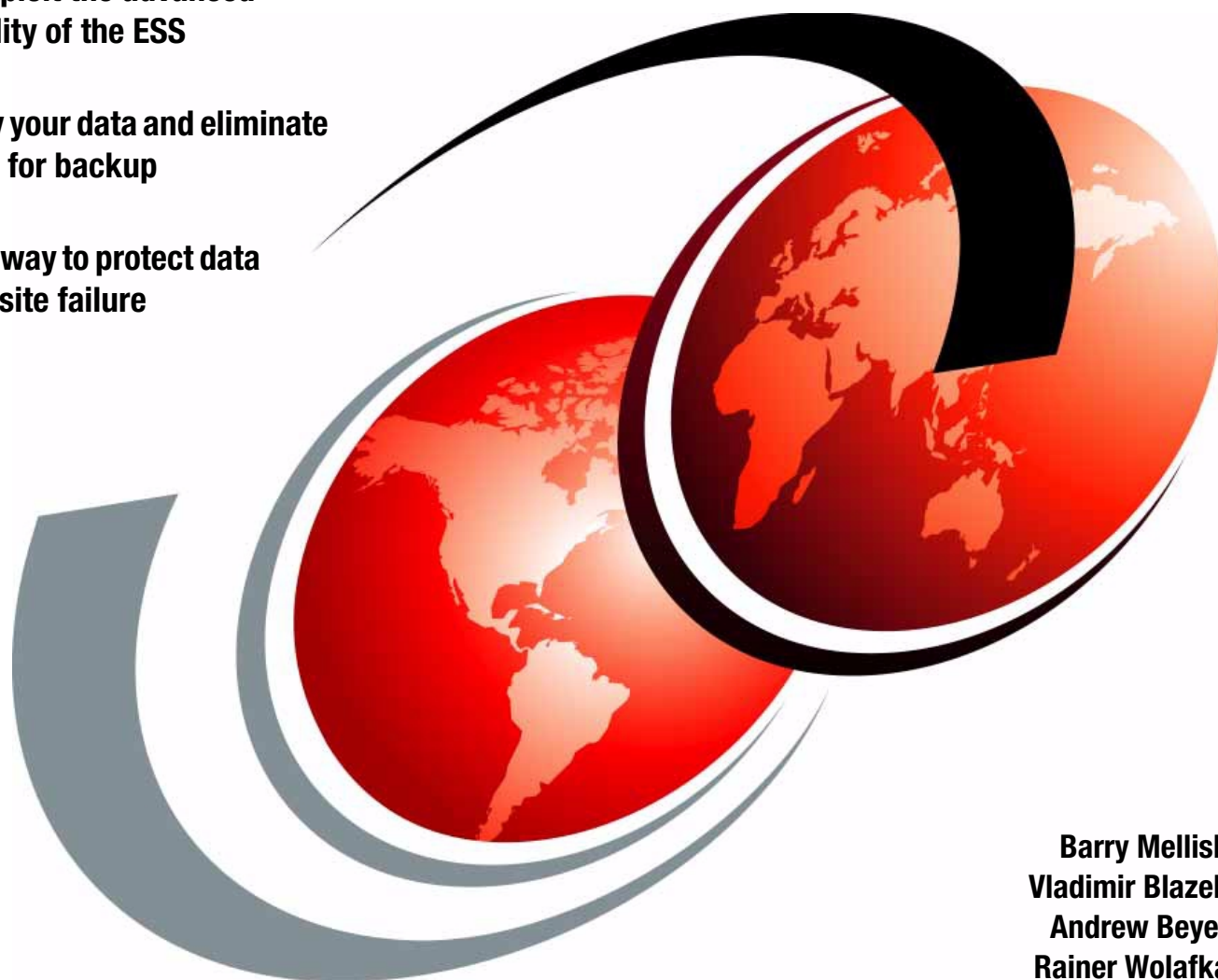


Implementing ESS Copy Services on UNIX and Windows NT/2000

How to exploit the advanced
functionality of the ESS

FlashCopy your data and eliminate
downtime for backup

PPRC, the way to protect data
against a site failure



Barry Mellish
Vladimir Blazek
Andrew Beyer
Rainer Wolafka

Redbooks



International Technical Support Organization

SG24-5757-00

**Implementing ESS Copy Services on UNIX
and Windows NT/2000**

February 2001

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix D, "Special notices" on page 193.

First Edition (February 2001)

This edition applies to the ESS, 2105 E and F models, and the Copy Services functions announced on March 29, 2000.

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

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

© Copyright International Business Machines Corporation 2001. All rights reserved.

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

Contents

Figures	vii
Tables	xi
Preface	xiii
The team that wrote this redbook	xiii
Comments welcome	xv
Chapter 1. Introduction	1
1.1 Seascape architecture	1
1.2 Enterprise Storage Server overview	1
1.3 ESS 1.1 announcement	4
1.4 Extensive capacity and scalability	4
1.5 ESS Advanced Copy Services	5
Chapter 2. Requirements and preliminary setup	7
2.1 Requirements for Copy Services	7
2.1.1 AIX	7
2.1.2 Windows NT	8
2.1.3 Sun	9
2.1.4 Hewlett Packard	10
2.1.5 NUMA-Q	10
2.2 Configuring primary and secondary Copy Services servers	10
2.2.1 Location of Copy Services server	13
2.2.2 How to switch to the backup server	14
2.3 Host authorization	17
Chapter 3. FlashCopy	19
3.1 Overview	19
3.2 Planning for FlashCopy on ESS	21
3.2.1 Hardware and software requirements	21
3.2.2 Configuration planning	21
3.2.3 Resource planning	22
3.2.4 Data consistency considerations	22
3.2.5 Test plan and disaster recovery plan	22
3.3 Operational considerations	22
3.3.1 Monitoring and managing FlashCopy pairs and volumes	22
3.3.2 Using a FlashCopy target volume	23
3.3.3 Automation	23
3.3.4 Microsoft Windows NT	24
3.3.5 Windows 2000	26
3.3.6 IBM AIX	27
3.3.7 Sun Solaris	34
3.3.8 HP-UX	36
3.4 Practical examples using FlashCopy	37
3.4.1 Moving and migrating data	37
3.4.2 Moving workload	37
3.4.3 Backup	38
3.4.4 Application testing	38
3.4.5 Other examples	38
3.5 Performance considerations	41

Chapter 4. Peer-to-Peer Remote Copy (PPRC)	43
4.1 Terminology	43
4.2 Overview	43
4.2.1 PPRC volume states	44
4.3 Planning for PPRC on an ESS	46
4.3.1 Hardware and software requirements	46
4.3.2 Configuration planning	47
4.3.3 Resource planning	48
4.3.4 Data consistency considerations	48
4.3.5 Test plan and disaster recovery plan	48
4.4 How to configure the ESS for PPRC	49
4.4.1 ESCON	49
4.4.2 ESS	49
4.5 How to set up PPRC	50
4.5.1 Recommendations	50
4.6 Operational considerations	50
4.6.1 Monitoring and managing volumes and paths	50
4.6.2 Operating system specifics	51
4.7 Moving and migrating data with PPRC	53
4.8 Using PPRC with FlashCopy	54
4.9 Practical example of PPRC and FlashCopy combination	56
4.10 Performance considerations	59
4.10.1 Optimized PPRC communication	59
4.10.2 Number of ESCON paths between Enterprise Storage Servers	59
4.10.3 Placement of the ESCON adapters used for PPRC	59
4.10.4 Grouping of physical and logical paths	59
4.10.5 Setup of the secondary ESS	60
Chapter 5. ESS Copy Services Web Interface	61
5.1 Overview and requirements	61
5.1.1 Volumes panel of the ESS Copy Services Web Interface	63
5.1.2 Storage Servers panel of the ESS Copy Services Web Interface	66
5.1.3 Paths panel of the ESS Copy Services Web Interface	69
5.1.4 Tasks panel of the ESS Copy Services Web Interface	72
5.1.5 Configuration panel of the ESS Copy Services Web Interface	75
5.1.6 Exiting from the ESS Copy Services Web Interface	77
5.2 Implementing FlashCopy with ESS Copy Services Web Interface	77
5.2.1 Establishing a FlashCopy pair	78
5.2.2 Getting information about a FlashCopy pair	80
5.2.3 Withdrawing a FlashCopy pair	81
5.2.4 Selecting multiple volumes for a FlashCopy task	82
5.2.5 Configuration tips	83
5.3 Implementing PPRC with the ESS Copy Services Web Interface	84
5.3.1 Setting up paths for PPRC	85
5.3.2 Establishing PPRC pairs	90
5.3.3 Terminating PPRC pairs	95
5.3.4 Establishing and terminating multiple PPRC pairs at once	98
5.3.5 Setting up FlashCopy and PPRC combinations	99
5.3.6 Suspending and resuming PPRC pairs	103
5.3.7 Manually suspending PPRC pairs	103
5.3.8 Configuration tips	103
5.4 ESS Copy Services Web Interface: tips for problem solving	105

Chapter 6. ESS Copy Services Command Line Interface	109
6.1 Requirements for Copy Services command line invocation	109
6.2 Installing the Command Line Interface (CLI)	110
6.2.1 Authorization of the CLI	110
6.2.2 Installing the CLI on an RS/6000 system	110
6.2.3 Installing the CLI on a SUN Solaris system	112
6.2.4 Installing the CLI on an HP-UX system	113
6.2.5 Installing the CLI on a Windows NT system	113
6.3 UNIX and NT Copy Services commands	114
6.3.1 UNIX command description	114
6.3.2 Scripting the Command Line Interface	121
6.3.3 Windows NT command description	124
6.3.4 Command line return codes	128
Chapter 7. High availability and disaster recovery	131
7.1 Automating site failover for AIX	131
7.1.1 Hardware and software requirements	131
7.1.2 Test environment	133
7.1.3 Single site recovery (two-node cluster)	134
7.1.4 Dual site recovery (two-node cluster) using PPRC	136
7.1.5 Implementation specifics	148
7.1.6 Creating the Copy Services tasks	152
7.1.7 Maintenance procedures	159
7.2 Implementing Microsoft Cluster Server with Copy Services	162
Appendix A. HACMP start and stop scripts	165
A.1 hacmpPPRC.vars	165
A.2 startESSapps	166
A.3 stopESSapps	174
Appendix B. Understanding logical subsystems	177
Appendix C. Operating system considerations	183
C.1 Definition of a "typical" backup solution	183
C.2 Preparation after the first FlashCopy	183
C.2.1 Further FlashCopy invocations	185
C.2.2 Invoke FlashCopy	185
C.2.3 Backup scripts	186
Appendix D. Special notices	193
Appendix E. Related publications	195
E.1 IBM Redbooks	195
E.2 IBM Redbooks collections	195
E.3 Other resources	195
E.4 Referenced Web sites	196
How to get IBM Redbooks	197
IBM Redbooks fax order form	198
Glossary	199
Index	209
IBM Redbooks review	211

Figures

1. Main Service Menu	11
2. Configuration Options Menu	12
3. Copy Services Menu	12
4. Primary and backup Copy Services server	13
5. ESS Storwatch Specialist Web User Interface front screen.	14
6. Task Help screen	15
7. Resetting the ESS Web Copy Services warnings	16
8. Available actions	16
9. Implementation of FlashCopy	20
10. Moving to a cluster environment using FlashCopy	40
11. PPRC write cycle	44
12. PPRC volume states	45
13. Ethernet and ESCON connection between Copy Services participating ESS	47
14. ESS logical paths	50
15. PPRC with FlashCopy	54
16. Asynchronous PPRC with FlashCopy	57
17. FlashCopy target as a PPRC primary volume	58
18. Main menu of the ESS Specialist.	62
19. Copy Services start message	62
20. Main menu of the ESS Copy Services Web Interface	63
21. Source and target area of the Volumes menu	64
22. Volume Information window	64
23. Find volume window.	65
24. Filter volumes window	65
25. Volumes display for AIX after running the rsPrimeServer command	66
26. Storage Servers window	67
27. LSS properties window	67
28. Information window of a logical subsystem	68
29. Find Storage Server.	68
30. Filter logical subsystems volume	69
31. Entry screen of the Paths menu.	70
32. SAID numbers of the ESS ESCON adapters.	70
33. Example of the Paths panel	71
34. Path information panel	72
35. ESCON adapter without and with defined path	72
36. Tasks panel of the Copy Services Web Interface	73
37. Task Information panel.	74
38. Configuration panel of the ESS Copy Services Web Interface	75
39. Example — ESS configuration and status summary	76
40. Log file of the Copy Services	76
41. Adding a user for the Copy Services Command Line Interface	77
42. Task Wizard window	78
43. Select copy options window	78
44. Define task window	79
45. FlashCopy Volume display.	80
46. Information panel of a FlashCopy source	80
47. Withdraw a FlashCopy pair	81
48. Define task window	81
49. Multiple volume selection and FlashCopy pairs.	82
50. A quick way to withdraw FlashCopy pairs	83

51. Withdraw FlashCopy Pair	83
52. Naming the task to withdraw	84
53. Paths window of the Copy Services Web Interface	85
54. Setup PPRC: Source selection	86
55. Set up PPRC: ESCON adapter selection	86
56. Set up PPRC: Target ESS selection	87
57. Set up PPRC: Select target LSS	87
58. Establish PPRC path	88
59. Establish PPRC path options	88
60. Define task window	89
61. Path successfully established (SAID0005)	89
62. Path information window	90
63. Select PPRC source and target LSS	91
64. Selecting PPRC source and target volume	91
65. Establish PPRC copy pair	92
66. PPRC copy options	92
67. Establish PPRC task window	93
68. PPRC relationship in progress	94
69. PPRC pair in full copy mode	94
70. Information window of a PPRC source volume	95
71. Terminate PPRC pair	96
72. Terminate PPRC pair task wizard	96
73. Terminate PPRC pair schedule task window	96
74. Save task window	97
75. Terminating PPRC pair	97
76. Establish multiple PPRC pairs	98
77. Multiple PPRC pairs in full copy mode	99
78. FlashCopy target used as PPRC source	100
79. FlashCopy / PPRC combination	100
80. Information window of a combined volume	101
81. PPRC target used as FlashCopy source	102
82. PPRC / FlashCopy combination	102
83. Quick way to suspend PPRC pairs	103
84. Suspend PPRC Copy Pair	104
85. Task to suspend PPRC Copy Pair	104
86. Tasks menu	105
87. Available actions	106
88. Example of rsList2105s.sh on AIX	116
89. Example of rsList2105s.sh on SUN Solaris	116
90. Example of rsList2105.sh on HP-UX	116
91. Example of rsList2105.sh showing vpaths in AIX	116
92. Example rsPrimeServer.sh: Source area: SUN Solaris, target area: AIX	118
93. Example of rsQuery.sh usage	119
94. Example of rsTestConnection.sh usage	121
95. csQueryVols.sh - an example script showing the use of the CLI	123
96. Example of output from csQueryVols.sh	124
97. Example of rsList2105s on NT system	125
98. Our test environment	134
99. Typical two-node cluster sharing a single ESS at one site	135
100. Two-node cluster spanning two sites with common network domains	138
101. Sample resource group for cluster spanning two sites	139
102. Two-node cluster spanning two sites with differing network domains	140
103. Resource group, cluster spanning two sites with differing network domains	141

104.Copy Services Specialist, Voulmes screen	153
105.Establish task, select task screen	153
106.Select copy options screen	154
107.Define task name.	154
108.List of defined tasks.	155
109.Selecting task to run	156
110.Select task type	156
111.Terminate PPRC	157
112.Execute task screen	157
113.Define task screen.	157
114.Task list screen	158
115.Selected task highlighted.	159
116.Microsoft Cluster Server	162
117.Challenge /defense protocol of MSCS.	163
118.Restrictions regarding source and target volumes	177
119.DA to LSS mapping.	178
120.Fully configured ESS A box showing 8 LSS mappings.	179
121.Fully configured ESS A box showing 16 LSS mappings.	180
122.Partially configured ESS	181
123.Output of lspv hdisk26.	184

Tables

1. Critical Heavy mode of PPRC pairs	93
2. Support for ESS on RS/6000 and RS/6000 SP	132
3. Output lsvg command issued against samplevg	184

Preface

This IBM Redbook will help you to install, tailor, and configure the new Copy Services functions of the IBM Enterprise Storage Server (ESS) on the UNIX, Windows NT, and Windows 2000 platforms.

The Copy Services functions include Peer-to-Peer Remote Copy (PPRC), FlashCopy, Extended Remote Copy (XRC), and Concurrent Copy (CC). It should be noted that the latter two Copy Services functions, XRC and CC, are not available on UNIX and NT platforms. They are only available on System/390.

This redbook provides a broad understanding of these functions, describes the prerequisites and corequisites, and then shows you how to implement each of the functions into your environment to ensure efficient usage and to maximize the benefits that these functions provide. This redbook also shows how to automate site failover using HACMP.

The team that wrote this redbook

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

Barry Mellish is a Project Leader at the International Technical Support Organization, San Jose Center. He has coauthored two previous redbooks and has taught many classes on storage subsystems. He joined IBM UK 16 years ago, and before joining the ITSO, he worked as a Senior Storage Specialist on the Disk Expert team in EMEA.

Vladimir Blazek is a Senior Engineer with Servodata, an IBM Business Partner in Prague, Czech Republic. He has 7 years of experience in supporting and servicing disk and tape subsystems. Currently he is designing and implementing data storage solutions for Open Systems. His areas of expertise include Storage Area Networks and data storage solutions for UNIX and Windows NT.

Andrew Beyer is an Advisory IT Specialist with the Enterprise Systems Division in Australia. Andrew has 19 years of experience in the IT industry, with 10 years working as an RS/6000 and AIX specialist. Andrew joined IBM 11 years ago and specializes in designing and implementing High Availability and DR solutions. He has co-authored a redbook on High Availability for the RS/6000 family.

Rainer Wolafka is working in the IBM EMEA Storage Competence Center in Mainz, Germany. He has 3 years of experience in supporting and implementing Open Systems storage solutions. His areas of expertise include various UNIX platforms, Windows NT, and the design and support of Storage Area Networks.

The team would also like to extend their thanks to the following people for their invaluable contributions to this project:

- Mike Downie, Senior Education Specialist, Learning Services
IBM Boulder
- Jack Flynn, Device Technical Office, VM/VSE/Open Systems
IBM San Jose

- Paula Collins Di Benedetto, Shark SLS Team Leader
IBM San Jose
- Paul Hurley, DST test
IBM San Jose
- Vladimir Atanaskovik, EMEA Storage Competency Center Mainz
IBM Germany
- Brian Smith, IO Subsystem Performance Evaluation
IBM San Jose
- Neena Cherian, IO Subsystem Performance Evaluation
IBM San Jose
- David Short, IO Subsystem Performance Evaluation
IBM San Jose
- John Aschoff, IO Subsystem Performance Evaluation
IBM San Jose
- Bill Micka, Subsystem Architecture
IBM Tucson
- Richard Kirchoffer, RAS Programmer
IBM San Jose
- Ann Rudy, Shark Program Management
IBM San Jose
- Charlie Burger, Technical Support for SSD Software Products
IBM San Jose
- Alison Pate, Technical Support for SSD Software Products
IBM San Jose
- Pat Blaney, Technical Support for SSD Software Products
IBM San Jose
- Ling Pong, Technical Support for SSD Software Products
IBM San Jose
- Edsel Carson, Information Development
IBM San Jose
- Henry Caudillo, Information Development
IBM San Jose
- Marilyn Pullen, Information Development
IBM San Jose
- EMEA Storage Competence Center Mainz
IBM Germany
- Mark Blunden, ITSO Open Storage Project Leader
IBM San Jose
- Sverre Torbjørn Bergum
IBM Norway
- Clemente Vaia
IBM Italy
- Jose Dovidauskas
IBM Brazil

Comments welcome

Your comments are important to us!

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

- Fax the evaluation form found in “IBM Redbooks review” on page 211 to the fax number shown on the form.
- Use the online evaluation form found at ibm.com/redbooks
- Send your comments in an Internet note to redbook@us.ibm.com

Chapter 1. Introduction

As businesses become more and more dependent on information technology to conduct their operations and stay competitive, the availability of their processing facilities becomes crucial. Today, most businesses require a high level of availability, which extends to continuous availability, 24 hours a day and seven days a week (24x7) operation. A lengthy outage could lead to significant financial losses, loss of credibility with customers, and maybe even a total failure of business. Therefore, the ability to provide continuous availability for the major applications is more often than not a necessity for business survival.

These important demands are fulfilled with the Enterprise Storage Server and its Advanced Copy Services. The ESS Copy Services provides replication of mission critical data; point-in-time with FlashCopy, dynamic synchronous mirroring to a remote with Peer-to-Peer Remote Copy (PPRC) and asynchronous copying to a remote site with Extended Remote Copying (XXRC). This last function is only available in a System/390 environment.

1.1 Seascope architecture

The introduction of the IBM Enterprise Storage Server (ESS) affirms IBM's leadership among disk storage providers. This advanced, SAN-ready disk storage system provides outstanding performance, scalability, and universal access across all major server platforms.

The ESS uses IBM's Seascope Architecture with advanced hardware and software technologies to deliver breakthrough levels of performance and maximize data sharing across the enterprise. The ESS provides customers the ultimate in scalability and flexibility across many platforms and configurations.

Seascope is a blueprint for comprehensive storage solutions optimized for a connected world. The Seascope architecture integrates leading technologies from IBM — including disk storage, tape storage, optical storage, powerful processors, and rich software function — to provide highly reliable, scalable, versatile, application-based storage solutions that span the range of servers from PCs to supercomputers.

At its heart, Seascope architecture uses an open, industry-standard storage server that can scale up exponentially in both power and performance. Since the storage servers can integrate snap-in building blocks and software upgrades, you can quickly deploy new or improved applications and rapidly accommodate new data and media types. In this way, Seascope storage servers become an essential element for storing, manipulating, and sharing data across the network.

1.2 Enterprise Storage Server overview

The IBM Enterprise Storage Server (ESS) is a member of the Seascope family of storage products. The ESS was announced in June 1999 and was generally available in September 1999. Since its announcement it has revolutionized the storage marketplace. It consists of a storage server and attached disk storage devices. The storage server provides integrated caching and RAID support for the attached disk devices. The disk devices are attached via a serial interface.

The ESS can be configured in a variety of ways to provide scalability in capacity and performance.

Redundancy within the ESS provides continuous availability. It is packaged in one or more enclosures, each with dual line cords and redundant power. The redundant power system allows the ESS to continue normal operation when one of the line cords is deactivated.

The ESS provides the image of a set of logical disk devices to the attached servers. The logical devices are configured to emulate disk device types that are compatible with the attached servers. The logical devices access a logical volume that is implemented using multiple disk drives.

The following host I/O interface attachments are supported:

- SCSI-3 Parallel Interface
- ESCON
- FC-AL

On SCSI-3 interfaces, the ESS emulates a variety of fixed-block devices with either 512 or 520 byte blocks. SCSI-3 is, in general, a superset of SCSI-2. A SCSI-3 disk device can be attached to a SCSI-2 initiator, provided that the cabling can be interfaced. Many SCSI-2 initiators attach directly to the cabling specified for the SCSI-3 parallel interface, but are referred to as SCSI-2 initiators because they limit their use of the command set to the SCSI-2 subset.

Host systems with SCSI-2 or SCSI-3 interfaces can attach to the ESS. The ESS provides multiple SCSI I/O interfaces (busses), each with multiple SCSI targets, and each with multiple disk logical units. The storage provided by the ESS for SCSI interfaces can be configured so that it is shared among multiple SCSI interfaces if desired.

On ESCON interfaces, the ESS emulates one or more IBM 3990 control units attaching variable size IBM 3390 devices in either 3390 or 3380 track format. The ESS provides multiple ESCON interfaces that provide a set of control unit images, each with multiple disk devices. The storage provided by the ESS for ESCON interfaces is configured so that it is accessible from any ESCON interface.

The ESS can be broken down into the following components:

- The storage server itself is composed of two clusters that provide the facilities with advanced functions to control and manage data transfer. Should one cluster fail, the remaining cluster can take over the functions of the failing cluster. A cluster is made up of the following subcomponents:
 - Host adapters: Each cluster has one or more host adapters (HAs). Each host adapter provides one or more host I/O interfaces. A host adapter can communicate with either cluster complex.
 - Device adapters: Each cluster has one or more device adapters (DAs). Each device adapter provides one or more storage device interfaces. Disk drives are attached to a pair of device adapters, one in each cluster, so that the drives are accessible from either cluster. At any given time, a disk drive is managed by only one device adapter

- Cluster complex: The cluster complex provides the management functions for the ESS. It consists of cluster processors, cluster memory, cache, nonvolatile storage (NVS) and related logic:
 - Cluster processors: The cluster complex contains four cluster processors (CP) configured as symmetrical multiprocessors (SMP). The cluster processors execute the licensed internal code that controls operation of the cluster.
 - Cluster memory / cache: These are used to store instructions and data for the cluster processors. The cache memory is used to store cached data from the disk drives. The cache memory is accessible by the local cluster complex, by device adapters in the local cluster, and by host adapters in either cluster.
 - Nonvolatile storage (NVS): This is used to store a nonvolatile copy of active written data. The NVS is accessible to either cluster-processor complex and to host adapters in either cluster. Data may also be transferred between the NVS and cache.
- The disk drives provide the primary nonvolatile storage medium for any host data stored within the ESS Storage devices. They are grouped into ranks and are managed by the clusters.

As a member of the IBM Seascope family, the ESS provides the outboard intelligence required by SAN solutions, off-loading key functions from host servers, which frees up valuable processing power for applications. As a comprehensive SAN-based storage solution, the ESS provides considerable management flexibility to meet the fast-paced requirements of the next century.

Among the many factors that make the IBM ESS an ideal SAN solution are:

- Support for all major server platforms, including S/390, AS/400, Windows NT, and many varieties of UNIX
- Fibre Channel attachment capability
- Extensive StorWatch management capabilities through a Web Interface
- Excellent scalability:
 - From 400 GB to over 11 TB
 - Simple selection from 16 standard configurations to meet capacity and performance needs
- Performance optimized to your heterogeneous environment needs:
 - High bandwidth and advanced transaction processing capabilities provide solutions for both online and batch applications
 - Innovations such as Parallel Access Volumes to reduce resource contention and dramatically improve performance
- Availability required to support e-business applications:
 - Non-disruptive access to data while making a copy using Concurrent Copy
 - Business continuity through remote copy services — PPRC and XRC
 - Rapid data duplication through FlashCopy, providing extensive capabilities to exploit, manage, and protect your information in a 24x7 environment

- Storage server availability through redundancy and nondisruptive service with design for no single point of failure or repair

1.3 ESS 1.1 announcement

The ESS already offers exceptional performance, extraordinary capacity scalability, heterogeneous server connectivity, and an extensive suite of advanced functions to support today's mission-critical, high-availability, multiplatform environments. It is now enhanced to support direct attachment to a SAN via native fiber connections and additional advanced storage management capabilities including FlashCopy for fast data duplication and Remote Copy Services for synchronous and asynchronous backup and disaster recovery. The Enterprise Storage Server is the most appropriate IBM disk solution for a majority of our customers' SAN disk system requirements, especially those that require:

- Multiple and/or heterogeneous server attachment
- High performance
- High availability
- Innovative function

1.4 Extensive capacity and scalability

Increased performance and scalability with two new models with 8 GB or 16 GB of cache

- Additional 16 standard capacity configurations
- Concurrent support for all your major server platforms, including S/390, OS/400, Windows NT, Windows 2000, NetWare, and many varieties of UNIX

Extensive storage management capabilities, including:

- FlashCopy for fast data duplication
- Advanced Remote Copy Services for your synchronous and asynchronous backup and disaster recovery needs
- Extensive StorWatch management capability through the Web
- Superior performance with options and innovations to meet your changing requirements
- High availability to support your e-business and other mission-critical applications

The ESS continues to deliver on its SAN strategy, as was previewed in the July 27, 1999, announcement of the ESS. The ESS now provides up to sixteen 100 MB/sec native Fibre Channel short-wave adapters. Each single port adapter supports Fibre Channel Protocol (FCP) in a direct point-to-point configuration, point-to-point to a switch (fabric) configuration, or Fibre Channel-Arbitrated Loop (FC-AL) in a private loop configuration. Fabric support includes the IBM SAN Fibre Channel switch1 (2109 Model S08 and S16), McDATA Enterprise Fibre Channel Director (ED1 5000), and IBM Fibre Channel Storage Hub (2103-H07).

1.5 ESS Advanced Copy Services

Copy Services is a separately sold feature of the Enterprise Storage Server. It brings powerful data copying and mirroring technologies to Open Systems environments previously available only for mainframe storage.

This publication deals with the two features of Copy Services for the Open Systems environment:

- Peer-to-Peer Remote Copy (PPRC)
- FlashCopy

PPRC is a synchronous protocol that allows real-time mirroring of data from one Logical Unit (LUN) to another LUN in another ESS. This secondary ESS can be located at another site some distance away; see Chapter 4, “Peer-to-Peer Remote Copy (PPRC)” on page 43.

PPRC is application independent. Because the copying function occurs at the disk subsystem level, the application has no knowledge of its existence.

The PPRC protocol guarantees that the secondary copy is up-to-date by ensuring that the primary copy will be written only if the primary receives acknowledgment that the secondary copy has been written.

FlashCopy makes a single point-in-time (T₀) copy of a LUN. The target copy is available once the FlashCopy command has been processed; see Chapter 3, “FlashCopy” on page 19. FlashCopy provides an instant or point-in-time copy of an ESS logical volume. Point-in-time copy functions give you an instantaneous copy, or “view”, of what the original data looked like at a specific point-in-time. This is known as the T₀ (time-zero) copy.

The point-in-time copy created by FlashCopy is typically used where you need a copy of production data to be produced with minimal application downtime. It can be used for online backup, testing of new applications, or for creating a database for data mining purposes. The copy looks exactly like the original source volume and is an instantly available, binary copy. See Figure 9 on page 20 for an illustration of FlashCopy concepts.

Copy Services provides a Command Line Interface (CLI) as well as a Web-based interface for setting up and managing its facilities; see Chapter 6, “ESS Copy Services Command Line Interface” on page 109. The CLI interface allows administrators to execute Java-based Copy Services commands from a command line. The Web-based interface, a part of the StorWatch ESS Specialist, allows storage administrators to manage Copy Services from a browser-equipped computer; see Chapter 5, “ESS Copy Services Web Interface” on page 61.

Although each has its specific features, PPRC and FlashCopy are typically used as data backup tools for creation of test data and for data migration. They can also be used in disaster recovery scenarios; see Chapter 7, “High availability and disaster recovery” on page 131.

Copy Services will be of great use to customers with large IT systems, big data volumes, and a requirement for round-the-clock IS availability.

Copy Services will provide the most benefit to the customer who:

- Needs to have disaster tolerant IT centers

- Is planning to migrate data between systems
- Is migrating workload often
- Has to backup large amounts of data
- Needs to reduce the time the server has to be taken off-line for backup
- Plans to test new applications
- Needs a copy of production data for data warehousing or data mining

Copy Services can be integrated with technologies such as Tivoli Storage Manager (formerly ADSM), Logical Volume Manager (LVM) mirroring, or SAN Data Gateway mirroring to solve a wide variety of business issues. Other companies will most likely be selling partial solutions as a means to solve these problems. IBM, however, with its broad portfolio of products in this industry, has many experts available to discuss the right solution for your business and to help you design and implement a solution that will give you the maximum business benefit.

Advanced solutions with ESS Copy Services have been endorsed by many Independent Software Vendors (ISV) worldwide.

Chapter 2. Requirements and preliminary setup

In this chapter, we describe the basic hardware and software requirements to set up Copy Services on an IBM Enterprise Storage Server. You must always check the supported servers website for the latest details as the list of supported servers and operating systems is constantly being expanded.

We then explain how to set up the primary and secondary Copy Services servers.

For the latest information regarding supported servers, adapters, and operating systems, please consult the following Web site for supported servers:

<http://www.storage.ibm.com/hardsoft/products/ess/supserver.htm>

2.1 Requirements for Copy Services

The following sections describe the various hardware and software needed by a host connected to an ESS to support Copy Services.

Complete and current information on all supported servers (including supported models), is available on the Internet at the following URL:

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

2.1.1 AIX

This section lists the requirements for AIX environments.

2.1.1.1 Hardware

ESS supports the following host server types:

- IBM RS/6000 types 7012, 7013, 7015, 1707, 7024, 7025, 7026
- IBM RS/6000 SP type 9076

With SCSI host adapter features:

- FC 2412
- FC 6207

With Fibre Channel features:

- Host adapters:
 - FC 6227
- Switch / Hub features:
 - IBM 2103 Fibre Channel Storage Hub Model H07

2.1.1.2 Software

The following operating system support is required for the ESS on RS/6000 and RS/6000 SP servers with SCSI attachment:

- AIX Version 4.2.1
- AIX Version 4.3.1 or higher

The following support is available for the ESS connected to RS/6000 and RS/6000 SP with Fibre Channel attach through the IBM SAN Data Gateway.

- AIX Version 4.3.3 or higher

For Copy Services Command Line support, the following products must be installed on the host server:

- Copy Services Command Line Interface for AIX
- Java Runtime for AIX at version 1.1.8 (Check the website for the latest information)

2.1.2 Windows NT

This section lists the requirements for Windows NT environments.

2.1.2.1 Hardware

ESS supports the following host server types:

Pentium Pro or later processors

- 200 MHz processor or faster
- 128 MB memory or greater

With SCSI host adapter features:

- Adaptec AHA-2944UW
- IBM P/N 59H3900 (for IBM Netfinity servers)
- QLogic QLA1041
- Symbios SYM8751D

With Fibre Channel features:

- Host adapters
 - P/N 01K7297 (for IBM Netfinity servers)
 - QLogic QLA2100F
- Switch / Hub features:
 - IBM 2109 SAN Fibre Channel Switch Models S08 and S16
 - IBM 2103 Fibre Channel Storage Hub Model H07

Note

In order to recognize a disk that has been added after boot, Windows NT has to be rebooted. Flushing the NT system disk cache to disk can be only achieved by unassigning the drive letter in NT Disk Administrator.

2.1.2.2 Software

The following operating system support is available:

- Microsoft Windows NT Server 4.0: requires Service Pack 4 or 5
- Microsoft Windows NT Server 4.0, Enterprise Edition: requires Service Pack 4 or 5

Windows NT supports a maximum of eight LUNs per SCSI target and a maximum of 255 LUNs per Fibre Channel port. The maximum number of logical drives is 22 per Windows NT system. You need to have free LUNs and logical drives available on the system that will use Copy Services target.

If you have installed IBM Subsystem Device Driver (which has superseded Data Path Optimizer) software, you have to reapply the Windows NT Service Pack you

are currently using on your system after the installation of the IBM Subsystem Device Driver.

2.1.3 Sun

This section lists the requirements for Sun environments.

2.1.3.1 Hardware

ESS supports the following host server types:

- Enterprise Server
- SPARCcenter
- SPARCserver
- Ultra Server

With SCSI host adapter features:

- X1062A
- X1065A
- X6541A

With Fibre Channel features:

- Host adapters
 - QLogic QLA2100F
 - JNI FC64-1063
 - JNI FCI-1063-N
 - Emulex LP8000
- Switch / Hub features:
 - IBM 2103 Fibre Channel Storage Hub Model H07

2.1.3.2 Software

The following operating system support is available:

- Solaris 2.6
- Solaris 7
- Solaris 8

Solaris 2.6 supports a maximum of eight LUNs per SCSI ID, however, if Solaris 2.6.1 or 7 is used and a PTF is applied, the number of LUNs per SCSI target increases to 32. Fibre Channel supports 255 LUNs. You need to have free LUNs available on the host adapter that will be used to connect the Copy Services volumes.

In Solaris, the ESS LUNs are represented as `/dev/dsk/cXtYdZsN` files. In order to recognize LUNs added after boot, you have to run the `devconfig` and `disks` commands. Sometimes the device files are not created correctly using this procedure and you may have to repeat it or reboot.

We recommend that you assign the Copy Services targets to their host ports in advance and create the device files by rebooting with the `boot -r OpenBoot` command. This ensures that the device files are created correctly, and that the host is prepared for routine Copy Services use.

You can then start using the Copy Services and mount the targets.

2.1.4 Hewlett Packard

This section lists the requirements for Hewlett Packard environments.

2.1.4.1 Hardware

ESS supports the following host server types:

- HP 9000 Enterprise Servers
- Enterprise Parallel Servers

With SCSI host adapter features:

- A2969A
- A4107A (HP-HSC)
- A4800A (PCI)
- 28696A (HP-PB)

HP 9000 hosts support up to eight LUNs per SCSI ID. You need to have free LUNs available on the host adapter that will be used to connect the Copy Services volume.

2.1.4.2 Software

The following operating system support is required:

- HP UX 10.20
- HP UX 11.00

HP-UX represents disks (and ESS LUNs) as `/dev/dsk/cXtYdZ` files. There is no need to reboot the host in order to see new LUNs.

You have to run `ioscan -f` in order to recognize disks and `insf -e` in order to create the special files for new LUNs. You can normally start using Copy Services and mount the targets then.

2.1.5 NUMA-Q

Shark attachment to NUMA-Q 1000 and 2000 servers is supported on DYNIX/ptx 4.4.7 and requires the use of the NUMA-Q Fibre Channel to SCSI Bridge. This support is available through NUMA-Q's Systems Integration Services (SIS) team. Native Fibre Channel support is now available using adapter IOC-0210-54. See the supported servers Web site:

<http://www.storage.ibm.com/hardsoft/products/ess/supserver.htm#5>

2.2 Configuring primary and secondary Copy Services servers

In order to use Copy Services you must configure one ESS to be the primary Copy Services Server. All information related to the Copy Services is stored in this ESS, such as volumes and their state, ESCON connectivity between ESS, and much more.

On each ESS that is configured to use Copy Services there is a client running. Whenever the Copy Services configurations changes this client notifies the Copy Services server of the changes.

Optionally, there could be one ESS defined as the backup Copy Services Server. In case the primary Copy Services Server is lost, the backup server could be used for controlling the ESS Copy Services. Once the primary ESS is up again, the backup server will notify all clients, and the clients will switch back all communication to the primary Copy Services Server.

We recommend that you have the primary and backup Copy Services Server on different sites if your ESS Storage Network spans multiple locations. In case the entire side with the primary Copy Services Server is going down for whatever reason, the backup Copy Services Server can keep Copy Services alive.

The information on the primary and backup Copy Services server has to be specified on each cluster of all the Enterprise Storage Servers that are going to be used for Copy Services. We recommend that you specify a backup for Copy Services whenever possible. This task is done by an IBM Customer Engineer (CE) via a service terminal connected to the ESS serial interface. From the Main Service Menu, select the **Configuration Options Menu** (Figure 1).

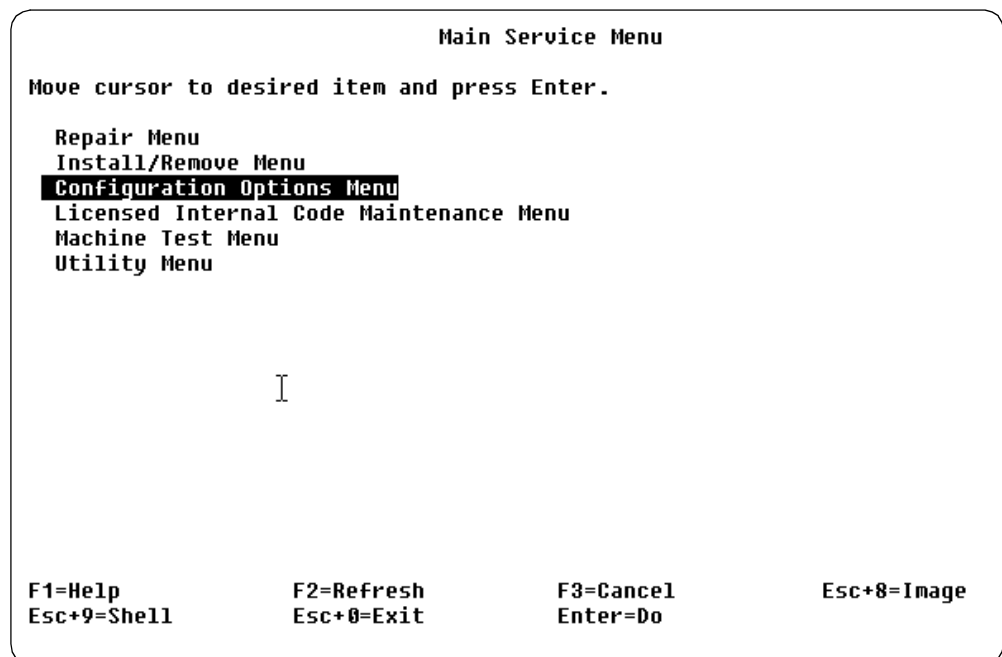


Figure 1. Main Service Menu

From the Configuration Options Menu, you select the **Copy Services Menu** (Figure 2).

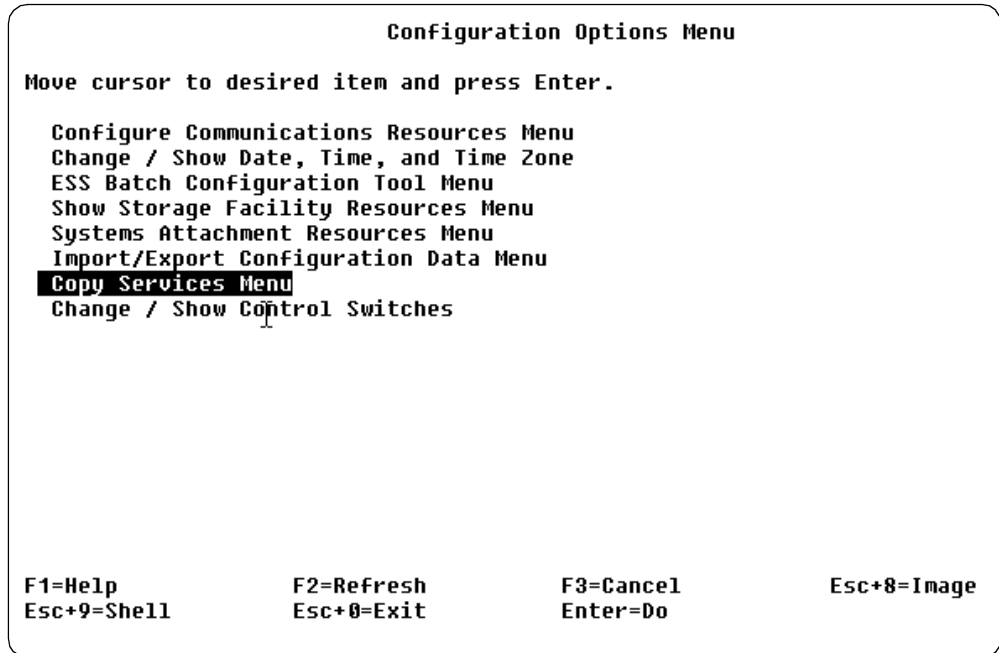


Figure 2. Configuration Options Menu

From the Copy Services Menu, you can change, show, and remove the primary (and optionally, the backup) Copy Services server (Figure 3).

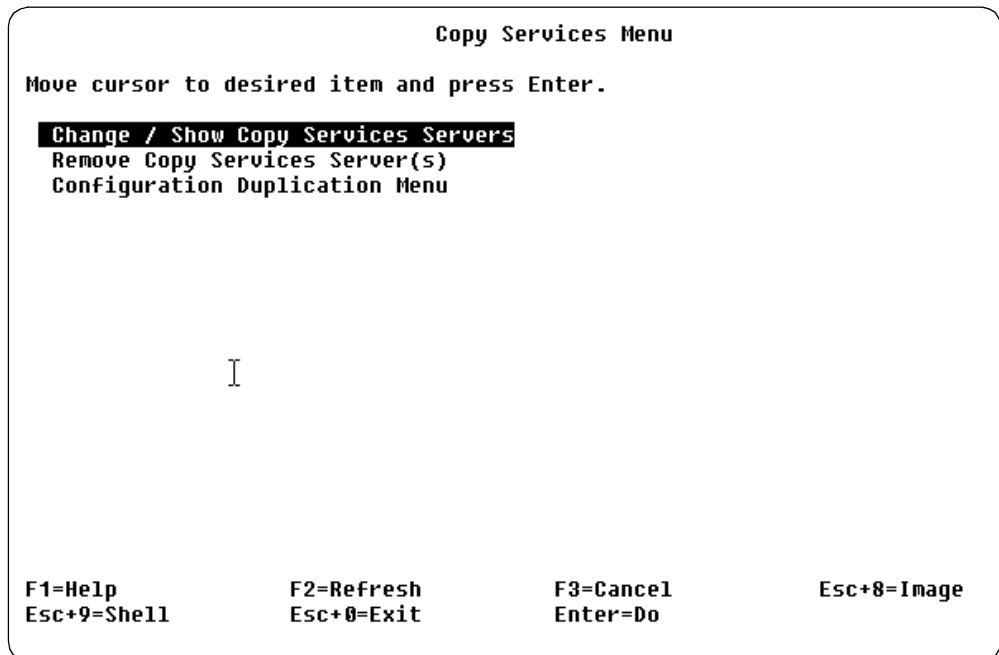


Figure 3. Copy Services Menu

The primary and backup Copy Services servers are specified either by the TCP/IP address or the hostname (Figure 4).

