Basically, each snapshot may be either a virtual copy, still reliant upon the contents of the source volume, or an independent duplicate (clone) of the source volume. See also replica and snapshot creation.

snapshot backup

See ZDB to tape, ZDB to disk, and ZDB to disk+tape.

snapshot creation (HP P4000 SAN Solutions, HP P6000 EVA Disk Array Family, HP P9000 XP Disk Array Family, and HP 3PAR StoreServ Storage specific term)

A replica creation process in which copies of the selected source volumes are created using storage virtualization technology. Such a replica is considered to be created at a particular point in time, and is immediately available for use. However, with certain snapshot types, a background data copying process continues to run on the disk array after the moment of the replica 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 P9000 XP Disk Array Family specific term)

A type of target volumes created using a specific replication technology. A split-mirror replica provides independent duplicates (clones) of the source volumes. See also replica and split mirror creation.

split mirror backup (EMC Symmetrix specific term)

See ZDB to tape.

split mirror backup (HP P9000 XP Disk Array Family specific term)

See ZDB to tape, ZDB to disk, and ZDB to disk+tape.

split mirror creation (EMC Symmetrix and HP P9000 XP Disk Array Family 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 P9000 XP Disk Array Family specific term)

A process in which data backed up in a ZDB-to-tape or a ZDB-to-disk+tape session is first copied from the backup media to a replica, and from the replica to the source volumes afterwards.

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 or registry (Informix Server specific term)

An Informix Server connectivity information file (on UNIX systems) or registry (on Windows systems) 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

A text file in the Unicode (UTF-16) format, generated during CONFIGURATION backup of a Windows system and stored on the Cell Manager. It contains system information required for installing and configuring the operating system on the target system in the case of a disaster. See also target system.

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 (SSEA)

See HP P9000 XP Agent.

sst.conf file

The file /usr/kernel/drv/sst.conf is required on each Data Protector 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.

Storage Group (Microsoft Exchange Server specific term)

A collection of mailbox stores and public folder stores that share a set of transaction log files. Exchange Server manages each storage group with a separate server process.

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.

Typically, these can be created or exist within a storage system such as a disk array.

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.

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. The 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.

SYMA

See EMC Symmetrix Agent.

synthetic backup

A 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.

synthetic full backup

The result of an object consolidation operation, where a restore chain of a backup objects is merged into a new, synthetic full version of this object. A synthetic full backup is equivalent to a conventional full backup in terms of restore speed.

System Backup to Tape (SBT) (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 Recovery Data file

See SRD file.

System State (Windows specific term)

The System State data comprises the Registry, COM+ Class Registration 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 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.

T**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 (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 database (Oracle specific term)

In RMAN, the target database is the database that you are backing up or restoring.

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 faulty 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.

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 (Microsoft 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 EMC 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

See 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.

transaction log table (Sybase specific term)

A system table in which all changes to the database are automatically recorded.

transportable snapshot (Microsoft 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).

U**unattended operation**

See lights-out operation.

user account (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 membership. This is checked whenever the user starts the Data Protector user interface or performs specific tasks.

User Account Control (UAC)

A security component in Windows Vista and Windows Server 2008 that limits application software to standard user privileges until an administrator authorizes an increase in privilege level.

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.

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.

user_restrictions file

A file that restricts specific user actions, which are available to Data Protector user groups according to the user rights assigned to them, to be performed only on specific systems of the Data Protector cell. Such restrictions apply only to Data Protector user groups other than admin and operator.

V**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.

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 P6000 EVA Disk Array Family specific term)

The firmware that manages all aspects of storage system operation, including communication with HP Command View EVA through the HSV controllers. See also HP Command View (CV) EVA.

Virtual Device Interface (Microsoft SQL Server specific term)

This is a Microsoft SQL Server programming interface that allows fast backup and restore of large databases.

virtual disk (HP P6000 EVA Disk Array Family specific term)

A unit of storage allocated from a storage pool of a disk array of the HP P6000 EVA Disk Array Family. A virtual disk is the entity that can be replicated using the snapshot functionality of such disk array. See also source volume and target volume.

virtual full backup

An efficient type of synthetic backup where data is consolidated using pointers instead of being copied. It is performed if all the backups (the full backup, incremental backups, and the resulting virtual full backup) are written to a single file library that uses distributed file medium format.

Virtual Library System (VLS)

A disk-based data storage device hosting one or more virtual tape libraries (VTLs).

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 resources. This way all requests for a particular virtual server are cached by a specific cluster node.

virtual tape

An archival storage technology that backs up data to disk drives in the same way as if it were being stored on tape. Benefits of virtual tape systems include improved backup and recovery speed and lower operating costs. See also Virtual Library System (VLS) and virtual tape library (VTL).

virtual tape library (VTL)

Data storage virtualization software that is able to emulate tape devices and libraries thus providing the functionality of traditional tape-based storage. See also Virtual Library System (VLS).

volser

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 (VSS).

VSS

See Microsoft Volume Shadow Copy Service (VSS).

VSS compliant mode (HP P9000 XP Disk Array Family VSS provider specific term)

One of two P9000 XP Array VSS hardware provider operation modes. When the P9000 XP Array provider is in the VSS compliant mode, the source volume (P-VOL) and its replica (S-VOL) are in simplex, unpaired state after a backup. Therefore the number of replicas (S-VOLs per a P-VOL) rotated is not limited. Restore from a backup in such a configuration is possible only by switching of the disks. See also resync mode, source volume, primary volume (P-VOL), replica, secondary volume (S-VOL), and replica set rotation.

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.

W

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, object copy, and object consolidation 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.

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 (Microsoft 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.

X**XBSA Interface (Informix Server specific term)**

ON-Bar and Data Protector communicate with each other through the X/Open Backup Services Application Programmer's Interface (XBSA).

Z**ZDB**

See zero downtime backup.

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 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 with specific disk array families, 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. With specific disk array families, split mirror restore can also be used. See also zero downtime backup (ZDB), ZDB to disk, ZDB to disk+tape, instant recovery, 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.

Index

A

- adding data to media during backups 137
- ADIC (EMASS/GRAU) AML 107
- alternative disaster recovery methods 98
 - operating system vendors 98
 - third-party tools 98
- ANSI X3.27 labels 135
- any-to-any connectivity 117
- application agents 19
- architecture
 - backup devices 18
 - Cell Managers 18
 - cells 18
- audience 13
- auditing 157
- autoloaders 107
- automated media copying 90
- automated object consolidation sessions 177
- automated object copy sessions 172, 175
- automated object verification sessions 179
- automated operation 16, 78

B

- backed up data
 - hiding from other users 44
 - visibility 44
- backing up data 70
 - procedure 70
- backup
 - IDB operation 149

- to disk 114
- backup agents 19
- backup concurrency 102-103
- backup configuration 74
- backup devices 25, 38
 - overview 100
- backup environment growth
 - database growth and performance key factors 152
- backup generations 61, 135
- backup interfaces 184
- backup media
 - verifying 91
- backup object 72
- backup objects
 - verifying 91
- backup overview 16
- backup ownership 45
- backup performance 103
- backup policies 21, 141
 - enterprise environment 21
- backup process
 - destination 16
 - source 16
- backup session
 - definition 73, 163
 - ownership 45
- Backup Session Manager 163
- backup sessions 20, 71, 74, 163
 - backup configuration 74
 - interactive 163
 - mount requests 167