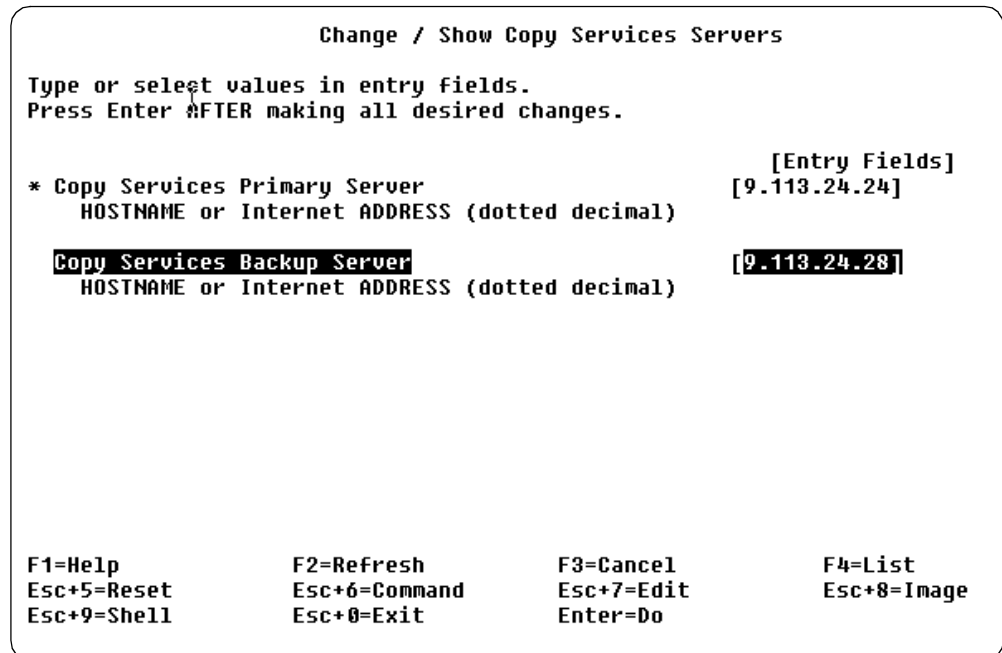


Figure 4. Primary and backup Copy Services server

It is possible to specify one cluster as the primary Copy Services server and the alternate cluster of the same ESS to be the backup Copy Services server. However, unless you only have a single ESS, this is not recommended, as the purpose of the backup Copy Services server is to be able to run Copy Services in the event of a total site failure.

2.2.1 Location of Copy Services server

For the following discussion we will assume that we have two ESSs: ESS1 and ESS2. These two ESSs are going to be configured as part of the same ESS WEB Copy Services group. This means that all four clusters will be using the same primary Web Copy Services primary server.

This would be the case if ESS1 and ESS2 are going to be used as PPRC peers. It is recommended to define a backup server. The switch to the backup server is a manual process and is described in 2.2.2, “How to switch to the backup server” on page 14. Planning for a disaster requires some additional steps for PPRC. First you need to decide where to place your primary and backup ESS Web Copy Services servers. There are two options:

1. The Primary Copy Services server is at your recovery site.

Choosing this configuration can gain you some efficiency in terms of recovering from a disaster. Because the primary server is at the recovery site, you do not have to perform manual recovery steps to switch to a different Web Copy Services server. If you have planned in advance and have created tasks for disaster recovery, then you can run those tasks and bring your production systems back up.

Note: In an Open Systems environment, this setup might have performance implications, because the communications with the Open Systems command line interface (CLI) and the ESS Web Copy Services server is done over TCP/IP links.

2. The Primary Copy Services server is at your production site.

This configuration might be the desirable setup in Open Systems environments. The reason for this is that the Open Systems CLI uses TCP/IP links to communicate with the host systems and the Web Copy Services Server, which may introduce network delays.

However, with this configuration, if your primary site fails, there can be delays in bringing up applications at the recovery site. Because the primary server is at the production site, you will need to perform manual procedures to enable the backup server as the active server at the recovery site.

2.2.2 How to switch to the backup server

In order to stop, start, or change the ESS Copy Services server, you will need the Admin userid and password for the ESS Storwatch Specialist. The buttons that enable you to do this are located within the Help text on the ESS Storwatch Specialist display. In order to access the **Help** panels, you need to click the ? (question mark) which can be seen in the top right-hand corner of Figure 5.

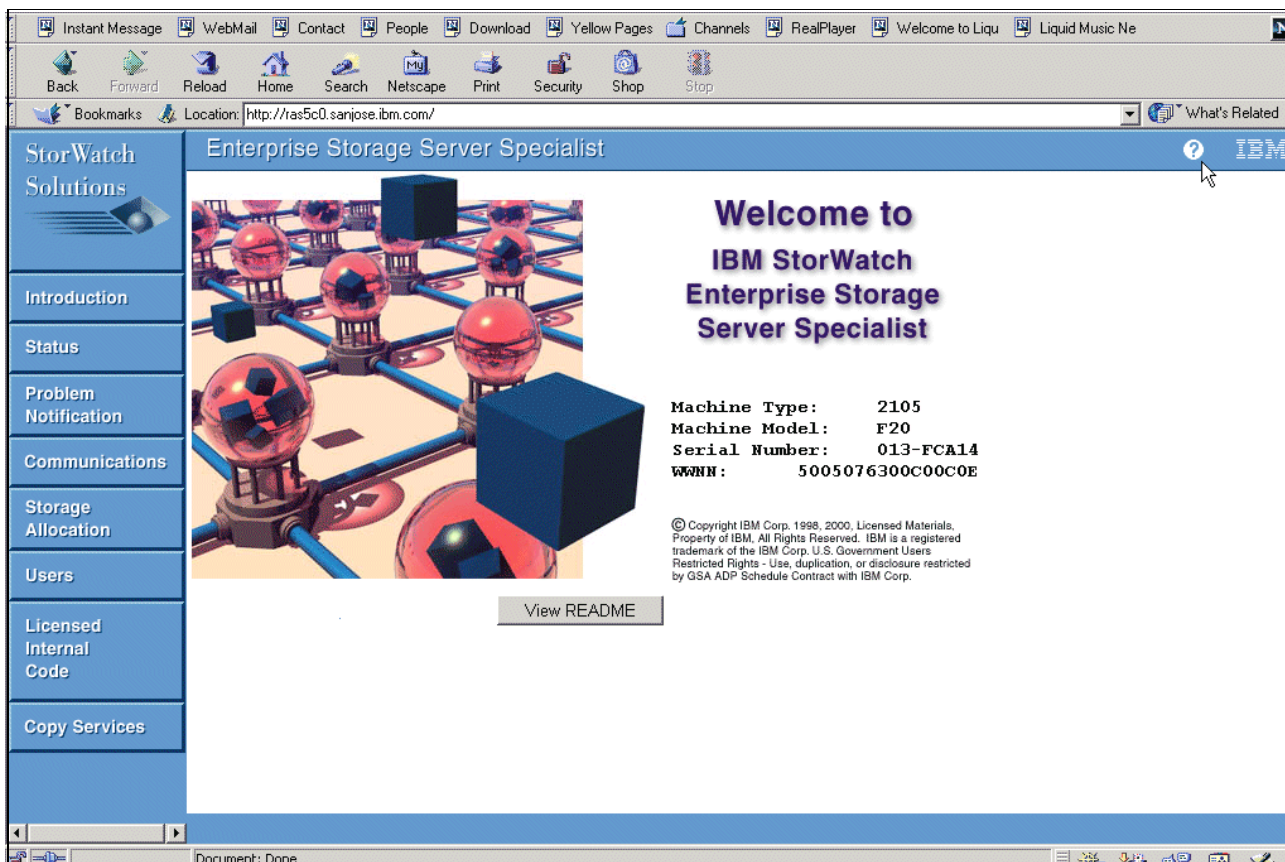


Figure 5. ESS Storwatch Specialist Web User Interface front screen

After you click the **Help** button, you enter the Help menus. You will need to select the **Task Help** flag and scroll down to the last item, **Copy Services troubleshooting and disaster recovery**, which can be seen in Figure 6. When you click this line, you will see the detailed Help text and be prompted to enter the Admin userid and password.

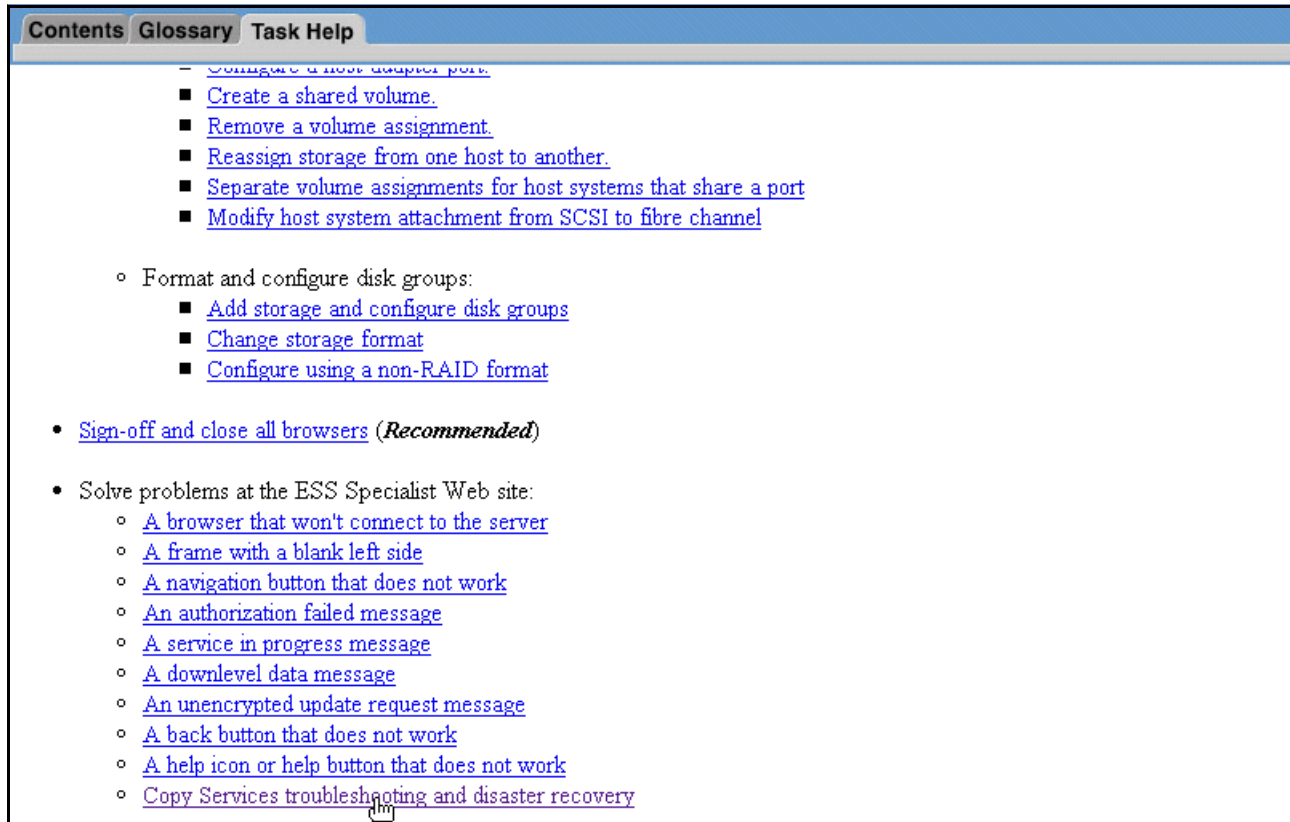


Figure 6. Task Help screen

Once the userid and password have been successfully entered you will enter the section Resetting the ESS Web Copy Services. You must pay particular attention to the warning and procedures that are at the start of this section, as shown in Figure 7. it is particularly important to notice that:

1. You will lose any PPRC or FlashCopy tasks for which you have not received a "successful" completion message.
2. You cannot submit any additional Command Line Interface tasks until ESS Web Copy Services has reinitialized.

Enterprise Storage Server Special

Contents
Glossary
Task Help

Resetting the ESS Web Copy Services

We do not recommend that you use this function, unless you are certain that there are no other recovery options. When you use this function, please be aware that:

- You will lose any PPRC or FlashCopy tasks for which you have not received a "successful" completion message.
- You cannot submit any additional Command-Line Interface tasks until ESS Web Copy Services has reinitialized.

Before resetting ESS Web Copy Services you may want to perform the following actions:

- Restart your web browser.
- Reboot the workstation that is running your browser.

If these actions do not resolve the problems that you are experiencing with the ESS Web Copy Services, you can attempt to use one of the **Reset** buttons, below. However, if you are not sure that you really want to do this, please press the **Cancel** button.

If after resetting ESS Web Copy Services you experience the same problems as before, you may have to contact your IBM Service Support Representative (SSR).

Figure 7. Resetting the ESS Web Copy Services warnings

The actions that you are able to perform are shown in Figure 8.

Available Actions:

Action	Description
Reset to Primary	Restart Copy Services with the primary server as active server.
Reset to Backup	Restart Copy Services with the backup server as active server.
Disable	Disable Copy Services.
Cancel	Return to main task-help, without performing any action.

NOTE: to ensure proper processing of these actions, your browser should be set to avoid caching these pages. For **Netscape**, select Edit, Preferences, Advanced, Cache, and select *Document in cache is compared to document on network every time*. For **Internet Explorer**, select Tools, Internet Options, General, Temporary Internet File Settings, and *Check for newer versions of stored pages on every visit to the page*.

Figure 8. Available actions

It is important to remember that you must disable Copy services before you start either the Primary Server or the Backup server and that Copy Services must be disabled on both ESS.

If we consider the scenario where we have two ESSs, ESS1 and ESS2 with the primary Copy Services Server defined on ESS1 and the backup Copy Services Server defined on ESS2. We wish to transfer control of Copy Services from the primary to the backup server. The steps that must be taken are these:

1. From the ESS Specialist on ESS2, disable Copy Services.
2. From the ESS Specialist on ESS1, disable Copy Services.
3. From the ESS Specialist on ESS2, start the Backup Server by clicking the **Reset to Backup** button, as shown in Figure 8.
4. From the ESS Specialist on ESS1, start Copy Services by clicking the **Reset to Backup** button, as shown in Figure 8.

It is important to note that you must first start Copy Services on the ESS that is running the Server, either primary or backup, that you wish to use. If you bring the client up before the server, problems will occur.

If you had wanted to bring up the primary server rather than the backup server, then you would have clicked the **Reset to Primary** button, and you would have started Copy services first on ESS1 and then on ESS2.

2.3 Host authorization

Copy Services offers a feature to control the usage of Copy Services issued from a host. You have the option to enable host password protection from the Copy Services Specialist Configuration panel. If you have password protection enabled, you must authorize a userid and password so that commands can be issued from host server.

The Specialist's standard userid and password, by default, is not included in the authorization list. You can authorize userids and passwords by clicking the **Authorize** button in the Configuration panel. Please see 5.1.5, "Configuration panel of the ESS Copy Services Web Interface" on page 75 for more information on the Copy Services Configuration.

Chapter 3. FlashCopy

Today, more than ever, organizations require their applications to be available 24 hours per day, seven days per week (24x7). They require high availability, minimal application downtime for maintenance, and the ability to perform data backups with the shortest possible application outage.

The prime reason for data backup is to provide protection in case of source data loss due to disaster, hardware failure, software failure or user errors.

Data copies can also be taken for the purposes of program testing, data mining by data base query applications. However, normal copy operations take a long time requiring the prime application to be off-line. In addition to the need for 24x7 data processing, it is also necessary to have an instant copy of the data.

FlashCopy allows you to move effectively towards such solutions. In this chapter, we introduce the new ESS function of FlashCopy.

3.1 Overview

FlashCopy provides an instant or point-in-time copy of an ESS logical volume. The point-in-time copy functions give you an instantaneous copy, or “view”, of what the original data looked like at a specific point-in-time. This is known as the T_0 (time-zero) copy.

When a FlashCopy is invoked, the command returns to the operating system as soon as the FlashCopy pair has been established and the necessary control bitmaps have been created. This process takes only a few seconds to complete. Thereafter, you have access to a T_0 copy of the source volume. As soon as the pair has been established, you can read and write to both the source and the target volumes.

The point-in-time copy created by FlashCopy is typically used where you need a copy of production data to be produced with minimal application downtime. It can be used for online backup, testing of new applications, or for creating a database for data-mining purposes. The copy looks exactly like the original source volume and is an instantly available, binary copy. See Figure 9 for an illustration of FlashCopy concepts.

FlashCopy provides a T₀ copy

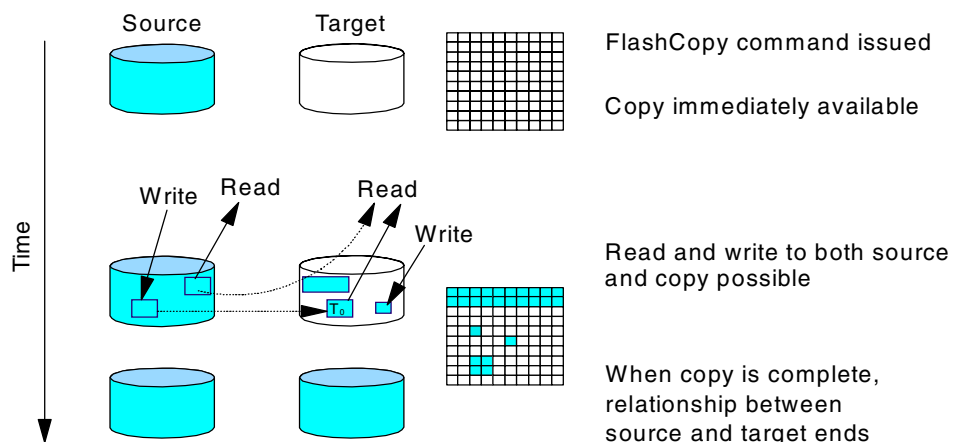


Figure 9. Implementation of FlashCopy

FlashCopy is possible only between disk volumes. It also requires a target volume to be defined within the same Logical Subsystem (LSS) as the source volume. A source volume and the target can be involved in only one FlashCopy relationship at a time. When you set up the copy, a relationship is established between the source and the target volume and a bitmap of the source volume is created.

Once this relationship is established and the bitmap created, the target volume copy can be accessed as though all the data had been physically copied. While a relationship between source and target volumes exists, a background task copies the tracks from the source to the target. The relationship ends when the physical background copy task has completed.

You can suppress the background copy task using the *Do not perform background copy (NOCOPY)* option. This may be useful if you need the copy only for a short time, such as making a backup to tape. If you start a FlashCopy with the *Do not perform background copy* option, you must withdraw the pair (a function you can select) to end the relationship between source and target.

At the time when FlashCopy is started, the target volume is basically empty. The background copy task copies data from the source to the target. The FlashCopy bitmap keeps track of which data has been copied from source to target. If an application wants to read some data from the target that has not yet been copied to the target, the data is read from the source; otherwise, the read is satisfied from the target volume. When the bitmap is updated for a particular piece of data, it signifies that source data has been copied to the target and updated on the source. Further updates to the same area are ignored by FlashCopy. This is the essence of the T₀ point-in-time copy mechanism.

Before an application can update a track on the source that has not yet been copied, the track is copied to the target volume. Reads that are subsequently directed to this track on the target volume are now satisfied from the target volume instead of the source volume. After some time, all tracks will have been copied to the target volume, and the FlashCopy relationship will end.

You cannot create a FlashCopy on one type of operating system and make it available to a different operating system. You can make the target available to another host running the same type of operating system.

3.2 Planning for FlashCopy on ESS

Because FlashCopy invariably will be used on production systems, you should carefully plan the setup of your environment and test it thoroughly. This is an important step to minimize the possible of error and potential rework.

3.2.1 Hardware and software requirements

If you want to use FlashCopy, you need to comply with the following prerequisites:

- Have a FlashCopy feature purchased and enabled on your Enterprise Storage Server (ESS) by the Customer Engineer (CE). The feature code is dependent on the total disk capacity of your ESS, rather than on the capacity of the volumes that will use FlashCopy.
- On the server that will have the FlashCopy target volumes attached, you need to have enough SCSI target IDs and/or SCSI or Fibre Channel LUNs available (not occupied by volumes). The ESS can have up to 15 SCSI target IDs each with up to 64 LUNs on one SCSI channel and up to 4095 LUNs on a Fibre Channel port.
- You need TCP/IP connectivity between the ESS and the host system that will initiate FlashCopy (usually this is the system that will access the FlashCopy target) in order to use Copy Services Command Line Interface (CLI). You can achieve that by connecting the ESS to the company intranet. You have to install the CLI on the host that will be using it.
- If you have Independent Software Vendor (ISV) software installed that writes directly to disk, you need to contact the ISV regarding their support for ESS Copy Services.
- Review your volume manager software considerations for Flash Copy:
 - AIX LVM
 - Veritas VxVM
 - HP SAM
 - Sun Solaris DiskSuite
- The IBM Subsystem Device Driver (which has superseded Data Path Optimizer) fully supports HACMP clusters in both concurrent and non-concurrent access modes. Subsystem Device Driver (SDD) works with FlashCopy volumes.

3.2.2 Configuration planning

The most important consideration that has to be taken into account is to have an available volume (LUN) in the logical subsystem (LSS) where the source volume

resides. The target LUN has to be of the same size as the source or bigger. The space for target data has to be available even if only the *Do not perform background copy* option will be used.

For an understanding of LSS concept see Appendix B, “Understanding logical subsystems” on page 177.

Note

You may need to review your configuration to have a target volume available in the source LSS. That will usually mean copying some of your data to another LSS.

3.2.3 Resource planning

When planning your ESS volume layout, it is important to consider the capacity you may need for FlashCopy targets. Bear in mind that the disk space you need is real disk space. You must also consider that a FlashCopy target is restricted to the same LSS as its source volume. So, when you allocate additional storage for FlashCopy targets, consider how much space in each LSS you need to leave unallocated for them.

You cannot initiate a FlashCopy session on a source and target that are already in a Flash Copy session. You need to wait for the FlashCopy task to complete, or until you can withdraw the pair manually. If you have used the *Do not perform background copy* option, you always need to withdraw the pair.

3.2.4 Data consistency considerations

It is very important to verify that the copy of the data you will be using is fully consistent by using a proper file system check procedure, as provided by your operating system. If you are going to automate your FlashCopy procedures, consider including this check each time when you make the FlashCopy target available to the host. In all cases, before starting the FlashCopy procedure, the target volume must be unmounted; this ensures that there is no data in any system buffers that could be flushed to the target and potentially could corrupt it.

3.2.5 Test plan and disaster recovery plan

If you plan to use FlashCopy, you need to test your setup. Do not forget that you are dealing with a binary copy of the data which was done out of control of your operating system. Prepare a test plan and, if you are using FlashCopy for backup/restore, a recovery plan.

3.3 Operational considerations

The following suggestions are intended to help you manage your FlashCopy pairs and, in particular, manage the target you create.

3.3.1 Monitoring and managing FlashCopy pairs and volumes

FlashCopy pairs and tasks can be managed by both the ESS Specialist Copy Services Specialist Web panel and the Command Line Interface (CLI) on the host.

The ESS Specialist Web Interface allows you to manage FlashCopy volumes and tasks. You can establish and withdraw a FlashCopy by clicking on the graphical representations of the volumes in the Copy Services Specialist. If you wish to perform a FlashCopy from the CLI, you must create a FlashCopy task within the Specialist and save it. You can either execute your tasks from the Specialist, or you call them with the `rsExecuteTask` command in the CLI.

Using the CLI with predefined tasks enables automation, and it minimizes the danger of a human error when handling physical volumes by their volume numbers or names from the ESS Specialist Copy Services Specialist Web panel.

3.3.2 Using a FlashCopy target volume

Remember that if you have established a FlashCopy with the *Do not perform background copy* option, you need to withdraw the FlashCopy pair after you have finished using the FlashCopy target volume. If you choose to perform a full copy, the relationship will be withdrawn automatically when the background copy task ends. The performance issues of using FlashCopy with full copy and *Do not perform background copy* options are discussed in 3.5, “Performance considerations” on page 41.

If you will be using FlashCopy for data backup purposes, change your recovery procedure so that you will be able to recover even when the data has been backed up by a different backup client than the original owner of the LUN, or it has been backed up from a different location in the filesystem (the target mount point).

Of course, you can perform the FlashCopy from the restored volume to the original LUN using the full copy option that will perform the actual data copy to the target volume.

Note

Do not attempt to defragment or optimize a FlashCopy source volume while the FlashCopy background copy task is running. This can significantly degrade your performance.

3.3.3 Automation

Different operating systems allow different levels of automation. The automation can be done using batch or script files executed before and after the application that uses the FlashCopy target.

In the batch file to be run before the application that will use the copy, you should include:

1. Quiescing of an application (switching on the backup mode). Proper quiesce procedure should be provided in your application’s documentation.
2. Flushing data to the source volume. This can be accomplished by unmounting the target but sometimes it may be necessary to shutdown the source server.
3. Establishment of FlashCopy pair(s) using the Copy Services Command Line Interface with the `rsExecuteTask` command.
4. Re-mounting the source.

5. Resuming an application (terminating the backup mode) using the procedure described in your application's documentation.
6. Hardware scan for new disks on a target system, in case their definitions are not already present.
7. Verifying the consistency of a FlashCopy target.
8. Mounting the target on a target system.

In the batch file to be run after backup is completed, you should include the following operations:

- Termination of the FlashCopy pair (in case the *Do not perform background copy* option was used) with an `rsExecuteTask` command.
- Unmounting the target from a target system.

3.3.4 Microsoft Windows NT

Windows NT handles disks in a way that is not similar to any other operating system covered in this book. The need to reboot a server in order to scan for new disks and the need to run a GUI-based Disk Administrator in order to manipulate the disks are the main factors that restrict the routine use of FlashCopy and make automation virtually impossible.

It is possible to automate the actions of the GUI-based Disk Administrator using third-party software to remotely reboot the server and to remotely assign the drive letter from the server that starts the FlashCopy task. This was not tested during our project. You can automate the invocation of FlashCopy using the ESS Copy Services Command Line Interface (CLI) which is described in 2.2, "Configuring primary and secondary Copy Services servers" on page 10. For example, most backup software will allow you to insert a batch file to be run before the backup and another one after the backup.

If you are going to create an automated script with Windows NT, you need to be very careful about data consistency. It may also mean that the script has to be run on a different server than the server that will read the FlashCopy target. Not all applications will allow this.

You have two options on how to make FlashCopy target available to the server, **with rebooting** and **without rebooting**. We recommend that you reboot the server, it is safer because then it is guaranteed that all the registry entries get created. However, using FlashCopy without rebooting is faster.

Note: If you have chosen the "without rebooting" method, go to 3.3.4.3, "Performing a FlashCopy" on page 25.

3.3.4.1 Making target volume registered with the server

If you are going to reboot the server, you do not have to make the target disks known to Windows NT before you do the FlashCopy. However, we recommend that you have them preassigned and registered in the server. The "assign disk and run FlashCopy" approach is useful for a *non-routine* FlashCopy, for example, for testing or migration.

For routine purposes we recommend having target disks already present in Disk Administrator with partitions created and partition information saved. You can

accomplish that by clicking **Start -> Programs -> Administrative Tools -> Disk Administrator**.

1. If the target disk has not been previously seen by the system, Disk Administrator will issue a pop-up message saying "No signature on Disk X. Should I write a signature?" where X is the number assigned to the newly present disk.
2. Click **OK** to save the signature on the target disk. The Disk Administrator will come up.
3. Left-click on the disk that is to be used as the FlashCopy target (it should be grey and marked as Free Space) and select **Create**.
4. Confirm the partition parameters and click **OK**. The partition appears as Unknown.
5. Left-click on the newly created partition and select **Commit Changes Now**.
6. Right-click on the partition and select **Assign Drive letter**.
7. Assign a drive letter and click **OK**. Exit Disk Administrator.

After this procedure the FlashCopy target is properly registered in the Windows NT. Continue with 3.3.4.2, "Bringing down the target server" on page 25 or with 3.3.4.3, "Performing a FlashCopy" on page 25 based on whether you have chosen to perform a reboot.

3.3.4.2 Bringing down the target server

Bring down the server that will use the target if you want to use the safer method.

If you have just assigned the target volume to the host server with ESS Specialist, you have to restart the Copy Services Server and check the volumes ESS Copy Services Specialist get refreshed. Also bear in mind that if you were assigning the volume to the host just before performing the FlashCopy, you will have to use the volume serial number for the FlashCopy target. You cannot use the `rsPrimeServer` CLI command to insert disk names into the ESS Copy Services Specialist as the server is down.

3.3.4.3 Performing a FlashCopy

Stop all applications using the source volume. Now you must flush the data to the source volume. You can accomplish that by clicking **Start -> Programs -> Administrative Tools -> Disk Administrator**.

1. Left-click on the disk that is to be used as the FlashCopy source (it should have a drive letter assigned and be formatted) and select **Assign Drive letter**.
2. In the pop-up box, select **Do not assign a drive letter** and click **OK**.

Now the data is flushed to the source and you can start the FlashCopy task from the ESS Copy Services Specialist or from any server CLI.

Observe the CLI or do the `rsQuery` to the volumes to see if the FlashCopy task had successfully started. Then you can reassign the drive letter to the source volume.

1. Left-click on the disk that is a FlashCopy source and select **Assign Drive Letter**.

2. Assign a drive letter and click **OK**. Exit Disk Administrator.

You can resume using the source volume.

If you are using the FlashCopy “without rebooting” method, go to 3.3.4.5, “Making FlashCopy targets available” on page 26.

3.3.4.4 Booting up target server

After that, you may boot up the target server. In this case, you have just assigned the target volumes to the host that will create the disk entry in the Windows NT registry. To verify that the registry entry is created, do the following:

1. Click **Start -> Settings -> Control Panel**.
2. In Control Panel, double-click **SCSI Adapters**.
3. Click on the adapter that has the target volume attached.

That opens a list of targets. Verify the list includes the target ID and LUN of the volume you have just made available to the server. If you are using SDD, you will see each disk entry several times, depending on how many paths to a volume you have defined.

You may also run the command `datapath query device` from the SDD command line to check if the FlashCopy targets are listed between the volumes. This command will also enable you to check volume serial numbers, and will give you a more understandable overview of the volumes and their paths.

3.3.4.5 Making FlashCopy targets available

Log in, start Windows NT Disk Administrator, write a signature if necessary, and assign a drive letter. To do that, click **Start -> Programs -> Administrative Tools -> Disk Administrator**.

1. If the disk has not been previously seen by this system, Disk Administrator will issue a pop-up message saying “No signature on Disk X. Should I write a signature?” where *X* is the number assigned to the newly present disk.
2. Click **OK** to save the signature on the target disk. The Disk Administrator will come up.
3. Left-click on the disk that is a FlashCopy target (you should see a formatted partition on it) and select **Assign Drive Letter**.
4. If you cannot assign a drive letter, the target is probably corrupt. Try repeating the whole process and consider the scenario that includes reboot.
5. Assign a drive letter and click **OK**. Exit Disk Administrator.
6. From a Windows NT command prompt run `chkdsk x: /f /r`, where *x* is the letter assigned to the FlashCopy target. An option is to run the disk check from the **Properties** of a disk in Windows NT Explorer.

After this procedure, the FlashCopy target is available to the Windows NT and can be handled like normal disk.

3.3.5 Windows 2000

Windows 2000 handles its disks differently than does Windows NT. Windows 2000 incorporates a stripped-down version of the Veritas Volume Manager, called the Logical Disk Manager (LDM).

With the LDM you are able to create logical partitions, perform disk mounts, and create dynamic volumes. There are five types of dynamic volumes: simple, spanned, mirrored, striped, and RAID-5.

On Windows NT the information relating to the disks was stored in the Windows NT registry. With Windows 2000 this information is stored on the disk drive itself in a partition called the LDM database, which is kept on the last few tracks of the disk. Each volume has its own 128 Bit Globally Unique Identifier (GUID). This is similar to the disk Physical Volume Identifier (PVID) in AIX. As the LDM is stored on the physical drive itself, with Windows 2000 it is possible to move disk drives between different computers.

3.3.5.1 FlashCopy limitations with Windows 2000

Having the drive information stored on the disk itself imposes some limitations when using Copy Services functionality on a Windows 2000 system:

1. The source and target volumes **must** be of the same physical size. Normally the target volume can be bigger than the source volume. With Windows 2000 this is not the case, for two reasons:
 - a. The LDM database holds information relating to the size of the volume. As this is copied from the source to the target, if the target volume is a different size from the source, then the database information will be incorrect, and the host system will return an exception.
 - b. The LDM database is stored at the end of the volume. The copy process is a track-by-track copy, unless the target is an identical size to the source the database will not be at the end of the target volume.
2. It is **not** possible to have the source and target FlashCopy Volume on the same Windows 2000 System, when they were created as Windows 2000 dynamic volumes. The reason is that each dynamic volume has to have its own 128 Bit GUID. As its name implies, the GUID must be unique on one system. When you perform FlashCopy, the GUID gets copied as well, so this means that if you tried to mount the source and target volume on the same host system, you would have two volumes with exactly the same GUID. This is not allowed, and you will not be able to mount the target volume.

3.3.5.2 Mounting a FlashCopy target volume

In order to see target volumes on a second Windows 2K host, you have to:

1. Perform the FlashCopy.
2. Reboot the host machine on which you wish to mount the target volume.
3. Open Computer Management, and then click **Disk Management**.
4. Find the Disk that is associated with your volume. There are two "panes" for each disk; the left one should read **Dynamic** and **Foreign**. It is likely that no drive letter will be associated with that volume.
5. Right-click on that pane, and select **Import Foreign Disks**. Select **OK**, then **OK** again. The volume now has a drive letter assigned to it, and is of Simple Layout and Dynamic Type. You can read/write to that volume.

3.3.6 IBM AIX

The FlashCopy functionality in ESS Copy Services copies the entire content of a source volume to a target volume. If the source volume is defined to the AIX

Logical Volume Manager (LVM), all of its data structures and identifiers are copied to the target volume, as well. This includes the Volume Group Descriptor Area (VGDA), which contains the Physical Volume Identifier (PVID) and Volume Group Identifier (VGID).

For AIX LVM, it is currently not possible to activate a volume group with a disk that contains a volume group ID (VGID) in the VGDA and a PVID that is already used in a active volume group. Even if the hdisk PVID is cleared and reassigned with the following two commands:

```
chdev -l <hdisk#> -a pv=clear,  
chdev -l <hdisk#> -a pv=yes,
```

It is currently not possible to modify the VGID without destroying the volume group. Therefore, the `importvg` command of the FlashCopy target will fail on the same server.

3.3.6.1 Accessing a FlashCopy target on another RS/6000

Accessing a FlashCopy target on another RS/6000 poses some problems. As a result of mirroring the entire disk contents, all the data structures and identifiers used by the LVM are also duplicated, thus causing conflicts. None of the existing set of LVM tools and scripts can be used to access the logical volumes, filesystems on this disk.

If your AIX level is 4.3.3, you may want to skip the section “Long method” on page 29 and concentrate on “New AIX command, recreatevg” on page 31. You will require a PTF to access the new functionality in AIX.

Currently, there are some limitations in LVM which make it necessary for you to undertake a non-trivial procedure to make a FlashCopy available on the same server as the source. See 3.3.6.2, “Source and target volumes on the same RS/6000” on page 29 for more information. LVM development is working on removing these limitations and providing new functionality to make the process much easier. For the time being, we recommend that you access the FlashCopy target volume on another RS/6000.

The following procedure makes the data of the FlashCopy target volume available to AIX.

1. If the target volume (hdisk) is new to AIX, run the Configuration Manager. If not, the device should first be removed using the `rmdev` command, and then the Configuration Manager should be run:

```
Type: cfmgr
```

2. Check which hdisk is your FlashCopy target:

```
Type: lsdev -Cc disk | grep 2105
```

3. Import the Volume Group:

```
Type: importvg -y <volume group name> <hdisk#>
```

4. Vary on the Volume Group:

```
Type: varyonvg <volume group name>
```

5. Verify consistency of all file systems on the FlashCopy target:

```
Type: fsck -y <file system name>
```

6. Mount the File System:

Type: `mount <file system name>`

The data is now available. For example, you can create a backup of the data on the FlashCopy Volume to a tape device. This procedure may be run once the relationship between FlashCopy source and target is established, even if data is still being copied from the source to the target in the background.

It may be the case that the disk containing the target volume was previously defined to an AIX system, for example, if you periodically do backups from the same volume. In this case, make sure you remove all the LVM information from the configuration database before you establish the next FlashCopy.

1. Unmount all file systems:

Type: `umount <file system>`

2. Vary off the volume group:

Type: `varyoffvg <volume group>`

3. Export the volume group:

Type: `exportvg <volume group>`

4. Clear the PVID of the target volume:

Type: `chdev -l <hdisk#> -a pv=clear`

5. Delete the target volume:

Type: `rmdev -dl <hdisk#>`

3.3.6.2 Source and target volumes on the same RS/6000

In this section we describe a method of accessing the FlashCopy target volume on a single AIX system while the source is active on the same server. The procedure is intended to be used as a guide and may not cover all scenarios.

The steps needed to access the target volume depend on your installed level of AIX. If your level is AIX 4.3.3 at the recommended maintenance level 05 (APAR IY10456) or higher, then skip over the next section, “Long method”, and use the procedure outlined in “New AIX command, recreatevg” on page 31. Otherwise, use the “Long method”.

Long method

In our example we have a volume group called *fc_source_vg* on *hdisk3*. On this volume group there is a logical volume defined *fc_source_lv* with a mounted file system on *fc_source_fs*. This volume will be our source for the FlashCopy operation. The target of our FlashCopy volume will be *hdisk4*, which is currently not in use.

Please make sure that the source for your FlashCopy is in a consistent state for the short period of time when establishing the FlashCopy pair. Use the following procedure to access the target volume:

1. If necessary, configure the FlashCopy target LUN to AIX:

Type: `cfgmgr`

This will make the target LUN known to AIX as an *hdisk*.

2. Get the source physical partition size:

```
Type: lsvg fc_source_vg | grep "PP SIZE"
```

In our example the physical partition size is 16MB.

3. Bring down applications that access the FlashCopy source and unmount the related file systems for the short period of FlashCopy establishment.
4. Establish the FlashCopy pair.
5. Once the FlashCopy pair is established, mount all file systems and restart the applications. This could be done even if data is still copied from the source to the target in the background.
6. Clear the PVID from the target hdisk to allow a new volume group to be made:

```
Type: chdev -l hdisk4 -a pv=clear
```

7. Make a new volume group on target hdisk using the physical partition size of the volume group containing the source logical volume:

```
Type: mkvg -y fc_target_vg -s 16 -f hdisk4
```

8. Generate the physical partition map for the AIX logical volume on the source, and write it to a file. The output of the `lslv -m` command needs to be reformatted for input to the `mklv` command. This step can be performed while FlashCopy process is copying data. For example:

```
Type: lslv -m fc_source_lv | awk '/hdisk/ {print $3 ":" $2 }' | sed  
's/\(:0*\)/:/' | sed "s/hdisk3/hdisk4/g" > /tmp/lv_map.out
```

You should note that the above command does not work for mirrored file systems. To generate the map file for a mirrored file system, use the following command. For example:

```
Type: lslv -m fc_source_lv | awk '/hdisk/ {print $3 ":" $2 " " $5 ":" $4}' |  
sed 's/\(:0*\)/:/' | sed "s/hdisk3:/hdisk4:/g" | sed  
"s/hdisk13:/hdisk14:/g"> /tmp/map.out
```

9. Determine the number of physical partitions in the logical volume. This value is used in Step 10:

```
Type: cat /tmp/lv_map.out | wc -l
```

10. Make the logical volume using the physical volume map file that was created from the source logical volume's partition map. Specify the new logical volume name, the map file name, the logical volume type, the volume group and the number of partitions in the new logical volume:

```
Type: mklv -y fc_target_lv -m /tmp/lv_map.out -t jfs fc_target_vg 50
```

11. Make a new JFS log logical volume. This step must be performed after all logical volumes have been created. Otherwise, creation of a subsequent logical volume may fail with a physical partition number conflict error.

```
Type: mklv -y fc_log -t jfslog fc_target_vg 1 hdisk4
```

12. Format the new logical volume for use as a JFS log.

```
Type: logform /dev/fc_log
```

13. Add a new stanza to `/etc/filesystems` to include the new file system attributes. Make sure the device entry points to the target logical volume and the log entry points to the new JFS log.


```
/fc_target_fs:
    dev          = /dev/fc_target_lv
    vfs          = jfs
    log          = /dev/fc_log
    mount        = false
    options      = rw
    account      = false
```

14. Create a new mount point for the target:

Type: `mkdir /fc_target_fs`

15. Check the new file system:

Type: `fsck -V jfs -y /fc_target_fs`

16. Mount the new file system:

Type: `mount /fc_target_fs`

Repeat steps 8 through 16 for each file system on the FlashCopy source volume, with the exception of steps 11 and 12. Of course, you need only one JFS log logical volume. You may wish to automate the steps by creating a shell script that recreates each file system.

Once the file system is mounted, you have access the data on the FlashCopy target volume.

Keep in mind that the procedure described above will, more than likely, need to be adapted to your configuration and specific objectives. For example, you may have multiple physical volumes in the volume group, mirrored file systems and so on.

New AIX command, `recreatevg`

You can see from the section, "Long method" on page 29, that the manual procedure in AIX to overcome the problems of making a FlashCopy target available, is not trivial.

FlashCopying a source volume's contents, with "Do not perform background copy" enabled or disabled, causes all of the data structures and identifiers used by AIX's Logical Volume Manager to be duplicated to the target volume. The duplicate definitions, in turn, cause conflicts within LVM. Until now, none of the existing set of LVM commands have had the capability to access the logical volumes or filesystems on the target disk. This problem is solved now with a new AIX command, `recreatevg`.

The `recreatevg` command is packaged as a PTF for AIX 4.3.3 in APAR IY10456 and higher. It is officially available in:

- AIX 4.3.3 Recommended Maintenance Level 05 (RML05)
- AIX 4.3.3 RML06

It is also being shipped in the beta code level AIX 5L.

The `recreatevg` command overcomes the problem of duplicated LVM data structures and identifiers caused by a disk copying process such as FlashCopy. It is used to recreate an AIX Volume Group (VG) on a set of disks that are copied

from a set of disks belonging to a specific VG. The command will allocate new physical volume identifiers (PVIDs) for the member disks and a new volume group identifier (VGID) to the volume group. The command also provides options to rename the logical volumes with a prefix you specify, and options to rename "labels" to specify different mount points for filesystems.

Here is the AIX *man* page synopsis.

```
recreatevg [ -y VGname ] [ -p ] [ -f ] [ -Y lv_prefix | -I LvNameFile ]  
[ -L label_prefix ] [ -n ] PVname...
```

Description:

This command can be used to recreate a VG on a set of disks that are mirrored from a set of disks belonging to a specific VG. This command will allocate new physical volume identifiers (PVID) for the member disks as the PVIDs will also be duplicated by the disk mirroring. Similarly, other LVM logical members that are duplicated will also be changed to new names with the specified prefixes.

Flags:

-y VolumeGroup: Specifies the volume group name rather than having the name generated automatically. Volume group names must be unique systemwide and can range from 1 to 15 characters. The name cannot begin with a prefix already defined in the PdDv class in the Device Configuration database for other devices. The volume group name created is sent to standard output.

-p disables the automatic generation of the new PVIDs. If -p flag is used, you must ensure that there are no duplicated PVIDs on the system. All the disks that were hardware mirrored must have had their PVIDs changed to an unique value.

-Y lv_prefix causes the logical volumes on the VG being recreated renamed with this prefix. The number of characters in the prefix should be such that the total length of the prefix and the logical volume name must be less than or equal to 15 characters. If the length exceeds 15 characters logical volume will be renamed with default name. The name cannot begin with a prefix already defined in the PdDv class in the Device Configuration Database for other devices, nor be a name already used by another device.

-L label_prefix causes the labels of logical volumes on the VG being recreated changed with this prefix. User must modify the /etc/filesystems stanza manually if a simple modification of the mount point is not enough to define the stanza uniquely.

-I LvNameFile entries in the LvNameFile must be in the format LV1:NEWLV1. After `recreatevg`, LV1 will be renamed with NEWLV1. All the logical volumes that are not included in the LvNameFile will be recreated with the default system generated name.

-f Allows a volume group to be recreated that does not have all disks available.

-n After `recreatevg`, the volume group is imported but varied off. Default is imported and vary on.

Notes:

1. To use this command, you must have root user authority.
2. All the member physical volumes of the volume group must be specified on the command line. The command will fail if the input list does not match with the list compiled from the Volume Group Descriptor Area (VGDA).

Examples:

1. To recreate a volume group that contains three physical volumes, enter this command:

```
recreatevg hdisk1 hdisk2 hdisk3
```

The volume group on hdisk1, hdisk2, and hdisk3 is recreated with an automatically generated name, which is displayed.

2. `recreatevg -y testvg hdisk1`

The volume group on hdisk1 is recreated with the new name testvg.

3. `recreatevg -Y newlv hdisk14`

The volume group on hdisk14 is recreated and all logical volumes in that volume group are recreated and renamed with the prefix newlv.

Short method using `recreatevg`

For example, we have a volume group containing two volumes (hdisks) and wish to FlashCopy the volumes for the purpose of creating a backup. To achieve this, we must have two target LUNs available of size equal to or greater than the sources in the same LSS.

The source volume group is `fc_source_vg` containing hdisk4 and hdisk5.

The target volume group will be `fc_target_vg` containing hdisk8 and hdisk9.

Perform the following tasks to create the FlashCopy and make the target volumes available to AIX.

1. Stop all applications that access the FlashCopy source volumes.
2. Unmount all related file systems for the short period of FlashCopy establishment.
3. Establish the FlashCopy pairs with the option to "Do not perform background copy". Use the Copy Services Specialist to establish the pairs or, if you have a task defined, use `rsExecuteTask.sh` in the CLI.
4. Mount all related filesystems.
5. Restart applications that access the FlashCopy source volumes.
6. The target volumes, hdisk8 and hdisk9, will now have the same volume group data structures as the source volumes hdisk4 and hdisk5.