- scheduled 163
 - timeout 167
- backup specifications 25, 71
- backup strategy 30
- backup strategy factors 32
- backup strategy planning 30
 - backup policies 33
 - catalog protection 33
 - data encryption 45
 - data protection 33
 - data types 32
 - defining requirements 30
 - definition 30
 - device configuration 33
 - media management 33
 - scheduling backups 33
 - system availability 32
- backup types 75, 193-194
 - full 40, 57-58
 - incremental 40, 57-58
 - planning performance 40
 - ZDB to disk 193
 - ZDB to disk+tape 194
 - ZDB to tape 193
- backup with disk discovery 168
- backups
 - adding data to media 137
 - automated 78
 - backup objects 72
 - backup specifications 71
 - configuring 40
 - devices 100
 - disk discovery vs. standard backup 168
 - disk image 40
 - filesystem 40
 - lights-out 78
 - local 38
 - network 38
 - scheduled 74
 - scheduling policies 74
 - sessions 74
 - staggering 75
 - standard backup vs. disk discovery 168
 - unattended 78
- barcode support 108
- barcodes 108
- benefits
 - disk backup 115
 - synthetic backup 62
 - Volume Shadow Copy Service 189
- benefits of online integrations 185
- block size
 - backup devices 104
 - default 104
 - devices 104
 - performance 104
- broadcasts 157
- browsing files 69
- BSM 163
- C**
- cache memory 41, 183

Catalog Database 147

- do not log any details 69
- filenames and file attributes size and growth 148
- log all detailed information 69
- log directory names only 69
- log level of information 72
- records 147
- size and growth for CDB Records other than filenames 148

Catalog Database growth factors

- catalog protection 69
- level of details 69

catalog protection 68

- as an IDB key tunable parameter 154
- backup generations 61
- browsing files 69
- expired 154
- IDB size and growth 145
- impact on backup performance 154
- restoring data when catalog protection expires 154

catalog protection as an IDB key tunable parameter 154

CDB records

- Catalog Database 147

CDB. 147

Cell Managers 35

- high availability 52
- optimizing the load 167

Cell Request Server 162

cells

- backup operation 19
- Cell Managers 19
- logical view 18
- mixed environment 37
- multiple 22, 34
- physical view 18
- planning 34
- planning security 43
- remote 37
- restore operation 19
- single-point management 23
- splitting 22
- UNIX environment 36
- Windows domains 36
- Windows environment 36
- Windows workgroups 37
- centralized licensing 24
- Centralized Media Management Database 24, 146
- checkpoints 184
- cleaning tape detection 108
- cleaning tape support 109
 - magazine devices 107
 - magazines 107
- client systems 19
- clients 19
 - installing 35
 - maintaining 35
- cluster (definition) 50
- cluster heartbeat 51

- cluster integrations
 - overview 52
- cluster node 51
- clustering 50
 - automatic restart 52
 - availability 52
 - device sharing 124
 - failover 51
 - floating drives 124
 - group 51
 - heartbeat 51
 - load balancing 52
 - MC/Service Guard 50
 - Microsoft Cluster Server 50
 - nodes 51
 - package 51
 - primary node 51
 - secondary node 51
 - shared disks 51
 - virtual cluster node backup 54-56
 - virtual server 51
- CMMDB 24
- CMMDB. 146
- collision 105
- commands
 - omniclus command 57
 - post-exec 165, 184
 - pre-exec 165, 184
- company backup policies 141
- comparison
 - disk-based devices 115
- compression
 - hardware 38, 40
 - software 40
- concurrency 102-103
- concurrent sessions
 - backup 164
 - media management 181
 - object consolidation 178
 - object copy 173, 176
 - restore 169
- configuring backup specifications 71
- configuring devices 100
 - large libraries 107
 - magazines 106
 - standalone devices 106
- consolidating a restore chain 85
- control files 183
- conventional incremental backup 58
- copying backed up data 80
- copying media 90
 - automated 90
- copying objects 81
 - for vaulting purposes 84
 - to consolidate a restore chain 85
 - to demultiplex a medium 84
 - to free a medium 84
 - to implement disk staging 85
 - to migrate to another media type 85
- creating backup specifications 71
- creating cells
 - mixed environment 37