Clear the PVIDs from the target hdisks to allow a new volume group to be made.

```
Type: chdev -l hdisk8 -a pv=clear
      chdev -l hdisk9 -a pv=clear
```

The output from `lspv` shows the result:

```
# lspv
hdisk4      000567992d4c9024    fc_source_vg
hdisk5      000567995abe005e    fc_source_vg
hdisk8      none                 None
hdisk9      none                 None
```

7. Create the target volume group and prefix all filesystem path names with */backup* and prefix all AIX logical volumes with *bkup*

Type: `recreatevg -y fc_target_vg -L /backup -Y bkup hdisk8 hdisk9`

You must specify the `hdisk` names of all disk volumes participating in the volume group. The output from `lspv` illustrates the new volume group definition.

```
# lspv
hdisk4      000567992d4c9024    fc_source_vg
hdisk5      000567995abe005e    fc_source_vg
hdisk8      000567995abdf345    fc_target_vg
hdisk9      00056799b2d831b9    fc_target_vg
```

An extract from `/etc/filesystems` shows how `recreatevg` generates a new filesystem stanza. The filesystem named `/u01` in the source volume group is renamed to `/backup/u01` in the target volume group. Also, the directory `/backup/u01` is created. Notice, also, that the logical volume and JFS log logical volume have been renamed. The remainder of the stanza is the same as the stanza for `/u01`:

```
/backup/u01:
dev          = /dev/bkupelv001
vfs          = jfs
log          = /dev/bkupelvlog001
mount       = true
check       = false
options     = rw
account     = false
```

8. Mount the new filesystems belonging to the target to make them accessible.

3.3.7 Sun Solaris

Making a FlashCopy target available to the same server or to another server is possible.

Be careful when you are adding a volume to a host and using it as a FlashCopy target at the same time. In such a case, a restart of the Copy Services Server may be necessary.

You can use the Copy Services CLI for automation and create scripts to automate your procedures. We recommend that you predefine the tasks to be run and test them thoroughly. Also prepare your target mount point.

The shell script to be run before the application that will use FlashCopy target should include the following operations:

```

#quiesce an application
# insert the quiescing script here
#unmounting the source
umount /source
#start FlashCopy task
rsExecuteTask.sh -s CopyServicesServer EstablishTaskName
#check if FlashCopy task is established
rsQuery.sh -f VolumeList -s CopyServicesServer
#if yes, you can mount the source back
mount /source
#and resume the application
# insert the resuming script here
#check the target for consistency
fsck -y /dev/rdisk/cXtYdZsN
#if OK mount it
mount /dev/dsk/cXtYdZsN /target

```

The shell script to be run after backup should be:

```

#unmount the target
umount /target
#terminate the FlashCopy pair if Do not perform background copy was used
rsExecuteTask.sh -s CopyServicesServer WithdrawTaskName

```

3.3.7.1 Accessing a FlashCopy target on SUN Solaris

In the following example we describe a method of accessing the FlashCopy target on a single SUN Solaris system. The FlashCopy source is active on that server at the same time.

In our example, there is a file system named `/source` on the source volume `c1t6d0s2` of the FlashCopy pair. The target of the FlashCopy will be `c1t6d1s2`.

You can display all available ESS LUNs using the `rsList2105s.sh` command of the Copy Services CLI.

```

./rsList2105s.sh
disk name          2105 serial number
-----
c1t6d0             500FCA24
c1t6d1             501FCA24
c1t6d2             502FCA24
c1t6d3             503FCA24
c1t6d4             504FCA24
c1t6d5             505FCA24

```

Please make sure that the source of your FlashCopy is in a consistent state for the short period of establishing the FlashCopy pair. Use the following procedure to access the target volume:

1. Bring down applications that access the FlashCopy source and unmount the related file systems for the short period of FlashCopy establishment:

Type: `umount /source`

2. Establish the FlashCopy using the Copy Services Web Interface or the Command Line Interface.

3. Once the FlashCopy pair is established, mount all file systems and restart the applications. This could be done even if data is still copied from the source to the target in the background.

Type: `mount /dev/dsk/c1t6d0s2 /source`

4. Check the file system on the target volume of the FlashCopy pair:

Type: `fsck -y /dev/rdisk/c1t6d1s2`

5. Create a mount point for the target file system and mount the file system:

Type: `mkdir /target`

Type: `mount /dev/dsk/c1t6d1s2 /target`

Now the data on the target could be accessed. Following is a part of the `mount -v` output which shows that the source and target are active at the same time.

```
/dev/dsk/c1t6d0s2 on /source type ufs read/write/setuid/largefiles...  
/dev/dsk/c1t6d1s2 on /target type ufs read/write/setuid/largefiles...
```

3.3.8 HP-UX

Both FlashCopy source and target volumes can be made available to the same server or to different servers.

You can use the Copy Services CLI for automation and create scripts to automate your procedures. If you are preparing scripts, you must also prepare your FlashCopy tasks and test them.

The shell script to be run before backup should look like the following:

1. Quiesce an application.
2. Unmount the source.
3. Start FlashCopy task with `rsExecuteTask.sh` CLI command or from Copy Services Web Interface Tasks panel.
4. Check if FlashCopy task is established with `rsQuery.sh` CLI command or from Copy Services Web Interface Volumes panel
5. If yes, you can mount the source back and resume the application.
6. Create a FlashCopy target mountpoint if you have not already done so.
7. Create a directory in `/dev` to place the imported volume group.
8. Create a node in the `/dev` using `mknod` choosing unused minor number.
9. Import the target volume group, activate it and back up the configuration.
10. Check the target for consistency.
11. If OK, mount it.

Following is a self-explanatory example of how the actions and their output should look:

```

# umount /test
# ./rsExecuteTask.sh -u storwatch -p specialist -s sls6c1 HPFlash
# ./rsQuery.sh -u storwatch -p specialist -f disklist -s sls6c1
26-Oct-00 3:42:16 PM rsClientImpl: Successfully connected to Name Server, port = 1703
26-Oct-00 3:42:17 PM rsClientImpl: Server identity = IBM2105 Copy Services/1.1.0
26-Oct-00 3:42:18 PM rsClientImpl: Received rsVSServer reference successfully
26-Oct-00 3:42:19 PM rsClientImpl: rsClientImpl registered successfully
*****Volume Information*****
Volume 20014744 found on 14744:12 as volume number 000
State=simplex, status=not_suspended, FlashCopy_state=source, Size=20.0_GB
*****Volume Information*****
Volume 20114744 found on 14744:12 as volume number 001
State=simplex, status=not_suspended, FlashCopy_state=target, Size=20.0_GB
*****
# mount /test
# mkdir /dev/vgibm
# mknod /dev/vgibm/group c 64 0x040000
# vgimport -v /dev/vgibm /dev/dsk/c0t6d2
Beginning the import process on Volume Group "/dev/vgibm".
Logical volume "/dev/vgibm/lv01" has been successfully created
with lv number 1.
Volume group "/dev/vgibm" has been successfully created.
Warning: A backup of this volume group may not exist on this machine.
Please remember to take a backup using the vgcfgbackup command after activating
the volume group.
# vgchange -a y /dev/vgibm
Activated volume group
Volume group "/dev/vgibm" has been successfully changed.
# vgcfgbackup /dev/vgibm
Volume Group configuration for /dev/vgibm has been saved in /etc/lvmconf/vgibm.c
onf
# fsck -F vxfs -p -y -o full /dev/vgibm/lv01
/dev/vgibm/lv01:log replay in progress
/dev/vgibm/lv01:pass0 - checking structural files
/dev/vgibm/lv01:pass1 - checking inode sanity and blocks
/dev/vgibm/lv01:pass2 - checking directory linkage
/dev/vgibm/lv01:pass3 - checking reference counts
/dev/vgibm/lv01:pass4 - checking resource maps
/dev/vgibm/lv01:OK to clear log? (ynq)y
/dev/vgibm/lv01:set state to CLEAN? (ynq)y
# mount /dev/vgibm/lv01 /ibm

```

The shell script to be run after backup should look like the following:

1. Unmount the target.
2. Terminate the FlashCopy pair if *Do not perform background copy* was used.

3.4 Practical examples using FlashCopy

In this section we discuss various practical examples using FlashCopy.

3.4.1 Moving and migrating data

Anytime you need to move data from one server to another, FlashCopy can be useful. Do not forget to quiesce disk access before making a FlashCopy and verifying the consistency of the data before attaching it to the target server.

3.4.2 Moving workload

In the same way that you can move data from one server to another, you can move workload between servers. To move workload across the channels, you can usually only reassign the volume in the ESS Specialist.

3.4.3 Backup

Flash copy does not usually speed up your backup, but it allows you to run your application while you back up, therefore the backup speed becomes less important for you. You may need fewer tape drives and a lower performance backup server. You will only need to shut down your application for the time the FlashCopy task is started, and can restart it almost immediately.

During the background copy process, the performance of the LSS in which the FlashCopy pair resides may be reduced. If you use the *Do not perform background copy* option, the performance is affected the whole time the FlashCopy relationship exists.

You can keep FlashCopy targets online after you back them up for some time, so you will be able to copy the files that need to be restored from the Flash Copy target rather than having to restore from tape.

FlashCopy also enables you to do backups whenever you want (not only during the off-shift period) because you do not need to wait for a lighter server load in order to do the backup.

You can use all your existing backup software to do backups, however, in case you intend to do full and incremental (differential) backups, you need to check how your software records the files that have been backed up. In some cases it marks the files as “archived” on the target disk. This change, of course, is not reflected in the FlashCopy source volume, so anytime you delete the FlashCopy target and you attempt incremental backup on the next FlashCopy of that source, you back up all files again.

There are three methods you can use to restore the data in case the target is still available:

1. Mapping the target to the source host and the original mountpoint/drive letter using ESS Specialist.
2. Copying the data back to source either using the standard operating system means or doing a FlashCopy back.
3. Creating a new FlashCopy of the target and assigning it to the source host.

3.4.4 Application testing

You can test new applications and new operating system releases against a FlashCopy of your production data. The risk of data corruption is eliminated, and your application does not need to be taken off-line for an extended period of time in order to perform the copy of the data.

3.4.5 Other examples

Data mining is a good example of an area where FlashCopy can help you. Data mining can now extract data without affecting your application.

In the following section we give examples of the usage of FlashCopy and PPRC. The examples are intended to show solutions that are possible with ESS Copy Services. They are not related to a specific operating system and could be used on all supported operating systems.

3.4.5.1 Moving from a single host to a cluster

In a clustered environment two or more servers are accessing the same resources, such as disk drives. There are different cluster models available which provide the following benefits. A cluster increases the availability of the data, as control of the resources from a server that is not available anymore is transferred to the remaining server(s). The failover, failback, and access to the resources is controlled by special cluster software running on all servers within a cluster. Depending on the cluster model, the workload may be shared between the nodes within a cluster, in addition.

If you are currently running your applications in a non-clustered environment and you are planning to move to a clustered environment, ESS FlashCopy functionality can be used to create a test environment for the cluster very quickly and easily with minimal impact on production.

In order to test the cluster, you need to prepare the new host systems of the cluster in advance. This is needed to create an identical environment from the host point of view.

Issue the FlashCopy command on all the volumes you want to use in the cluster environment later on. The FlashCopy target volumes are shared with the new hosts of the cluster. Disk sharing with the Enterprise Storage Server is very easy, as the FlashCopy target LUNs only need to be assigned to multiple host adapters. This is done with the ESS Specialist Web Interface. It would even be possible to move to Fibre Channel connectivity in the clustered environment if the shared target volumes are connected to a Fibre Channel host port.

Once the FlashCopy pair is established, the shared target volumes with the T_0 copy of the production data could be tested in the cluster environment. However, we recommend waiting until the data copy from the source to the target is finished. This will ensure that there is minimal impact on production I/O when testing on the FlashCopy targets.

An example is illustrated in Figure 10.

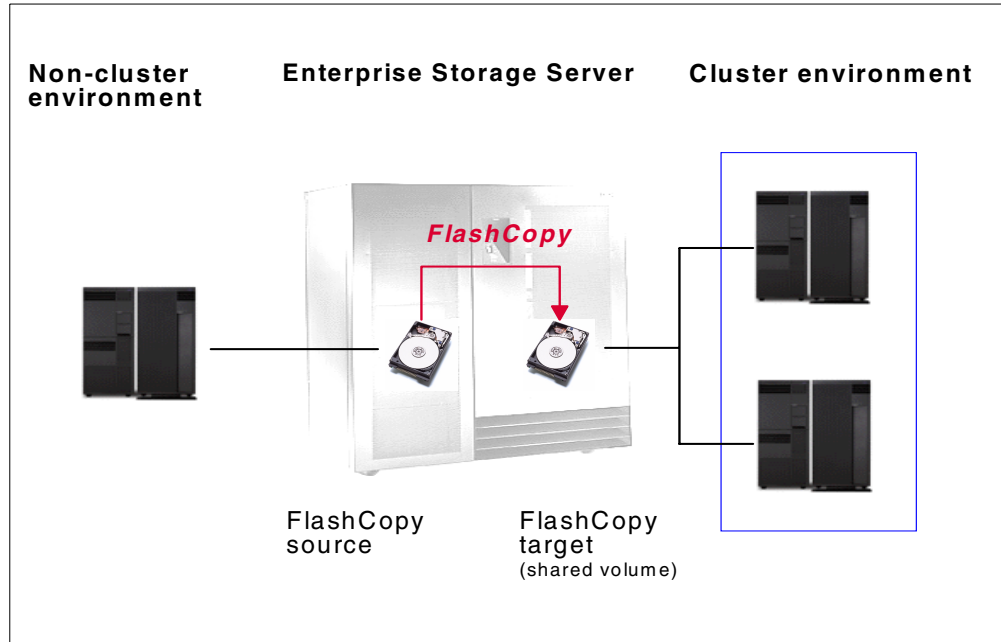


Figure 10. Moving to a cluster environment using FlashCopy

3.4.5.2 Testing a clustered environment

If there is already a clustered environment, you can use FlashCopy to duplicate the shared data on the target volume(s). In this case, both the FlashCopy source and target volumes are LUNs that are shared on the ESS. This may be applicable if you want to test your cluster without having any impact on the production environment.

3.4.5.3 Moving between connectivity methods (SCSI to FC)

You can use ESS FlashCopy functionality to move from one connectivity method to another in the Open Systems environment. The source and target volumes for FlashCopy are not required to be the same kind of connectivity.

An example would be a source volume that is connected to a SCSI port and a FlashCopy target volume that is attached to a Fibre Channel port.

Please keep in mind that although the data on the target is identical to the source, there is a different hardware path to the data. Therefore, modifications to your application or operating system may be required, depending on the software used.

3.4.5.4 Data backup with minimal impact on production

ESS FlashCopy functionality can be used to do online backups of production data. As the data on the FlashCopy target is immediately available after the FlashCopy relationship is established, you need to ensure a consistent state of the data on the source volume for a few seconds only. After the short time of FlashCopy establishment, the data on the target could be written to backup media immediately. This decreases the time of backup windows of production data significantly.

3.5 Performance considerations

This section is intended to give you an idea of the considerations involved when setting up Copy Services of the Enterprise Storage Server (ESS) in order to achieve best performance.

It should assist you to understand the performance impact of ESS Copy Services. As there are many different parameters that have influence on performance, such as applications, kind of workload and configuration of the Enterprise Storage Server, the information should serve as a guideline when planning ESS Copy Services.

Please keep in mind that the general ESS performance considerations such as the volume placement or amount of storage per host adapter still apply when planning for FlashCopy.

3.5.0.1 Placement of source and target volume

As we explained earlier in this book, the source and target volume of a FlashCopy pair must be in the same logical subsystem (LSS). In certain configurations of the ESS, a single LSS includes more than one RAID5 Array (rank). In such configurations, we recommend that you use a target and source volume from different ranks for your FlashCopy pair.

Furthermore, we recommend that you use ranks for your source and target volumes that reside on different SSA loops of the Device Adapters (DA) if possible in your ESS configuration. This will distribute the I/O load over more disks if data is copied from the source to the target volume.

If you are making a FlashCopy with the full copy option when all data from the source is physically copied to the target (default) and you do not have to work with the target volume directly after the FlashCopy was issued, we recommend that you wait until the background copy is finished. This will give you better performance for both source and target volume as host I/O requests do not interfere with I/O of the FlashCopy task. The progress of the FlashCopy background copy process could be determined with the Copy Services Web Interface, and the completion of the process with the Command Line Interface.

3.5.0.2 Do not perform background copy option

Please consider whether you want to select the **Do not perform background copy** option for FlashCopy or not. This option is actually presented by the Select Copy Options panel of the Task Wizard as **Do not perform background copy if checked**.

If you mainly have read access to the source and the target of your FlashCopy pair, we recommend that you use the **Do not perform background copy** option to minimize I/O traffic to the RAID arrays. Keep in mind that when selecting this option, the relationship between the FlashCopy source and target stays until the pair has to be withdrawn manually.

3.5.0.3 Number of simultaneous FlashCopy pairs

Also, you need to consider the number of FlashCopy pairs you have active at the same time. The time for a single FlashCopy pair to finish will increase with the amount of FlashCopy pairs you have established, at the same time as data is copied in between all pairs. Try to logically group FlashCopy pairs together and execute the different groups one after the other.

When grouping multiple FlashCopy pairs into a single task, keep in mind that all of these pairs will be processed in parallel.

Chapter 4. Peer-to-Peer Remote Copy (PPRC)

PPRC is an established data mirroring technology that has been used for many years in Mainframe environments. It is used primarily to protect an organization's data against disk subsystem loss or, in the worst case, complete site failure.

In this chapter, we describe PPRC in detail. We explain how to set it up and give some practical examples of its implementation.

4.1 Terminology

Throughout our book, we use the following terms:

- A system where the production applications run is referred to as the primary site or application site.
- An Enterprise Storage Server where the production data resides, and which is the primary member of a PPRC pair, is referred to as a primary ESS or application ESS.
- A system where the recovery or test applications run is referred to as a secondary site or recovery site.
- An Enterprise Storage Server where copies of production data reside, which is used to keep the data current, and which is the secondary member of a PPRC pair, is referred to as a secondary ESS or recovery ESS.

4.2 Overview

PPRC is a synchronous protocol that allows real-time mirroring of data from one Logical Unit (LUN) to another LUN in a second ESS. The secondary ESS can be located at another site some distance away. PPRC is application independent. Because the copying function occurs at the disk subsystem level, the application has no knowledge of its existence.

The PPRC protocol guarantees that the secondary copy is up-to-date by ensuring that the primary copy will be written only if the primary system receives acknowledgement that the secondary copy has been written.

Figure 11 shows the sequence of events.

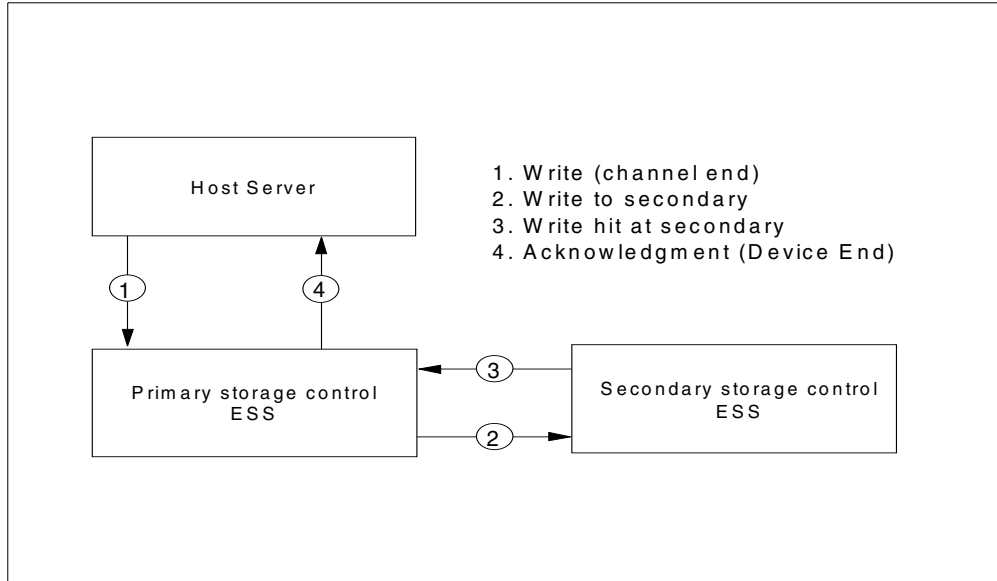


Figure 11. PPRC write cycle

1. The host server requests a write I/O to the primary ESS. The write is staged through cache into Non-Volatile Storage (NVS).
2. PPRC dispatches the write over an ESCON channel to the secondary ESS. The write hits the secondary ESS's NVS.
3. The primary then expects acknowledgment of the remote write. If the secondary write fails, the write does not return to the host server and is eventually "aged" from NVS.
4. The write returns to the host server's application.

Once acknowledgement of the write has been received by the primary, both the primary and secondary write I/Os are eligible for destage to disk. Destage from the cache to the disk drives on both the primary and the secondary ESS is performed asynchronously.

If acknowledgement of the remote write is not received within a fixed period of time, the write is considered to have failed, and is rendered ineligible for destage to disk. At this point, the application receives an I/O error, and in due course, the failed write I/O is "aged-out" of each NVS.

4.2.1 PPRC volume states

Volumes within the Enterprise Storage Server used for PPRC can be found in one of the following states (Figure 12):

Simplex

The Simplex state is the initial state of the volumes before they are used in any PPRC relationship, or after the PPRC relationship has been withdrawn. Both volumes are accessible only when in Simplex state.

Duplex Pending

Volumes are in Duplex Pending state after the PPRC copy relationship was established, but the source and target volume are still out of sync. In that case, data still needs to be copied from the source to the target volume of a PPRC pair. That may be the case either after the PPRC relationship was just established (or reestablished for suspended volumes), or in case the PPRC volume pair reestablishes after a storage subsystem failure. The PPRC secondary volume is not accessible when the pair is in Duplex Pending state.

Duplex

This is the state of a volume pair that is in sync; that is, both source and target volume containing exactly the same data. Sometimes this state is also referred as the full copy mode. The PPRC secondary volume is not accessible when the pair is in Duplex state.

Suspended

Volumes are in Suspended state when the source and target storage subsystems cannot communicate anymore, and therefore the PPRC pair could not be kept in sync, or when the PPRC pair was suspended manually. During the Suspended state, the primary volume's storage server keeps track of all updates to the source volume for reestablishment of the PPRC pair later on. The PPRC secondary volume is not accessible when the pair is in Suspended state.

You can FlashCopy the PPRC secondary when it is in Suspended mode to another volume and use it the way a FlashCopy target can be used. It is necessary to comply with the FlashCopy source consistency requirements, as documented in Chapter 3, "FlashCopy" on page 19. The possible usage scenario of a FlashCopy from suspended PPRC secondary is described in 4.8, "Using PPRC with FlashCopy" on page 54.

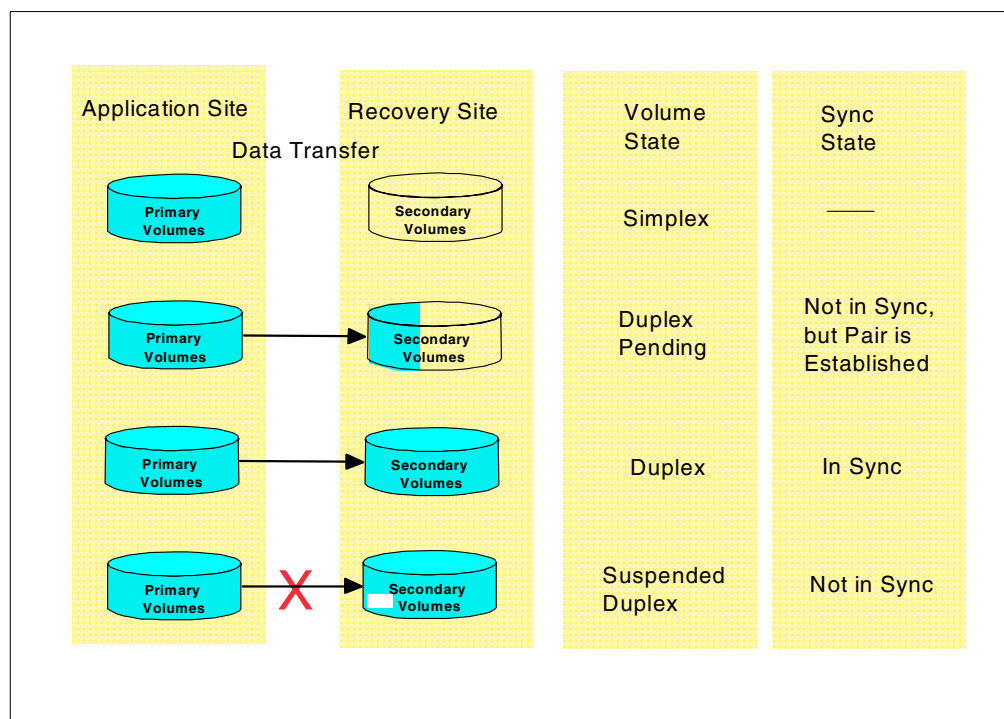


Figure 12. PPRC volume states

4.3 Planning for PPRC on an ESS

The following sections describe the important areas you should consider when planning for PPRC on an Enterprise Storage Server.

4.3.1 Hardware and software requirements

PPRC is possible only between Enterprise Storage Servers. Other disk storage units that support PPRC can also communicate to the same type of unit only.

You need to have the PPRC feature purchased and PPRC-capable microcode activated on all ESS units that will be used for PPRC.

PPRC operates at a volume level from one LSS to another LSS. That means you need to have the target volumes available on the secondary ESS, and you need to identify the LSSs where the primary and secondary volumes are located.

ESCON connections have to be configured between the units (see 4.4.1, “ESCON” on page 49). There can be up to eight ESCON links between the subsystems. A primary ESS can communicate with up to four secondary ESSs (Figure 13). A secondary ESS can be connected to any number of ESS primary subsystems.

You will need to purchase ESCON cables and possibly some other equipment, depending on the distance between the primary and the secondary ESS. However, each of the Copy Services Servers can only control two ESSs. Therefore, to have one primary ESS communicate with more than one secondary ESS, this will mean that multiple ESS Copy Services Servers will have to be configured.

The ESS units involved in PPRC must be connected with their standard Ethernet and TCP/IP to the units that are the primary and backup Copy Services servers. All ESS cluster hostnames, including its own, must be added to the cluster hostname list, the `/etc/hosts` file, during installation. This is configured by the IBM CE during the installation of the PPRC feature.

A browser for the ESS Specialist has to be installed on the machines that will be used to control PPRC with the Web Interface. See the IBM Redbook, *Implementing the Enterprise Storage Server in Your Environment*, SG24-5420, for the browser recommendations.

If you plan to use CLI, install the Java Developers Kit (JDK) on the machines that will run the CLI commands. Check the Host Attachment Guide for the recommended JDK revision.

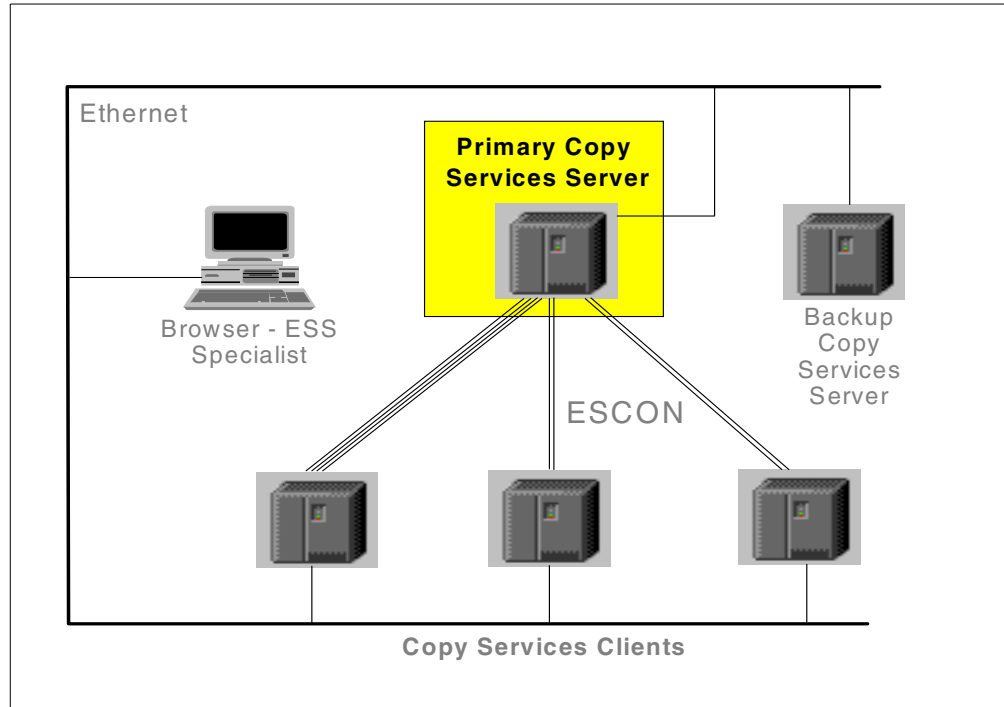


Figure 13. Ethernet and ESCON connection between Copy Services participating ESS

4.3.2 Configuration planning

An important setting is the CRIT parameter in the ESS Vital Product Data (VPD). This is a parameter that determines the behavior of the PPRC pairs or consistency groups after a failure in communication between the primary and secondary ESS, when all the paths between a pair are lost. This option is set by the CE at Copy Services activation time. You can change it, but the change will take effect at the next ESS power cycle.

4.3.2.1 CRIT

CRIT (NO) specifies that, following an I/O completion error to the secondary volume, PPRC allows subsequent write requests to the PPRC primary volume. The primary ESS control will perform change recording. The default is NO.

CRIT (YES) specifies that if an I/O error to the secondary volume occurs, PPRC either allows or does not allow subsequent writes to the primary, depending on how the storage subsystem is configured. The PPRC pair then remains in a suspended state until you correct the problem and either issue a command to resync the PPRC pair or delete the PPRC pair.

The implementation of CRIT (YES) on the ESS is similar to the implementation on the IBM 3990. There is an option that can be set by the CE in the VPD of the ESS which determines how this CRIT(YES) setting will behave in an error situation:

- CRIT=YES - Paths (light version)
 - Suspend the pair and do not accept any further writes if the control units can no longer communicate.
 - Suspend the pair and accept further writes if the control units still can communicate with each other. The reason for not being able to copy the

data to the remote volume is probably only a device problem on the secondary site and not a disaster. Therefore we continue with write operations to the primary volume. The ESS records which cylinders have changed. After investigation of the problem and after it has been solved, you can re-synchronize source and target volume again.

- Suspend the pair and do not accept any further writes to the primary volume if data cannot be sent to the secondary volume.

4.3.3 Resource planning

When planning your secondary ESS volume layout for PPRC, optimize your disk capacity. It is important to realize that the capacity needed on the secondary ESS for disaster recovery may not have to be initially as large as the primary ESS. A disaster recovery plan (DRP) requires significant investment financially in technology, people, and process. Every company will be different, but the I/T components of a disaster recovery plan are essentially driven by the applications and data you require for business continuity, should a disaster occur. Some applications and data will be more critical than others. An organization will typically require its core business systems to be available in a short time, whereas less critical systems quite possibly could be restored over a number of days.

Bearing in mind that the disk space you need for PPRC secondary volumes is real disk space, size your secondary ESS based on your critical business requirements, possibly with some headroom for applications of intermediate importance. Create PPRC pairs for the critical data so that is copied in real time. Then, if a disaster happens, you will have the core systems available on the secondary copies. After the initial recovery priorities have been handled, you can add more disks ranks for the applications of lower importance and restore them from tape.

4.3.4 Data consistency considerations

In any recovery situation, including disaster recovery scenarios, you may be exposed to so-called *lost writes*. These are the “in-flight” transactions that have not been committed from memory to the ESS’s NVS. You should expect that uncommitted transactions will be lost. On the other hand, data that was transferred to the ESS and confirmed back as written into the NVS (of the secondary ESS in case of PPRC), will be destaged to disk.

Invariably, a host server will check its file systems after recovering from a crash. At a DR site, the host servers may be operational when the primary site fails. So it is important to perform a full file system check on all PPRC secondary volumes before you start using them. Of course, rebooting the servers will achieve the same result.

When your database restarts, normal database recovery commences and any partially committed transactions will be rolled back.

4.3.5 Test plan and disaster recovery plan

DRP is complex — nothing can understate the importance of rehearsals and testing your environment. You only get one shot at getting it right when a real disaster hits.

Carefully set up your PPRC tasks for establishing and terminating pairs. Ensure that they are well tested and documented. Prepare your documentation as if it were intended for someone else; you may not be around when a disaster strikes. Make sure you understand any operating system specific issues related to bringing your PPRC secondaries online. Have them well documented in your recovery operations control book. See 4.5, “How to set up PPRC” on page 50.

4.4 How to configure the ESS for PPRC

In this section, we describe how to set up an ESS in preparation for PPRC.

4.4.1 ESCON

For PPRC primary to secondary unit (channel to control unit) communication, a maximum of eight ESCON links using modified ESCON protocol can be configured. ESCON channels provide 160 Mbps point-to-point links. While PPRC can be bi-directional, these links are uni-directional. The primary unit ESCON port (the one in channel mode) has to be dedicated for PPRC. The ESCON port on the secondary unit can also be used for S/390 host attachment, provided the ESCON director is used, and the host is connected to it.

ESCON links support distances of up to 2 km with 50 micron multimode optical cables and up to 3 km with 62.5 micron multimode cables. The ESCON channel performance is a function of distance.

By using up to two ESCON Directors, you can extend these distances. You can use Extended Distance Feature (XDF) ports with single mode fiber optic cables, the XDF maximum distance being 20 km. The total maximum distance is 103 km between the two ESSs.

Various channel extenders can also be used to increase the distance between ESS servers. IBM 9729 Optical Wavelength Division Multiplexer (MuxMaster) enables a 50 km distance between MuxMaster units. ESCON is used to attach ESS to it.

There is a new product, IBM 2029 Fiber Saver, also known as Dense Wavelength Division Multiplexer (DWDM), that supports ESCON, FICON, Fibre Channel, and many more protocols, enabling up to a 50 km distance between Fiber Saver units that are ESCON attached to the primary and secondary ESS. You can use it for Fibre Channel, network, and telephone links between the sites as well.

4.4.2 ESS

PPRC requires logical paths to be established between the primary and the secondary ESS logical control units (or LSS). Each LSS in the primary ESS that will use PPRC requires at least one path to be set to the LSS in the secondary ESS that holds the secondary volumes (Figure 14).

Each ESCON link supports 64 logical paths, so even with 16 LSS defined you are able to set up a logical path from each LSS to each LSS with only four ESCON PPRC links. We always recommend that you use all available links for each LSS pair used for PPRC. That gives you maximum protection against ESCON link failure.

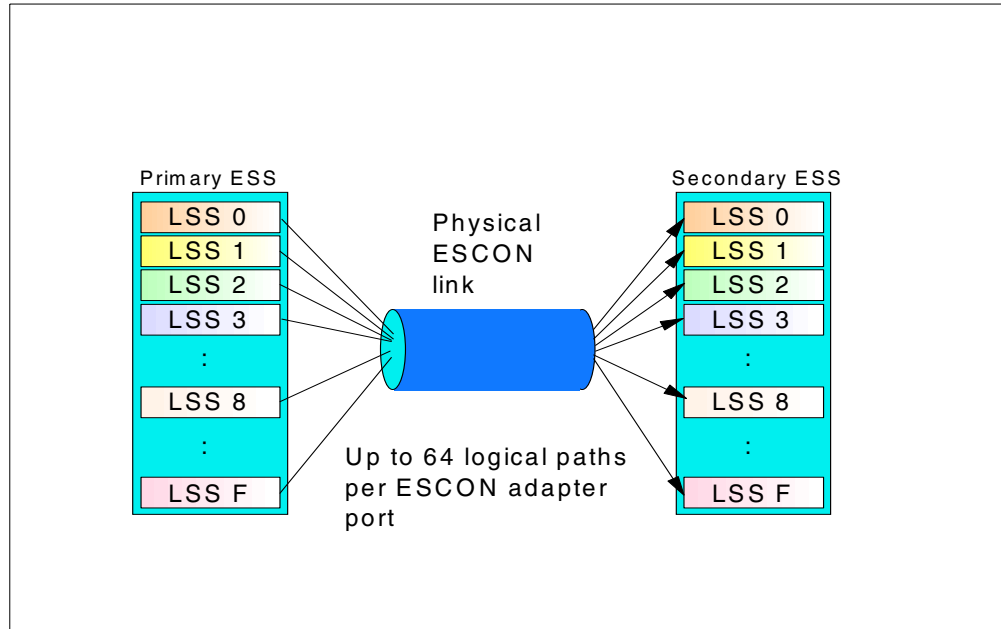


Figure 14. ESS logical paths

4.5 How to set up PPRC

In this section we describe how to set up PPRC.

4.5.1 Recommendations

We recommend that you use Storwatch ESS Specialist Copy Services to perform one-time unique tasks, like path setup, task setup, and one-time PPRC invocation for data migration and workload migration.

Once your tasks are set up, you can use CLI for their invocation.

4.6 Operational considerations

We already know that an ESS can be connected to a variety of host platforms. While configuration and management of Copy Services at the ESS is the same regardless of host platform, each platform's operating system has some differences in the way it presents and manages Copy Services volumes.

The following sections describe a range of general and specific operational considerations relating to AIX, NT, Sun, and HP.

4.6.1 Monitoring and managing volumes and paths

PPRC paths, pairs, and tasks can be managed by both the ESS Web Interface and the Command Line Interface (CLI).

The ESS Web Interface will allow you to manage paths, PPRC volumes, and tasks. We recommend that you set up paths and preset, establish, suspend, and terminate tasks in advance before you start using PPRC. You can use the tasks from the Web Interface, or you can start them with the CLI.

In the CLI, only the invocation of the predefined tasks is possible. Using CLI with predefined tasks minimizes the danger of a human error and it enables automation. You can run tasks with CLI and check their status, but the graphical view in the Web Interface may be a useful source of information about the state of the tasks.

4.6.2 Operating system specifics

In this section we discuss some of the specifics related to the operating system you are using.

4.6.2.1 Microsoft NT

You can manage PPRC pairs with both the ESS Specialist Copy Services Web Interface or with the Command Line Interface (CLI).

Automation of PPRC with Windows NT server is complicated, due to some Windows NT limitations. As explained earlier in this chapter, after a failure in the primary ESS, the PPRC volumes become suspended. In order to bring them online to the host ports on the secondary ESS, a withdraw pair command has to be issued. Then the NT server may have to be rebooted, and the Disk Administrator has to be run in order to register the disks and assign them drive letters.

4.6.2.2 IBM AIX

Both the ESS Specialist Copy Services Web Interface and the CLI can be used to manage PPRC.

To access PPRC secondaries on a p-Server or an RS/6000, each established PPRC pair must be terminated first. If the volume groups containing the secondary volumes are being used for the first time, they must be imported onto the server. Hence, ODM and /etc/filesystems are populated with the volume definitions. If the target disk volumes were previously known to AIX as hdisks or vpaths, then they must be removed (`rmdev`) and re-defined (`cfgmgr`) before running `importvg`. If this is not done first, `importvg` will import the volume group improperly. The volume group data structures (PVIDs, and so on) in ODM will differ from the data structures in the VGDA's and disk volume super blocks. The filesystems will not be accessible.

Once the volume groups have been imported, the PPRC secondary volumes will appear to AIX as Available hdisks on the server. If the secondary volumes have the state of Target and their status is not_suspended and if you reboot the server, then their hdisks will be configured to AIX again.

The reason for this situation is as follows: AIX knows that these Physical Volumes already exist with entries in the Configuration Database. However, when the configuration manager runs during reboot, it cannot read their PVIDs because, as PPRC targets, they are locked by the ESS. This results in AIX causing the original hdisks to be configured to a Defined state, and new (phantom) hdisks being configured and placed in an Available state. This is an undesirable condition that must be remedied before the secondary volumes can be accessed.

To access the Secondary volumes, the phantom hdisks must be removed and the "real" or original hdisks must be changed from a Defined state to an Available state.

For example, hdisk6 through hdisk9 are assigned to a volume group, evg001. Each of the disk volumes is currently participating as a secondary volume in a PPRC pair. If the server is rebooted, four new hdisks are configured to AIX. These phantom disks, hdisk13 through hdisk16, appear in the output from `lspv` as follows:

```
# lspv
hdisk6      000567992d4c9024    evg001
hdisk7      000567995abe005e    evg001
hdisk8      000567995abdf345    evg001
hdisk9      00056799b2d831b9    evg001
hdisk13     none                 None
hdisk14     none                 None
hdisk15     none                 None
hdisk16     none                 None
```

When you execute `lsdev -Cc disk`, you can observe the states of each disk volume.

```
# lsdev -Cc disk
hdisk6 Defined 20-58-01 IBM FC 2105F20
hdisk7 Defined 20-58-01 IBM FC 2105F20
hdisk8 Defined 20-58-01 IBM FC 2105F20
hdisk9 Defined 20-58-01 IBM FC 2105F20
hdisk13 Available 20-58-01 IBM FC 2105F20
hdisk14 Available 20-58-01 IBM FC 2105F20
hdisk15 Available 20-58-01 IBM FC 2105F20
hdisk16 Available 20-58-01 IBM FC 2105F20
```

It is important to execute both the `lspv` and `lsdev` commands back-to-back, so that you can be certain which disks are the phantoms. From the `lspv` output, the phantom disks will have no PVIDs and will not be assigned to a volume group. From the `lsdev` output, the phantom will be in an Available state. The original disks will have PVIDs, be assigned to a volume group, and be marked in a Defined state.

To remove the phantom hdisks, run the `rmdev` command on each phantom disk device.

```
# for i in 13 14 15 16
do
rmdev -dl hdisk$i
done
```

Set the original hdisks to an Available state with the `mkdev` command.

```
# for i in 6 7 8 9
do
mkdev -l hdisk$i
done
```

Now you can activate the volume group, evg001, and mount its filesystems.

4.6.2.3 Sun Solaris

Both ESS Specialist Copy Services Web Interface and the CLI can be used to manage PPRC.

When you set up tasks to manage PPRC pairs, you can easily automate disaster recovery functions, data migration, and split mirror backup and recovery.

4.6.2.4 HP UX

For HP-UX, both ESS Specialist and CLI can be used to manage PPRC.

When you set up tasks to manage PPRC pairs, you can easily automate disaster recovery functions, data migration, and split mirror backup and recovery.

4.7 Moving and migrating data with PPRC

Apart from disaster recovery, you can use PPRC for tasks assisting you with direct disk to tape copy, data migration, and workload migration.

Direct disk-to-tape copy from PPRC secondary volumes can be done in order to minimize the impact on the primary system.

1. Query the PPRC pair status to see if volumes are in duplex.
2. Make sure the data on the primary disk is consistent (stop I/O, unmount), that will make the secondary consistent as well.
3. Click on both volumes (source and target) and do a "suspend pair" or run the task from CLI. The target volume is still not accessible, being in Suspended state.
4. Click on the suspended target volume only and do a "terminate pair", which will turn the target volume to the simplex state (the secondary volume is now accessible to the host on the remote site.)
5. Bring online the secondary volumes and check the file system.
6. Mount secondary volumes.
7. Copy them to tape.
8. To re-establish the PPRC pair, click on both volumes (first on the suspended source, second on the simplex target) and establish a PPRC copy pair with the "copy out-of-sync cylinders only".

Note: Care must be exercised that the secondary volume is not corrupted while it is in simplex state.

To migrate data from one ESS to another, you may implement a process similar to the following:

1. Vary secondary volumes offline from the host that is using them.
2. Establish PPRC paths.
3. Establish PPRC pairs with COPY option.
4. Query the PPRC pair status to see if volumes are in duplex.
5. When copy is complete (pairs are in duplex) you can switch to secondary.
6. Stop all write I/O to primary and unmount.
7. Withdraw the PPRC pairs so volumes will return to simplex - you may delete paths if you wish.

8. If volumes are attached to the same host, you have to vary the primary volumes offline at the host.
9. Vary the secondary volumes online, check the file system and mount.

Then you can start the applications using secondary volumes.

To move workload between hosts or sites and then use the original primary site as the secondary site, you should perform the steps for migrating data to the other host and then do the following additional actions. This procedure can also be used when you wish to make a temporary move, for example when you know that the primary site is going to be unavailable due to maintenance.

1. Establish paths and pairs with NOCOPY in reverse direction.
2. Suspend the pairs.
3. Start applications on secondary volumes.
4. When the old primary site is available you can establish the pair with the RESYNC option.

4.8 Using PPRC with FlashCopy

One good example of combined PPRC with FlashCopy is the split mirror solution for backup and recovery. This solution has been tested by SAP for SAP R/3 and S/390 with ESS, and its description may also help you to design your own split mirror solutions for other applications.

The basic idea of the split mirror solution is managing a consistent copy of the database in a certain point of time on a secondary site (the t_0 copy) while production continues on a primary site on the t_2 copy. On the secondary site, a Flash Copy of the consistent t_1 copy is created, referred to as the t_0 copy, that enables off-line backup to tape and application testing (see Figure 15).

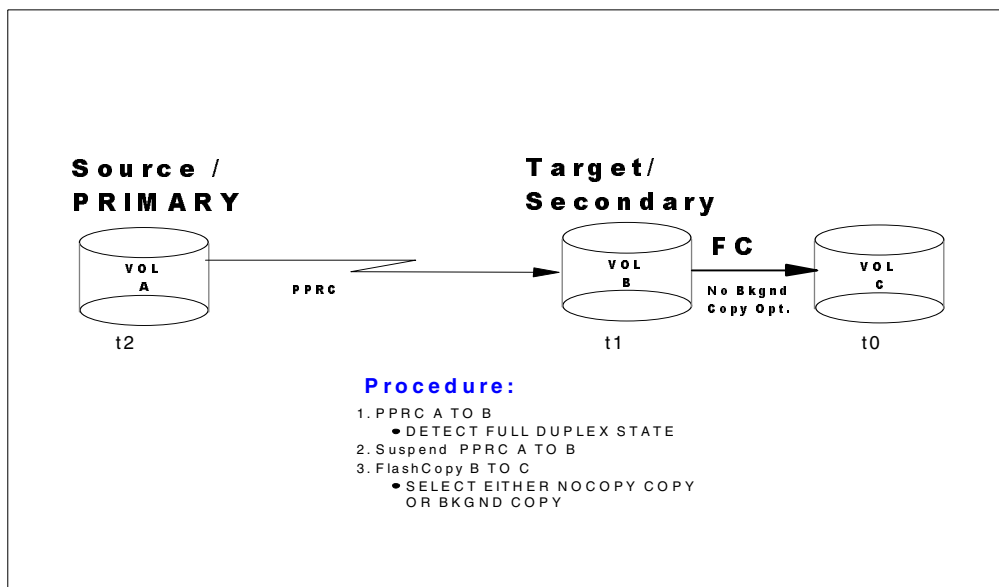


Figure 15. PPRC with FlashCopy

The PPRC pairs are normally suspended, and the mirror is split, preventing propagation of logical errors to the t_1 secondary copy. At certain checkpoints, the pairs are resumed and resynchronized; this means that the database is copied to secondary. This allows restart of the application on the last known consistent copy, or roll forward to a certain point of time using logfiles. The database active and archive logs are copied constantly; the volumes they are at are not suspended.