- UNIX environment 36
- Windows domains 36
- Windows environment 36
- Windows workgroups 37
- creating replicas 193
- CRS 162
- D**
- data
 - hiding from other users 44
 - visibility 44
- data encoding 45
- data encryption 45
- data files 183
- data protection 68
- Data Protector architecture
 - cell 18
 - Cell Managers 18
 - client systems 18
 - devices 18
 - logical view 18
 - physical view 18
- Data Protector concepts
 - Cell Managers 18
 - cells 18
 - clients 18
 - devices 18
- Data Protector features 14
- Data Protector functionality 14
- Data Protector GUI 26
- Data Protector Inet 162
- Data Protector operation 162
- Data Protector processes
 - Cell Request Server 162
 - Inet 162
 - Key Management Server 162
 - Media Management Daemon 162
- Data Protector services
 - Cell Request Server 162
 - hdpd-idb 162
 - hdpd-idb-as 162
 - hdpd-idb-cp 162
 - Inet 162
 - Key Management Server 162
 - Media Management Daemon 162
- Data Protector setup 27
- Data Protector user interfaces 19, 26
- data security 73
- database
 - advantages 145
 - architecture 146
 - Catalog Database 147
 - catalog protection 145
 - Detail Catalog Binary Files 148
 - growth and performance 152
 - IDB management 151
 - in the Manager-of-Managers environment 146
 - Media Management Database 147
 - operation 149
 - Session Messages Binary Files 149
 - size and growth 145
- database architecture 146

- database growth and performance key factors 152
 - backup environment growth 152
 - filesystem dynamics 152
- database growth and performance key tunable parameters 153
 - catalog protection 154
 - logging level 153
 - usage of logging level and catalog protection 154
- database in the MoM environment 146
 - Centralized Media Management Database 146
- Database Library 185
- database operation 182
- databases 182
 - backup interfaces 184
 - cache memory 183
 - Centralized Media Management Database 24
 - checkpoints 184
 - control files 183
 - data files 183
 - dbspaces 182
 - files 182
 - online backups 184
 - segments 182
 - tables 182
 - tablespaces 182
 - transaction logs 183
- dbspaces 182
- DC binary file
 - Detail Catalog Binary Files 148
- IDB operation 150
- DC directory
 - Detail Catalog Binary Files 148
- DCBF information
 - Detail Catalog Binary Files 148
- DCBF size and growth
 - Detail Catalog Binary Files 148
- DCBF. 148
- default block size 104
- default media pools 127
- demultiplexing media 84
- Detail Catalog Binary Files 148
 - DC binary file 148
 - DC directory 148
 - DCBF size and growth 148
 - information 148
- device chaining 102
- device chains 106
- device collision 105
- device configuration 100
- device lists 101
- device locking 105
- device sharing in clusters 124
- device sharing in SAN 120
 - drives 122
 - robotics 122
- device streaming (definition) 102
- devices 25, 38, 100
 - ADIC (EMASS/GRAU) AML 107
 - autoloaders 107
 - cleaning tape support 109

- concurrency 102
- configuring 100
- device chaining 102
- device filtering 103
- device lists 101
- device locking 105
- device streaming 102
- disk-based 115
- exchangers 107
- filtering 103
- GRAU/EMASS 107
- jukeboxes 107
- library management console,
support 100
- load balancing 101
- lock names 105
- multiple devices 101
- number of buffers 105
- overview 100
- physical device collision 105
- planning performance 38
- SCSI libraries 107
- segment size 103
- selecting for restore 94
- standalone 106
- StorageTek/ACSL 107
- TapeAlert support 101
- direct library access 123
- dirty drive detection 109
- disaster 96
 - disaster recovery 97
 - alternative 98
 - alternative methods 98
 - Disaster Recovery
 - concepts 96
 - overview 96
 - Phase 0 97
 - Phase 1 97
 - Phase 2 97
 - Phase 3 97
 - disk-based devices
 - comparison 115
 - Disk Agent concurrency 102-103
 - Disk Agents 19
 - disk backup 114
 - benefits 115
 - disk discovery (definition) 168
 - disk discovery vs. standard backup 168
 - disk fragmentation 41
 - disk image backups 40, 42
 - disk image vs. filesystem backups 40
 - disk performance 41
 - cache memory 41
 - compression 42
 - disk image backups 42
 - disk staging 85
 - disk virtualization 192
 - do not log any details
 - Catalog Database 69
 - drive servers 19