In case of an application logical error, which is more frequent than a total site failure, the production resumes on the secondary database, while the primary can be analyzed. In case of a hardware failure, you can switch quickly to the secondary site.

The implementation planning should include:

1. Identification of logical to physical volume mapping.
2. Preparation of both the primary and secondary site ESSs.
3. Setup of PPRC ESCON channels and paths.
4. Setup of PPRC tasks.
5. Setup of recovery site hosts.
6. Testing.

Identification of mapping will include creating a list of physical volumes that contain the database and the logs, as well as checking that they are placed on physical volumes properly in such a way that they can be suspended and resumed. For example, the database cannot be on the same physical volume the logs are on.

The preparation of the ESSs means purchasing the PPRC features and installing ESCON adapters, checking that disk space is available on the selected LSS, and installing Java on hosts in order to use CLI.

You need to have PPRC-dedicated ESCON channels available between the primary and secondary sites.

Tasks have to be set up using the Web Interface; and then they are ready to be executed from the CLI. You also need to group the tasks into task groups to be able to carry out the point-in-time critical operations (establish, suspend, resync, FlashCopy) into task groups.

The host server on the remote site has to be set up so it can take over in case of a primary site failure.

All the PPRC tasks, backup operations, and site takeover have to be tested regularly.

Before you start using split mirror routinely, you have to perform the initialization steps, either with the Web Interface or CLI.

1. Establish the paths between primary and secondary ESS
2. Establish the pairs for all necessary volumes (establish, copy entire volume)
3. Query the pairs if establish is complete
4. Suspend write I/O to the primary t_2 volumes

5. Withdraw the pairs between primary and secondary ESS (suspend pairs)
6. Resume write I/O to the primary t_2 volumes
7. FlashCopy the t_1 volumes on the secondary ESS to t_0 copy
8. Resynchronize the primary t_2 and secondary t_1 volumes (establish, copy changes only)
9. Dump or backup the t_0 copy to tape

Repeated routine steps will include these:

1. Create safety copy if resynchronization fails (suspend all, recover t_1 and FlashCopy t_1 to t_0)
2. Resynchronize the primary t_2 and secondary t_1 volumes
3. Query the status of resynchronize process
4. Suspend write I/O to the primary t_2 volumes
5. Withdraw the pairs
6. Resume the write I/O
7. FlashCopy the t_1 volumes to t_2
8. Resynchronize the primary and secondary volumes for BSDS, active and archive logs.
9. Dump or backup the t_0 copy to tape

4.9 Practical example of PPRC and FlashCopy combination

Asynchronous PPRC with FlashCopy

PPRC on the ESS is a *synchronous* copy process. This ensures that the data on the source and target is always in sync. As PPRC most commonly is used for disaster recovery, you want to have a synchronous copy to ensure that no transaction written to the primary volumes is lost on the secondary volumes. There is a permanent relationship between the source and the target.

However, with a combination of FlashCopy and PPRC, you can achieve a kind of *asynchronous* PPRC. In this solution, the data on the secondary side represents the data of the primary to a specific point-in-time (t_0). In case the primary side is going down, you can recover on the secondary side with the t_0 data. The administrator has to take care to periodically update the data on the secondary side. You can think of the data on the secondary side to be a shadow of the primary side. The delta of the primary and its shadow on the secondary is based on the time difference in updating the data on the secondary (using PPRC). There is no permanent relationship between the source and the target.

The setup of such a configuration is shown in Figure 16.

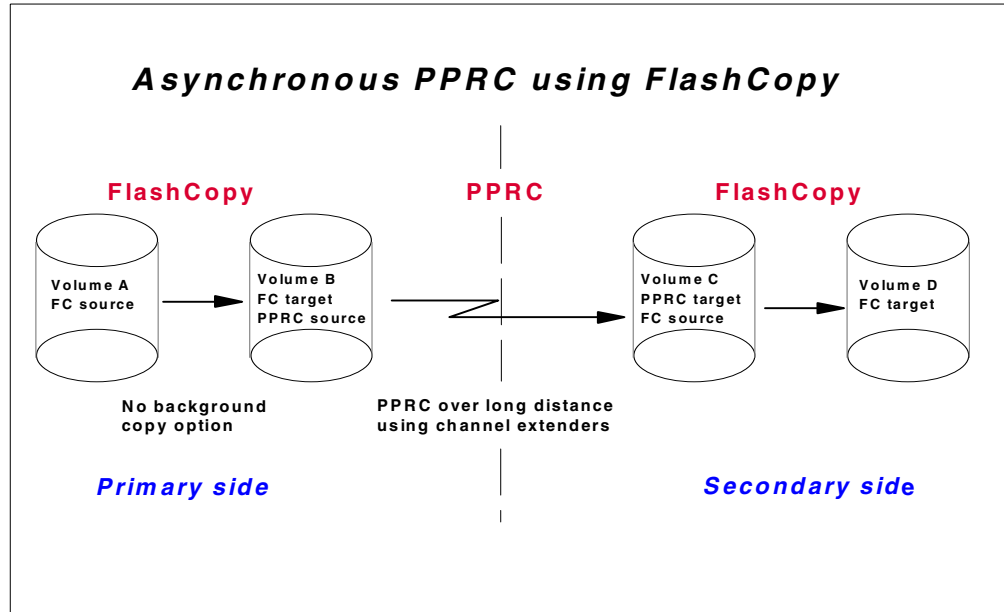


Figure 16. Asynchronous PPRC with FlashCopy

Procedure for asynchronous PPRC with FlashCopy

The procedure for asynchronous PPRC is as follows:

1. Establish a FlashCopy pair on the primary side between volume A and volume B. To minimize I/O going to the disks, select the FlashCopy NOCOPY option. Make sure the data on the primary side is in a consistent state for the short period of FlashCopy establishment.
2. Once the FlashCopy relationship is established, create a PPRC pair between volume B and volume C. Make sure to copy the entire volume between the two sites. See Figure 17 for an illustration of how the volume information should look like.

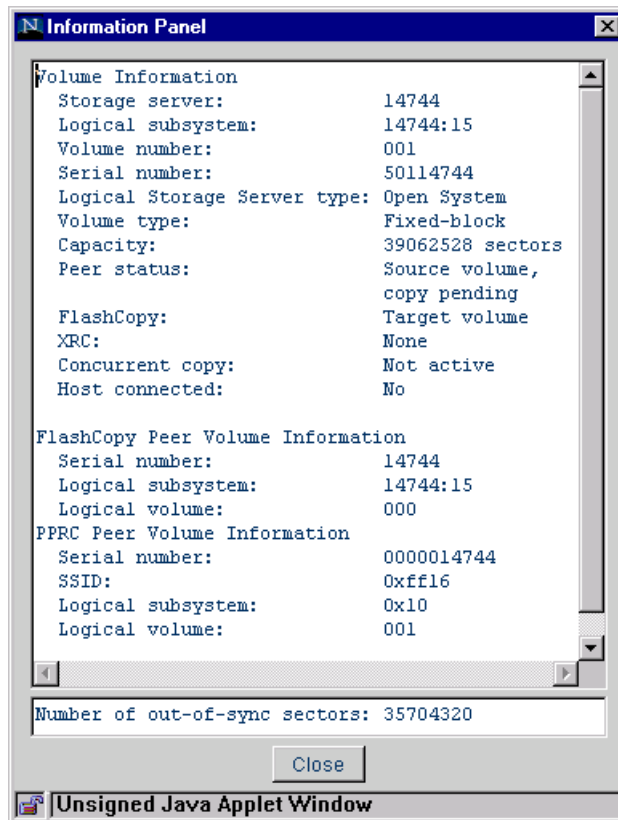


Figure 17. FlashCopy target as a PPRC primary volume

3. Wait until the PPRC pair is in full copy mode (full duplex state) to ensure that all data is transferred to the secondary side.
4. Terminate the PPRC pair between volume B and volume C.
5. Withdraw the FlashCopy pair between volume A and volume B. There is no need for the relationship anymore, as all T_0 data is transferred to the secondary side.
6. Establish a FlashCopy pair between volume C and volume D on the secondary side.
7. Repeat steps 1 to 6 when you require the data on the secondary side to be updated.

In case the primary side goes down, you can restart your applications from the data of the FlashCopy target on the secondary side and continue production from t_0 time on.

The update of the secondary site and the recovery in case of a failure with the primary side could be automated using the Command Line Interface of the ESS Copy Services. It is possible to create the FlashCopy and PPRC tasks once within the Copy Services Web Interface and to periodically execute these tasks from the host's command line (for example, within scheduled UNIX `cron` jobs).

4.10 Performance considerations

The following sections are intended to show you the considerations involved when setting up Copy Services of the Enterprise Storage Server (ESS) in order to achieve better performance.

This should help you to understand the performance impact of ESS Copy Services. As there are many different parameters that have an influence on performance, such as applications, type of workload, and configuration of the Enterprise Storage Server, the information should serve as a guideline when planning ESS Copy Services.

Please keep in mind that the general ESS performance considerations, such as volume placement or amount of storage per host adapter, still apply when planning for PPRC.

4.10.1 Optimized PPRC communication

There were certain modifications made to the ESCON protocol used for PPRC communication of the Enterprise Storage Server; in particular:

- A larger frame size, which results in less overhead during PPRC communication.
- Less handshaking between the two communicating ESSs, which makes transfer of data more efficient. The handshake was reduced from 6 down to 3 exchanges.

4.10.2 Number of ESCON paths between Enterprise Storage Servers

When a host system is sending a write request to a primary volume of a PPRC pair, an I/O complete will be returned to this host once the data is written to the source and target ESS (synchronous copy). This will have some performance impact upon the write I/O operations of the application.

Always make sure that you are using an appropriate number of ESCON paths for PPRC between the source and the target ESS. Increasing the number of the physical ESCON links will increase the maximum overall bandwidth for updating the targets. Using multiple ESCON connections for a PPRC pair (maximum of 8) will improve the response time of an I/O request from the host. Keep in mind that too few paths may result in a bottleneck. A minimum of four paths between the primary and secondary ESS are recommended.

4.10.3 Placement of the ESCON adapters used for PPRC

Distribute the ESCON adapters used for PPRC evenly across the two clusters and the host adapter bays of the ESS. This will distribute the PPRC workload over multiple busses and both clusters.

For example, if there are four ESCON adapters used for PPRC between two Enterprise Storage Servers, place one ESCON adapter in each of the host adapter bays.

4.10.4 Grouping of physical and logical paths

A physical path describes the physical ESCON connection between two Enterprise Storage Servers. A logical path is the connection used for the PPRC

copy pair, either between two volumes or two logical subsystems. There could be multiple logical connections established over a single physical connection. This will be most likely the case in a real environment.

Also, consider that multiple logical paths will share the bandwidth of the ESCON path(s) between each other. If there are critical PPRC pairs, we recommend that you separate them on dedicated physical paths so that the I/O traffic of the data copy from the primary to the secondary side will not interfere with I/O traffic of lower critical PPRC pairs.

4.10.5 Setup of the secondary ESS

For disaster recovery reasons, you may be doing PPRC between two or more different Enterprise Storage Servers. Under normal operating conditions, you always have a source (primary side) and a target (secondary side) of a PPRC pair.

One single ESS could have up to four secondary ESSs. However, the number of primary servers of a single secondary server is only limited by the number of available ECSON links. So it may be the case that different primary storage servers are connected to the same secondary ESS. In that case, the I/O traffic of multiple primaries has to be computed by a single secondary ESS.

Furthermore, it may be possible that secondary volumes from different primary storage servers are placed on the same disks within the same Array (rank). In that case, the response time of each primary storage server will increase if other primaries are doing I/O at the same time as all requests are handled simultaneously.

Therefore, when planning your Enterprise Storage Server network, keep in mind how many primary storage servers are connected to the same secondary. Distribute the I/O load evenly across the secondary storage server.

Try to distribute the volumes used for PPRC pairs of multiple primaries across all available RAID arrays within the secondary ESS.

Chapter 5. ESS Copy Services Web Interface

There are two different methods of using the ESS Copy Services in the Open Systems environment:

- A Web-based Interface
- A Java-based Command Line Interface (CLI)

In this chapter we explain how to use and set up the ESS Copy Services Web Interface. The usage of the Command Line Interface is described in Chapter 6, “ESS Copy Services Command Line Interface” on page 109.

5.1 Overview and requirements

The ESS Copy Services run inside the Enterprise Storage Server. One ESS has to be defined as the Copy Services server, and it holds all Copy Services related information. Optionally, there could be a second ESS defined to be the backup Copy Services server. On each ESS that is intended to use Copy Services, there is a Copy Services client running who communicates to the server.

Access to the Copy Services is provided through an Internet browser. Using a Web browser gives the possibility to easily control the ESS copy functionality over the network from any platform for which the browser is supported.

The ESS Copy Services require one of the following Internet browsers:

- Netscape Communicator 4.6 or above
- Microsoft Internet Explorer (MSIE) 5.0 or above

You enter the ESS Copy Services main menu by selecting the **Copy Services** button of the ESS Specialist (Figure 18). This connects the browser to the ESS that is specified as the Copy Services Server. If you have not previously selected any of the other ESS Specialist buttons, you will be prompted for the user name and the password before starting the Copy Services Web screen.

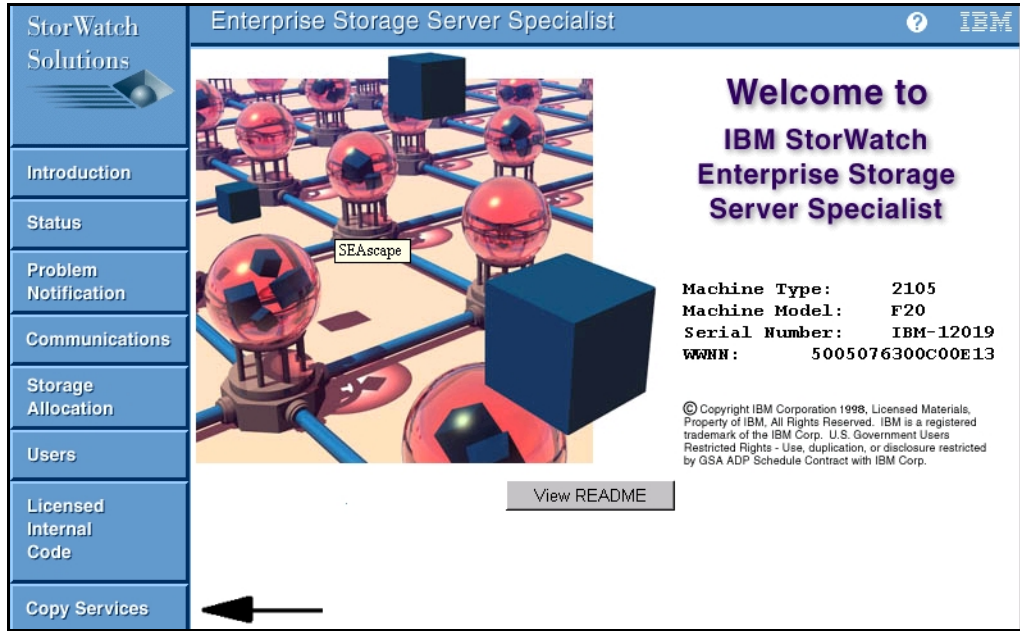


Figure 18. Main menu of the ESS Specialist

The message window shown in Figure 19 will be displayed while connecting to the Copy Services Server.

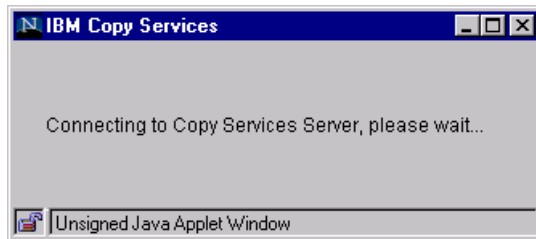


Figure 19. Copy Services start message

Once the connection to the Copy Services server is successful, the main menu of the Copy Services Web Interface will be displayed (Figure 20). From here you can access all Copy Services menus by selecting one of the buttons on the left side.



Figure 20. Main menu of the ESS Copy Services Web Interface

5.1.1 Volumes panel of the ESS Copy Services Web Interface

Volumes are defined with the ESS Specialist in order to provide a fixed storage capacity to the connected host system. They are the base components of each data copy task. The ESS assigns each volume a unique 8 digit identifier (ID). This identifier is used to address each volume within the ESS.

From the **Volumes** menu you will be able to:

- Get information and status about a volume defined in a logical subsystem (LSS) of the ESS.
- Select source and target volume for a PPRC or FlashCopy task.
- Filter the output of the volume display to a selected range.
- Search for a specific volume based on its unique volume ID.
- Establish, terminate and suspend PPRC Copy pairs and optionally save the operation as a task.
- Establish and withdraw FlashCopy pairs and optionally save the operation as a task.
- Enter the multiple selection mode for PPRC and FlashCopy.

Figure 21 shows the entry window of the **Volumes** panel. On the left side you select the source LSS and on the right side the target LSS for your copy task.

The source and target logical subsystems are specified as follows:

Device type (4 digits):ESS Serial number(5 digits):LSS number (2 digits). An example would be a logical subsystem that is addressed by 2105:12019:20.

The **Volumes** menu shows all volumes defined within the LSS. Below the volume you will find its unique serial number. The color of the volume indicates if it is used in any copy relationship (source or target), or if it is not part of a copy pair at all.

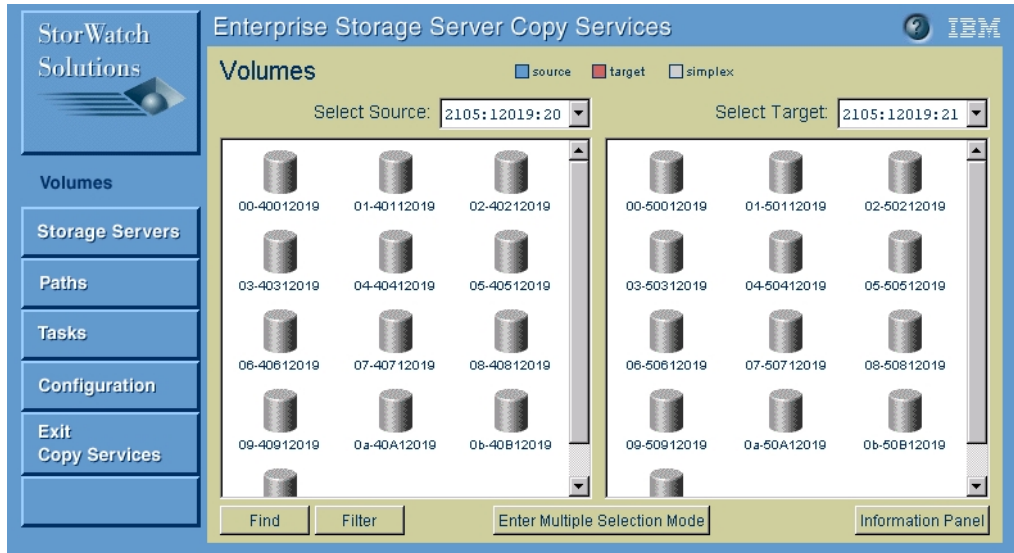


Figure 21. Source and target area of the Volumes menu

Note

You cannot display the same logical subsystem in both the source and the target area. Therefore, select two different logical subsystems as source and target, or only one LSS, in either the source or target area.

5.1.1.1 Working with volumes

You can get more detailed information about a single volume by selecting the volume and clicking the **Information Panel** button, as shown in Figure 22.

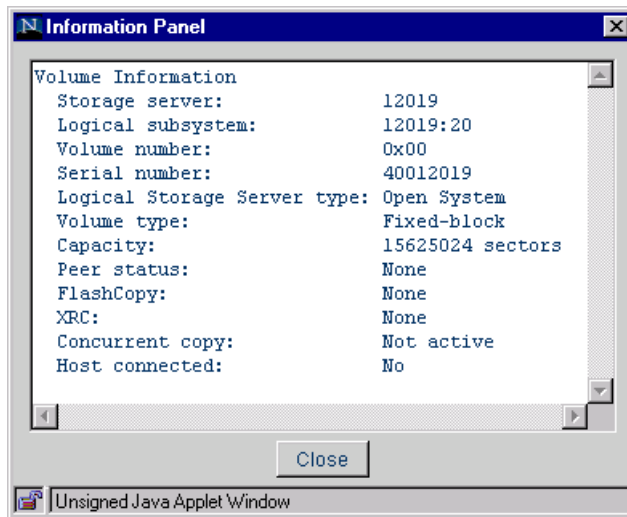


Figure 22. Volume Information window

With the **Find** button there is the possibility to search for a specific volume. The volume is specified with its 8-digit ID. See Figure 23.

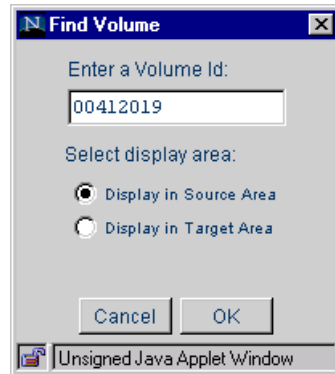


Figure 23. Find volume window

In addition, you can filter the output of the volume display to a selected range by clicking the **Filter** button and selecting the **Filter volumes** option.

In the example shown in ,Figure 24 we want to display Open Systems volumes only. The volumes should be in the simplex state; that means they are currently not in use in any Copy relationship.

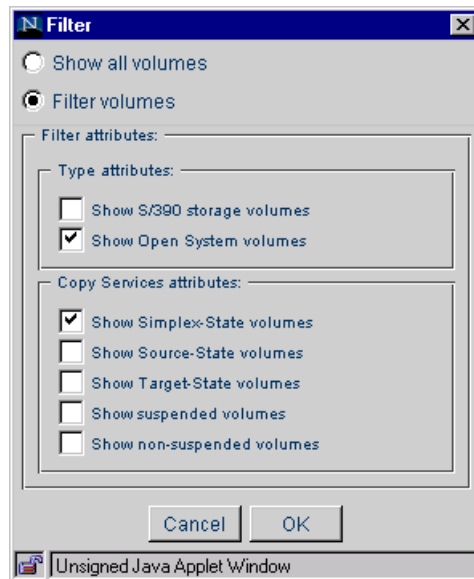


Figure 24. Filter volumes window

5.1.1.2 Using host device names

After running the `rsPrimeServer` command you have the possibility to display the host device names instead of the volumes serial number on the **Volumes** output. The `rsPrimeServer` command is part of the Copy Services Command Line Interface and have to be executed from the host that is connected to the storage of the ESS.

In the example shown in Figure 25 we have executed the `rsPrimeServer` command on a AIX system named `fastlynx`. This will show the `hdisk` numbers of all ESS LUNs connected to the server.

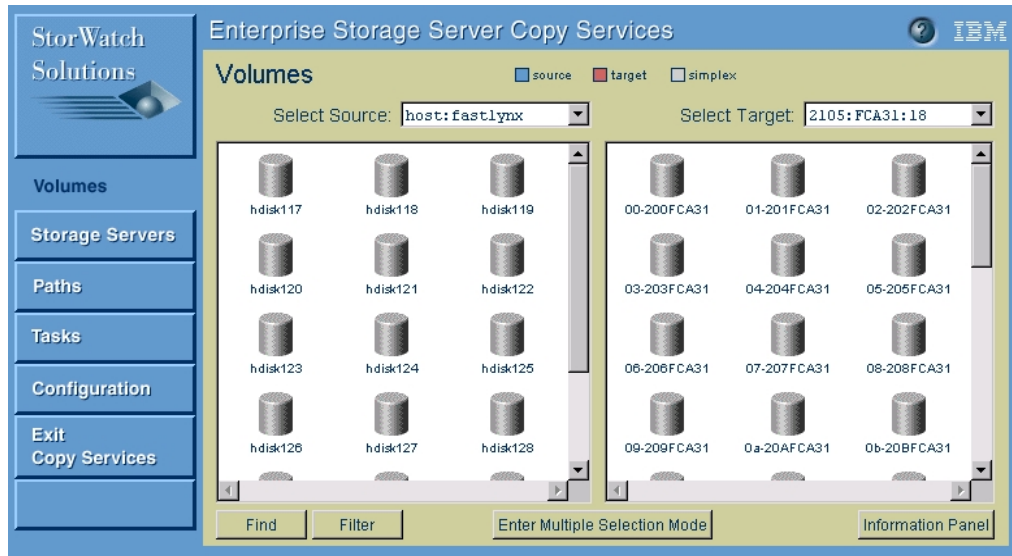


Figure 25. Volumes display for AIX after running the `rsPrimeServer` command

5.1.2 Storage Servers panel of the ESS Copy Services Web Interface

The **Storage Servers** panel displays the Enterprise Storage Servers and the logical subsystems within the storage network. The storage network includes all Enterprise Storage Servers that are configured to use the same Copy Services Server. Each of the logical subsystems is specified by the serial number of the ESS it belongs to and its 2-digit LSS number within the ESS.

With the **Storage Servers** menu, you will be able to:

- View all Enterprise Storage Servers within the storage network.
- View all logical subsystems within the storage network.
- Get information about a logical subsystem and its status.
- View and modify the copy properties of a logical subsystem.
- Filter the output to a selected range.
- Search for a specific logical subsystem based on its unique address.
- Establish, terminate, and suspend PPRC Copy pairs and optionally save them as a task.

In Figure 26 you can see the Storage Servers output for a selected ESS. The color indicates the state of the LSS, whether it contains volumes that are currently in any copy relationship (source, target, or mixed), or that are not part of a copy pair at all.

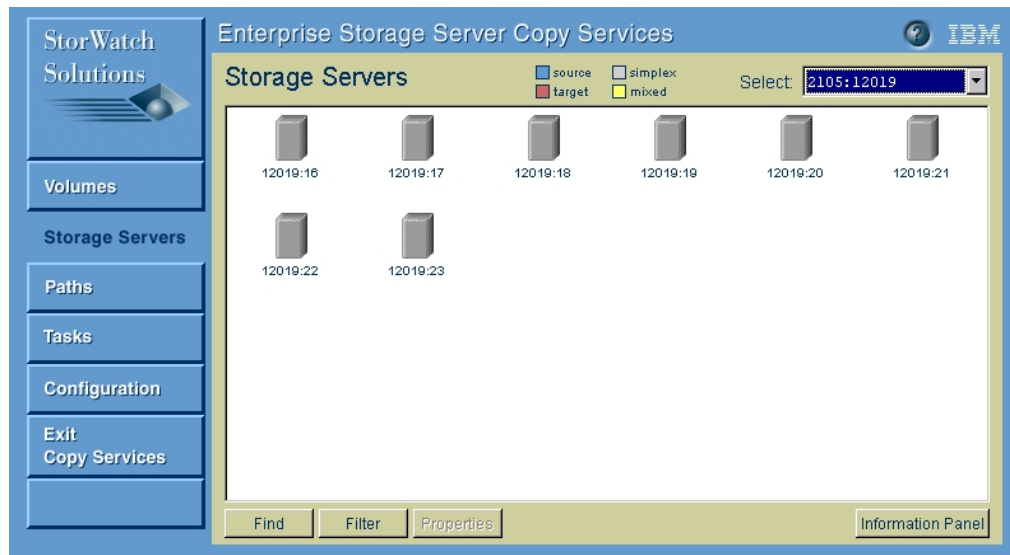


Figure 26. Storage Servers window

5.1.2.1 Properties

By selecting one LSS and clicking the **Properties** button, you can view or change the copy properties of the entire LSS (see Figure 27).

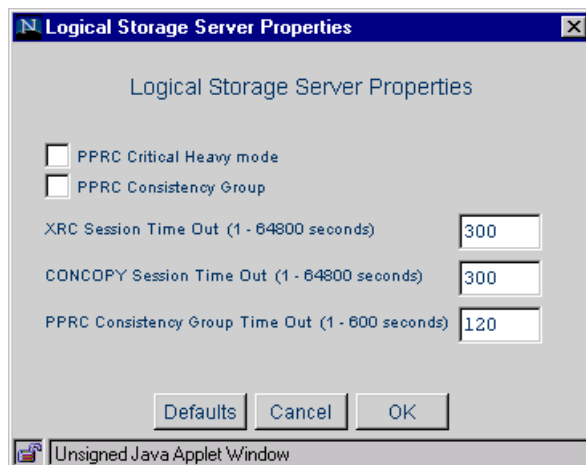


Figure 27. LSS properties window

PPRC Critical Heavy mode

This parameter works in conjunction with the **Critical Volume mode parameter** when establishing a PPRC pair. It means that updates to the pairs of volumes involved is critical to the operation of a system or database. See Table 1 on page 93 for more information.

PPRC Consistency Group

A consistency group is a set of volumes paired for PPRC. When a failure occurs on any device within such a group all I/O to volumes within the same group can be frozen by automation software. A consistency group can include multiple LSSs. This option enables that function for the entire logical subsystem.

XRC Session Time Out / CONCOPY Session Time Out

This parameter does not apply to Open Systems.

PPRC Consistency Time Out

This parameter indicates the amount of time that write I/O is withheld from the devices of a consistency group. This time out enables automation software to detect that an error has occurred and to issue commands to freeze all other members of the consistency group.

5.1.2.2 Working with logical subsystems

You can get more detailed information about a single logical subsystem by selecting the LSS and clicking the **Information Panel** button. In our example in Figure 28, we have selected LSS 16, which is specified for Open Systems. This LSS contains 13 volumes. None of these volumes are currently part of a FlashCopy or PPRC pair.

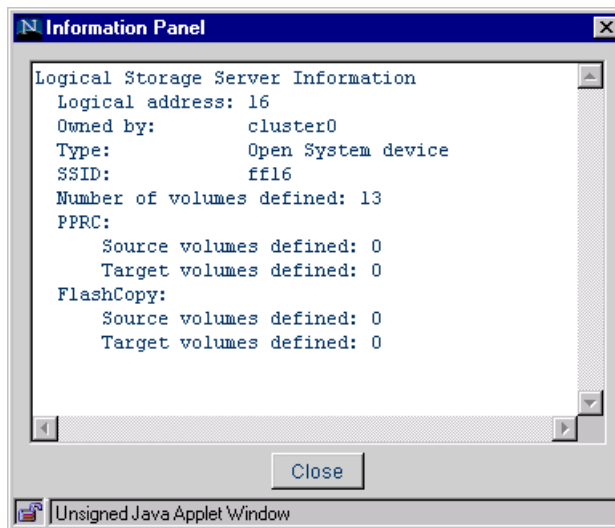


Figure 28. Information window of a logical subsystem

You can search for a specific LSS based on its address by selecting the **Find** button of the Storage Servers screen. Figure 29 shows an example where we want to find the logical subsystem 22 of the Enterprise Storage Server with the Serial Number 12019.



Figure 29. Find Storage Server

In addition, you can limit the output of the volume display to a selected range by clicking the **Filter** button and selecting the **Filter** devices option.

In our example (Figure 30) we want to display physical and logical storage servers only. In addition, we only want to display Open Systems devices that contain volumes which are currently part of a Copy relationship.

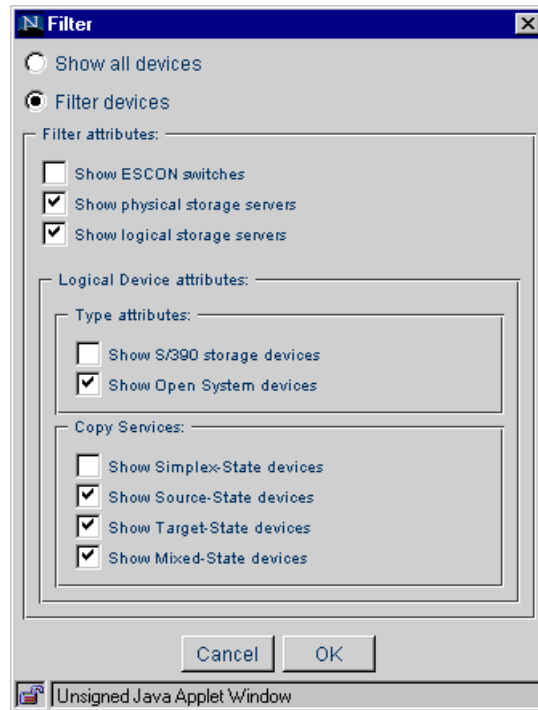


Figure 30. Filter logical subsystems volume

5.1.3 Paths panel of the ESS Copy Services Web Interface

A path is used to send data between the source and target of PPRC pairs. The physical path consists of the ESCON connection between two Enterprise Storage Servers while a logical path describes the connection of the PPRC source and targets. There could be multiple logical paths established over a single physical path.

From the **Paths** panel you will be able to:

- Establish PPRC paths.
- Remove PPRC paths.
- View information about PPRC paths.

A path is always specified between two logical subsystems within Enterprise Storage Servers. Volumes on these logical subsystem can use the paths defined to transfer PPRC data. For availability and performance reasons, we recommend that you define multiple physical paths between PPRC source and target.

Figure 31 shows the entry screen of the **Paths** panel.

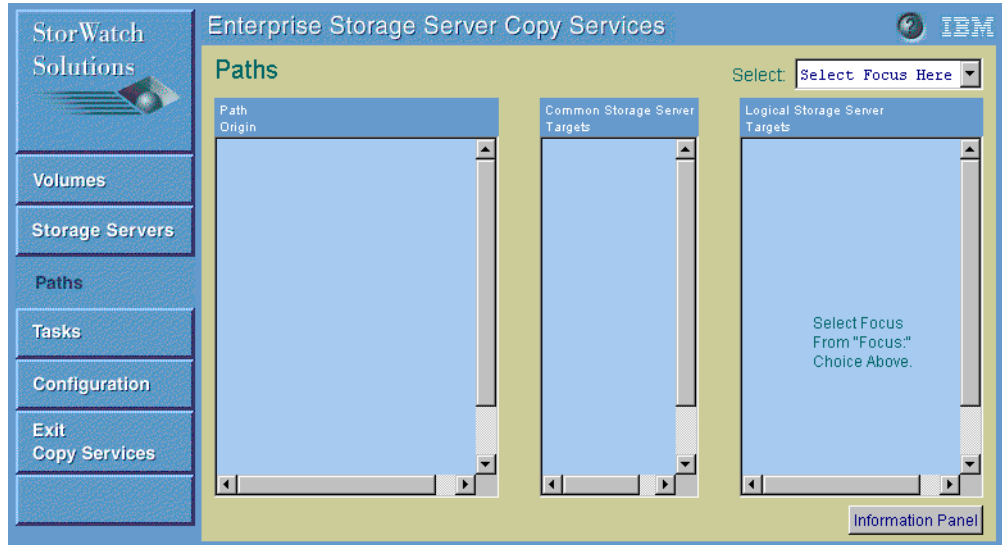


Figure 31. Entry screen of the Paths menu

On the top of the **Paths** panel, you select the source of the PPRC path, which is done by ESS serial number and LSS number. This will show all configured ESCON adapters of the selected source in the **Path Origin** area.

The ESCON adapters are specified by their system adapter ID (SAID). Figure 32 shows the SAID of all ESS ESCON adapters.

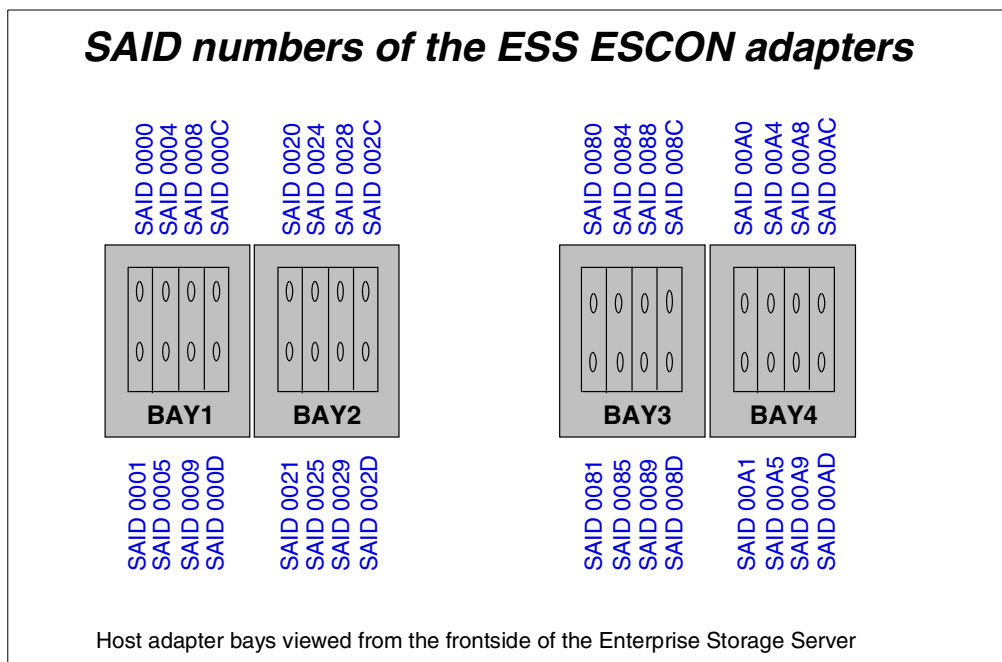


Figure 32. SAID numbers of the ESS ESCON adapters

Once an ESCON adapter is selected, all Enterprise Storage Servers that are connected to this adapter are displayed in the **Common Storage Server Targets** area. All logical subsystems that are available on a particular ESS will be listed in the **Logical Storage Server Targets** area if one of the Storage Servers is selected.

In the example shown in Figure 33 we have selected the ESS 2105:FCA24:16, that is the LSS number 16 of ESS with serial number FCA24. The path origin is the ESCON adapter with the SAID0001.

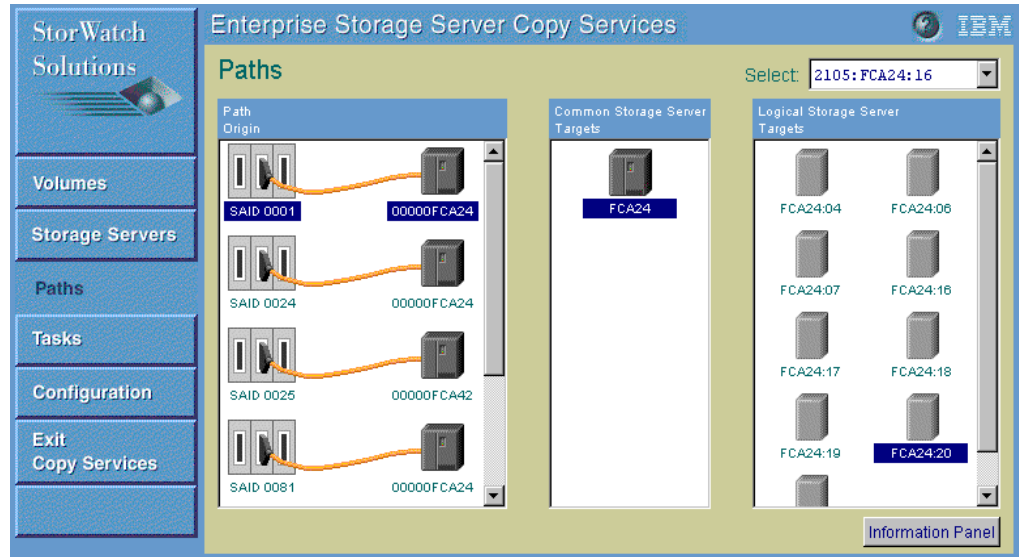


Figure 33. Example of the Paths panel

5.1.3.1 Getting path information

Once an ESCON adapter is selected, you can get more information about the paths by clicking the **Information** button at the bottom of the path menu.

The example shown in Figure 34 shows a path defined between source LSS17 and target LSS18 using the ESCON adapters SAID0005 and SAID00a1 on the same ESS.

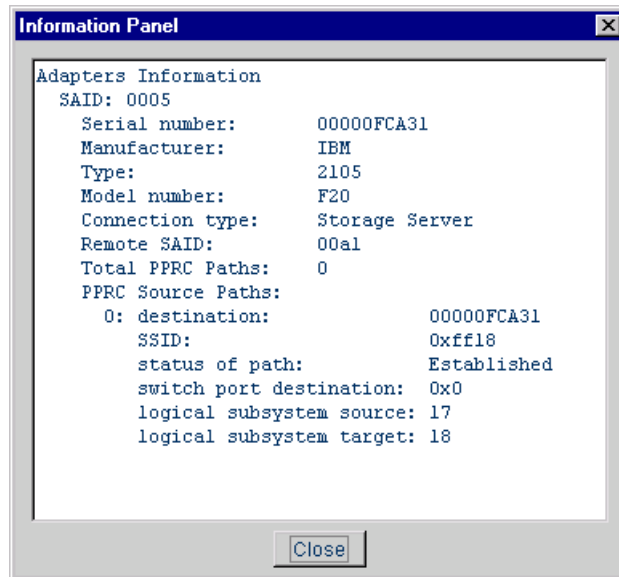


Figure 34. Path information panel

If there are paths defined on an ESCON adapter, you will find three blue asterisks right below the adapter in the Path Origin.

Figure 35 shows ESCON adapter SAID0005 without and with the defined path.

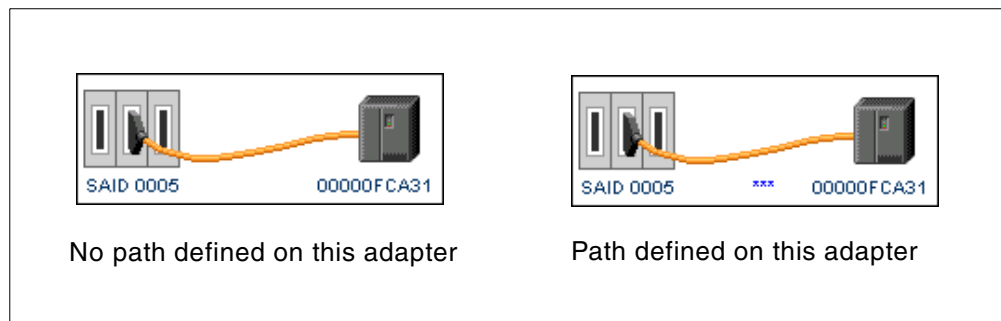


Figure 35. ESCON adapter without and with defined path

5.1.4 Tasks panel of the ESS Copy Services Web Interface

With the ESS Copy Services you have the possibility to save the setup of any data copy action within a Task. This could be any kind of FlashCopy, PPRC, and path operation.

In addition, multiple tasks can be grouped together into a single task group. This could be the case if multiple FlashCopy pairs from different logical subsystems have to be established at the same time in order to do a backup. All tasks within a task group will be processed in parallel.

With the **Tasks** menu you will be able to:

- View all specified tasks.
- Run previously created tasks.
- Display and modify properties of a task.
- Group or ungroup tasks.
- Remove tasks.
- Verify if the task has run successfully, or if it has failed.
- Display information about a task.

Figure 36 shows the Tasks panel of the ESS Copy Services. For each task, the name, a description, and the last status of execution is displayed.

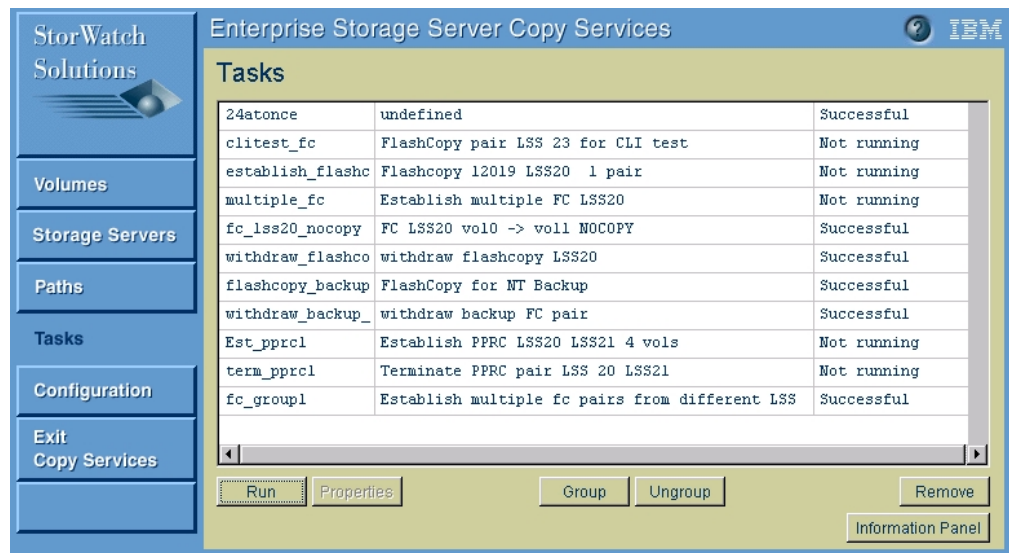


Figure 36. Tasks panel of the Copy Services Web Interface

5.1.4.1 Working with tasks

To create a group task, click the single tasks you want to group together while holding the Shift key. Once you are finished, click the **Group** button and specify the group name. It is not possible to include a group into another task group.

An example for the usage of a task group would be multiple FlashCopy pairs from different logical subsystems that need to be issued all at the same time in order to do a backup.

You can get more information about the setup of a task. Select the task and click the **Information** button at the lower right.

In our example in Figure 37, we have a grouped task named fc_group1. This group contains two single tasks named fc1 and fc2. Both of the tasks establish a FlashCopy pair within LSS 20 and LSS18.