- drives 122
 - connecting to multiple systems 110
 - floating 124
 - static 124

- duplicating backed up data 80

E

- e-mail 157

- encoding 45

- encryption 45
 - drive-based 45-46
 - encrypted control communication 47, 49
 - encryption key 45
 - Key Management Server 45
 - software-based 45

- encryption key
 - Key Management Server 45

- enhanced incremental backup 58

- enterprise environment 21

- enterprise reporting 24

- environment
 - enterprise 21
 - Manager-of-Managers 21
 - mixed 37
 - network 17
 - UNIX 36
 - Windows 36

- examples
 - media pool usage 131
 - reporting and notification 159
 - scheduling policies 76
 - using data provided by 160

- vaulting usage 141

- examples of media usage policies 138

- exchangers 107

- expired catalog protection 154

- exporting media 70

- IDB operation 151

- removed objects 151

F

- factors affecting restore duration 93

- factors influencing backup strategies 32

- failover 51-52

- FC-AL 119

- features of Data Protector 14

- fibre channel

- planning performance 42

- Fibre Channel (definition) 118

- Fibre Channel Arbitrated Loop 119

- Fibre Channel topologies 119

- loop topology 119

- point-to-point 119

- switched topology 120

- file jukebox device 116

- file library device 116

- filenames and file attributes size and growth
 - Catalog Database 148

- filesystem backup 40

- Volume Shadow Copy Service 189-190

- filesystem dynamics

- database growth and performance key factors 152

- filesystem vs. disk image backups 40

- filtering 103
- floating drives 124
- formatting media 126
- fragmentation 41
- freeing media 84
- full and incremental backups 57
- full backups 40
 - staggering 75
- functionality of Data Protector 14

G

- General Media Agent 110
- geographically remote cells 37
- GRAU/EMASS 107
- group 51

H

- hardware compression 38, 40
- heartbeat 51
- high availability 15, 52
- hot-backup mode 192
- HTML 157

I

- IDB 145
 - advantages 145
 - architecture 146
 - Catalog Database 147
 - Detail Catalog Binary Files 148
 - in the Manager-of-Managers environment 146
 - internal encoding used 145
 - location 145
 - management 151

- Media Management Database 147
- operation 149
- Session Messages Binary Files 149
- size and growth 145
- IDB advantages 145
- IDB architecture 146
 - Catalog Database 147
 - Detail Catalog Binary Files 148
 - IDB parts 146
 - IDB parts scheme 147
 - Media Management Database 147
 - Session Messages Binary Files 149
- IDB configuration
 - creating a backup specification for the IDB backup 151
 - IDB management 151
- IDB growth and performance 152
 - backups as key factors 152
 - key factors 152
 - key tunable parameters 153
- IDB in the MoM environment
 - Centralized Media Management Database 146
- IDB maintenance
 - IDB management 152
- IDB management
 - IDB configuration 151
 - IDB maintenance 152
 - IDB recovery 152
 - overview 151
 - setting up backup environment 151

- IDB operation 149
 - backup 149
 - DC binary file 150
 - exporting media 151
 - media position record 150
 - restore 150
 - session messages binary files 150
 - verification 150
- IDB parts
 - architecture 146
- IDB parts scheme
 - IDB architecture 147
- IDB recovery
 - IDB management 152
- IDB size and growth 145
 - catalog protection 145
 - logging level 145
- incremental backup types
 - conventional incremental backups 58
 - enhanced incremental backups 58
 - leveled incremental backups 59
- incremental backups 40
 - Change Log Provider 59
 - types 59
- indirect library access 123
 - library access 122
 - Storage Are Networks 122
- initializing media 126
 - media ID 135
- Installation Servers 19
- instant recovery
 - advantages 192
 - introduction 194
- integration with database applications 16, 182
- integrations
 - Volume Shadow Copy Service 190
- interactive backup sessions 163
- interactive object consolidation sessions 177
- interactive object copy sessions 172
- interactive object verification sessions 179
- interactive replication sessions 175
- Internal Database 145
- J**
 - java-based online reporting 160
 - java reporting 160
 - jukeboxes 107
- K**
 - Key Management Server 45, 162
 - KMS 45
- L**
 - labeling media 135
 - labels 135
 - LAN-free backups 120
 - large libraries 107
 - leveled incremental backups 59
 - libraries 24
 - barcode support 108
 - cleaning tape support 109
 - connecting to multiple systems 110
 - drives 110