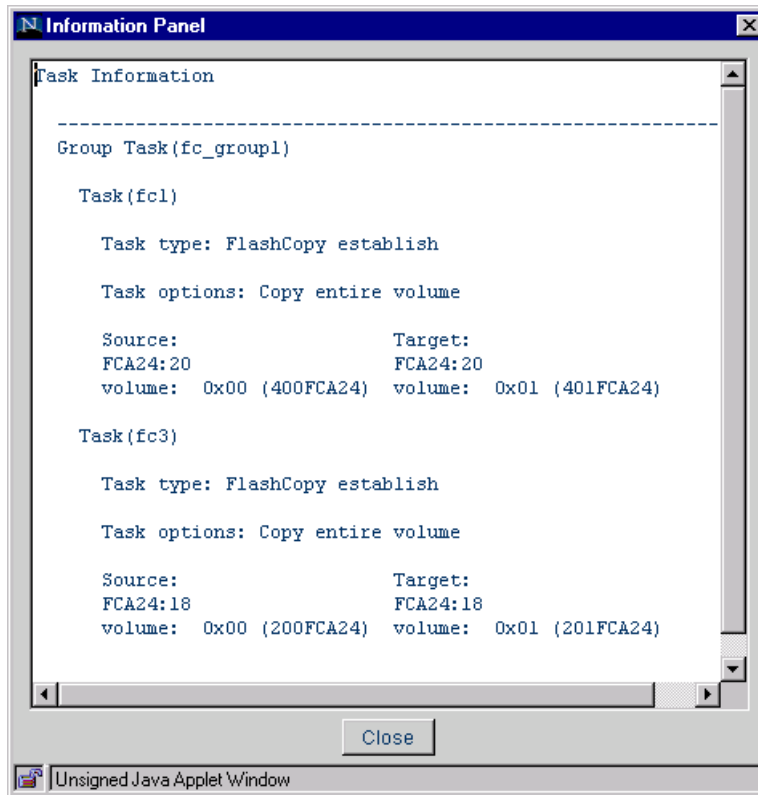


Figure 37. Task Information panel

To display or change the setup of a task, select the task you want to modify and click the **Properties button**. You can review or change any of the parameters of this task. Once you have completed your changes you have the possibility to replace the existing task, create a new additional task or delete the task if needed.

Note:

It is not possible to change the properties of a grouped task. You first have to ungroup the task, make the changes and group the single task together again when finished.

We recommend that you always use a task when setting up the ESS Copy Services. This will:

- Simplify the usage of data copies that have to be done periodically.
- Prevent user mistakes, once the task has been created correctly.

5.1.4.2 Considerations when using Copy Services CLI

A predefined task could be executed from the hosts command line with the `rsExecuteTask` command. There are two things to consider when creating tasks that are intended to be executed from the command line:

1. Use unique names for your tasks. If you have specified more than one task with the same name, the `rsExecuteTask` command will execute the task that is first found on the Copy Services server. This may not be the task you want to run.
2. Do not use any blank characters in the name of your tasks. These could not be executed through the command line.

5.1.5 Configuration panel of the ESS Copy Services Web Interface

With the **Configuration** menu you will be able to:

- Identify the Copy Services Server.
- View the date of the last configuration update.
- Send a summary of the ESS configuration, via the **Email** button.
- View and modify the Email address of the summary recipient.
- View and clear the Copy Services log file.
- Enable and disable password protection for the host commands.
- Administrate users for the host command authorization.

Figure 38 shows the **Configuration** panel.

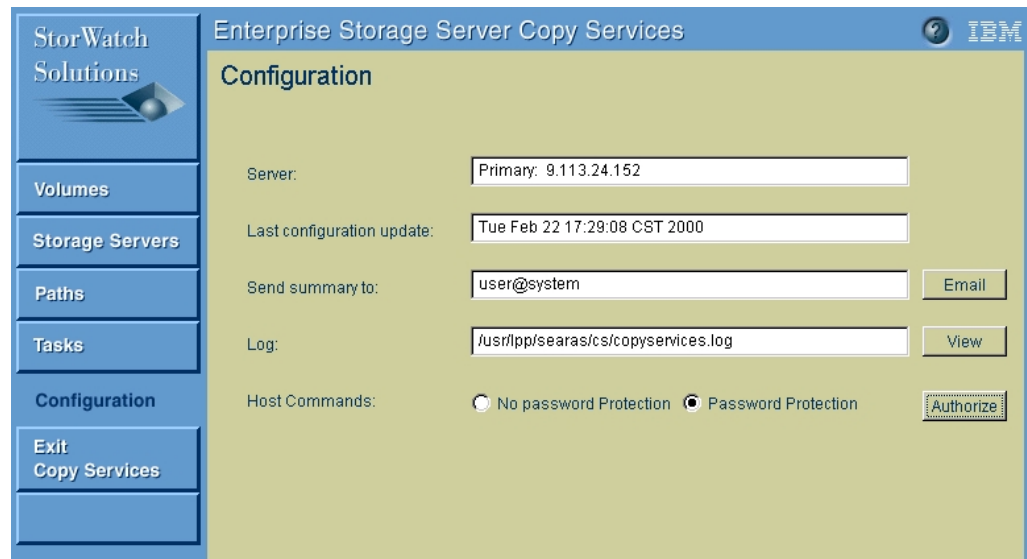


Figure 38. Configuration panel of the ESS Copy Services Web Interface

If you click the **Email** button, a summary of the ESS configuration and status is sent to the address specified at the configuration panel. An example is shown in Figure 39.

```
Enterprise Storage Server Copy Services
Server Configuration & Status
Tue Mar 07 11:10:27 PST 2000

Physical Storage Server: 12019
  Logical Storage Server 16 (SSID=0xff16)
    000 00012019 nolabel simplex, notFlashCopy, not_suspended
    001 00112019 nolabel simplex, notFlashCopy, not_suspended
    002 00212019 nolabel simplex, notFlashCopy, not_suspended
    ...
```

Figure 39. Example — ESS configuration and status summary

The log of the Copy Services provides useful information. If you select the **View** button, a new window with the content of the log file is displayed (see Figure 40).

From the log, you can get various kinds of information, such as the execution status of tasks, or the time needed to complete a Copy Services operation. If you experience any problems in performing a copy operation, we recommend that you always check the log file.

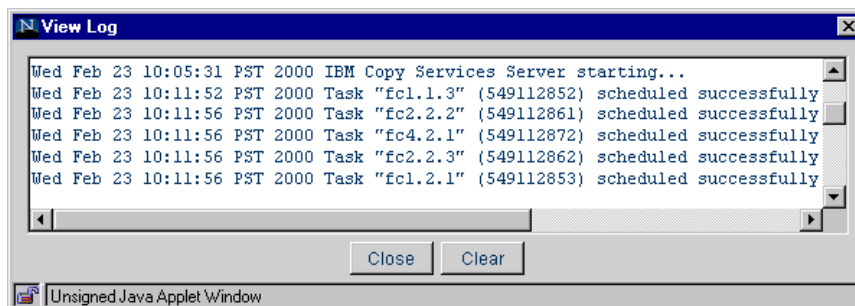


Figure 40. Log file of the Copy Services

Once the entries in the log file are not of interest anymore, you can clear the entire log by clicking the **Clear** button.

5.1.5.1 Command Line Interface authorization

From the Configuration panel you also select whether you want to authorize the usage of the Command Line Interface (CLI) or not. Per the default, **Password protection** is enabled, which means you have to specify a user name and password at the time you invoke a host command. If **No Password protection** is enabled, every user is able to work with the Copy Services from the command line.

If you want to add or remove users for the Copy Services CLI, select the **Authorize** button. When adding a new user, you have to specify their name and password (see Figure 41).

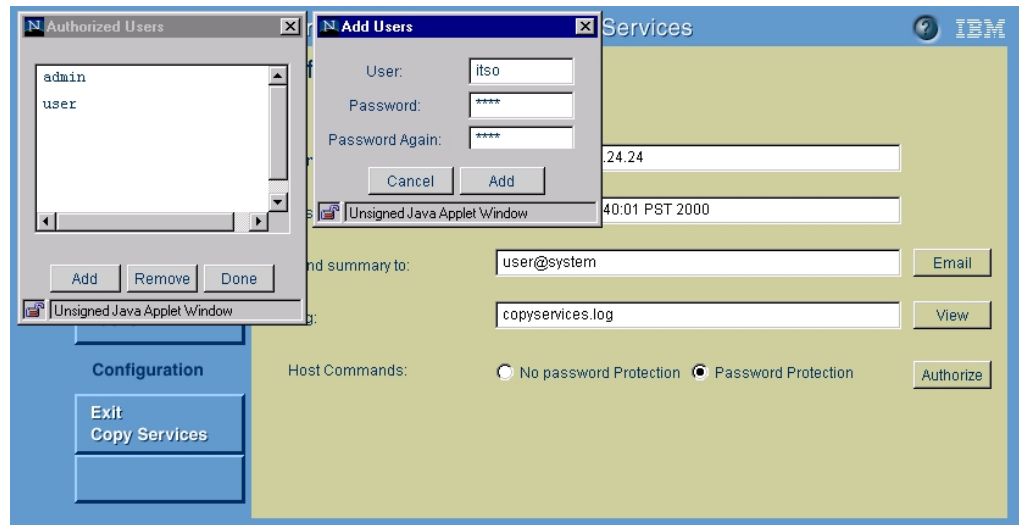


Figure 41. Adding a user for the Copy Services Command Line Interface

The usage of the Command Line Interface is described in detail in Chapter 6, “ESS Copy Services Command Line Interface” on page 109.

5.1.6 Exiting from the ESS Copy Services Web Interface

To exit from the ESS Copy Services Web Interface, select the **Exit Copy Services** button. This will cause the current browser window to return to the main ESS Specialist screen.

5.2 Implementing FlashCopy with ESS Copy Services Web Interface

In this section, we explain how to set up FlashCopy using the ESS Copy Services Web Interface.

Please make sure you are aware of the requirements of the FlashCopy functionality:

- The source and target volume have to be in the same LSS.
- The target volume has to be the same size as the source volume or larger.
- A volume can be only in one FlashCopy relationship at a time.

Note:

We recommend that you demount the target volume from all hosts systems before performing a FlashCopy. Be aware that the FlashCopy process is a destructive operation to the target and will overwrite the data on the target volume.

There are two different ways of establishing a FlashCopy pair:

- From the **Volumes** panel
- From the **Tasks** panel (once a task for FlashCopy is created)

5.2.1 Establishing a FlashCopy pair

Use the **Volumes** panel to establish a FlashCopy pair. Select the LSS within which you want to perform the FlashCopy. This can be either done in the source or target area of the volumes panel.

You always need to have two components to establish a FlashCopy pair: a source and a target. With a left-click you select the source volume and with a right-click the target. If you have selected the wrong source or target volume, just left-click on the correct source volume again.

Once you have selected the source and the target you do a second right-click on the target to bring up the Task Wizard (Figure 42). Select the **Establish FlashCopy pair** option and click **Next**.

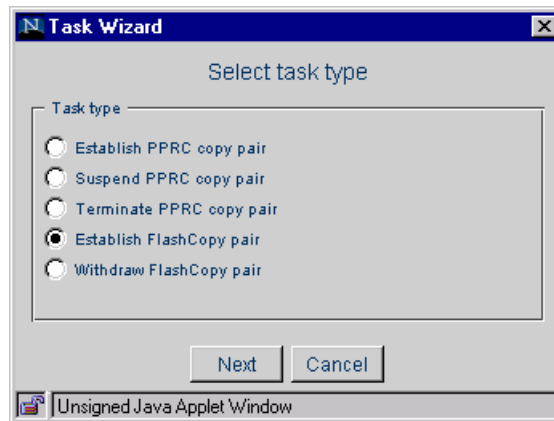


Figure 42. Task Wizard window

Within the next window you can specify the copy options of the FlashCopy pair (Figure 43).

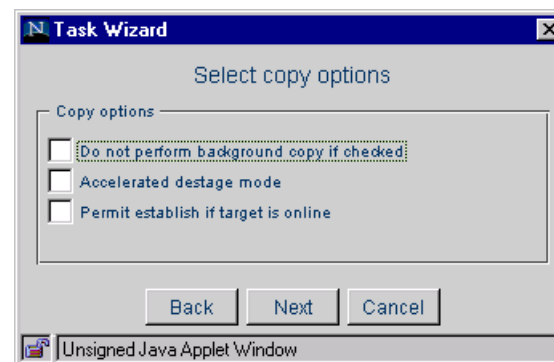


Figure 43. Select copy options window

Do not perform background copy if checked

If this option is checked, only the tracks that are modified on the source volume are copied to the target volume. The relationship between source and target volume remains forever and has to be broken manually. By default, this option is not selected and all data is copied from the source to the target volume of a FlashCopy pair. Once all data is copied, this relationship ends automatically.

Accelerate destage mode

If this option is checked, I/O of the FlashCopy process gets a higher priority than other I/O requests at the same time. Therefore the data from the FlashCopy process that is staged into cache will be destaged to the disk sooner than with the normal destage algorithm.

Permit if target is online

If this option is checked, the FlashCopy will be performed even if the target volume is in use of an operating system. This is the case, for example, if a AIX volume group is active on the target.

From the next window you can either **Save**, **Run**, or **Cancel** the copy task, shown in Figure 44. Once a task is saved, it can be executed from the Task panel at any time. Optionally, a name and description of the task can be specified. Even if you do not want to save the task you have created, we recommend that you specify a name and description. This will help with the interpretation of the Copy Services log file later on. An example would be to retrieve the execution time of the background copy of a FlashCopy pair.

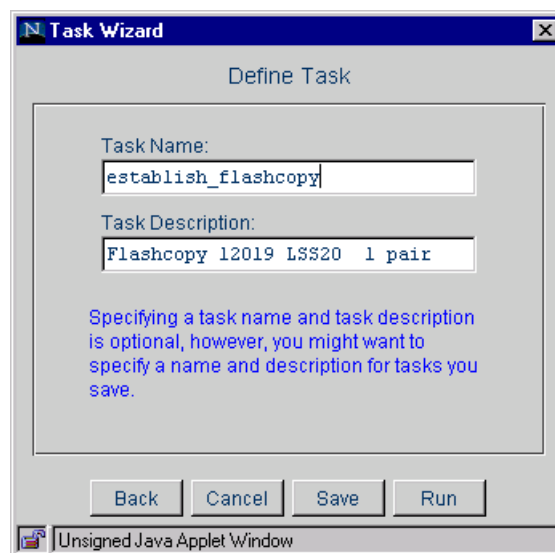


Figure 44. Define task window

If a FlashCopy is issued, a bitmap is created for the data copy from the source to the target. The time to establish the FlashCopy relationship is only a few seconds. After this period the source is immediately available to the host system, and data will be copied from the source to the target in the background.

Once a FlashCopy is started the display of the source and target volume from the Volumes panel changes. Two triangles within the source and target volume will be displayed. The color of the triangles defines whether it is a source or a target. The legend of the color is shown at the top of the volumes panel.

During the short period where the copy pair is created, only one of the triangles is filled. Once the relationship has been established successfully, both of the triangles will be solid-colored. This is illustrated in Figure 45.

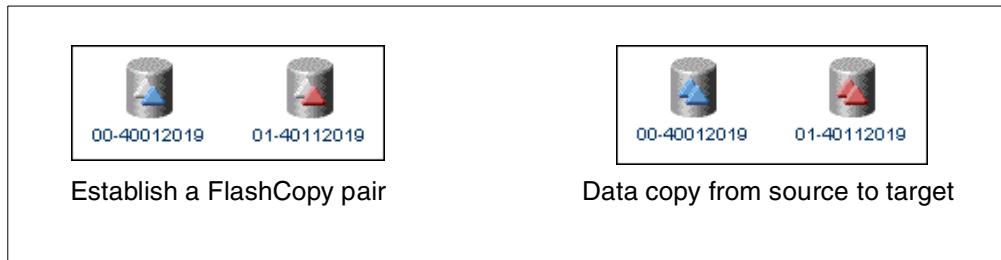


Figure 45. FlashCopy Volume display

5.2.2 Getting information about a FlashCopy pair

By selecting one of the volumes of a FlashCopy pair and clicking the **Information** button, you get information about this particular pair. If you have selected the source volume, you will also see how many tracks still have to be copied to the target volume (Figure 46).

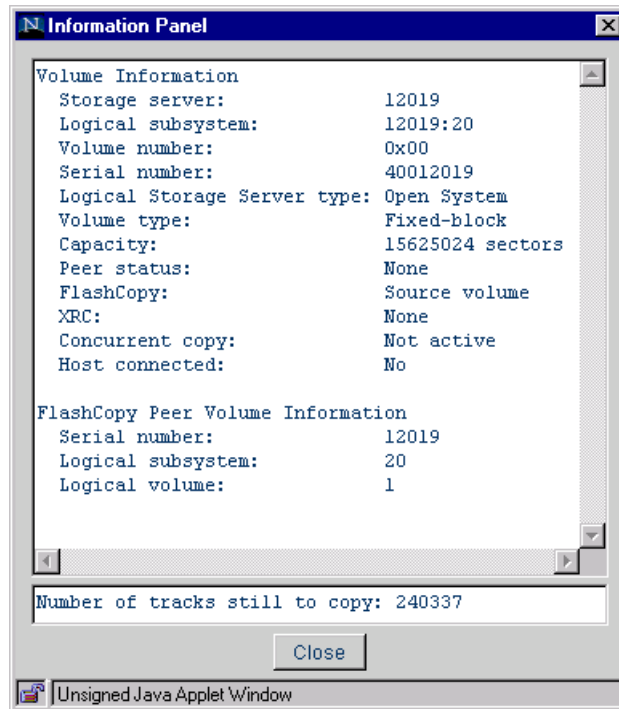


Figure 46. Information panel of a FlashCopy source

5.2.3 Withdrawing a FlashCopy pair

In the following cases, you need to withdraw a FlashCopy pair:

- If a FlashCopy pair is not needed anymore, but it has not yet finished the background copy.
- If a FlashCopy pair that was created with the NOCOPY option is not needed anymore.

To withdraw a FlashCopy pair, select either the source or the target volume with the left-click and start the Task Wizard by right-clicking one of the volumes of FlashCopy pair. This will bring up the window shown in Figure 47.

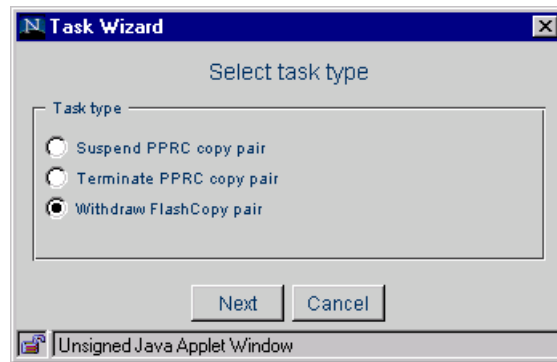


Figure 47. Withdraw a FlashCopy pair

Select the Withdraw FlashCopy pair option and decide whether to Save, Run, or Cancel the task (Figure 48).

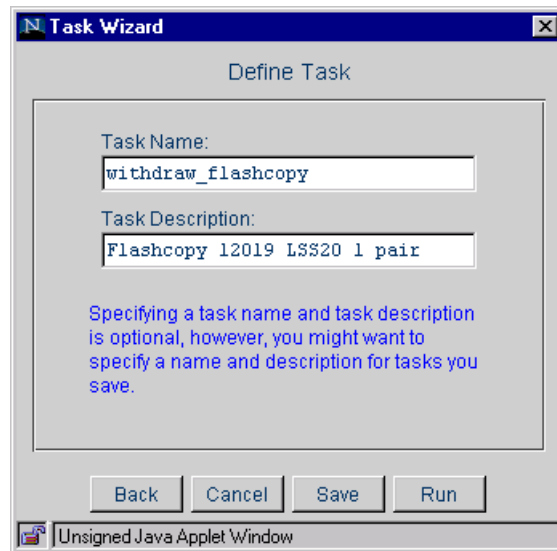


Figure 48. Define task window

5.2.4 Selecting multiple volumes for a FlashCopy task

In some cases you may want to establish or withdraw multiple FlashCopy pairs at the same time. Therefore you have the possibility to define multiple FlashCopy pairs within the volume panel. Click the **Enter multiple selection mode** button at the bottom of the volumes panel. Select one pair at a time beginning with a left-click for the source and a right-click for the target. Once you are finished with the selection, right-click again on the last target volume, which will start the Task Wizard. Continue setting up the FlashCopy task as described in the previous section.

In the example shown in Figure 49 we have created 6 FlashCopy pairs within the same LSS. After running the task for these pairs, the copy process for all pairs is started at the same time.

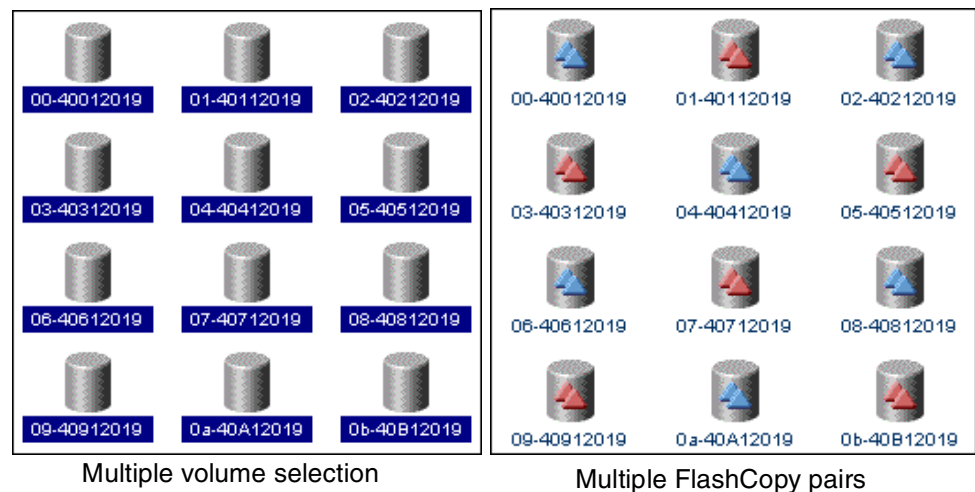


Figure 49. Multiple volume selection and FlashCopy pairs

If you have made a mistake during the volume selection, you have to exit the multiple selection mode by clicking **Exit multiple selection mode** and enter the multiple selection mode again afterwards.

Note:

The multiple selection mode is limited to one LSS within the source and target area. Once you have entered this mode, the drop-down menu of the selection area will be disabled.

5.2.5 Configuration tips

If you are creating a FlashCopy task that involves multiple source and target volumes, there is a quick way to create the task to withdraw the FlashCopy pairs. See Figure 50.

At the Tasks panel:

1. Click the task you created that establishes the FlashCopy relationships.
2. Click the **Properties** button.

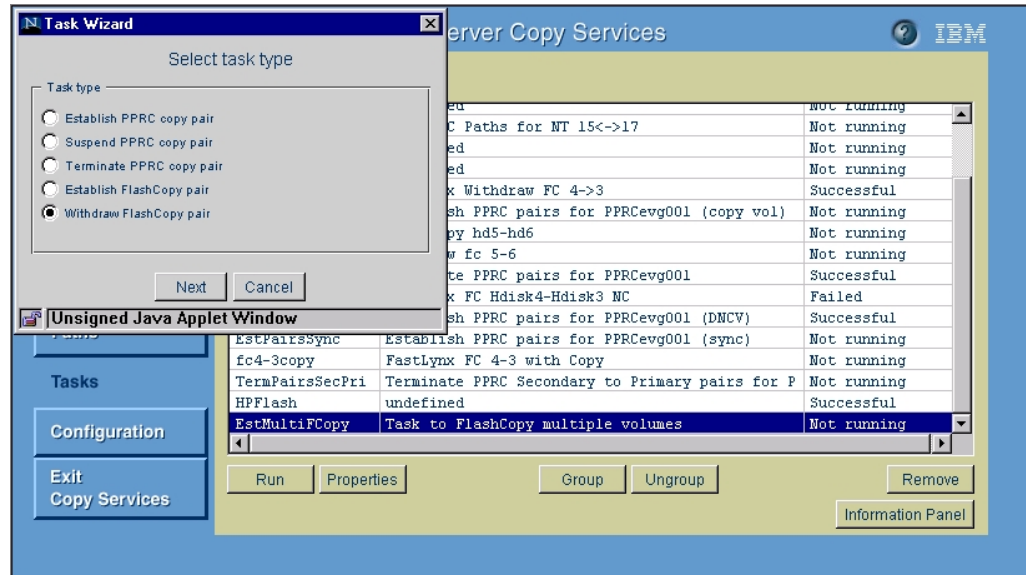


Figure 50. A quick way to withdraw FlashCopy pairs

3. Click the **Withdraw FlashCopy Pair** button, then click **Next**. See Figure 51.

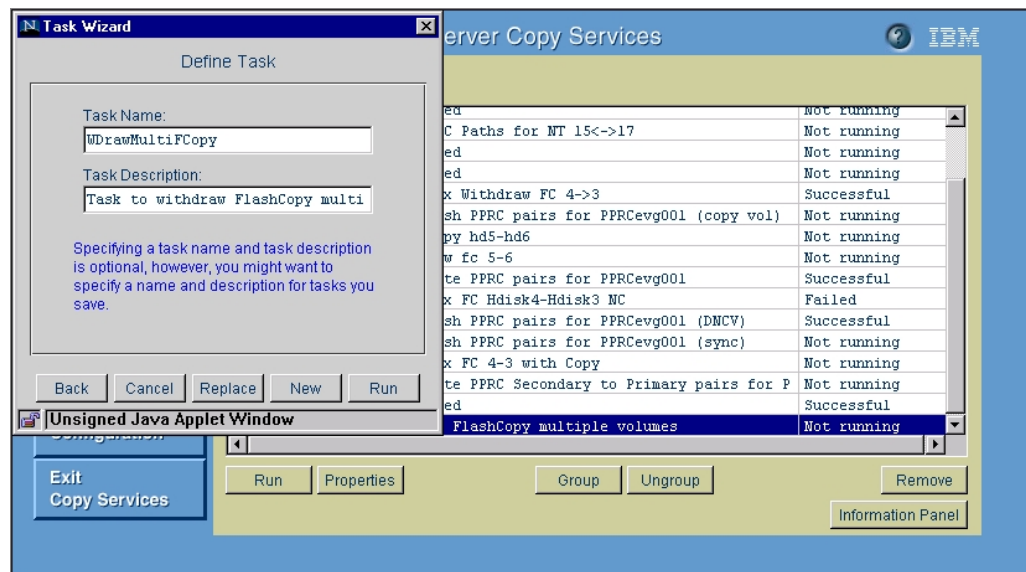


Figure 51. Withdraw FlashCopy Pair

4. Type the new Task Name and Task Description, then click **New**. See Figure 52.

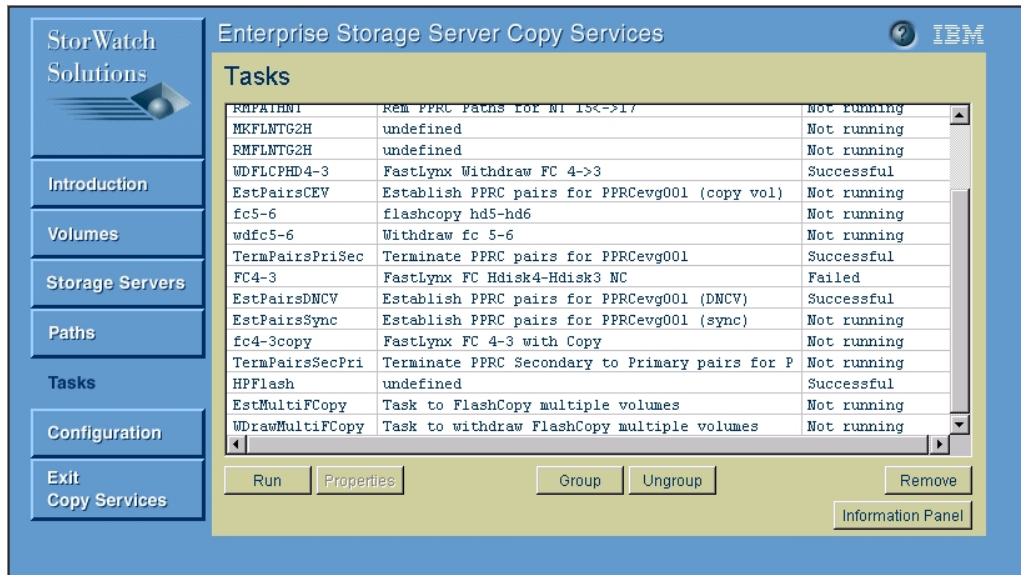


Figure 52. Naming the task to withdraw

5. Now you have a task to withdraw the FlashCopy Pairs.

5.3 Implementing PPRC with the ESS Copy Services Web Interface

In this section, we explain how to set up PPRC using the ESS Copy Services Web Interface. In general, there are two steps needed to successfully establish PPRC:

- Setting up paths between PPRC source and target.
- Establishing the PPRC pairs, either single volumes or entire logical subsystems.

Please make sure you are aware of the requirements of the PPRC functionality:

- Paths for PPRC must be available and need to be defined first.
- All PPRC ESCON links are unidirectional.
- The target volume must be the same size as the source, or larger.
- One primary ESS can have up to four secondary ESSs.

Note:

You must demount the target volume from all hosts systems before establishing a PPRC pair. Be aware that the PPRC process is a destructive operation to the target and will overwrite the data on the target volume.

There are three different ways of establishing a PPRC pair:

- From the **Volumes** panel (based on volumes)
- From the **Storage Servers** panel (based on entire logical subsystems)
- From the **Tasks** panel (once a task for PPRC is created)

5.3.1 Setting up paths for PPRC

Before you can establish any PPRC pairs, you first have to set up the paths between the source and the targets. The paths are needed for communication between PPRC pairs and to copy data from the source to the target.

Note:

In our example we have used only one ESS to set up PPRC. That is possible, as the ESS contains both source and target volumes at the same time. However, for high availability and disaster recovery configurations, two or even more Enterprise Storage Servers are required.

Use the **Paths** panel of the ESS Copy Services Web Interface to set up paths for PPRC (Figure 53).

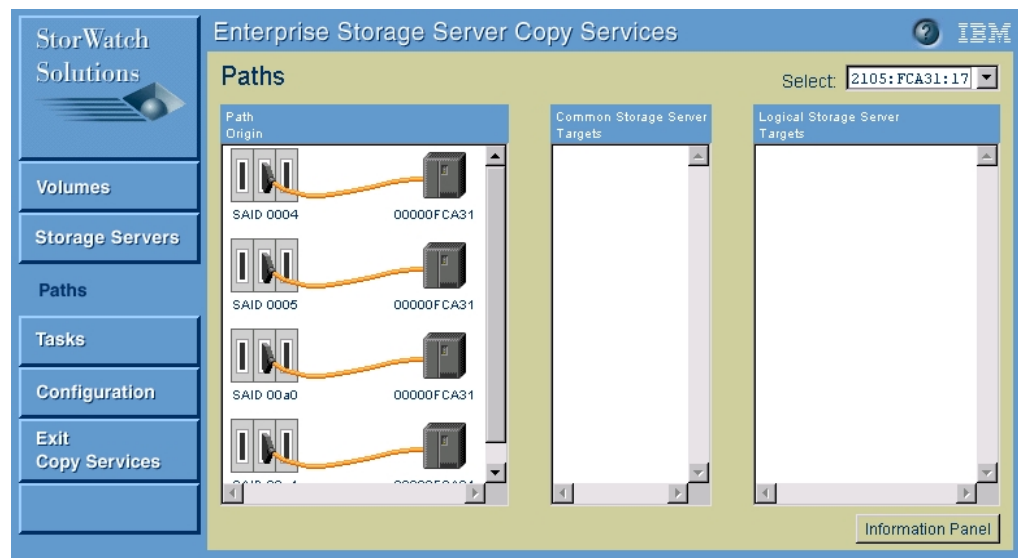


Figure 53. Paths window of the Copy Services Web Interface

Select the source of the PPRC relationship. This is done with the drop-down menu of the Select box. All available ESCON adapters for the source will automatically be displayed in the **Path Origin** area (see Figure 54).

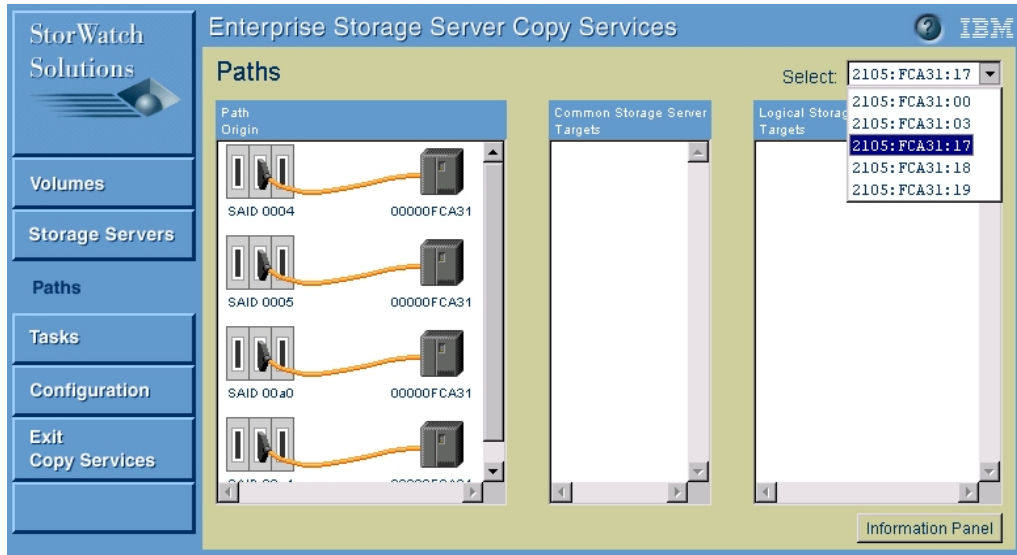


Figure 54. Setup PPRC: Source selection

Next, select the ESCON adapters you want to use for the PPRC which will automatically show the Enterprise Storage Servers that are connected to this ESCON adapter. Select one ESCON adapter with a left-click. Multiple adapters could be selected with a right-click after the first ESCON adapter was selected. If you have chosen the wrong adapters, just left-click the correct ESCON adapter again to delete the selection.

The Enterprise Storage Servers that are connected to the adapters will be automatically displayed in the **Common Storage Server Target** area.

Figure 55 of our example shows that we have selected the ESCON adapter with system adapter ID SAID0005.

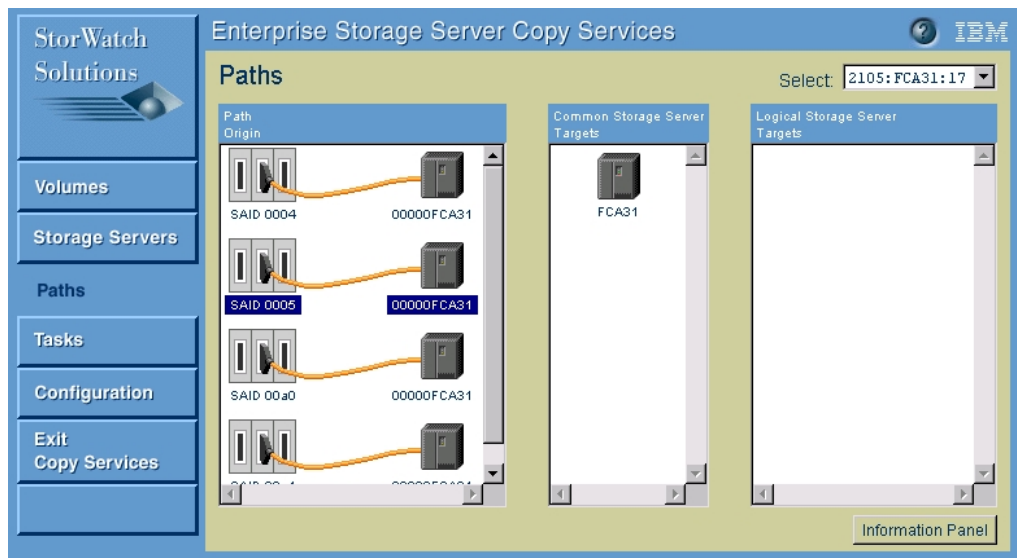


Figure 55. Set up PPRC: ESCON adapter selection

Next, you left-click the target Storage Server. All logical subsystems available on the target ESS will be displayed in the **Logical Storage Server Targets** area (see Figure 56).

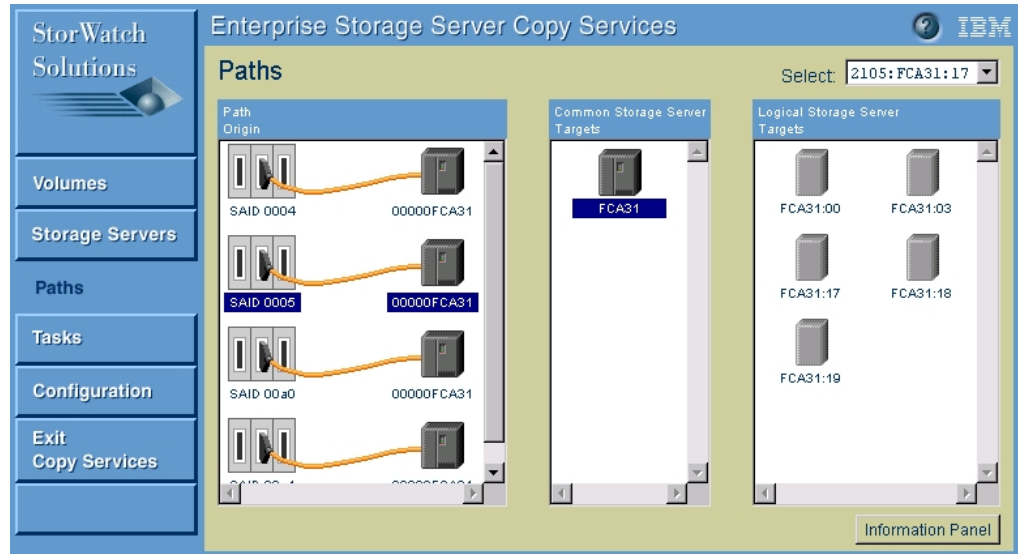


Figure 56. Set up PPRC: Target ESS selection

Within the Logical Storage Server targets area, select the target LSS of your PPRC path. Select the target LSS with a left-click. Multiple LSS could be selected with a right-click after the first target LSS was selected. If you have chosen the wrong target, just left-click the correct LSS again to delete the selection.

In the example shown in Figure 57 we have selected LSS 18 to be the target logical subsystem.

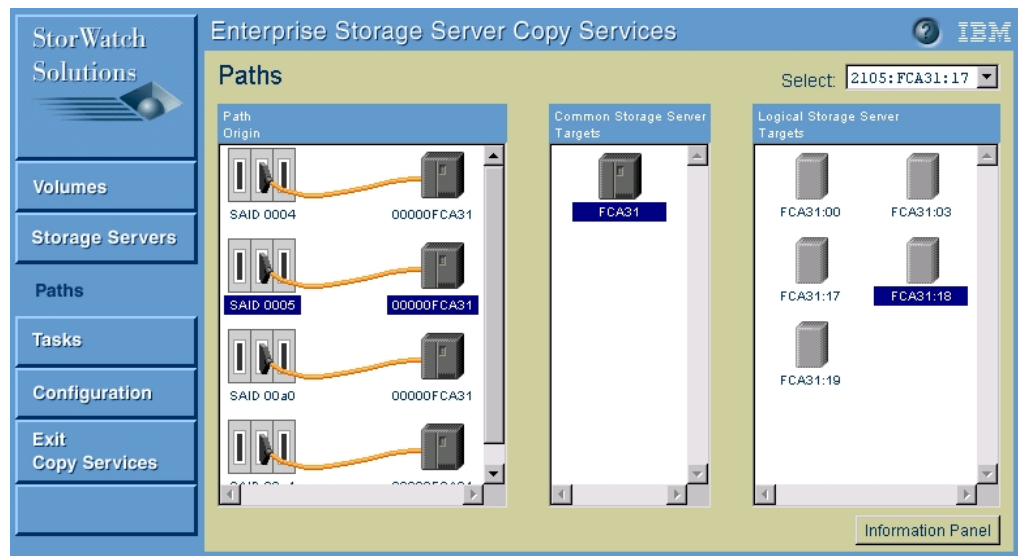


Figure 57. Set up PPRC: Select target LSS

Once the target and source of the PPRC path have been selected, right-click one of the highlighted target LSSs to bring up the Task Wizard (Figure 58). Select the **Establish Path** option and click **Next**.

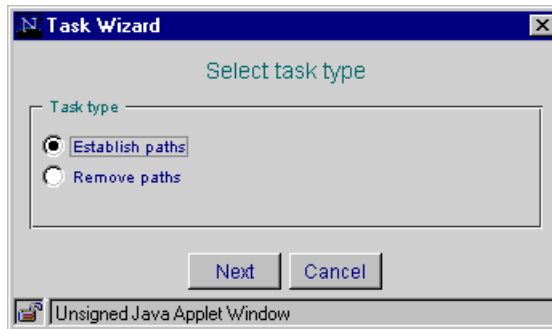


Figure 58. Establish PPRC path

Within the next window you can specify the path options (Figure 59).

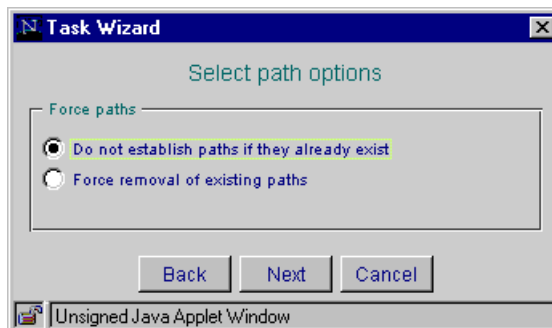


Figure 59. Establish PPRC path options

Do not establish paths if they already exist

If this option is checked and there is already a path defined from the source to the target, the operation of establishing the path will not be executed.

Force removal of existing paths

If this option is checked and there are already paths defined between the selected source and the target, these paths will be removed prior to establish the new paths.

From the next window, you can either **Save**, **Run**, or **Cancel** the path task, shown in Figure 60. Once a task is saved, it can be executed from the Task panel at any time. Optionally, a name and description of the task can be specified. Even if you do not want to save the task you have created, we recommend that you specify a name and description. This will help with the interpretation of the Copy Services log file later on. An example would be to retrieve the execution time needed to establish a PPRC pair.

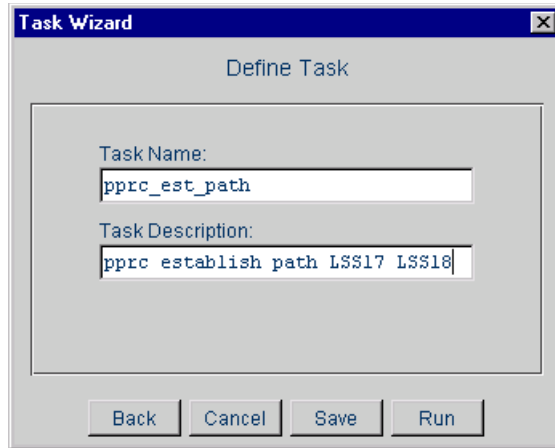


Figure 60. Define task window

Once the path has been successfully established, you will see three blue asterisks directly below the ESCON adapter. This is shown in Figure 61.

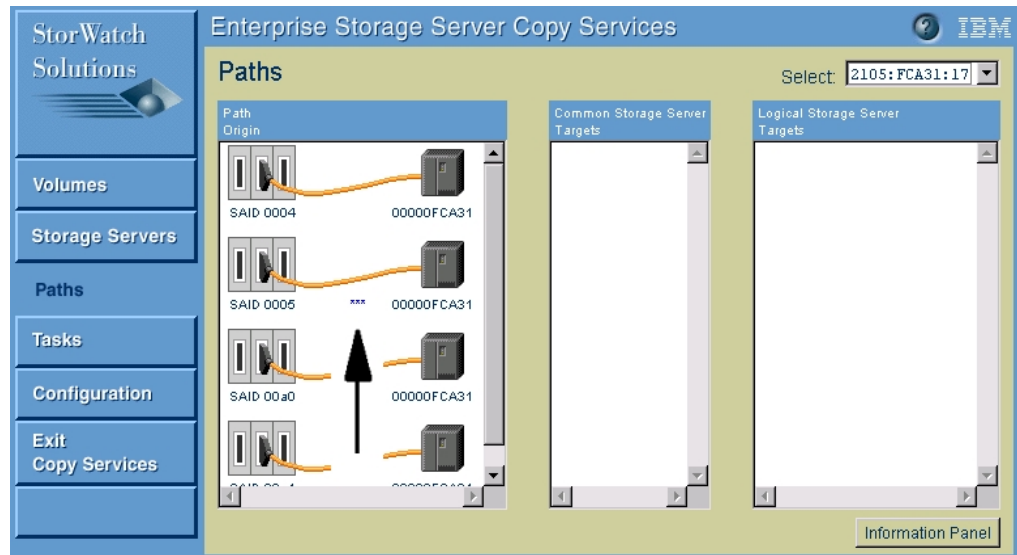


Figure 61. Path successfully established (SAID0005)

To get the path information for an ESCON adapter, select the adapter and click the **Information Panel**. A window will be displayed showing all path information for the selected adapter. Figure 62 shows the path we have created between LSS 17 and LSS 18.

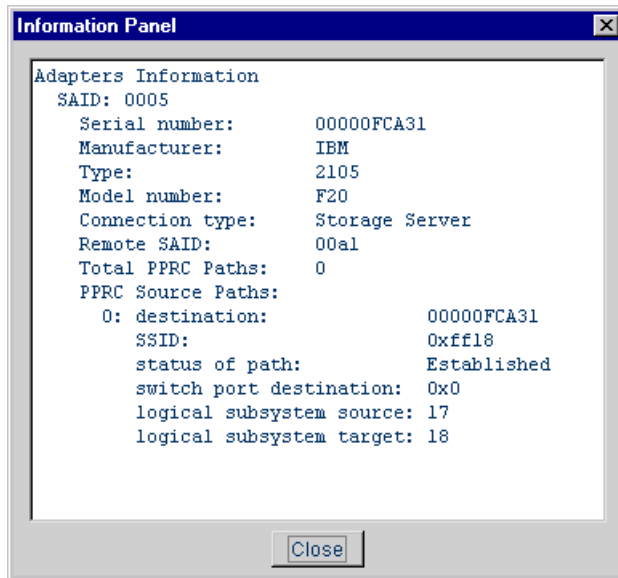


Figure 62. Path information window

5.3.2 Establishing PPRC pairs

Use the **Volumes** panel to establish PPRC pairs. On the left side you select the source LSS and on the right side the target LSS. This is done using the drop-down menu at the top of the **Volumes** menu.