- entering and ejecting mail slots 108
- management console, support 100
- media handling 107
- multiple slots 108
- sharing 108
- silo 107
- size 107
- slot range 107
- slots 107
- library access
 - direct 123
- library management console, support 100
- library sharing 109
- library size 107
- life cycle, media 126
- lights-out operation 16, 78
- LIP 119
- load balancing 39, 52, 72, 101
- load balancing (definition) 101
- location fields 136
- lock names 105, 122
- log all detailed information
 - Catalog Database 69
- log directory names only
 - Catalog Database 69
- log level of information 72
- logging level
 - enabling restore 153
 - IDB size and growth 145
 - impact on ability to browse for restore 153
 - impact on backup speed 153
 - impact on restore speed 154
 - Log All 153
 - Log Directories 153
 - Log File 153
 - No Log 153
- Loop Initialization Primitive (Protocol) 119
- loop topology 119
- M**
- magazine devices
 - cleaning 107
- management console 100
- Manager-of-Managers 23
 - enterprise reporting 24
 - remote cells 38
 - sharing libraries 24
- MC/Service Guard 50
- media
 - age 140
 - barcode support 108
 - barcodes 108
 - catalog segments 103
 - cleaning tape support 109
 - copying 90
 - copying, automated 90
 - data segments 103
 - device errors 140
 - ejecting mail slots 108
 - encrypting 46
 - entering mail slots 108
 - estimating quantity of needed media 134

- exporting 70
- file marks 103
- formatting 126
- header segments 103
- initializing 126, 135
- labeling 108, 135
- location fields 136
- mail slots 108
- number of overwrites 139
- object distribution 41
- preparing 126
- retiring 126
- selecting for backup 137
- selecting for restore 94
- vaulting 126, 140
- Media Agents 19
 - General Media Agent 110
 - NDMP Media Agent 110
- media allocation policies 127, 134, 137
 - loose 137
 - strict 137
- media condition 139
 - calculating 139
 - fair 137
 - good 137
 - poor 137
- media condition factors 139
- media copies 90
- media description 135
- media handling 107, 134
- media life cycle 126
- media location 135
- media location priority 94
- media management 24, 100
 - adding data to media 137
 - copies 90
 - copying media 90
 - labeling media 135
 - media allocation policies 137
 - media condition 137
 - media copies 90
 - media life cycle 126
 - media pools 25, 126
 - media rotation policies 133
 - pre-allocation policies 137
 - selecting media 137
 - vaulting 140
- media management after backing up 140
- media management before backing up 135
- media management concepts 24
- Media Management Daemon 162
- Media Management Database 147
 - records 147
 - size and growth 147
- media management during backing up 136
- media management functionality 24, 125
- media management session (definition) 180
- media pool properties
 - append incrementals only 127
 - appendable 127
 - media allocation policy 127

- media pool usage examples 131
 - large library configuration 132
 - multiple devices/multiple pools 133
 - multiple devices/single pool 132
 - one device/one pool 131
- media pools 25, 126
 - default 127
 - definition 126
 - properties 126
 - usage examples 127, 131
- media recognition 108
- media rotation policies 133
- media rotation policy (definition) 133
- Media Session Managers 180
- media set
 - definition 74
 - selection algorithm 94
- media usage 126
- media usage policies 137
 - appendable 137
 - appendable of incrementals only 137
 - examples 138
 - non-appendable 137
- media vaulting 126
- Microsoft Cluster Server 50
- migrating to another media type 85
- mirroring objects 88
- mixed environment 37
- MMD 162
- MMDB records
 - Media Management Database 147
- MMDB size and growth
 - Media Management Database 147
- MMDB. 147
- MoM 23
- monitoring 16, 158
- mount prompt handling 79
- mount requests 167, 174, 178
 - automating 167
 - notification 167
 - responding 167, 170
- mount requests (restore sessions) 170
- MSM 180
- multiple cells 22, 34
- multiple devices 101
- multiple slots 108
- N**
- NDMP Media Agent 110
- network environment 17
- node
 - cluster 51
 - primary 51
 - secondary 51
- notifications 16
- number of buffers 105
- number of cells 34
 - considerations 34
- number of concurrent sessions
 - backup 164
 - media management 181
 - object consolidation 178
 - object copy 173, 176

restore 169

O

object consolidation sessions 176

mount requests 178

queuing 178

object copy sessions 172

mount requests 174

queuing 174, 176

object copy tasks 83

object copying 81

object distribution to media 41

object mirroring 88

object verification

session flow 179

object verification sessions 178

omniclus command 57

online backup of databases 184

online integrations 185

online reporting 160

optimizing the load on Cell Managers 167

overview

backup 16

Disaster Recovery 96

IDB management 151

restore 17

synthetic backup 62

Volume Shadow Copy Service 187

ownership 45

backup sessions 45

restore sessions 45

P

package 51

parallel restore vs standard restore 170

parallel restores 170

parallelism 39

physical device collision 105

planning cells 34

Cell Managers 35

number of cells 34

planning performance 38

backup types 40

cache memory 41

compression 38, 42

devices 38

disk fragmentation 41

disk performance 41

fibre channel 42

hardware compression 40

infrastructure 38

load balancing 39

local backups 38

network backups 38

parallelism 39

software compression 40

planning security 42

cells 43

data encoding 45

encrypted control communication 47, 49

user accounts 43

user groups 44

visibility of backed up data 44

- point-to-point topology 119
- post-backup media copying 90
- post-backup object copying 81
- post-backup replication 87
- post-exec commands 165, 184
- post-exec scripts 73
- pre-exec and post-exec scripts 165
- pre-exec commands 165, 184
- pre-exec scripts 73
- predefined user groups 143
- preparing a backup strategy plan 32
- preparing media 126
- preventing collision 105
- primary node 51
- processes 162
 - backup 16
 - Backup Session Manager 163
 - restore 17
 - Restore Session Managers 168
- properties of media pools 126
- protection types
 - catalog 68
 - data 68
- Q**
- queuing
 - object consolidation sessions 178
 - object copy sessions 174, 176
 - restore sessions 169
- R**
- recovery 97
 - disaster recovery 97
- recycling media 126
- remote cells 37-38
- replicas
 - creating 193
 - introduction 193
- replication sessions 174
- reporting 16, 158
- reporting and notification
 - broadcasts 157
 - e-mail 157
 - examples 159
 - HTML 157
 - SNMP 157
- restore chain 66
- restore duration 93
 - factors affecting 93
 - parallel restore 93
- restore from ZDB
 - instant recovery 194
 - split mirror restore 194
 - standard Data Protector restore 194
- restore overview 17
- restore policies 93
 - end users 96
 - operators 95
- Restore Session Managers 168
- restore sessions 21, 45, 168
 - definition 168
 - mount requests 170
 - queuing 169
 - timeout 169

- restores 93, 168
 - configuring 40
 - duration 93
 - end users 96
 - IDB operation 150
 - media location priority 94
 - operators 95
 - optimizing 75
 - parallel 170
 - selecting devices 94
 - selecting media 94
 - vaulting 141
- restoring data 93
- restoring from media in a vault 141
- retiring media 126
- robotics 122
- RSM 168
- S**
- SAN 117
- scheduled backup sessions 163
- scheduled backups 74
- scheduled media copying 90
- scheduled object copying 82
- scheduled replication 87
- scheduling
 - backup configuration 74
- scheduling policies 74-75
- scheduling policy examples 76
- scheduling tips and tricks 75
- scripts
 - post-exec 73
 - pre-exec 73
 - pre-exec and post-exec 165
- secondary node 51
- security
 - data encoding 143
 - definition 42
 - unauthorized access of data 143
 - user-related 143
 - user groups 143
 - visibility of backed up data 143
- security features 43
- segment size 103
- segments 182
- selecting backup objects 72
- selecting media for backup 137
- service management 157
 - monitor 158
 - notification 158
 - reporting 158
- service management applications 157
- service management examples 160
- services 162
- Session Messages Binary Files 149
 - records 149
 - size and growth 149
- sessions
 - backup 20, 163
 - media management 180
 - object consolidation 176
 - object copy 172
 - object verification 178