The source and target logical subsystems are specified as follows: Device type (4 digits):ESS Serial number(5 digits):LSS number (2 digits).

In the example shown in Figure 63, we have selected 2105:FCA31:17 as source and 2105:FCA31:18 as target of our PPRC pair.

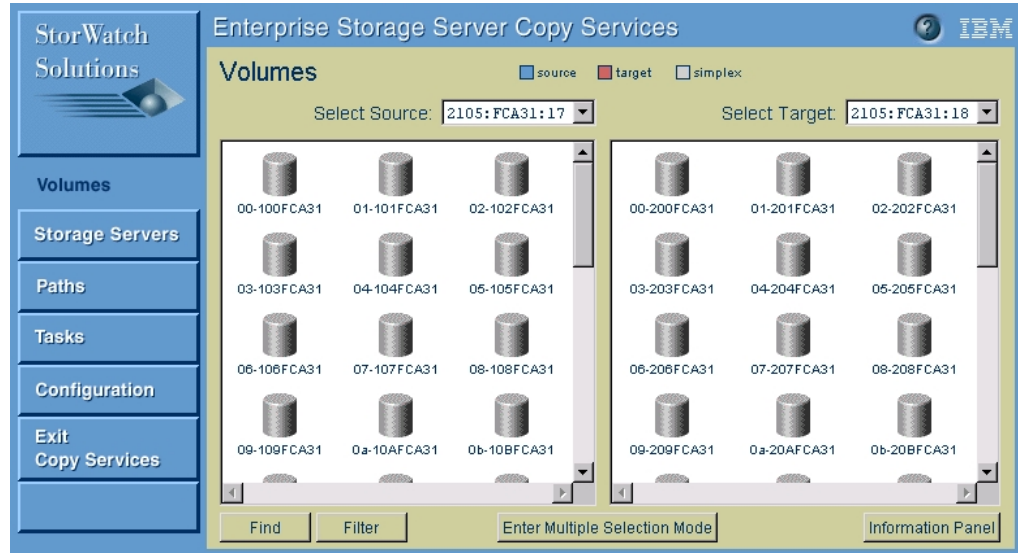


Figure 63. Select PPRC source and target LSS

You always need to have two components to establish a PPRC pair, a source and a target. With a left-click you select the source volume and with a right-click the target. If you have selected the wrong source or target volume just left-click on the correct source volume again to clear the selection (Figure 64).

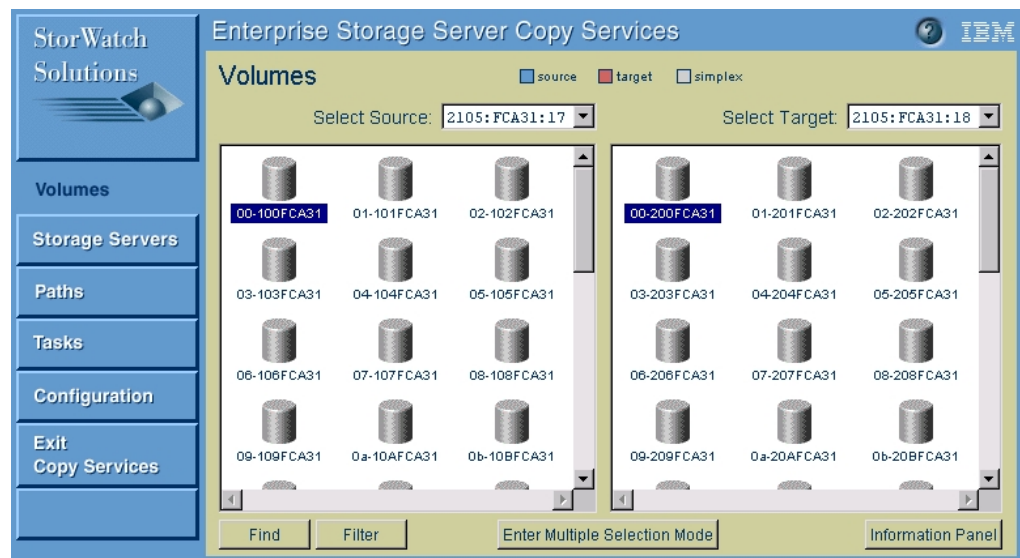


Figure 64. Selecting PPRC source and target volume

Once you have selected the source and the target, you do a second right-click the target to bring up the Task Wizard (Figure 65). Select the **Establish PPRC copy pair** option and click **Next**.

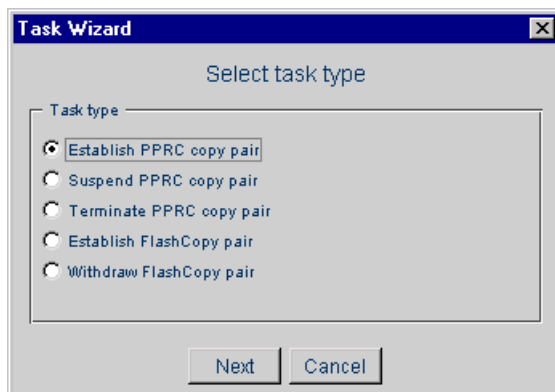


Figure 65. Establish PPRC copy pair

Within the next window you can specify the copy options of the PPRC pair (Figure 66). Click **Next** when you have finished the selection.

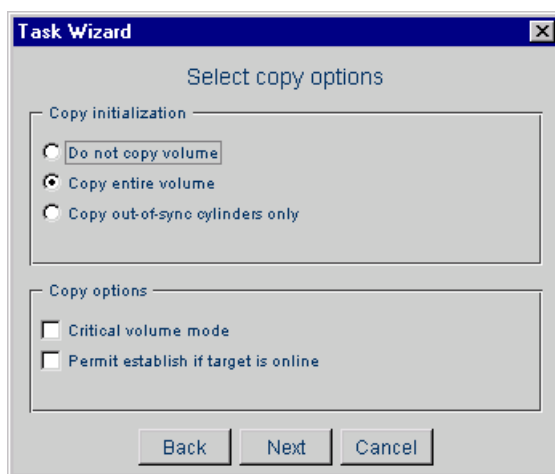


Figure 66. PPRC copy options

Do not copy volume

If this option is checked, the PPRC pair relationship is established without copying any data from the source to the target. This option is used when source and target are in sync; that means, they contain exactly the same data.

Copy entire volume

If this option is checked, all data is copied from the source to the target volume. This option has to be used the first time a PPRC relationship is going to be established and is needed to guarantee that source and target contain the same data.

Copy out-of-sync cylinders only

This option copies only the data that was updated on the target volume since a PPRC copy pair was suspended. The option is used to resynchronize a PPRC pair.

Critical Volume mode

This check-box works in conjunction with the Critical Heavy mode check-box of the LSS (see Table 1 on page 93).

Permit if target is online

If this option is checked, the PPRC operation will be performed even if the target volume is mounted on the host system. This is the case, for example, if a UNIX file system is open on the target.

Table 1. Critical Heavy mode of PPRC pairs

Critical Volume Checkbox	Critical Heavy mode not checked	Critical Heavy mode checked
Not checked	When secondary volume cannot be updated, the pair suspends. Updates to the primary are allowed.	When secondary volume cannot be updated, the pair suspends. Updates to the primary are allowed.
Checked	When secondary volume cannot be updated, the pair suspends. The primary volume is write inhibited only after the last path to the secondary is lost.	When secondary volume cannot be updated, the primary and secondary volumes are write inhibited.

From the next window you can either **Save**, **Run**, or **Cancel** the copy task, as shown in Figure 67. Once a task is saved, it can be executed from the Task panel at any time. Optionally, a name and description of the task can be specified. Even if you do not want to save the task you have created, we recommend that you specify a name and description. This will help with the interpretation of the Copy Services log file later on.

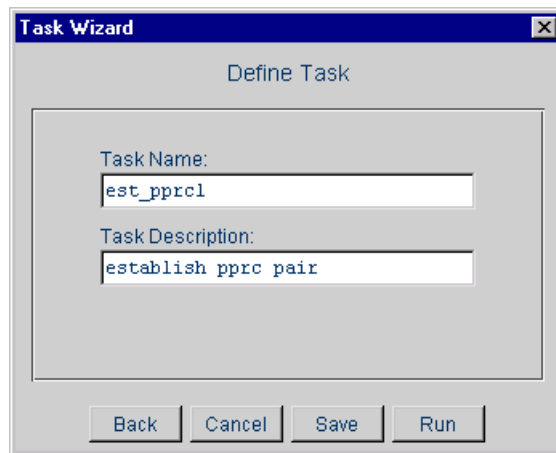


Figure 67. Establish PPRC task window

During the short period during which the PPRC relationship between source and target is established, there is a solid-colored triangle displayed within the volumes (Figure 68). Once the relationship has been successfully established, the source and target volume will change its color, indicating the status of the volume (Figure 69).

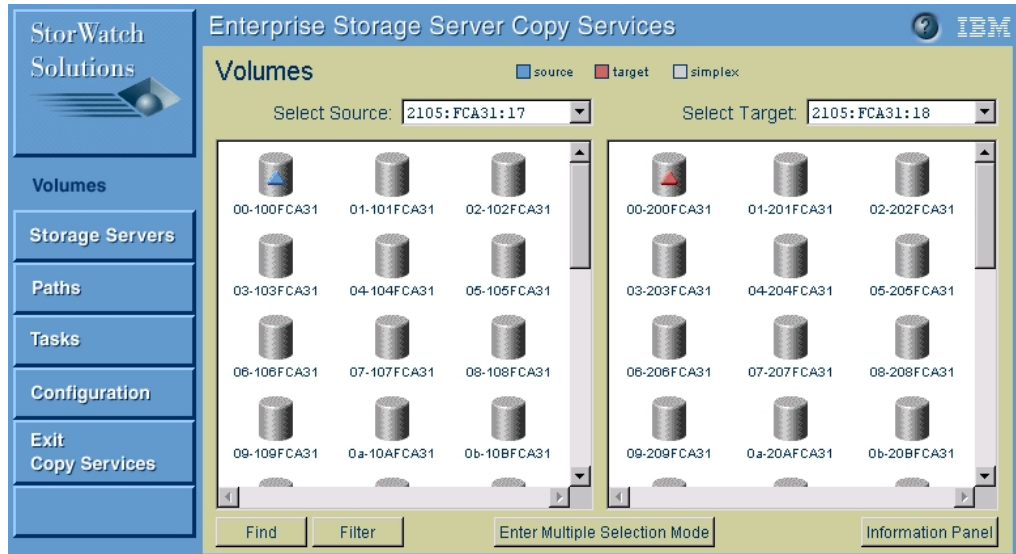


Figure 68. PPRC relationship in progress

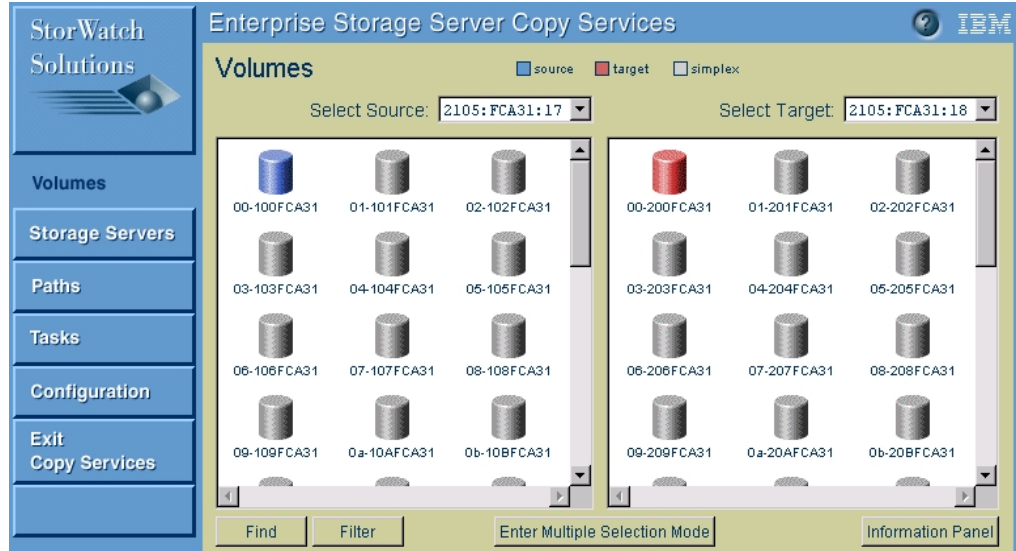


Figure 69. PPRC pair in full copy mode

Select a volume and click the **Information** button to retrieve more information about the status. If the source of a PPRC pair is selected, the number of out-of-sync cylinders that are still left to copy are displayed in addition. Those are the tracks that need be copied from the source to the target to achieve full copy mode (Figure 70).

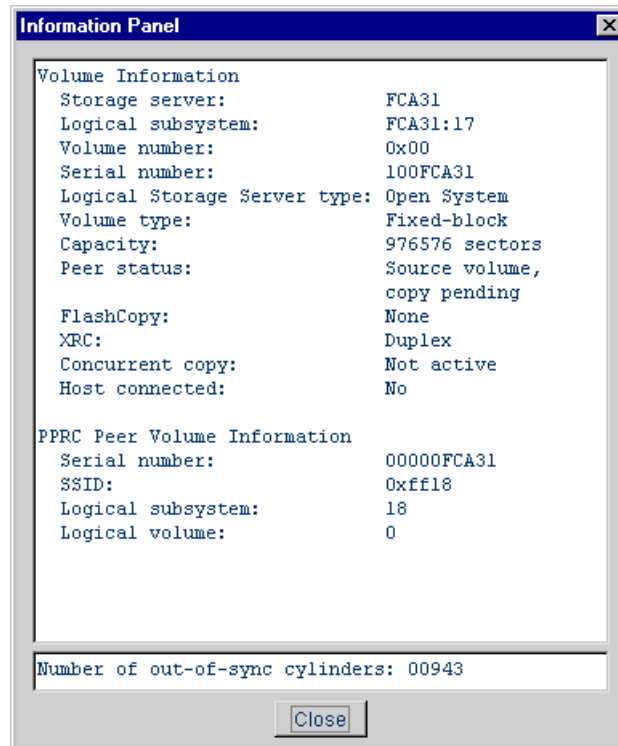


Figure 70. Information window of a PPRC source volume

Note:

Once a PPRC pair has been established, the target volume will not be accessible on any host system until the PPRC relationship has been terminated.

5.3.3 Terminating PPRC pairs

To end the PPRC relationship, you must manually terminate the PPRC pair. After a PPRC pair has been terminated, the target volume will be accessible from a host system again.

To terminate a PPRC pair, select the pair you want to terminate by left-clicking the source and right-clicking the target (Figure 71).

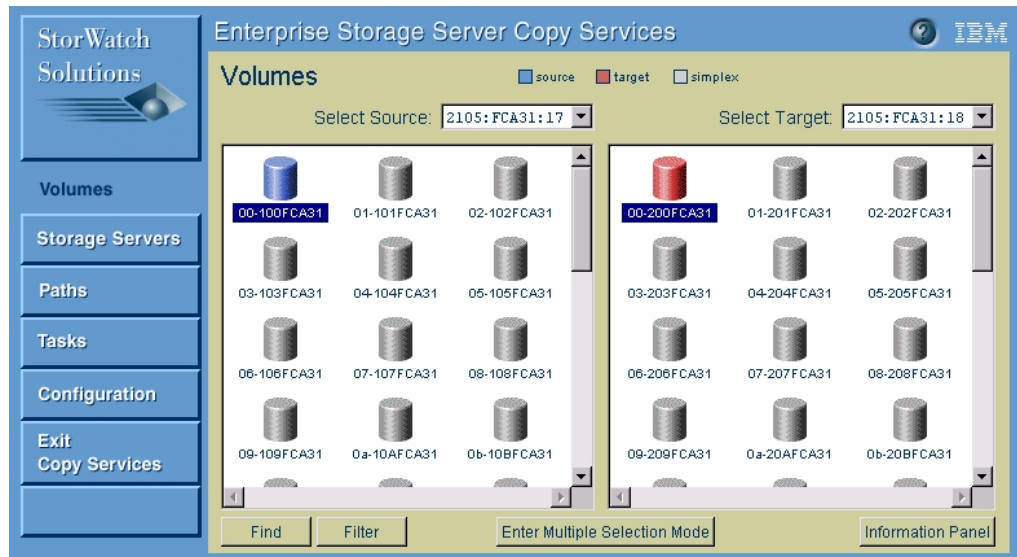


Figure 71. Terminate PPRC pair

Next, left-click one of the volumes of the pair to bring up the window, as shown in Figure 72. From here, select **Terminate PPRC copy pair**.

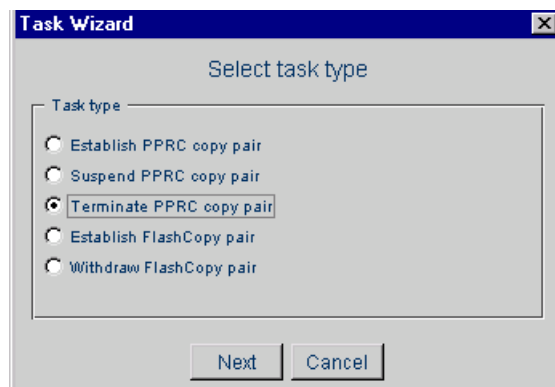


Figure 72. Terminate PPRC pair task wizard

Within the next window you have to choose from which ESS you want to schedule the task — the source or the target storage server (Figure 73). This means choosing which Enterprise Storage Server should execute the task. An example would be to schedule a termination of a PPRC pair with the target storage server in case the source ESS did not happen to be available.

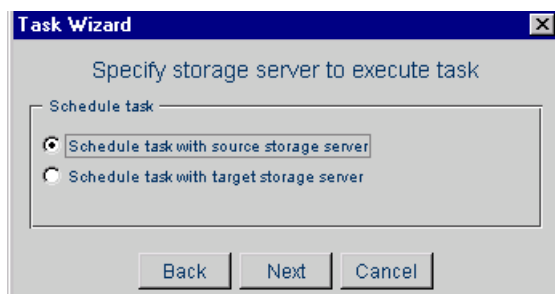


Figure 73. Terminate PPRC pair schedule task window

From the next window, you can either **Save**, **Run**, or **Cancel** the copy task, as shown in Figure 74. Once a task is saved, it can be executed from the Task panel at any time. Optionally, a name and description of the task can be specified. Even if you do not want to save the task you have created, we recommend that you specify a name and description. This will help with the interpretation of the Copy Services log file later on.

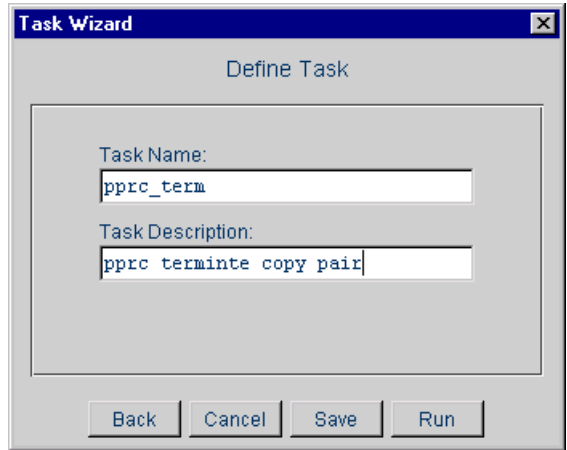


Figure 74. Save task window

During the short period of terminating the PPRC pair, you will see a gray triangle within the source and target volume (Figure 75). Once the pair has been terminated successfully, the volumes will be in the simplex state again.

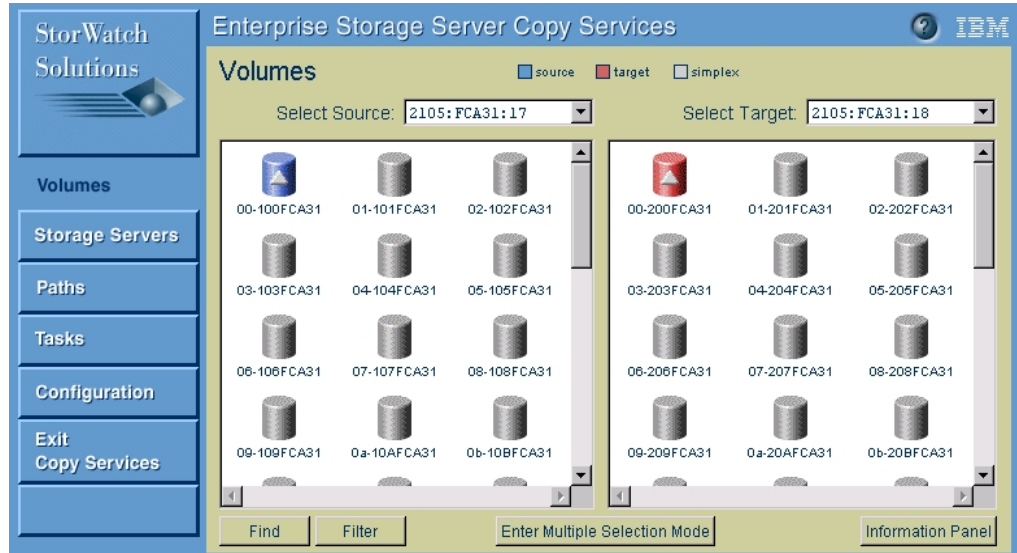


Figure 75. Terminating PPRC pair

5.3.4 Establishing and terminating multiple PPRC pairs at once

There are two possibilities for creating multiple PPRC pairs at the same time:

- Selecting the entire source and target LSS from the **Storage Servers** menu
- Using the **Multiple Selection Mode** from the Volumes menu

When using the Storage Servers panel to select entire logical subsystems, just treat the LSS like a single volume from the Volumes panel when performing a PPRC operation.

Next we will give an example of the usage of the **Multiple Selection Mode**.

Go to the **Volumes** menu and click the **Enter Multiple Selection Mode** button at the bottom of the window. Select one pair at a time beginning with a left-click for the source and a right-click for the target. Once you are finished with the selection, right-click again on the last target volume, which will start the Task Wizard. Continue setting up the PPRC task as described in the previous section.

In the example shown in Figure 76, we have selected 12 PPRC pairs. After running the task for these pairs, the copy process for all pairs is started at the same time, as shown in Figure 77.

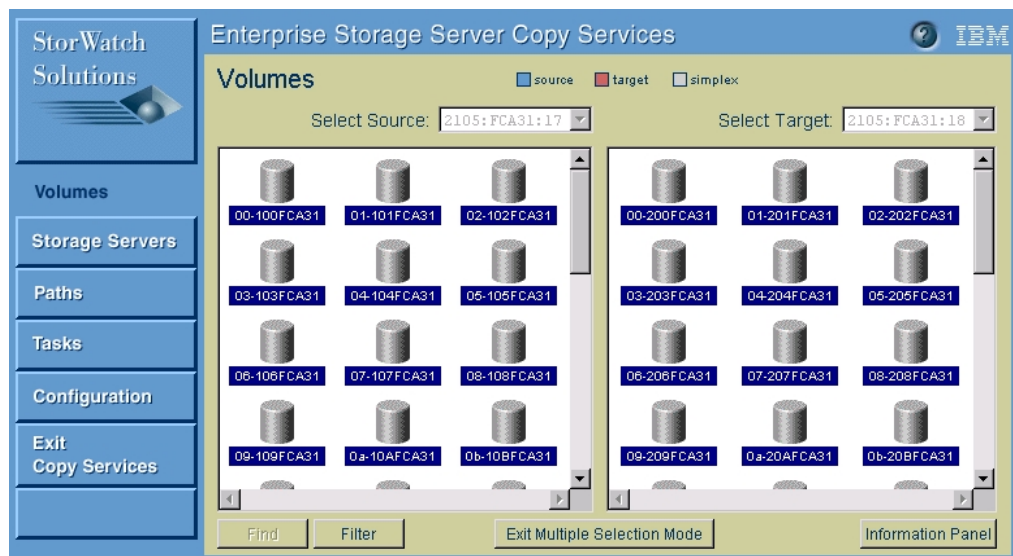


Figure 76. Establish multiple PPRC pairs

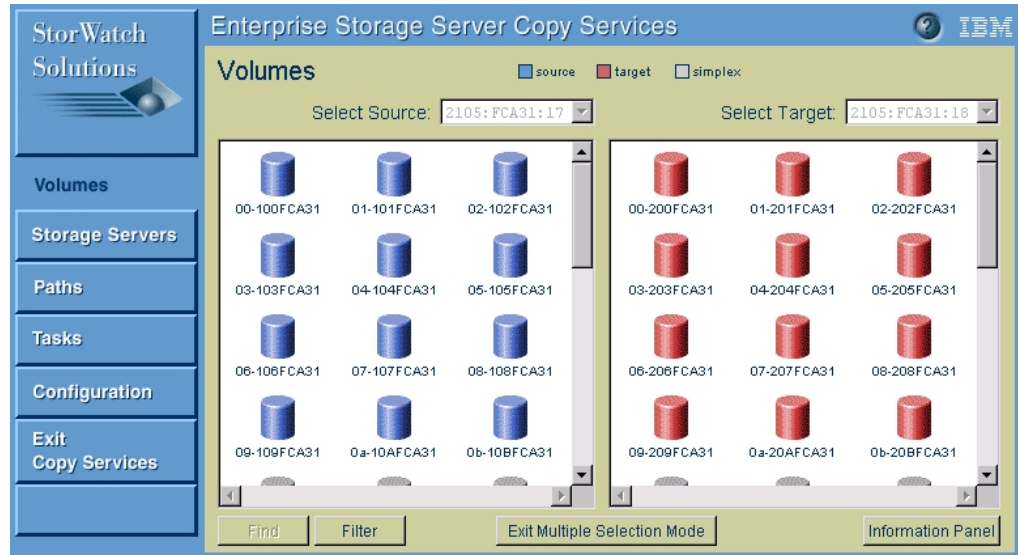


Figure 77. Multiple PPRC pairs in full copy mode

The selection of multiple PPRC volume pairs for termination is done the same way.

Another alternative to select multiple volumes for PPRC at the same time is provided through the **Storage Server** panel. Select the source LSS with a left-click. From the drop-down menu at the top of the window select the target ESS and right-click on the target LSS. A second right-click on the target LSS starts the task wizard for the PPRC copy task. When selecting an entire LSS, all volumes within this LSS are automatically selected for the PPRC copy task.

5.3.5 Setting up FlashCopy and PPRC combinations

A volume of the ESS could only be in one FlashCopy relationship at the same time. However, it is possible to combine PPRC and FlashCopy functionality. That could one of the following cases:

- A FlashCopy target volume is the source of a PPRC pair.
- A PPRC target volume is the source of a FlashCopy pair.

An solution example of such a combination is explained in “Asynchronous PPRC with FlashCopy” on page 56.

Note:

When setting up combinations of PPRC and FlashCopy pairs, you have to wait until the first relationship has successfully completed before using this relationship as a source for another copy pair.

Example 1:

In the first example, we use the FlashCopy Target as the source for a PPRC pair.

First establish the FlashCopy pair as described in this chapter. This can be done with or without the NOCOPY option. Wait until the FlashCopy pair has been created successfully (Figure 78).

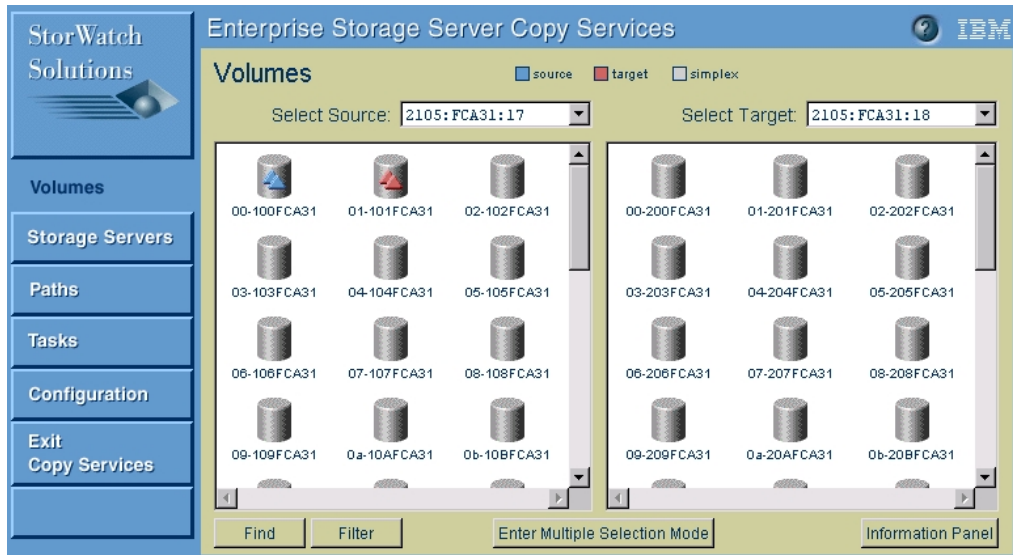


Figure 78. FlashCopy target used as PPRC source

Once the FlashCopy pair has been established, the target volume will be the source of a PPRC pair. The PPRC pair could be established even if there are still tracks left to copy from the FlashCopy source to its target (Figure 79).

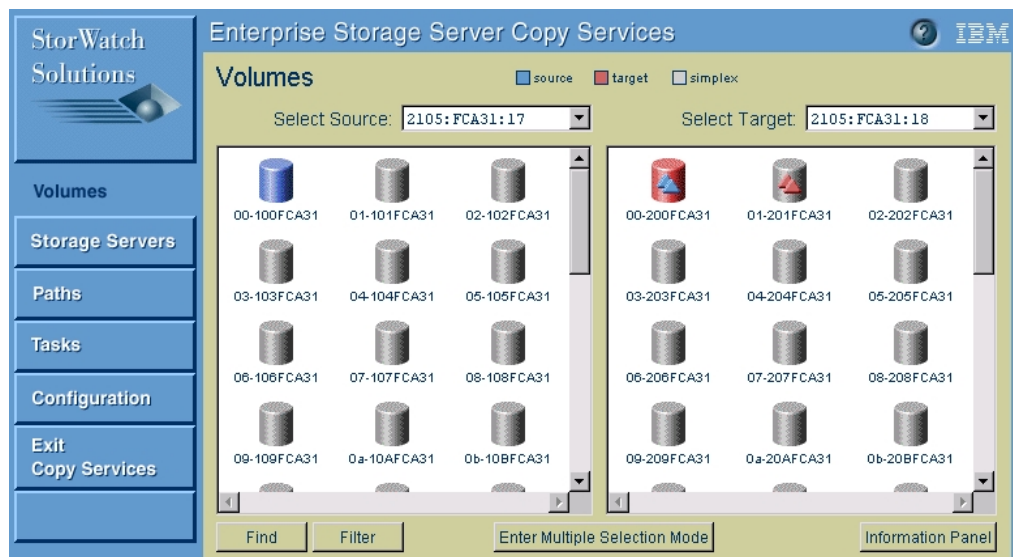


Figure 79. FlashCopy / PPRC combination

The status of a selected volume is displayed if you click the **Information Panel** button of the window. Figure 80 shows the status information of the FlashCopy source / PPRC target volume.

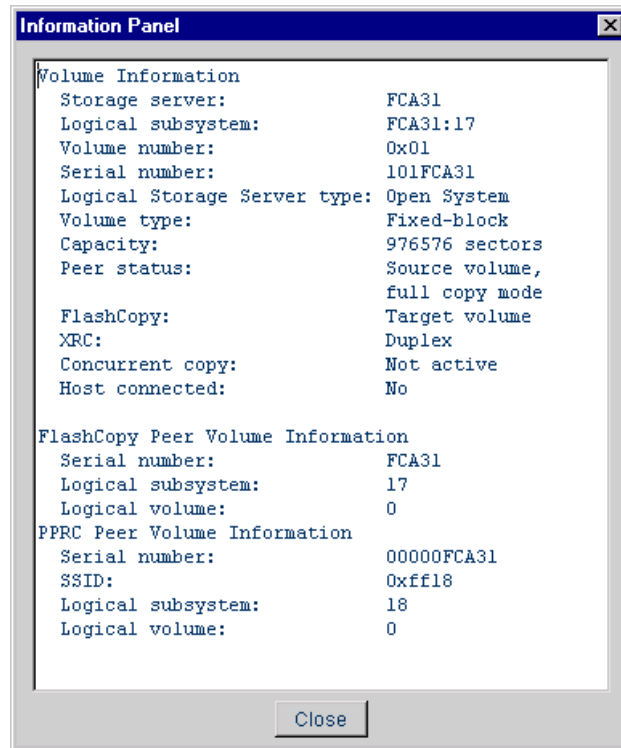


Figure 80. Information window of a combined volume

Example 2:

Next we will give an example of a Copy Services combination using a PPRC target volume as the source of a FlashCopy pair.

First set up the PPRC pair as described in this chapter. Wait until the PPRC pair was successfully established (Figure 81).

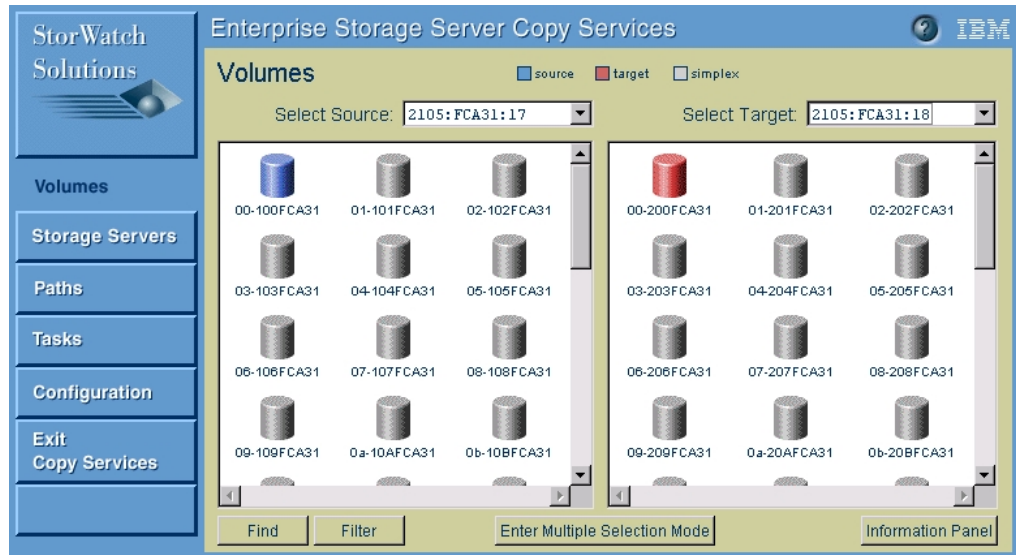


Figure 81. PPRC target used as FlashCopy source

Once the PPRC pair has been established, the target volume will be the source of a FlashCopy pair. The FlashCopy could be created even if there are still tracks left to copy from the PPRC source to its target.

Create the FlashCopy pair as described in this chapter. Keep in mind that FlashCopy source and target have to be in the same logical subsystem. Figure 82 shows the output of the volumes panel once the FlashCopy pair has been established.

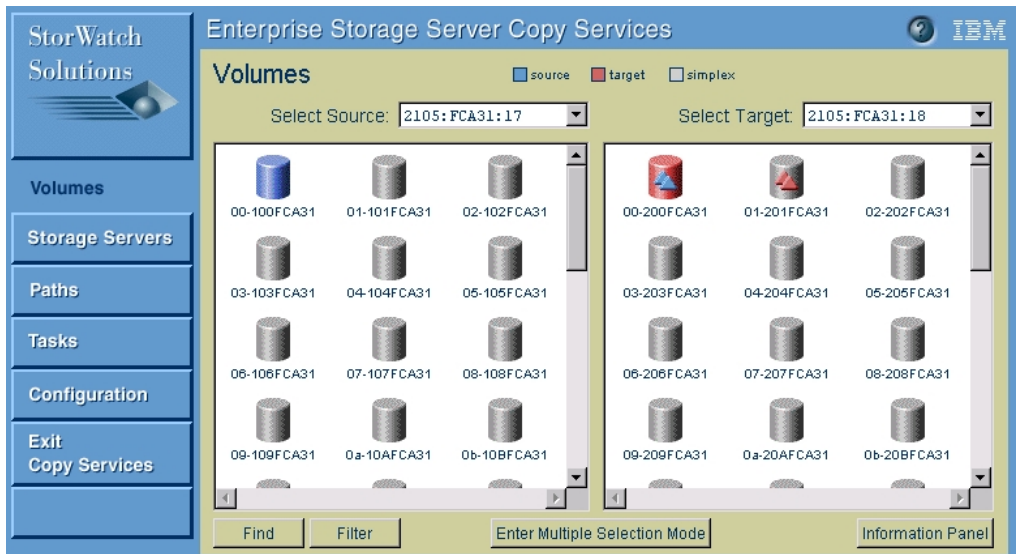


Figure 82. PPRC / FlashCopy combination

5.3.6 Suspending and resuming PPRC pairs

There are two possible reasons for a PPRC pair to be in the suspended mode:

- The source and target of a PPRC can not communicate to each other anymore. This could be the case if the source or target volume is not available or the ESCON link in between the pair is offline.
- The administrator manually have suspended the PPRC pair.

If a volume is in the suspended mode it is not accessible from the host side even if the volume is connected to a host adapter.

5.3.7 Manually suspending PPRC pairs

The process of suspending a PPRC pair is similar to the termination process, except that the Suspend action is selected instead of terminate.

5.3.8 Configuration tips

If you are creating a PPRC task that involves multiple source and target volumes, there is a quick way to create the tasks to suspend and terminate the PPRC pairs.

At the Tasks panel in the Copy Services Specialist:

1. Click the task you created that establishes the PPRC pairs.
2. Click the **Properties** button (Figure 83).

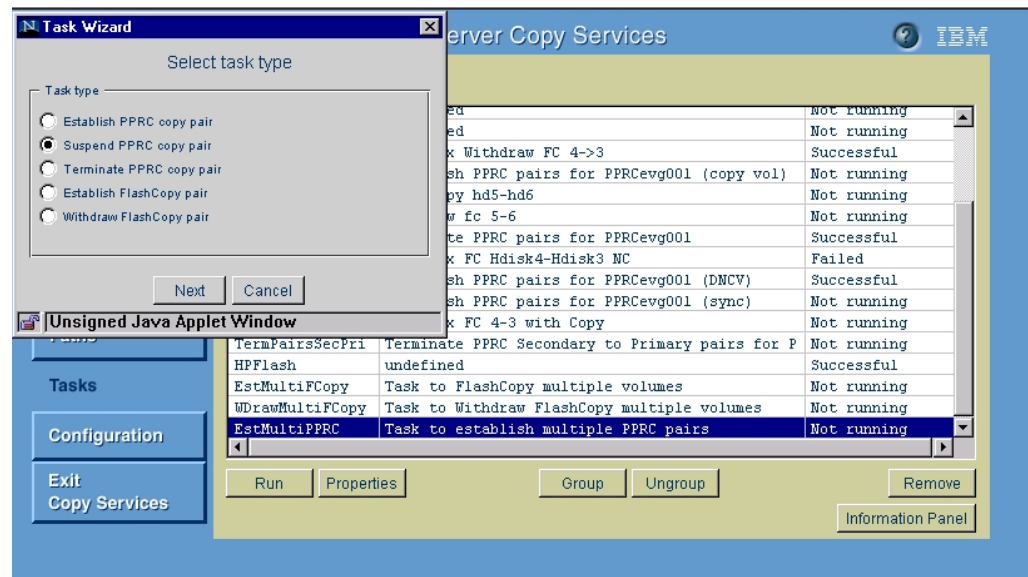


Figure 83. Quick way to suspend PPRC pairs

3. Click the **Suspend PPRC Copy Pair** button, then click **Next** (Figure 84).

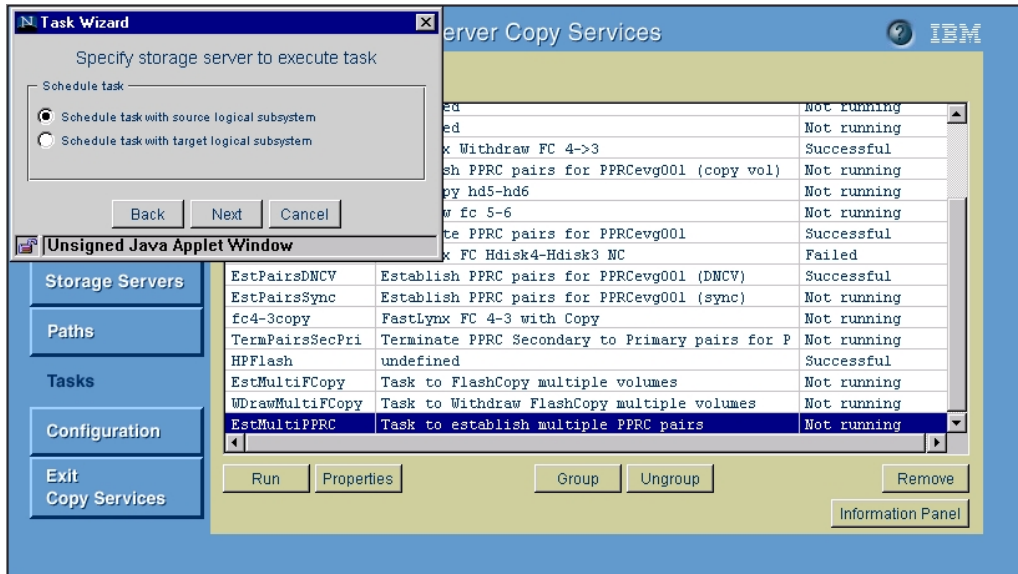


Figure 84. Suspend PPRC Copy Pair

4. Specify the storage server to execute the task, then click **Next** (Figure 85).

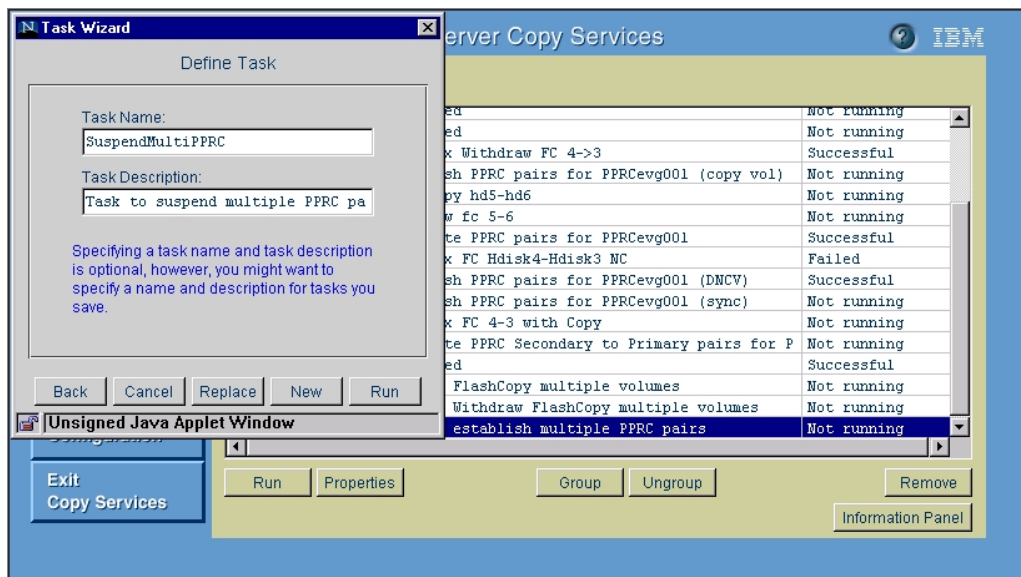


Figure 85. Task to suspend PPRC Copy Pair

5. Type the new Task Name and Task Description, then click **New** (Figure 86).

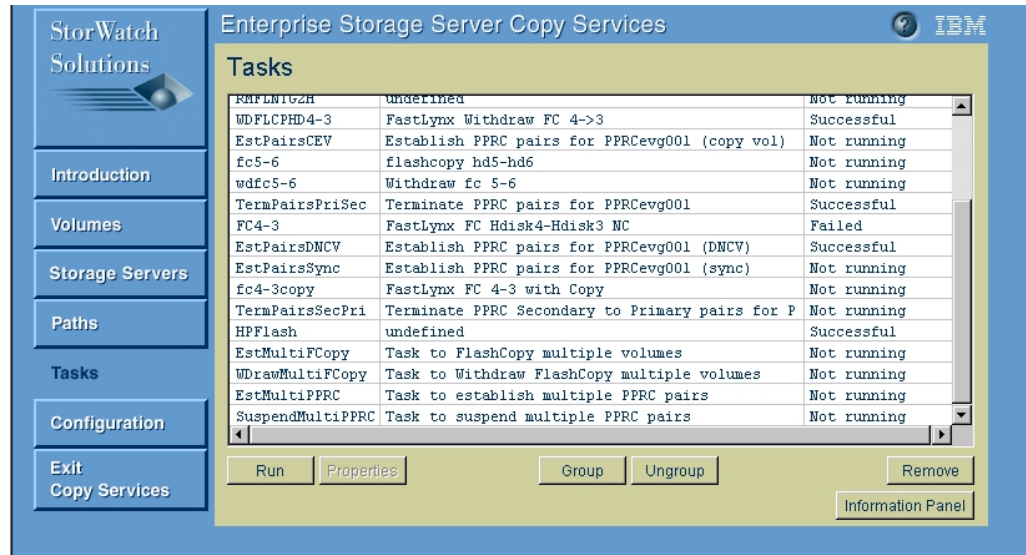


Figure 86. Tasks menu

6. Now you have a task to suspend the PPRC Pairs.

You can repeat the procedure in steps 1 through 6 to create a task that terminates the PPRC pairs.

5.4 ESS Copy Services Web Interface: tips for problem solving

The Copy Services client program is responsible for notifying the primary CopyServices server when a status change occurs. However, in rare instances, the primary CopyServices server does not respond to newly initiated activity or does not have complete information from all Copy Services clients.

If you cannot access ESS Web Copy Services when you click **Copy Services** on the ESS Specialist Welcome panel (see Figure 32 on page 67), do the following:

1. Sign off from ESS Specialist and close the browser window. Open the browser and signon to ESS Specialist again.
2. If step 1 is unsuccessful, shut down and restart the workstation that is running your browser.
3. Open the browser and access the ESS Specialist again.
4. Click **Copy Services** from the ESS Specialist Welcome panel.

If you still cannot access Copy Services, it is possible that the Copy Services server functions are not operational. You might want to reset ESS Web Copy Services. Be aware that you need to have administrator authority to do this. To reset ESS Web Copy Services, do the following:

1. From the ESS Specialist, Click the **Help** icon (?) in the upper right corner of the panel. This links you to user assistance (Help).
2. Click the **Task Help** tab.
3. Scroll all the way to the end of the Task Help list to the last item in the Solve problems at the ESS Specialist Web site tasks.
4. Click **Unable to connect to Copy Services Server message** to link to the reset Help panel. The Copy Services Trouble Shooting Help panel is displayed. Be sure you understand the results of implementing a reset before proceeding:

Resetting the ESS Web Copy Services .It is not recommended that you use this function, unless you are certain that there are no other recovery options. When you use this function, please be aware that:

- You will lose any PPRC or FlashCopy tasks for which you have not received a successful completion message.
- Established PPRC and FlashCopy relationships are maintained.
- You cannot submit any additional command-line interface (CLI) tasks until ESS Web Copy Services has reinitialized.

The available actions that you can take are shown in Figure 87

Available Actions:

Action	Description
Reset to Primary	Restart Copy Services with the primary server as active server.
Reset to Backup	Restart Copy Services with the backup server as active server.
Disable	Disable Copy Services.
Cancel	Return to main task-help, without performing any action.

NOTE: to ensure proper processing of these actions, your browser should be set to avoid caching these pages. For **Netscape**, select Edit, Preferences, Advanced, Cache, and select *Document in cache is compared to document on network every time*. For **Internet Explorer**, select Tools, Internet Options, General, Temporary Internet File Settings, and *Check for newer versions of stored pages on every visit to the page*.

Figure 87. Available actions

Note: It is the responsibility of the user with administrator authority to ensure that no one is using Copy Services before doing a reset or disable action to Copy Services.

1. If you decide not to reset ESS Web Copy Services, click **Cancel** .
2. Click **Reset to Primary** to restart Copy Services with the primary server as the active server.
3. Click **Reset to Secondary** to restart Copy Services with the secondary server as the active server.

Note: This is the action you would take if the primary Copy Services server was unavailable because of a disaster or emergency situation.

4. Click **Disable** to disable Copy Services.

Notes:

1. To ensure proper processing of the above actions, your browser should be set to avoid caching these pages.
2. For Netscape, select **Edit Preferences Advanced Cache** and then select **Document in cache is compared to document on network every time**
3. For Internet Explorer, select **Tools Internet Options General Temporary Internet File Settings** and then **Check for newer versions of stored pages on every visit to the page**

You must take this action on all the ESSs that are controlled by this Copy Services Server. You must always invoke ESS Copy Services on the ESS that will run the server code, be it primary or backup server, before you invoke it on the client machine. If you start the client before starting the server you will have problems with Copy Services.

Chapter 6. ESS Copy Services Command Line Interface

The Copy Services of the Enterprise Storage Server (ESS) provide a Command Line Interface (CLI) for the different host platforms (UNIX and Windows NT). Using the Command Line Interface, you are able to communicate with the ESS Copy Services server from the host's command line. An example would be to automate tasks like doing a FlashCopy by invoking the Copy Services commands within customized scripts.

The Command Line Interface is available for the following operating systems:

- AIX
- SUN Solaris
- HP-UX
- Windows NT 4.0 and Windows 2000

We will describe how to install and use the Command Line Interface on these operating systems throughout this chapter.

6.1 Requirements for Copy Services command line invocation

The Copy Services Command Line Interface is Java based, and therefore the Java runtime environment needs to be installed on each host system from which you want to issue the commands.

The CLI code level must be at the same release level as the microcode that is installed in the ESS clusters.

The host system does not necessarily need to be connected to storage assigned on one of the host ports of the ESS. The only requirement is that the server from where you want to invoke the commands is connected to the ESS that is defined as the primary Copy Services server via a local area network (LAN). However the options of commands where the hosts physical volume name instead of the volume serial number is used and the `rsList2105s.sh` or `.bat` command will only work on a host system that is physically connected to the ESS storage.

At the time of writing, the use of the Command Line Interface requires a Java level 1.1.8. You must check the CLI documentation for current recommendations before you begin the installation procedure. If Java is already installed on the host system, you can determine the current level by running the following command:

```
# java -version
```

Please check the Web site of the vendor of your operating systems for other additional software requirements in order to use Java 1.1.8.

Because the copy commands need to communicate with the Copy Services server, you have to identify a copy server before using any of the commands.

Optionally it is possible to authorize the usage of the Copy Services commands by specifying a user and its password when invoking the commands. The administration of these users is done from the Enterprise Copy Services panel of the ESS Specialist.

Note:

In order to enable user authorization for the Copy Services Command Line Interface you have to create new users for that purpose. This is done from the Copy Services configuration panel of the ES Specialist. The accounts already created for the ESS Specialist to administrate the ESS and its storage could not be used for the Copy Services commands.

6.2 Installing the Command Line Interface (CLI)

In this section we describe the installation of the Command Line Interface for the different operating systems.

It is not required to reboot the host system after installation of the Command Line Interface. You can invoke the commands once the installation process is finished.

Note:

Do not move any of the core files from the installation directory to another location. The shell scripts (UNIX) and batch files (NT) depend on the directory tree. The commands will fail if files are moved to other locations. However, the shell scripts and batch files itself could be moved to another location.

Within the batch files there are two variables: `INSTALL` and `JAVA_HOME`. The `INSTALL` variable is set to the directory where the command line files are installed. The `JAVA_HOME` variable specifies the location of the Java runtime environment. For both variables make sure that the correct location is specified. If necessary, modify these variables in each of the shell scripts or batch files according the locations on your host system.

6.2.1 Authorization of the CLI

Using the Copy Services Specialist Configuration panel, you can decide if you wish to have password protection enabled or not. If you have password protection enabled, you must authorize a userid and password so that Copy Services commands can be issued from the host server.

The Specialist's standard userid and password, by default, is not included in the authorization list. You can authorize userids and passwords by clicking the Authorize button in the Configuration panel. Please see 5.1.5, "Configuration panel of the ESS Copy Services Web Interface" on page 75 for more information on the Copy Services Configuration.

6.2.2 Installing the CLI on an RS/6000 system

The required AIX level for the Command Line Interface is AIX 4.3.3. or higher with a Java level of 1.1.8 to be installed on the server.

The fileset for the Command Line Interface will be provided on a CD-ROM and could be directly installed from this CD.

The name of the installation fileset is `ibm2105cli.rte` and the destination of the files is set to `/usr/lpp/ibm2105cli`.

If Java is installed on another directory than /usr/jdk_base, you have to modify each of the shell scripts in /usr/lpp/ibm2105cli. Shell scripts are all the files that have the ending *.sh. In that case, change the line that reads:

```
export JAVA_HOME=/usr/jdk_base
```

to:

```
export JAVA_HOME=/your_java_installation_directory
```

You can verify that the Command Line Interface fileset was installed properly with the following command:

```
# lslpp -L ibm2105cli.rte
```

If the Command Line Interface is installed correctly, you will get output similar to this example:

```
Fileset                Level  State  Description
-----
ibm2105cli.rte        1.1.0.0  C      IBM 2105 Command Line Interface
                        for AIX

State Codes:
A -- Applied.
B -- Broken.
C -- Committed.
O -- Obsolete. (partially migrated to newer version)
? -- Inconsistent State...Run lppchk -v.
```

Once you have installed the CLI, you should view the readme file

/usr/opt/ibm2105cli/README.aix for any issues or updates related to the specific version.

Make sure you have installed the latest AIX version of the 2105 host installation files (2105inst script). This will configure all ESS attached devices (hdisk) to be configured as 2105 devices. The Copy Services commands that are using the host related device types, instead of the volume ID, depend on 2105 configured hdisk.

You can verify that the host installation files are installed with this command:

```
# lslpp -L ibm2105.rte
```

This will give output similar to this example:

```
Fileset                Level  State  Description
-----
ibm2105.rte           32.6.100.4  C      IBM 2105 runtime for AIX

State Codes:
A -- Applied.
B -- Broken.
C -- Committed.
O -- Obsolete. (partially migrated to newer version)
? -- Inconsistent State...Run lppchk -v.
```

6.2.3 Installing the CLI on a SUN Solaris system

Pre-installation requirements:

1. Install Java 1.1.8 on your system.

Installation:

2. Log in as the root user.
3. Load the CDR into the selected drive.

Note

A window display might pop-up for the CD-ROM drive. You do not need to be concerned about the new window.

4. Display the devices mounted to the system at a terminal window. To display type `mount`. This will show you the path to all the devices connected to the Sun system.
5. Look for the path for the CD-ROM device. If you do not see the CDdevice path, repeat step 3 again.
6. Change your directory to the full path for the CD-ROM device. Example:
`cd /cdrom/unnamed_cdrom/`
7. Type `pkgadd -d solaris7` for Solaris 7 or `pkgadd -d solaris6` for Solaris 6
8. Select **1** just to install the `ibm2105cli` program.
9. Answer **yes** to the next few questions, unless you would like to custom install the program. If you choose to do a custom install, answer the question according to your needs.
10. Select **q** when the installation program return to the option menu. This will exit you from the program.
11. To remove the CD-R, type `eject cdrom` at a terminal window. If the system responds with a busy statement, type `cd` and then `eject cdrom.`

Post-installation procedure:

12. If you have not done it yet, install Java 1.1.8 on your system.
13. The CDR install process will install this package in the `/opt/ibm2105cli` directory.
14. If for any reason, the package was installed into a location other than `/opt/ibm2105cli`, then do the following:
 - a. Edit each shell script.
 - a. Change the line that reads: `export INSTALL=/opt/ibm2105cli` to:
`export INSTALL=/your_install_directory`
15. If Java 1.1.8 is installed in a location other than `/usr/java` then do the following:
 - a. Edit each shell script or bat file.
 - a. Change the line that reads: `export JAVA_HOME=/usr/java` to: `export JAVA_HOME=/your_java_directory`
16. Execute the following additional installation steps:

If you installed this package as recommended, then execute:

```
ln -s /opt/ibm2105cli/libioser.so /usr/lib/libioser.so
```

Otherwise, execute:

```
ln -s /your_install_directory/libioser.so /usr/lib/libioser.so
```

6.2.4 Installing the CLI on an HP-UX system

Pre-installation procedure:

1. Install Java 1.1.8 on your system. The default install directory is `/opt/java`.

To install from a CDR:

2. Log in as the `root` user.
3. Load the CDR into the selected drive.
4. Mount the CDR to `SD_CDROM`.
5. Type `swinstall -s /SD_CDROM/hpux/IBMcli.depot`. This will bring up the graphical interface.
6. Select `IBMcli.tag`. On the top menu bar, select **Actions** and then **Install**.
7. Follow the installation procedure.

Post-installation procedure:

8. If you have not done it yet, install Java 1.1.8 on your system. The default install directory for Java is `/opt/java`.
9. The CDR install process will install this package in the `/opt/ibm2105cli` directory.
10. If for any reason, the package was installed into a location other than `/opt/ibm2105cli`, then do the following:
 11. Edit each shell script.
 12. Change the line that reads: `export INSTALL=/opt/ibm2105cli` to: `export INSTALL=/your_install_directory`
13. If Java 1.1.8 is installed in a location other than `/opt/java`, then do the following:
 - a. Edit each shell script or bat file.
 - b. Change the line that reads: `export JAVA_HOME=/opt/java` to: `export JAVA_HOME=/your_java_directory`

6.2.5 Installing the CLI on a Windows NT system

The required NT version for the Command Line Interface is NT 4.0 with Service Pack 4 or higher applied. Make sure that the recommended Java level is installed on the host system.

The Install Shield Wizard will guide through the installation of the Command Line Interface for NT. You can select the destination location of the files during installation.

Within the batch files there are two variables `INSTALL` and `JAVA_HOME`. The `INSTALL` variable is set to the directory where the command line files are installed. The `JAVA_HOME` variable specifies the location of the Java runtime environment. For

both variables, make sure that the correct location is specified. If necessary modify these variables of each batch file accordingly

If Java is installed on another directory than `c:\jdk1.1.8`, you have to modify each of the batch files. In that case, change the line that reads:

```
set JAVA_HOME=c:\jdk1.1.8
```

to:

```
set JAVA_HOME=:\your_java_installation_directory
```

6.3 UNIX and NT Copy Services commands

For both UNIX and NT the same Copy Services commands are available. The command set for UNIX operating systems consists of shell scripts with the *.sh ending; the commands for Windows NT are batch files with the *.bat ending. Functionally they are identical, but there may be some differences in the parameters you can specify when invoking the commands.

rsExecuteTask (.sh .bat)

Accepts and executes one or more pre-defined Copy Services Server tasks. Waits for these tasks to complete execution.

rsList2105s (.sh .bat)

Displays the mapping of host physical volume name to 2105 (ESS) volume serial number.

rsPrimeServer (.sh .bat)

Notifies the Copy Services Server of the mapping of host disk name to 2105 (ESS) volume serial number. This command is useful when the Copy Services Web screens are used to perform FlashCopy and/or PPRC functions. Collects the mapping of the host physical volume names to the ESS Specialist Copy Services server. This permits a host volume view from the ESS Specialist Copy Services Web screen.

rsQuery (.sh .bat)

Queries the FlashCopy and PPRC status of one or more volumes.

rsQueryComplete (.sh .bat)

Accepts a pre-defined Copy Services Server task name and determines whether all volumes defined in that task have completed their PPRC copy initialization. If not, this command waits for that initialization to complete.

rsTestConnection (.sh .bat)

Determines whether the Copy Services Server can successfully be contacted.

6.3.1 UNIX command description

A Copy Services server must be identified and configured before you can use the Command Line Interface.

6.3.1.1 rsExecuteTask

**rsExecuteTask.sh [-v] -u username -p password -s primaryserver
-b backupserver taskNames**

-v = Verbose (optional).

-u = Username defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

-p = Password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

-s = IP address or complete host name of Copy Services server.

-b = IP address or complete host name of Copy Services Backup server.

taskNames = Names of one or more Copy Services pre-defined tasks to execute.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

No task names specified

You must specify at least one task to run. A task can be defined using the Enterprise Storage Specialist Copy Services Web screens

Exit Status: 0 = Successful completion and >0 = an error occurred.

6.3.1.2 rsList2105s.sh

rsList2105s.sh

Error messages: Not applicable.

Exit Status: 0 = Successful completion and >0 = an error occurred.

Examples of rsList2105s.sh on various platforms are shown in Figure 88, Figure 89, and Figure 90.

```

Example:
# rsList2105s.sh 9.113.24.146
disk name      2105 serial number
-----
hdisk3         22312088
hdisk4         22412088
hdisk5         22512088
hdisk6         22612088
hdisk7         22712088
hdisk8         22812088
hdisk9         22912088
hdisk10        22A12088
...

```

Figure 88. Example of rsList2105s.sh on AIX

```

Example:
# ./rsList2105s.sh
disk name      2105 serial number
-----
c1t6d0        500FCA24
c1t6d1        501FCA24
c1t6d2        502FCA24
c1t6d3        503FCA24
c1t6d4        504FCA24
c1t6d5        505FCA24

```

Figure 89. Example of rsList2105s.sh on SUN Solaris

```

# rsList2105s.sh
disk name      2105 serial number
-----
c0t6d0        005FCA24
c0t6d1        006FCA24
c0t6d2        007FCA24
c0t6d3        008FCA24

```

Figure 90. Example of rsList2105.sh on HP-UX

If you have the Subsystem Device Driver (formerly Data Path Optimizer) installed, the output of rsList2105s.sh shows the vpath names associated with the disk names (Figure 91).

```

# rsList2105s.sh
VpathName      Serial      VolumeNames
-----
vpath0         30114744   hdisk4
vpath1         30014744   hdisk4
vpath2         60014744   hdisk7
vpath3         60014744   hdisk7

```

Figure 91. Example of rsList2105.sh showing vpaths in AIX

6.3.1.3 rsPrimeServer

**rsPrimeServer.sh [-v] -u username -p password -d hostname
-s primaryserver -b backupserver**

-v = Verbose (optional).

-u = Username defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

-p = Password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

-d = IP or complete name of hostname to be removed (Give this flag only when removing the hostname)

This parameter allows you to “un-prime” a specified host from Copy Services’s definitions.

-s = IP address or complete host name of Copy Services server.

-b = IP address or complete host name of Copy Services Backup server.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Example: **rsPrimeServer.sh** ... your.copyservices.server

Exit Status: 0 = Successful completion and >0 = an error occurred.

Figure 92 shows the Volumes panel of the ESS Copy Services Web Interface. On the left side within the source area you will see volumes from a SUN Solaris system (`cxydz`), on the right side within the target area there are volumes from a AIX system (`hdisks`).

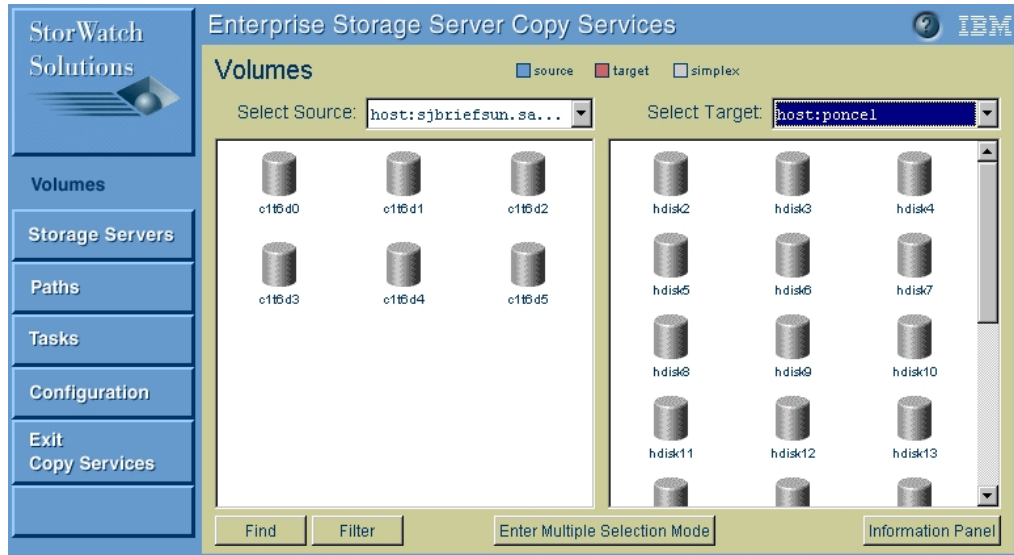


Figure 92. Example `rsPrimeServer.sh`: Source area: SUN Solaris, target area: AIX

6.3.1.4 `rsQuery.sh`

`rsQuery.sh [-v] -m -u username -p password -q volume | -f filename -s primaryserver -b backupserver`

-v = Verbose (optional).

-m = Optional. Map all disk names specified to 2105 volume serial numbers. The user of this parameter permits host disk names to be used as volume parameters.

-u = username defined at the Copy Services server

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

-p = password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

-f = name of a file containing pairs of volume serial numbers, all of which are to be queried. This parameter is required if the **-v** parameter is not used. The format of this file may be either:

```
sourceVolume1 targetVolume1
sourceVolume2 targetVolume2
...
sourceVolumeN targetVolumeN or:
```

```
volume1
volume2
volume 3
volumeN
...
```


-q = single volume serial number. Not valid if **-f** option is specified. This parameter is required if the **-f** parameter is not used.

-s = IP address or complete host name of Copy Services Primary Server.

-b = IP address or complete host name of Copy Services Backup Server.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Example: **rsQuery.sh ... your.copyservices.server**

Missing parameters

One or more of the following parameters was specified but the argument to the parameter was not: **-u**, **-p**, **-f**, **-v**

Example: **rsQuery.sh -u** userName **-p** password ...

Conflicting parameters

The **-f** and **-v** parameters were both specified

Example: **rsQuery.sh -q** volume ... or **rsQuery.sh -f** filename

Volume list not specified

No volumes were specified. You must at least one volume to query using either the **-q** parameter or the **-f** parameter.

Volume list file does not exist

The file specified with the **-f** parameter above does not exist. You must create it before running this command.

Exit Status: 0 = Successful completion and >0 = an error occurred.

An example of **rsQuery.sh** usage is shown in Figure 93.

```
Example:
rsQuery.sh -q 108FCA31 -s 9.113.24.146

Volume 108FCA31 found on FCA31:17 as volume number 8
State=simplex, status=not suspended, FlashCopy state=source

rsQuery.sh -q 109FCA31 -s 9.113.24.146

Volume 109FCA31 found on FCA31:17 as volume number 9
State=simplex, status=not suspended, FlashCopy state=target
```

Figure 93. Example of **rsQuery.sh** usage

6.3.1.5 rsQueryComplete

rsQueryComplete.sh [-v] -u username -p password [-m minutes] [-t threshold] -s primaryserver -b backupserver taskNames

-v = Verbose (optional).

-u = Username defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and password which are authorized to use this command.

-p = Password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and password which are authorized to use this command.

-m = Optional. Number of minutes to wait between checks of whether peer-to-peer Remote Copy initialization has completed. Default is 1 minute/

-t = Optional. Threshold percentage which defines completion. Default is 100%

-s = IP address or complete host name of Copy Services Primary Server

-b = IP address or complete host name of Copy Services Backup Server

taskNames = name of a pre-defined peer-to-peer Remote Copy Establish Pair task defined at the Copy Services server.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified

Example: **rsQueryComplete.sh** ... your.copyservices.server

Exit Status: 0 = Successful completion and >0 = an error occurred.

6.3.1.6 rsTestConnection.sh

rsTestConnection.sh [-v] -s serverName

-v = Verbose (optional).

-s serverName = IP address or complete host name of Copy Services server

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Exit Status: 0 = Successful completion and >0 = an error occurred.

An example of rsTestConnection.sh usage is shown in Figure 94.

```
Example:
# rsTestConnection.sh -v -s 9.113.24.146
rsWebTest: Using 9.113.24.146 as server name
rsWebTest: Application ORB Initialized
rsWebTest: rsVSServer reference obtained successfully
rsWebTest: rsVSServer reference narrowed successfully
rsWebTest: HeartBeat to the server was successful.
rsWebTest: command successful
```

Figure 94. Example of rsTestConnection.sh usage

6.3.2 Scripting the Command Line Interface

You can enhance the functionality of the Copy Services Command Line Interface by incorporating its use in your own customized scripts. Common applications of the CLI might include batch, automation, and custom utilities.

The following example script, shown in Figure 95, accepts the name of the Copy Services server and one more AIX volume group names (including pattern matching) and returns the hdisk and volume identifier, volume state and status, and the volume's status if it exists in a FlashCopy pair.

The script will execute on an RS/6000 server and with little modification, could be adapted to run on HP-UX or Sun Solaris. You need to be root user to execute the script. Also, if your CLI has been authorized to use a username and password, it is a good idea to set the script's permissions to 700 (-rwx-----) to protect your password.

```

#!/usr/bin/ksh

#####
#
# Name:csQueryVols.sh
#
# Description:This script takes one or more AIX Volume Groups that have
#been defined over Copy Services hdisks, or one or more
#disk names, and displays their status information.
#
# Called by:shell
#
# Arguments:ServerName, [vgname... [vgname*]] | [hdiskname... [hdisk*]]
#
# Author:Andrew Beyer, IBM Australia
#
#####

if [ $# -lt 2 ]
then
print " Usage: csQueryVols.sh ServerName vgname1... [ vgname* ]\n" \
      "      csQueryVols.sh ServerName hdiskname1... [ hdisk* ]\n" \
      "where: ServerName is the IP address of an ESS CS server\n" \
      "      vgname1, ..., vgnameN is an AIX Volume Group name\n" \
      "      hdiskname1, ..., hdisknameN is an AIX Physical Volume name"
exit 1
fi

CLI_CMDS="/usr/opt/ibm2105cli"
userid="storwatch"
password="specialist"

csserver=$1
hdisk_list=""
vg_list=""

if [ $(print $2|grep hdisk) ]
then
while [[ $2 != "" ]]
do# Parse command line arguments
if [ $(print $2|grep "*") ]
then
hdisk_list=$(lsdev -Cc disk|grep 2105|cut -f1 -d" ")
shift
else
hdisk_list="${hdisk_list} ${2}"
shift
fi
done
else
while [[ ${2} != "" ]]
do# Parse command line arguments
if [ $(print ${2}|grep "*") ]
then
substr=$(print ${2}|sed "s/\*/g")
vg_list="${vg_list} $(lsvg|grep ${substr}|sort|uniq)"
shift
else
vg_list="${vg_list} ${2}"
shift
fi
done
fi

```

```

echo "\nTesting connection to Copy Services server... \c"
if ( ${CLI_CMDS}/rsTestConnection.sh -s ${csserver} >/dev/null 2>&1 )
then
print "OK"
else
print "Failed"
exit 1
fi

echo "Querying Copy Services server..."
echo "Results may take a few moments to complete...\n"

if [ "${hdisk_list}" = "" ]
then
echo "Volume Group\tDisk:Volume ID    State\t\tStatus\t\tFCopy State"
echo "-----\t-----"
for vg in ${vg_list}
do
hdisk_list=$(lsvg -p ${vg}|awk '/hdisk/ {print $1}')

for hdisk in ${hdisk_list}
do
# Get the volume id from the configuration inventory
vol=$(lscfg -pvl ${hdisk}|grep "Serial Number"|sed "s/.*[a-z]\.*/g")
print "${vg}\t\t${hdisk}:${vol}    \c"

# Filter and reformat rsQuery.sh output
${CLI_CMDS}/rsQuery.sh -u ${userid} -p ${password} -q ${vol} -s ${csserver}|awk -F
"State=|status=|FlashCopy_state=|, " '/State/ {print $2 "\t" $4 "\t" $6}'
done
done
else
echo "Disk:Volume ID    State\t\tStatus\t\tFCopy State"
echo "-----"
for hdisk in ${hdisk_list}
do
# Get the volume id from the configuration inventory
vol=$(lscfg -pvl ${hdisk}|grep "Serial Number"|sed "s/.*[a-z]\.*/g")
print "${hdisk}:${vol}    \c"

# Filter and reformat rsQuery.sh output
${CLI_CMDS}/rsQuery.sh -u ${userid} -p ${password} -q ${vol} -s ${csserver}|awk -F
"State=|status=|FlashCopy_state=|, " '/State/ {print $2 "\t" $4 "\t" $6}'
done
done
fi

exit 0

```

Figure 95. csQueryVols.sh - an example script showing the use of the CLI

A typical output from `csQueryVols.sh` is shown in Figure 96.

```
# ./csQueryVols.sh shark6c1 evg* tvg003

Testing connection to Copy Services server... OK
Querying Copy Services server...
Results may take a few moments to complete...

Volume Group      Disk:Volume ID      State      Status      FCopy State
-----
evg001             hdisk1:40014744     source     not_suspended  none
evg001             hdisk2:40114744     source     not_suspended  none
evg003             hdisk3:30014744     simplex    not_suspended  none
evg004             hdisk6:10114744     simplex    not_suspended  none
tvvg003           hdisk4:30114744     simplex    not_suspended  none
#
```

Figure 96. Example of output from `csQueryVols.sh`

6.3.3 Windows NT command description

6.3.3.1 `rsExecuteTask.bat`

`rsExecuteTask [/v][/u username /p password] /s serverName [/b backup serverName] taskNames`

`/v` = Verbose (optional).

`/u` = Username defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

`/p` = Password defined at the Copy Services server

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

`/s` = IP or complete host name of Copy Services server

`/b` = IP or complete host name of Copy Services Backup server

Before this command can be used, a Copy Services server must be identified and configured.

`taskNames` = Name of one or more Copy Services pre-defined tasks to execute.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

No task names specified

You must specify at least one task to run. A task can be defined using the Enterprise Storage Specialist Copy Services Web screens.

Exit Status: 0 = Successful completion and >0 = an error occurred. See 6.3.4, “Command line return codes” on page 128 for error descriptions.

6.3.3.2 rsList2105s.bat rsList2105s

Error messages: None.

Exit Status: 0 = Successful completion and >0 = an error occurred (Figure 97).

```
C:\ibm2105cli>rsList2105s
VpathName      Serial VolumeNames
-----
Disk0          30114744   Disk0
Disk1          30014744   Disk1
Disk2          60014744   Disk2
Disk3          60014744   Disk3
C:\ibm2105cli>
```

Figure 97. Example of rsList2105s on NT system

6.3.3.3 rsPrimeServer.bat

rsPrimeServer [/v][/u username /p password] [/d hostname] /s serverName
[/b backup serverName]

/v = Verbose (optional).

/u = Username defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

/p = Password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

/d = IP or complete name of host to be removed from ESS Copy Services Web Interface (give this flag only when removing the hostname)

/s = IP or complete host name of Copy Services server.

/b = IP or complete host name of Copy Services Backup server

Before this command can be used, a Copy Services server must be identified and configured.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Example: **rsPrimeServer** ... your.copyservices.server

Exit Status: 0 = Successful completion and >0 = an error occurred. See 6.3.4, "Command line return codes" on page 128 for error descriptions.

6.3.3.4 rsQuery.bat

rsQuery [/v] [/m] [/u username /p password] [/q volume | /f filename] /s
serverName [/b backupServerName]

/v = Verbose (optional).

/m = Optional. Map all disk names specified to 2105 volume serial numbers. The user of this parameter permits host disk names to be used as volume parameters.

/u = username defined at the Copy Services server

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

/p = password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

/f = name of a file containing pairs of volume serial numbers, all of which are to be queried. This parameter is required if the **/v** parameter is not used. The format of this file may be either:

```
sourceVolume1 targetVolume1  
sourceVolume2 targetVolume2  
...  
sourceVolumeN targetVolumeN or,
```

```
volume1  
volume2  
volume3  
volumeN  
...
```

/q = single volume serial number. Not valid if **/f** option is specified. This parameter is required if the **/f** parameter is not used.

/s = IP address or complete host name of Copy Services Primary Server.

/b = IP address or complete host name of Copy Services Backup Server.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Example: **rsQuery** ... your.copyservices.server

Missing parameters

One or more of the following parameters was specified but the argument to the parameter was not: /u, /p, /f, /v

Example: **rsQuery /u** userName **/p** password ...

Conflicting parameters

The /f and /v parameters were both specified

Example: **rsQuery /q** volume ... or **rsQuery /f** filename

Volume list not specified

No volumes were specified. You must at least one volume to query using either the /q parameter or the /f parameter.

Volume list file does not exist

The file specified with the /f parameter above does not exist. You must create it before running this command.

Exit Status: 0 = Successful completion and >0 = an error occurred. See 6.3.4, "Command line return codes" on page 128 for error descriptions.

6.3.3.5 rsQueryComplete.bat

rsQueryComplete [/v] [/u username /p password] [/m minutes] [/t threshold] /s serverName [/b backup serverName] taskName

/v = Verbose (optional).

/u = Username defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and password which are authorized to use this command.

/p = Password defined at the Copy Services server.

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and password which are authorized to use this command.

/m = Optional. Number of minutes to wait between checks of whether peer-to-peer Remote Copy initialization has completed.

/t = Optional. Threshold percentage which defines completion.

/s = complete IP or host name of Copy Services server

/b = IP or complete host name of Copy Services Backup server

Before this command can be used, a Copy Services server must be identified and configured.

taskName = name of a pre-defined peer-to-peer Remote Copy Establish Pair task defined at the Copy Services server.

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Example: **rsQueryComplete** ... your.copyservices.server

Task name not specified

At least one task name must be specified to be executed.

Exit Status: 0 = Successful completion and >0 = an error occurred. See 6.3.4, "Command line return codes" on page 128 for error descriptions.

6.3.3.6 rsTestConnection.bat**rsTestConnection [/v][/u username /p password] /s serverName**

/u = username defined at the Copy Services server

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command

/p = Password defined at the Copy Services server

This parameter is required if the Enterprise Storage Server administrator requires it. The Enterprise Storage Specialist Copy Services Web screens can be used to define users and passwords which are authorized to use this command.

/s = IP or complete host name of Copy Services server

Error messages:

Server name not specified

The IP address or complete host name of the Copy Services server is required but was not specified.

Exit Status: 0 = Successful completion and >0 = an error occurred. See 6.3.4, "Command line return codes" on page 128 for error descriptions.

6.3.4 Command line return codes**Successful completion return codes:**

0 Task Completed or Command Successful

Missing parameter return codes:

- 1 User name not specified or Missing parameter user name
- 2 Password not specified or Missing parameter password
- 3 Server name not specified

- 4 No task names specified
- 5 Missing parameter disk serial numbers
- 6 Mismatching number of disk pairs specified
- 7 Missing parameter local disks
- 8 No disk pairs specified
- 9 Missing parameter hostname to be deleted
- 10 Missing parameters
- 11 Missing parameter minutes between samples
- 12 Missing parameter threshold
- 13 No 2105 volumes found
- 14 Volume list not specified
- 15 Conflicting parameters
- 16 More volume/s per line
- 17 Volume list file "\$FILENAME" does not exist
- 18 No volumes specified or No volumes specified in given file

Server connection return codes:

- 40 Failed to connect to server
- 41 The primary server down and there is no backup server specified. The task is not completed.
- 42 The primary and backup server are down. Task not completed.
- 43 failed while creating communication to server
- 44 failed while disconnecting from Server

Exception related return codes:

- 60 System exception occurred

Command failed return codes:

- 80 Command failed

Chapter 7. High availability and disaster recovery

Today, more than ever, organizations are seeking to distribute mirrored disk resources geographically to establish their computing resources as part of an overall disaster recovery plan. As a result, organizations rely on products and services to achieve this end.

This chapter describes how you can incorporate ESS Copy Services into highly available, host cluster configurations. In particular, we focus on how to customize standard software products that can be used to automate the functions of Copy Services for high availability and disaster recovery.

We look at some configuration examples in detail and give code examples to include in your implementations.

7.1 Automating site failover for AIX

HACMP for AIX provides the means to automate rapid recovery of application services by allowing a workload that was running on one host server to be taken over by another host server. In most cases, the cluster nodes are located at the same site, but they can also be distributed over multiple sites.

Please note that at this time HACMP does not formally support the use of PPRC. The scenarios which we describe in this chapter demonstrate ways that HACMP can be used to automate the recovery to a standby site in the event of a major primary site failure.

The information contained in this chapter has not been submitted to any formal test and is distributed as-is. The examples described are provided for guidance to aid in understanding the principles behind integrating PPRC with HACMP and automating failover to a standby site.

First, we review a basic two-node cluster using an ESS without Copy Services. Next, we examine ways in which HACMP can be used to automate the management of PPRC, so that recovery time is minimized after an outage or a disaster.

We also assume that you have experience implementing HACMP environments. Thus, the emphasis is on the process of integrating ESS Copy Services with HACMP and not on the details of implementing HACMP itself.

As is the case for any complex implementation, it is worth emphasizing that detailed and thorough planning and testing will be paramount to the success of your implementation of HACMP and ESS Copy Services.

7.1.1 Hardware and software requirements

As of October 3, 2000 the following support is available for the ESS on RS/6000 and RS/6000 SP (see Table 2).

Table 2. Support for ESS on RS/6000 and RS/6000 SP

HACMP/ESS Hardware and Software Support	HACMP 4.2.2 AIX 4.2.1	HACMP 4.2.2 AIX 4.3.3	HACMP 4.3.1 AIX 4.3.3	HACMP 4.4 AIX 4.3.3
IBM Enterprise Storage Server 2105-E10, 2105-E20 Letter 199-188	HACMP APAR IY04403(1) SCSI only	HACMP APAR IY04403(1) SCSI only	HACMP APAR IY03438(1) IY11560 IY11564 IY12021 IY12056	HACMP APAR IY11563 IY11565 IY12022 IY12057
IBM Enterprise Storage Server 2105-F10, 2105-F20 Letter 100-089	Not supported	Not supported	HACMP APAR IY03438(1) IY11560 IY11564 IY12021 IY12056 IY08933	HACMP APAR IY11480 IY11563 IY11565 IY12022 IY12057
IBM 2032 McData ED-5000 Enterprise FC Director 2032-001 Letter 100-082	Not supported	Not supported	X	X
IBM SAN Fibre Channel Switch 2109-S08, 2109-S16 Letter 199-167	Not supported	Not supported	X	X
IBM SAN Data Gateway Model G07 2108-G07 Fibre Attachment Letter 100-086	Not supported	Not supported	X (2) HACMP APAR IY07313 IY09595(3)	X (2) HACMP APAR IY10564(3)

1. Required for the CRM or ESCRM feature codes only.
2. This now includes the features 2214, 2313, and 2319. When using two Fibre Channel adapters to a single SAN Data Gateway in an HACMP environment, the SAN Data Gateway must be run in "split mode". Each Fibre Channel adapter must be connected to a separate port on the SAN Data Gateway, and two separate ports on the SAN Data Gateway must be used to attach to two separate SCSI ports on the ESS. The SAN Data Gateway will not be supported in a switched environment when HACMP support for switched Fibre Channel becomes available.
3. Required when attaching to the ESS model E10.

For Copy Services support, the following products must be installed on each cluster node:

- Copy Services Command Line Interface for AIX
- Java Runtime for AIX at version 1.1.8

7.1.2 Test environment

The examples described are provided for guidance to aid in understanding the principles behind integrating PPRC with HACMP.

Our test environment included the following components.

Hardware

2105 Enterprise Storage Server Model F20:

- OS Level 4.3.2.11.1
- Code EC SC01029
- SEA.rte level = 2.6.302.1082
- SEA.ras level = 2.6.302.1082

Two 7025 RS/6000 Model F50s:

- OS Level 4.3.3

Software

HACMP Version 4.3.1 for AIX.

PPRC over an ESCON connection between two LSSs in the same ESS was used to simulate PPRC connection between two geographically separated ESSs. Its use in our scenario was purely for proof of concept testing.

A single SCSI connection provided the I/O path from each of the two RS/6000 F50s.

Figure 98 shows the configuration we used for the test scenarios created in this chapter.

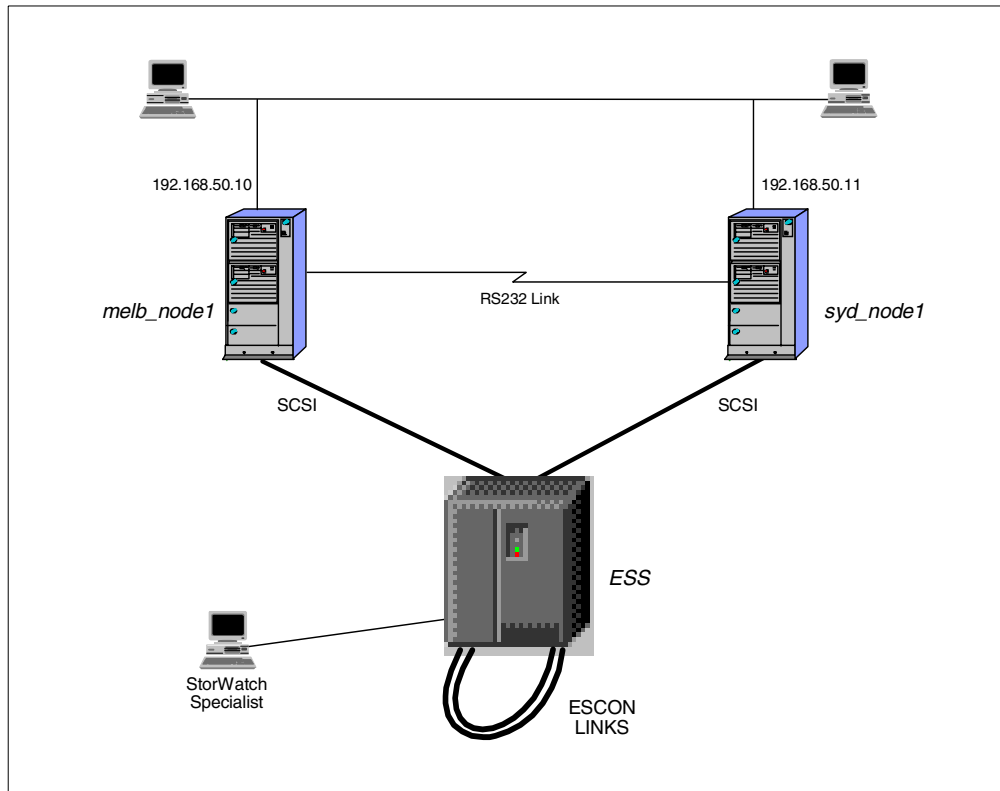


Figure 98. Our test environment

7.1.3 Single site recovery (two-node cluster)

To begin, let us review setting up HACMP without Copy Services. Configuring HACMP without Copy Services is similar to configuring HACMP with any other disk subsystem.

Here, we assume that you are establishing an HACMP cluster using an ESS with the following attributes:

- The ESS has SCSI or FC host bay adapters cabled to the ^ pSeries servers or RS/6000s
- The storage LUNs are already defined.

The ESS presents an hdisk device to AIX in the same way that any other disk or disk subsystem is presented to AIX. Because the data for most application environments require the use of file systems or raw logical volumes, HACMP only needs to know about file system names or volume group names in order to move shared disk resources from one host server to another. Therefore, the underlying disk medium is more or less transparent to HACMP.

Unfortunately, at the time of writing this book, some implementation constraints existed which restrict the way you can define storage to multiple hosts using a given LSS. Please see 7.1.5, "Implementation specifics" on page 148 for current restrictions and workarounds.

Let us now look at some examples in detail.

Figure 99 shows a typical implementation of ESS in a two-node cluster. However, the theory and practice applies equally well to clusters consisting of larger quantities of nodes. HACMP/ES currently supports a maximum 32 nodes. A practical maximum will be affected largely by the I/O adapter configuration in your ESS.

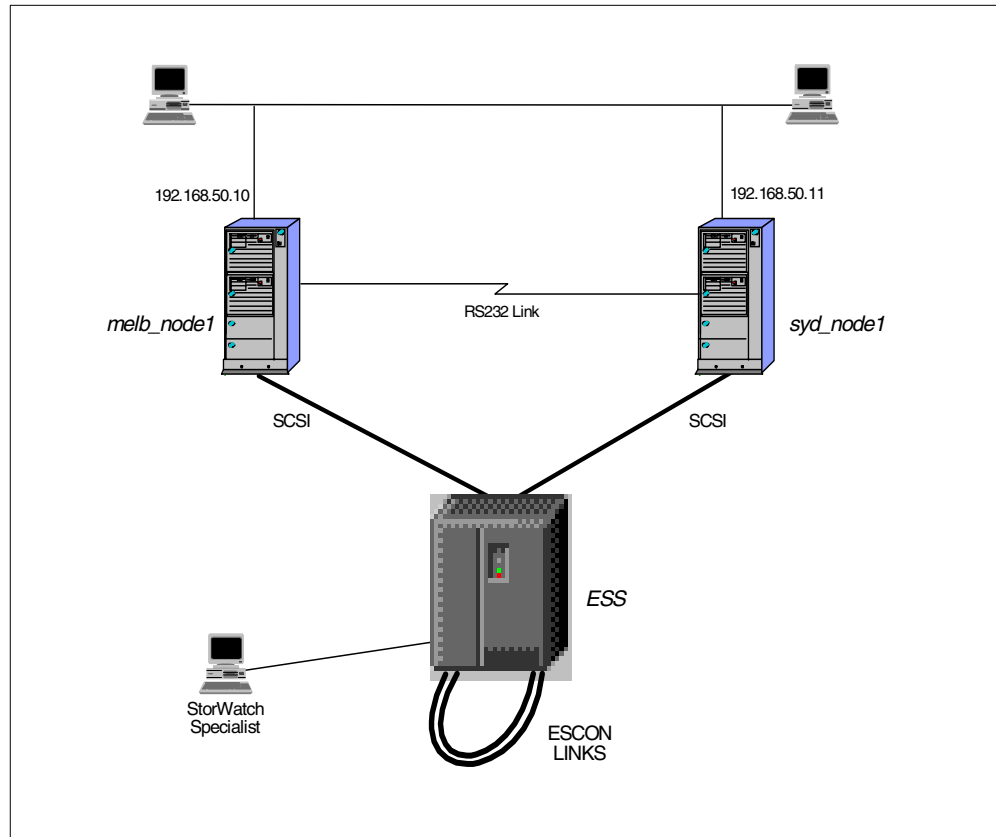


Figure 99. Typical two-node cluster sharing a single ESS at one site

Each cluster node is attached to the ESS via a SCSI or Fibre Channel connection. The ESS presents an hdisk device to AIX in the same way that any other disk or disk subsystem is presented to AIX. You may notice it is possible for the hdisk names or SCSI connection IDs on each cluster node to be different. You do not need to worry about this. HACMP always resolves the Physical Volume ID (PVID) which will be the same for a given physical disk resource, regardless of the differences in hdisk names on each cluster node. Nevertheless, on a given host server, the hdisk names must be unique

You may configure resource groups containing ESS disk volumes in all of the usual ways:

- Cascading
- Rotating
- Concurrent

7.1.4 Dual site recovery (two-node cluster) using PPRC

Now let us look at ways to customize HACMP and PPRC together. Although PPRC can be established between volumes within an ESS, when we speak of implementing PPRC, we almost implicitly are speaking of implementing an environment involving more than one site. This type of environment could typically form part of an organization's disaster recovery plan (DRP).