- replication 174
- restore 21, 168
- setting catalog protection
 - usage of logging level and catalog protection 154
- setting up backup environment
 - IDB management 151
- setting up Data Protector (overview) 27
- shadow copy 187
- shadow copy provider 188
- shadow copy set 187
- shared disks 51
- sharing devices in SAN 120
 - drives 122
 - robotics 122
- sharing libraries 24, 107-109
- silo libraries 107
- single file restore 171
- size
 - libraries 107
- size and growth for CDB Records other than filenames
 - Catalog Database 148
- slot range 107
- slots 107
- SMBF records
 - Session Messages Binary Files 149
- SMBF size and growth
 - Session Messages Binary Files 149
- SMBF. 149
- SNMP 157
- software compression 40
- source volumes 193
- stacker devices 106
- staggering full backups 75
- standalone devices 106
- standalone file device 116
- standard backup vs disk discovery 168
- standard restore vs parallel restore 170
- static drives 124
- Storage Area Networks 117
 - any-to-any connectivity 117
 - concepts 117
 - device sharing 120
 - device sharing in clusters 124
 - direct library access 123
 - Fibre Channel 118
 - Fibre Channel topologies 119
 - indirect library access 122
 - LAN-free backups 120, 122
 - lock names 122
 - sharing devices 120
- storage duration of backed up data 68
- StorageTek/ACSL 107
- switched topology 120
- synthetic backup 62
 - benefits 62
 - media space consumption 64
 - operation 62
 - restore 64
- synthetic full backup 62
- systems to be backed up 19
- systems with backup devices 19

T

- tablespaces 182
- TapeAlert support 101
- Target System 96
- target volumes 193
- timeout 167
- timeout (restore sessions) 169
- transaction logs 183, 192, 194
- types of incremental backups 59
 - leveled incremental backups 59

U

- unattended operation 16, 78, 106
- usage of logging level and catalog protection 154
 - setting catalog protection 154
 - specifics for large cells 155
 - specifics for small cells 155
 - using different logging levels in the same cell 155
- usage of media pools 127
- user-related security 143
- user accounts 43
- user groups 44, 143
 - predefined 143
- user interfaces 19, 26
 - GUI 26
- user rights 143-144
- user rights (definition) 44
- users 143
- users and user groups 143

V

- vaulting 126, 140
 - definition 140
 - restoring 141
- vaulting usage example 141
- verification
 - IDB operation 150
- Verifying backup media and backup objects 91
- virtual cluster nodes 54-56
- virtual full backup 63
- virtual server 51
- virtualization 192
- visibility of backed up data 44, 143
- Volume Shadow Copy service (VSS)
 - shadow copy provider 188
- Volume Shadow Copy Service (VSS)
 - backup 190
 - backup model 188
 - benefits 189
 - filesystem backup 189
 - filesystem backup and restore 190
 - integration with Data Protector 189
 - overview 187
 - shadow copy 187
 - shadow copy set 187
 - writer 187
- VSS 187
- VSS backup 190
- VSS backup model 188

W

Windows domains 36

Windows workgroups 37

writer 187

Z

ZDB, backup types 193-194

- ZDB to disk 193

- ZDB to disk+tape 194

- ZDB to tape 193

ZDB, introduction 192

- advantages 192

- backup types 193

- concepts 192

- replicas 192

- replication 192

- snapshot backup 193

- source volumes 193

- split mirror backup 193

- target volumes 193

zero downtime backup

- ZDB 192

We appreciate your feedback!

If you have comments about this document, you can [contact the documentation team](#) by email. If an email client is configured on this system, click the link above and an email window opens with the following information in the subject line:

Feedback on Concepts Guide (Data Protector 8.10)

Just add your feedback to the email and click send.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to AutonomyTPFeedback@hpe.com.