PPRC can be integrated with HACMP to provide automated failover from a primary site to a standby site.

However, to make site failover work, the configuration is a little unconventional. The following problem arises from a limitation in HACMP's current design.

HACMP is designed to handle the failover of any disk resource between cluster nodes. It is not currently designed to handle the case where one server is connected to disk resources and another server is connected to different disk resources which have the same volume group, logical volume and even filesystem names. However, by placing the code to activate or deactivate the storage elements (VGs, LVs, and FSs) in the start and stop scripts, it is possible to customize HACMP to handle a two node cluster where one server attached to an ESS at one site and one server attached to an ESS at another site. See Figure 100 on page 138 and Figure 102 on page 140 for cluster illustrations.

HAGEO extends the functionality of HACMP and handles site failover by allowing you to define a site resource group that provides the definitions needed for site failover. Extensive scripts have been written for HAGEO to handle local failover and re-integration, and site failover and re-integration. The logic runs as pre and post events to HACMP events, and when a node failure is processed, it is able to determine if the resources can be moved to a standby local peer node, or in the case where there are no more local peers (a site failure), start the resources at a standby site.

Unfortunately, it is not feasible to use HAGEO to solve our multi-node geo-cluster problem. This is because HAGEO's functionality was designed to manage Geo Mirror Devices at each site and not ordinary filesystems. So, even though HAGEO provides the fundamentals for site failover, it cannot be adapted readily to handle ordinary filesystem takeover from one site to another.

With the current design of HACMP and the customization methodology described throughout the remainder of this section, configurations of more than two cluster nodes will not work.

Consequently, the configuration assumes an HACMP cluster with following attributes.

The cluster consists of no more than two nodes (^ pSeries servers or RS/6000s)

At least one ESS is connected to each node (server)

No filesystems or volume groups are configured in a resource group

No NFS filesystems are configured in a resource group

Other requirements include these:

- The ESS has SCSI or FC host bay adapters cabled to each node.
- The storage LUNs are already defined.
- The Copy Services microcode is installed and configured.
- There is a serial RS232 connection between the servers at each site.
- At least two ESCON paths have been configured between the two ESSs for path redundancy (SSG recommends at least eight paths for a 1.6GB data transfer).

Figure 100 shows a two-node cluster, where each cluster node is located at a separate site. This is a basic configuration that can be set up to automate recovery, should a disaster occur at one of the two sites.

Notice that the disk resources are not shared between cluster nodes. Instead, each node is attached to a separate ESS.

There are two ESCON links between each ESS to provide a basic level of redundancy and performance requirements would indicate if more links are needed. Note that the maximum distance for ESCON connection is 103km using Fibre Savers or Wave Division Multiplexers. You can expect performance degradation at this distance.

Also, the IP domains are common to both sites.

Important Note

IP address takeover is supported only if the IP subnetwork for the service network adapters is available at both sites and if the IP subnetwork for the standby network adapters is available at both sites. HACMP does not support IP address takeover across different subnetworks.

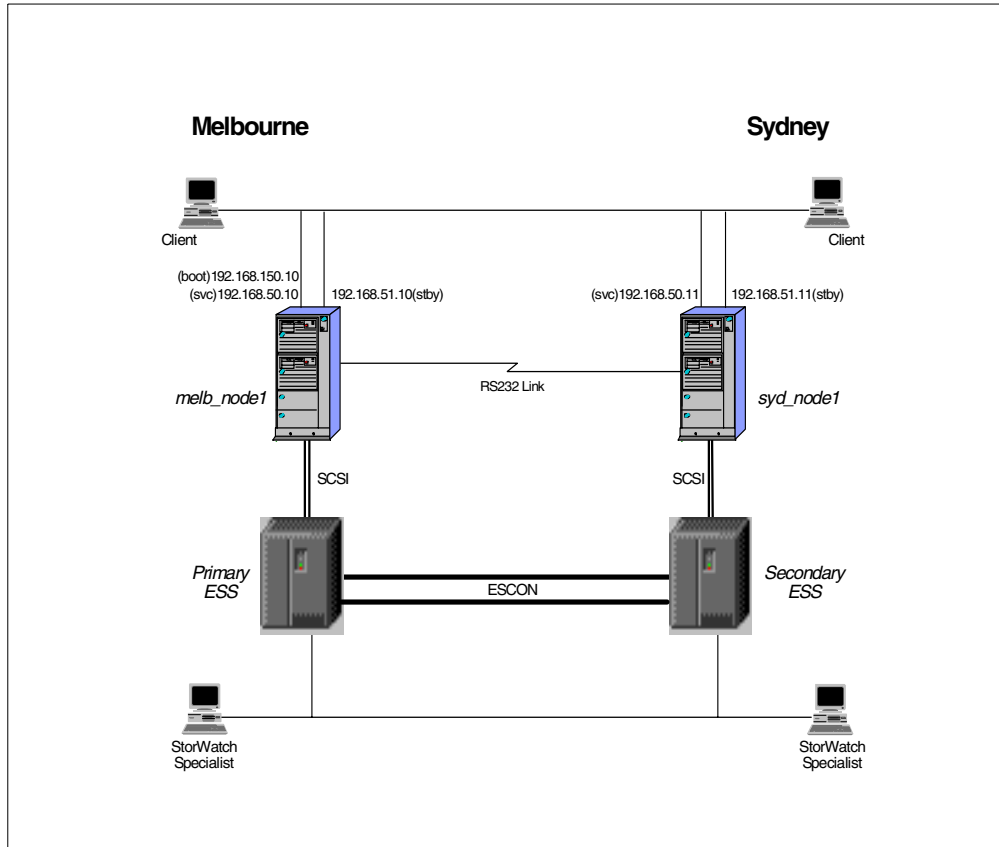


Figure 100. Two-node cluster spanning two sites with common network domains

Configuring HACMP with PPRC involves customizing the cluster resources in a different way to the customizing the cluster for shared disk resources. Because there are no shared disk resources, we cannot place filesystem and volume group names in the resource group. These resources are handled by the HACMP Application Server start and stop scripts. Therefore, a service IP label and an Application Server are the only resources added to the resource group.

In summary, HACMP is configured to:

- Facilitate IP address takeover (IPAT) and release within each cluster node, and IPAT from the primary to the standby node
- Manage application services via the Application Server resource, in order to:
 - Activate volume groups, mount filesystems, and start applications
 - Stop applications, unmount filesystems, and activate volume groups

Several ESS volumes are mirrored via PPRC from the primary site to the standby site. When node or site failure occurs, access to disk resources are not passed from one node to another. Instead, all highly available applications are restarted at the standby site using the data copy residing on the secondary volumes.

HACMP resources

Figure 101 is a sample Resource Group showing the resources for a cascading node relationship. Notice that it has two resources only:

- Service IP label

- Application Server

```

                                Configure Resources for a Resource Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
Resource Group Name                  geographic
Node Relationship                     cascading
Participating Node Names             melb_node1 syd_node1

Service IP label                      [melb_node1_svc]      +
Filesystems                          []                    +
Filesystems Consistency Check        fsck                  +
Filesystems Recovery Method          sequential            +
Filesystems to Export                []                    +
Filesystems to NFS mount             []                    +
Volume Groups                        []                    +
Concurrent Volume groups             []                    +
Raw Disk PVIDs                      []                    +
AIX Connections Services             []                    +
AIX Fast Connect Services            []                    +
Application Servers                  [pprc]               +
Highly Available Communication Links []                    +
Miscellaneous Data                  []                    +

Inactive Takeover Activated          false                 +
9333 Disk Fencing Activated          false                 +
SSA Disk Fencing Activated           false                 +
Filesystems mounted before IP        false                 +
configured

[BOTTOM]

F1=Help          F2=Refresh      F3=Cancel      F4=List
F5=Reset         F6=Command      F7=Edit       F8=Image
F9=Shell         F10=Exit        Enter=Do

```

Figure 101. Sample resource group for cluster spanning two sites

When node failover occurs, the service IP label and Application Server are the only resources taken over by the standby node. There are no shared disk resources to take over, because each site has its own local disk resources mirrored in real-time by PPRC to an ESS at the standby site.

Figure 102 also shows a two-node cluster, where each cluster node is located at a separate site. However, each site exists with different subnets. This is an important difference between the environment shown here and the one shown in Figure 100 on page 138, where the subnets are common to both sites.

HACMP does not support IP address takeover across different subnetworks, so no services address is available for IP address takeover. Unfortunately, this means that a service interface cannot be protected against network adapter failure, thus introducing a single point of failure. Still, we can define the subnetted domain as a service network for keepalive heartbeat traffic.

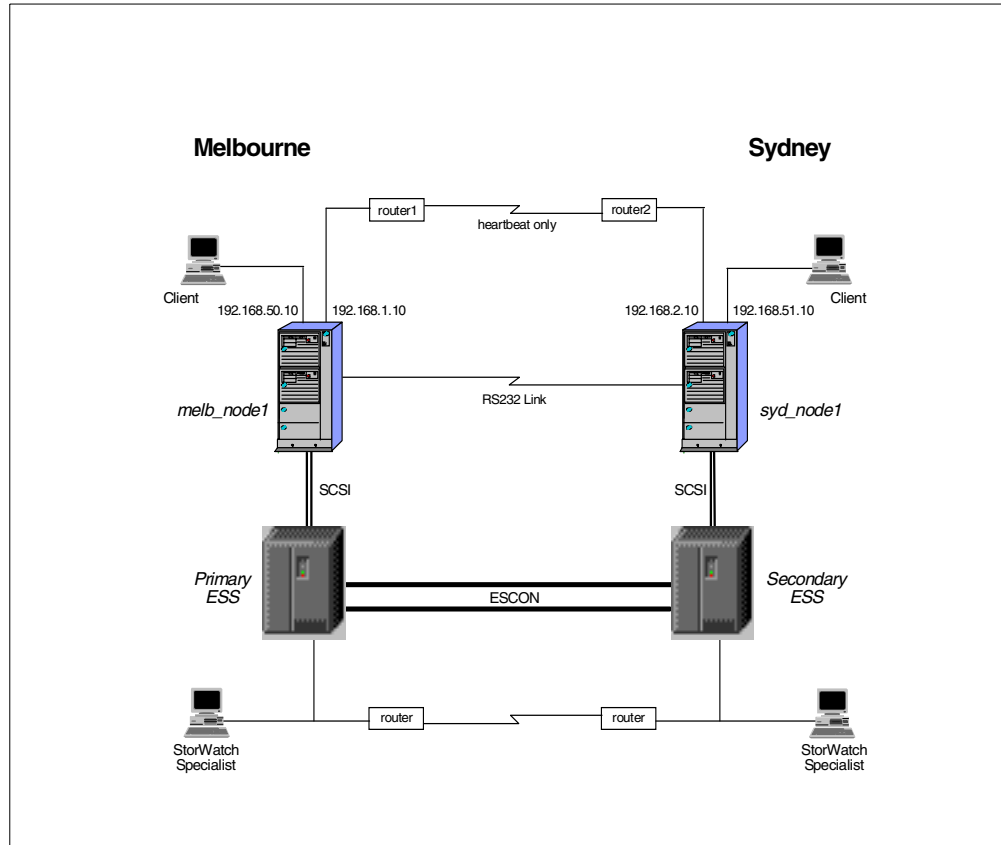


Figure 102. Two-node cluster spanning two sites with differing network domains

Each cluster node is set up with static routes to provide connectivity for an HACMP service network. This network cannot provide IP address takeover. For a heartbeat ring to be formed, router1 and router2 must permit the passage of UDP broadcast packets. See 7.1.5.1, “Using different subnets at each site” on page 148 for more information.

The primary node, melb_node1, at the primary site Melbourne, has the following static route to reach the network in use at Sydney.

```
route add -net 192.168.2.0 192.168.1.254 -netmask 255.255.255.0
```

The standby node, syd_node1, at the standby site Sydney, has one static route to reach the network in use at Melbourne.

```
route add -net 192.168.1.0 192.168.2.254 -netmask 255.255.255.0
```

Router 1 has 192.168.1.254 as the main address on its ethernet interface.

Router 2 has 192.168.2.254 as the main address on its ethernet interface.

HACMP resources

Figure 103 is a sample Resource Group showing the resources for a cascading node relationship. Notice that the resource group has **one** resource only: the Application Server.

```

                                Configure Resources for a Resource Group

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                [Entry Fields]
Resource Group Name                  geographic
Node Relationship                     cascading
Participating Node Names             melb_node1 syd_node1

Service IP label                      [] +
Filesystems                          [] +
Filesystems Consistency Check        fsck +
Filesystems Recovery Method          sequential +
Filesystems to Export                [] +
Filesystems to NFS mount             [] +
Volume Groups                        [] +
Concurrent Volume groups             [] +
Raw Disk PVIDs                      [] +
AIX Connections Services             [] +
AIX Fast Connect Services            [] +
Application Servers                  [pprc] +
Highly Available Communication Links [] +
Miscellaneous Data                   []

Inactive Takeover Activated          false +
9333 Disk Fencing Activated          false +
SSA Disk Fencing Activated           false +
Filesystems mounted before IP configured false +
[BOTTOM]

F1=Help          F2=Refresh          F3=Cancel          F4=List
F5=Reset         F6=Command          F7=Edit           F8=Image
F9=Shell         F10=Exit            Enter=Do

```

Figure 103. Resource group, cluster spanning two sites with differing network domains

We mentioned earlier that for cluster configurations where the disk resources are shared, the cluster manager controls activation of volume groups, and checking and mounting of file systems. For a dual-site configuration, where each server manages its own storage, the logic needed to accomplish these tasks must be provided elsewhere. Additional logic is also required to manage PPRC pairs. A good place to specify the logic is within the start and stop scripts of an HACMP Application Server definition.

7.1.4.1 HACMP Application Server

An HACMP Application Server consists of two main shell scripts, defined as the start and stop scripts. Copies of the scripts must reside on all cluster nodes. It is good practice to design the start and stop scripts so they will be identical on all cluster nodes.

Start script design

The start scripts should contain the site specific logic for site or node failover, re-integration and PPRC pair establishment and termination. It needs to perform the following main tasks:

1. Check for the existence of a reachable Copy Services server and use the appropriate server address.
2. Check that all ESS volumes (hdisks) are available.

3. At the **Primary Site** only:

- If the cluster is starting normally, then:
 1. Check if all required primary-to-secondary PPRC pairs have been established.
 2. Attempt to establish any suspended/terminated pairs which are part of the application's data set. If the pairs are suspended, they can be established with the option to "Copy out-of-sync cylinders". If the pairs are *simplex* source and targets, then pairs can be established with either the option to "Copy the entire volume" or the option to "Do not copy volume" but not with the option to "Copy out-of-sync cylinders".

Note that in our example start script, it is assumed that all pairs are synchronized before the cluster is started. Therefore, all pairs are established with the option to "Do not copy volume".
- If the primary node/site is being reintegrated into the cluster, then:
 1. Assume that the data resynchronization has been started manually by the storage administrator.
 2. Check for the completion of re-synchronization and continue only if resynchronization of all pairs is complete.

In our example start script, it is assumed that all pairs are synchronized before the cluster is started.
 3. When re-synchronization is complete, terminate the secondary-to-primary pairs.
 4. Re-establish all primary-to-secondary pairs.

4. At the **Standby Site only**: At failover, terminate all suspended primary-to-secondary PPRC pairs so that the secondary copy can be made available.

5. Activate volume groups.
6. Mount file systems.
7. Start applications.

Establishing and terminating PPRC pairs is initiated by executing a task from the Copy Services Command Line Interface. However, the Task or Task Group must first be created within Copy Services Specialist. See 7.1.6, "Creating the Copy Services tasks" on page 152 for details.

Stop script design

The stop script includes the logic to:

1. Stop all applications. Remember, you may need to include logic to determine if application-related processes still remain in the process list, and if so, then take the appropriate action to terminate them. Otherwise, your script may have difficulty unmounting filesystems.
2. Unmount file systems.
3. Deactivate volume groups.

Earlier we mentioned the need for your start and stop scripts to contain logic to manage the PPRC pairs. To manage the pairs, you need to create a set of pre-defined tasks that can be executed by CLI commands in your scripts. You will need to create tasks for establishing and terminating PPRC pairs.

The following section describes the Copy Services tasks required for automated management of the PPRC pairs.

7.1.4.2 Tasks for cluster start-up, failover and re-integration

Tasks are needed to handle different conditions during the following modes of cluster operation.

- Start-up
- Failover
- Re-integration

Task names can be no more than 16 characters in length. The following list of tasks uses an arbitrary naming convention. However, if you choose not to use these ones, you will need to modify the scripts in Appendix A, “HACMP start and stop scripts” on page 165.

The following list describes the Copy Services tasks needed for the different modes of cluster operation:

- **<VGname>EstDNCV** — This task establishes PPRC pairs for all disk volume pairs in the AIX volume group *VGname* with the condition “Do not copy volume”.
- **<VGname>EstSync** — This task establishes PPRC pairs for all disk volume pairs in the AIX volume group *VGname* with the condition “Copy out-of-sync cylinders”. This task can only be invoked on pairs that are in a suspended state.
- **<VGname>TermPriSec** — This task terminates PPRC pairs for all disk volume pairs in the AIX volume group *VGname* where remote copying is occurring from volumes in the primary ESS to volumes in the secondary ESS. In a real disaster, the primary ESS may be destroyed, so the task should be scheduled with the Target Storage Server.
- **<VGname>TermSecPri** — This task terminates PPRC pairs for all disk volume pairs in the AIX volume group *VGname* where remote copying is occurring from volumes in the secondary ESS to volumes in the primary ESS. The task should be scheduled with the Source Storage Server.

To see examples of creating the tasks, please refer to Section 7.1.6, “Creating the Copy Services tasks” on page 152.

Start-up

Before starting cluster services on the primary node, we assume that the data on the disk volume pairs is synchronized or that the pair is in a suspended state. The tasks to handle these conditions are:

- **<VGname>EstDNCV** to establish pairs with *Do not copy volume* option
- **<VGname>EstSync** to establish pairs with *Copy out-of-sync cylinders* option

If the cluster is starting normally or the primary node is joining the cluster with standby node down, check that all PPRC pairs are established. If not, then establish them.

You can see how these tasks are executed in the example start script shown in A.2, “startESSapps” on page 166,

Important Note

If you have any doubts about the synchronization state of any of the pairs, you must check them manually and synchronize them if necessary before you start Cluster Services.

Failover

When the primary node fails, the standby node needs to access the data on the secondary ESS. To make the secondary copies available to AIX, the secondaries must be terminated. You need to tasks that will *Terminate PPRC copy pair* and put the secondaries in a *simplex* state. An example of how this task is defined is shown in 7.1.6, “Creating the Copy Services tasks” on page 152.

The task `<VGname>TermPriSec` is executed to terminate the secondary volumes.

You can see how the `TermPriSec` task is executed in the start script shown in A.2, “startESSapps” on page 166.

Re-integration

As mentioned earlier, in our examples we have assumed that the resynchronization process will not be handled by the cluster’s recovery automation. Instead you need to complete this process manually.

Complete resynchronization of all ESS volumes, after re-integrating the repaired primary cluster node, ESS or reconstructed site, may be quite time consuming. While the resynchronization process is occurring, your applications will still be running at the standby site. If you think resynchronization will require too much time, you might consider rebuilding the volumes from tape.

Once your data has been synchronized, the primary to secondary pairs need to be re-established.

First, each pair is terminated to allow the primary volumes to become simplex. Use `<VGname>TermSecPri` for this purpose.

Then, because we know that each volume has been synchronized, all pairs should be re-established with the option *Do not copy Volume*. Use `<VGname>EstDNCV` for this purpose.

You can see how these tasks are executed in the start script shown in A.2, “startESSapps” on page 166,

It would be possible to allow HACMP to manage the resynchronization of data from the secondary ESS to the primary ESS. To achieve this, additional logic could be placed in the start script. If you allow the start script to handle the resynchronization of data, then it should also check for the completion of the resynchronization task. Our example start script does not contain the logic to achieve this function.

7.1.4.3 Example start and stop scripts

There are three scripts and they are described in this section. Copy the three scripts to the same directory on each of the two cluster nodes. For example, `/usr/es/sbin/cluster/local` is a suitable place.

1. `hacmpPPRC.vars`
2. `startESSapps`
3. `stopESSapps`

Please note the following important features and restrictions that occur as a consequence of the customization necessary to handle non-shared storage.

- You can add, change, or delete logical volumes or filesystems at the source ESS (primary site) and the updated definitions will be propagated to the target ESS (standby site) at failover time. This works because the start script calls `clvaryonvg`, which will cause volume group changes will be learned by the standby node on failover.
- You cannot use Cluster System Management commands (CSPOC) to manage the storage components of this environment.
- It is likely that the PPRC secondary volumes will appear to AIX as *Available* hdisks on the standby cluster node. If you reboot the standby node, the AIX Configuration Manager knows that these physical volumes already exist with entries in the Configuration Database. However, it cannot read their PVIDs because they are PPRC targets locked by the ESS. This results in AIX causing the original hdisks to be configured to a *Defined* state and new (phantom) hdisks are configured and made *Available*.

To make use of the secondary volumes, the phantom hdisks must be removed and the “real” hdisks must be changed from a *Defined* state to an *Available* state. At failover, the start script calls HACMP utilities to handle this undesirable condition. It changes the PPRC secondaries to a *simplex* state, removes the phantom hdisks and sets the original hdisks to an *Available* state.

- Since no filesystem, volume group or disk definitions appear in the resource group, a cluster verification will not provide information about the consistency of their definitions at each cluster node. You must play the role of watch-dog to ensure that all volume group and filesystem characteristics meet HACMP’s requirements.
- No specific limitation is placed on number of volume groups and filesystems.
- This script does NOT handle SDD (DPO) vpaths, though it could be modified to do so.

Please see Appendix A, “HACMP start and stop scripts” on page 165 for a complete listing of the start and stop scripts used in this chapter.

hacmpPPRC.vars

This file is called by `startESSapps` and `stopESSapps` to add the variables to the shell environment. It contains the profile to set up HACMP/PPRC dependencies and variables. Most of the variables that you need to change, and there are not many of them, are defined within `hacmpPPRC.vars`:

- AIX filesystems to be remote copied from the primary storage server to the secondary storage server.
- AIX Volume groups containing raw LVs to be remote copied from the primary storage server to the secondary storage server. If there are no raw LVs, then do not define any volume groups.
- Node names of your primary and standby cluster nodes.

- Hostnames of your cluster nodes.
- IP addresses of your primary and secondary ESSs.
- Username and password of your Copy Services server.

startESSapps

This is a shell script called by HACMP's `start_server` sub-event when cluster services are started on the primary node (site), or on the standby node (site) when the primary node fails (failover). It is defined to HACMP as the Start Script in an Application Server. It contains the shell code to manage PPRC pairs, activate volume groups, mount filesystems, and start applications. Because it relies heavily on calls to HACMP scripts and code, it will not work unless HACMP is installed on the hosts.

Important Note

Before `startESSapps` will function properly with HACMP, the following setup conditions must be met.

- All of the Volume Groups initially must be participating in PPRC pairs.
- Each PPRC pair needs to be terminated and the target Volume Groups must be imported on the standby cluster node so that ODM and `/etc/filesystems` have been populated.
- If the Target disk volumes were previously known to AIX, then they must be removed and re-configured before the running `importvg`. If this is not done first, `importvg` will import the volume group improperly. The volume group data structures (PVIDs, and so on) in ODM will differ from the data structures in the VGDA's and disk volume super blocks.

Place-holder variables are defined for names of the programs or scripts you wish to start. You can add more if you wish. Apart from the program or script names, `startESSapps` should not need further modification.

stopESSapps

This is a shell script called by the cluster manager when a node on which the cluster resources resides, is stopping. It is defined to HACMP as the Stop Script in the same Application Server. It contains the shell code to stop applications, deactivate volume groups and unmount of filesystems. Because it relies heavily on calls HACMP scripts and code, it will not work unless HACMP is installed on the hosts.

Place-holder variables are defined for names of the programs or scripts you wish to stop. You can add more if you wish. Apart from the program or script names, `stopESSapps` should not need further modification.

7.1.4.4 Additional script requirements

The two scripts, `startESSapps` and `stopESSapps`, rely on several HACMP utilities to check disk availability, activate and deactivate volume groups, mount and unmount filesystems, and so on. Remember that we had to externalize these functions from the resource group to handle disk resources that are not shared. Two of the utilities used in the scripts are:

- `/usr/es/sbin/cluster/utilities/cl_activate_fs`

- /usr/es/sbin/cluster/utilities/cl_deactivate_fs

Calling these scripts, when there are no shared filesystems, causes them to fail. The problem arises when the scripts attempt to determine the resource group name from the *HACMPresource* object class. The `odmget` command, shown in point 2 below, expects to find a shared filesystem name as an attribute in `/etc/objrepos/HACMPresource`. To work around this problem, do the following:

1. Copy `cl_activate_fs` to `/usr/sbin/cluster/local/esscl_activate_fs` and copy `cl_deactivate_fs` to `/usr/sbin/cluster/local/esscl_deactivate_fs`.
2. Make the following changes in `esscl_activate_fs` and `esscl_deactivate_fs`. Comment out the following line:

```
#RES_GRP=`odmget -q "name=FILESYSTEM AND value=$FILE1" \
HACMPresource | grep group | awk '{print $3}' | sed 's///g'`
Then add the following line under the line you just commented out.
RES_GRP=${GROUPNAME}# GROUPNAME is exported by
# node_up_complete and node_down_complete
```

Important Note

The modifications to `esscl_activate_fs` and `esscl_deactivate_fs` have been tested for a cluster where only **one** resource group is configured for failover of nodes in a given resource chain; for example, failover from melbourne to sydney [`melb_node1` `syd_node1`]. The modifications do not cater to more than one resource group.

An alternative to using the modified HACMP utilities is to replace the `esscl_activate_fs` and `esscl_deactivate_fs` script calls in `startESSapps` and `stopESSapps` with your own filesystem check, `mount`, and `umount` commands.

7.1.4.5 Cluster reconfiguration duration

For some application environments, cluster event processing may take some minutes to complete. The time required to process events may cause the cluster to think it has been reconfiguring itself for too long. Under these circumstances, you could expect the cluster managers to call `config_too_long`, thus generating a message similar to the following:

```
WARNING: Cluster geographic has been running recovery program start_server for
540 seconds. Please check cluster status.
```

The cluster will begin generating this message after the default timeout of 360 seconds. Unless another problem has occurred, you can safely ignore this warning. The warning is cancelled once the `node_up` event completes successfully.

If you would like the cluster managers to avoid generating the warning message unnecessarily, you can change the value of the threshold time, at which point the cluster managers will begin executing `config_too_long`. The following command changes the threshold value:

```
chssys -s clstmgr -a "-u <milli-seconds_to_wait>"
```

For example, to change the threshold value to 10 minutes, login as root user and type the following command:

```
chssys -s clstmgr -a "-u 600000"
```

Important Note

Only increase the cluster manager threshold timer value if absolutely necessary. Set it to a value (plus headroom) that accommodates reasonable event processing duration. Otherwise you may be waiting for an unreasonable duration to discover if an event has failed.

7.1.4.6 Error notification

To cover total failure of the primary ESS, HACMP could promote its failure to a node/site failure. This allows rapid recovery of application services through the secondary ESS at the standby site.

The idea is that the primary node is brought down gracefully, then the standby takes over the workload.

Unfortunately, we were unable to set up and test the idea. Nevertheless, the approach should be technically feasible.

7.1.4.7 Dual site recovery with three or more cluster nodes

Configuring a cluster containing three or more cluster nodes may be possible, but it will require a significant amount of customization that may not be feasible or practical.

Issues that arise will specifically relate to configuring storage. As mentioned earlier, HACMP is designed to handle the failover of any disk resource between cluster nodes. However, HACMP is not currently designed to handle the case where one server is connected to disk resources and another server is connected to different disk resources which have the same volume group, logical volume and even filesystem names.

7.1.5 Implementation specifics

Before you implement Copy Services in a high availability environment, we strongly recommend that you read the following sub-sections.

7.1.5.1 Using different subnets at each site

Many DR environments have the requirement for different subnets at each site. When HAGEO for AIX was announced in 1996, support in HACMP to allow two different subnets to be connected, followed shortly thereafter. This feature means that you are able to configure a service network of two different subnets for the passage of keep alive heartbeats. Now, in general, routers do not pass the UDP broadcast packets that HACMP nodes use to discover each other. Fortunately, most modern routers have the capability to configure broadcast packet handling in a controlled way that avoids the potential broadcast storm that could otherwise result.

The current broadcast address standard provides specific addressing schemes for forwarding broadcasts. Detailed discussions of broadcast issues in general can be found in RFC919, "Broadcasting Internet Datagrams", and RFC 922, "Broadcasting IP Datagrams in the Presence of Subnets".

Modern router software usually provides facilities to allow UDP broadcasts. For example, Cisco routers provide the following functions:

- Enabling forwarding of IP directed broadcasts on an interface where the broadcast becomes a physical broadcast.
- Specifying a UDP destination port to control which UDP services are forwarded. HACMP uses specific ports for its heartbeat protocols.
- Allowing IP broadcasts to be flooded throughout your network in a controlled fashion using the database created by the spanning-tree protocol.

7.1.5.2 Failure detection

Because a dual site configuration requires network connections over larger distances, a higher level of latency may exist in the network compared to a LAN. This may cause delays in the transmission and receipt of heartbeat keepalives between cluster nodes. HACMP provides the means to vary the Failure Detection Rate for the range of networks it supports, thus allowing the detection of network failure to be tuned to suit the behavior of the network. The standard settings of Slow, Normal, and High vary depending on the network type — Ethernet, ATM, FDDI Token Ring, and so on.

However, the minimum, standard settings of Slow may be insufficient for your network environment. If the Failure Detection Rate is too sensitive for your environment, unnecessary and undesirable *swap_adapter* events can occur. The adapter swap will compound any network delays caused by network latency because of the finite time it takes for the event to complete. You can tune the sensitivity of failure detection outside of the standard values presented in `smit`, by altering the HACMPnim object class.

To make failure detection less sensitive, use the following procedure. This is an example taken from the *High Availability Cluster Multi-Processing for AIX, Troubleshooting Guide, Version 4.3.1*.

In the HACMPnim ODM class, the Failure Detection Rate is made up of two components:

- Heartbeat rate (`hbrate`): the number of microseconds between heartbeats
- Cycles to fail (`cycle`): the number of heartbeats that must be missed before detecting a failure

Together, these two values determine the Failure Detection Rate. For example, for a NIM with an `hbrate` value of 1,000,000 microseconds and a `cycle` value of 12, the Failure Detection Rate would be 12 (1 second x 12 cycles).

Before altering the Failure Detection Rate, note the following:

- Before altering the NIM, you should give careful thought to how much time you want to elapse before a real node failure is detected by the other nodes, and the subsequent takeover is initiated.
- It is recommended that you first set the Failure Detection Rate to Slow to change the `hbrate`, and then adjust the `cycle` value as needed to reach the desired Failure Detection Rate.
- The Failure Detection Rate must be set equally for the corresponding NIM on each cluster node. Therefore, the change must be synchronized across

cluster nodes. The new values will become active the next time cluster services are started.

To alter the Failure Detection Rate outside of the definitions provided by `smit`:

1. Identify the NIMs to be modified. All NIMs used in the cluster should be included.

To determine the NIMs in use, check the output from the command:

```
/usr/sbin/cluster/utilities/cllsif . For example:
```

```
# cllsif -cS|cut -d':' -f4|sort -u
ether
rs232
```

2. For each NIM you are modifying, save the HACMPnim information as follows:

```
# odmget -q name=NetworkType HACMPnim > /tmp/NetworkType.out
```

For example:

```
# odmget -q name=ether HACMPnim > /tmp/ether.out
# odmget -q name=rs232 HACMPnim > /tmp/rs232.out
```

3. Edit the generated files to alter the `cycle` values for the NIMs. Increase the `cycle=` value so that multiplying it by the corresponding `hbrate` value produces the desired failure detection rate.

In this example, to increase the failure detection rate to 30 seconds, you change the `cycle` for `ether` to 30 and the `cycle` for `rs232` to 10.

```
/tmp/ether.out:
HACMPnim:
  name = "ether"
  desc = "Ethernet Protocol"
  addrtype = 0
  path = "/usr/sbin/cluster/nims/nim_ether"
  para = ""
  grace = 30
  hbrate = 1000000
  cycle = 30
```

```
/tmp/rs232.out
HACMPnim:
  name = "rs232"
  desc = "RS232 Serial Protocol"
  addrtype = 1
  path = "/usr/sbin/cluster/nims/nim_sl"
  para = ""
  grace = 30
  hbrate = 3000000
  cycle = 10
```

- For each NIM, change the HACMPnim ODM class on one node using `odmchange -o HACMPnim -q name=NetworkType < /tmp/NetworkType.out`

For example:

```
# odmchange -o HACMPnim -q name=ether < /tmp/ether.out
# odmchange -o HACMPnim -q name=rs232 < /tmp/rs232.out
```

4. Synchronize the Cluster Topology and resources from the node on which the change occurred to the other nodes in the cluster.
5. Run `clverify` to ensure that the change was propagated.

6. Restart the cluster services on all nodes to make the changes active.

7.1.5.3 Multiple LUNs per SCSI ID

We included this section because, at the time this book went to print, a PTF was not available to address the following condition.

Once you have read this section and determined whether it applies to your environment, you should check the current status of any fixes and review any associated documentation.

Under certain, very limited circumstances, HACMP configurations with multiple LUNs per SCSI ID could experience data loss or corruption.

It should be emphasized that this is not a problem with ESS or other devices, nor is it a general problem with HACMP configurations.

It was discovered that HACMP does not properly handle configurations with multiple LUNs per SCSI ID, in the case where the LUNs correspond to disks that are in more than one AIX Volume Group. When HACMP breaks the SCSI reserve on a LUN (hdisk) during failover, the reserve is broken on all LUNs on that SCSI ID. If the hdisks corresponding to separate LUNs on a given SCSI ID are part of separate Volume Groups, it is possible that data loss or corruption can occur, as resets from one cluster node interfere with the I/O from other non-cluster nodes. If a “SCSI reserve” is absent, then a disk can be left unprotected from accidental access by other host servers. It is expected that RVSD has the same vulnerability.

The problem could manifest itself in two ways:

- Data lost as it is written to disk
- Disks left unreserved, and hence vulnerable to inadvertent modification

The first form of the problem could occur if **all** of the following conditions are present:

- An HACMP mutual takeover configuration (concurrent mode and hot standby configurations are not exposed to this problem).
- A mapping of disks to LUNs such that, for a given SCSI ID, the disks in the LUN for that ID are not all part of the same volume group.
- More disks on the given SCSI ID that are part of a single volume group than twice the queue depth for the adapter.
- I/O in progress to disks on that SCSI ID that HACMP is not taking over at the same time as HACMP is taking over other disks.

The second form of the problem could occur if **all** of the following conditions are present:

- Three or more systems connected to the same SCSI disk enclosure
- A mapping of disks to LUNs such that, for a given SCSI ID, the disks in the LUN for that ID are not all part of the same volume group.
- Human error on one of those systems in the window of time between when a node failure occurs, and the other node — the one that “rightly” owns the disks — once again does I/O.

When either form of the problem occurs, the error log will record a write failure to disk. For configurations that are potentially exposed, both forms of this problem can be completely avoided by ensuring that the disk configuration is such that all the disks that are LUNs on a given SCSI ID are part of the same volume group.

7.1.6 Creating the Copy Services tasks

The start and stop scripts, `startESSapps` and `stopESSapps`, rely heavily on the Copy Services Command Line Interface (CLI). They manage PPRC functions as part of the automation required to provide unattended failover or reintegration of application services. A Copy Services function must be defined as a Task in Copy Services before it can be executed by the CLI. The Copy Services Specialist is the only interface through which you can create tasks for execution by the CLI.

7.1.6.1 Tasks to establish PPRC pairs

We have two AIX volume groups: `evg001` and `evg002`. Volume group `evg001` consists of two hdisks and volume group `evg002` has one hdisk. Each hdisk is defined by an ESS logical volume or LUN. These are the primary volumes.

In this example, we create a task to establish the PPRC pairs for the AIX volume group `evg001`. The task defines Copy Initialization with the option “Do not copy volume”. One task is defined for copying `evg001` and another for `evg002`.

The following figures guide you through the steps required to create a Task for establishing the PPRC pairs for `evg001` only. Simply repeat the steps for `evg002`.

Important Note

You must ensure that all Volume Assignments have been performed before you start Copy Services on the ESS. If you have made changes to Volume Assignments since starting Copy Services, you need to log on to the ESS then stop and restart Copy Services.

1. Start the Copy Services Specialist, then select the **Volumes** panel as shown in Figure 104.
2. Select the required Source LSS from the left-side pull-down and select the required Target LSS from the right-side pull-down.

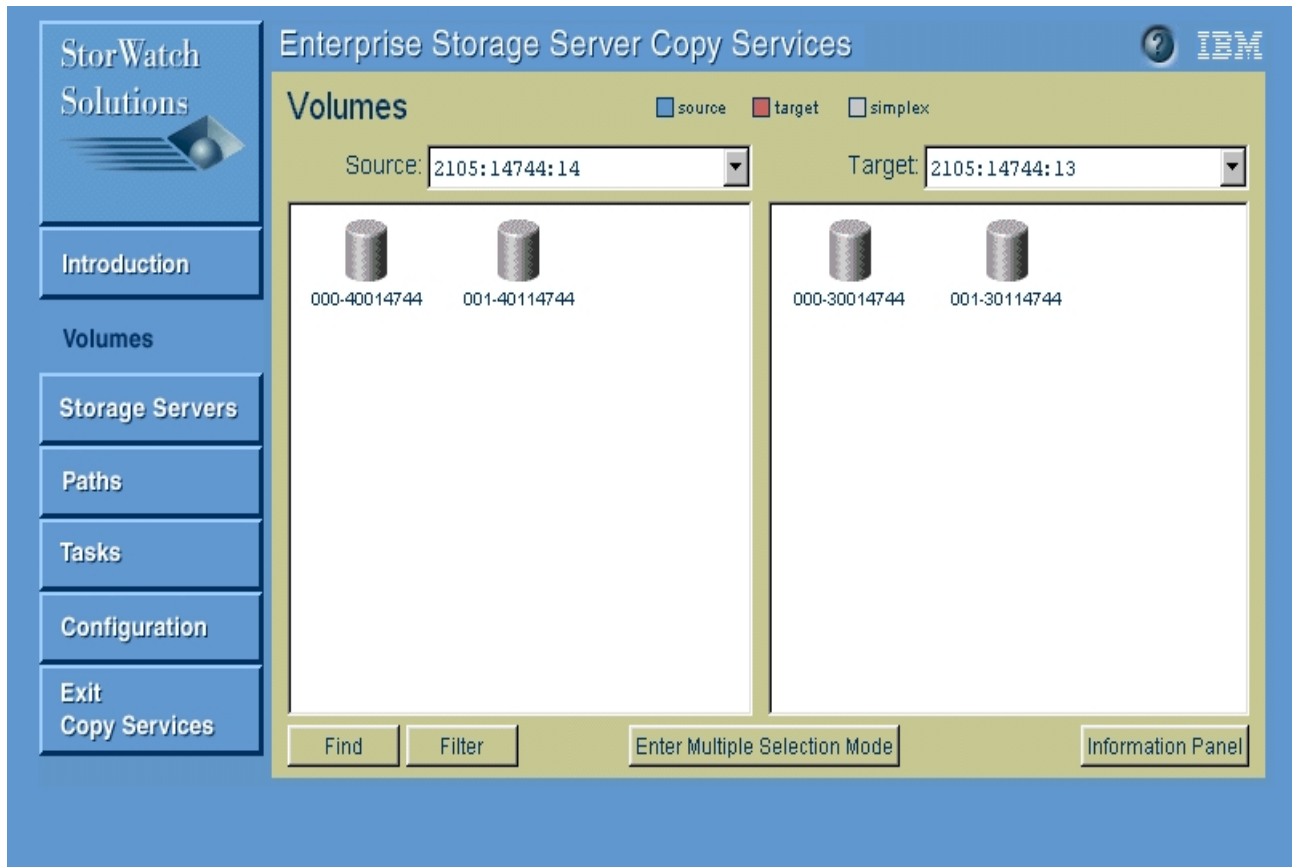


Figure 104. Copy Services Specialist, Volumes screen

3. Click **Enter Multiple Selection Mode**. The button is replaced by **Exit Multiple Selection Mode**.
4. For the first volume pair, left-click on the source volume (40014744), then right-click on the target volume (30014744). Repeat this for the second volume pair in the volume group (40114744 and 30114744). A second right-click on the target volume (30114744) causes the Task Wizard to be displayed. Select **Establish PPRC copy pair**, then click **Next**. (See Figure 105.)

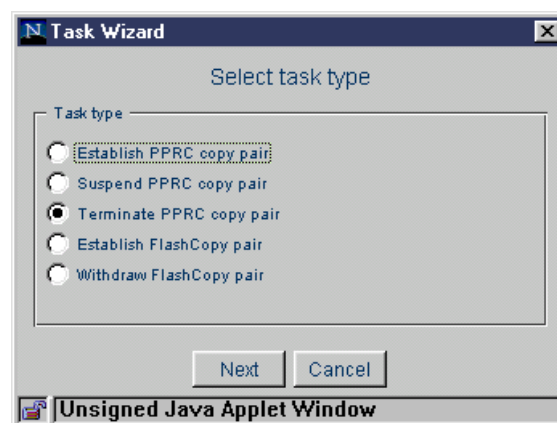


Figure 105. Establish task, select task screen

- When the Select Copy Options panel opens, click on Do not copy volume, then click Next. (See Figure 106.)

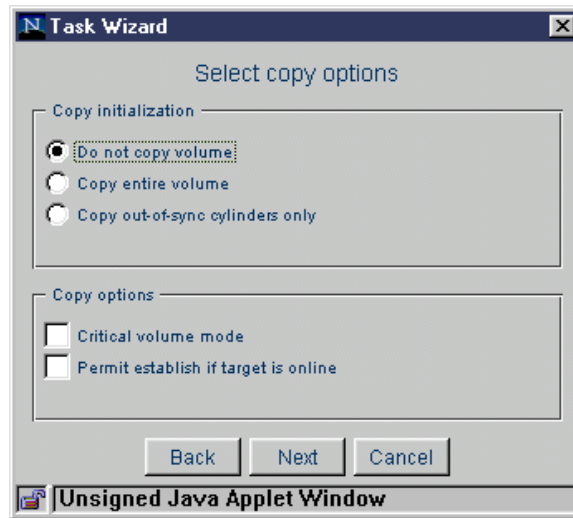


Figure 106. Select copy options screen

- It is always a good idea to name the tasks based on a meaningful naming convention. Even though it really does not matter how you name the tasks, we highly recommend using a consistent, documented approach that will assist you and other system administrators in the future. Here, we have used the AIX volume group name containing the volumes, *evg001*, and a descriptor for the type of task performed, *EstDNCV*.

Type in a Task Name and a Task Description, then click **Save**. (See Figure 107.)

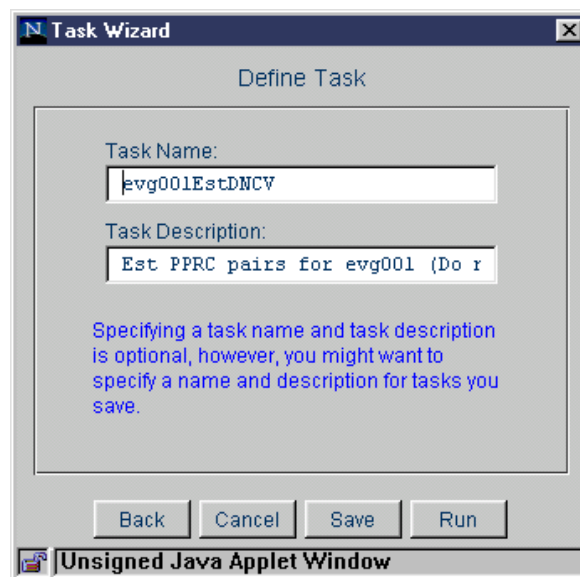


Figure 107. Define task name

- Now you have a task that will establish PPRC pairs for the volumes in evg001. (See Figure 108.)

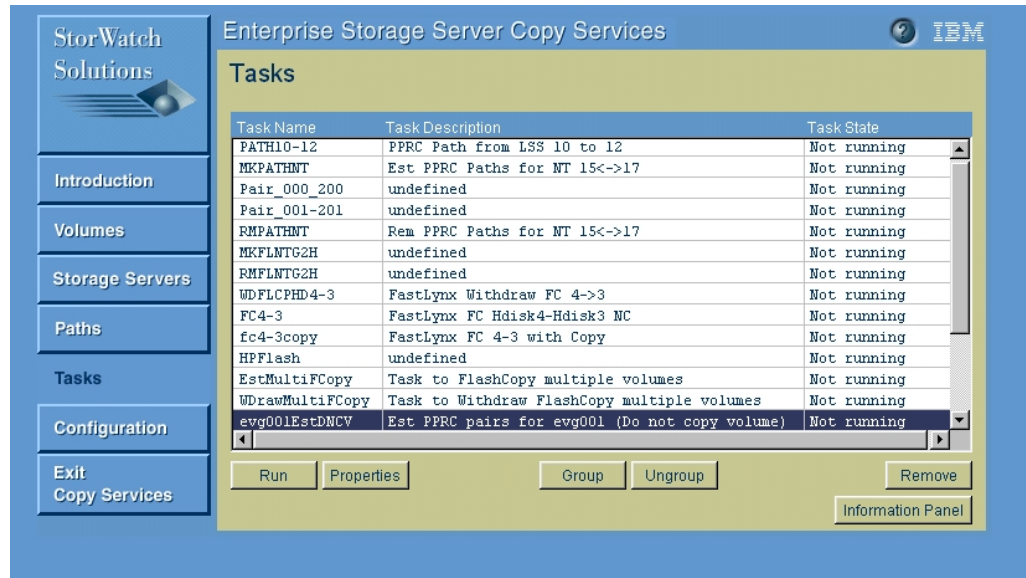


Figure 108. List of defined tasks

Repeat step 2. on page 147 through step 7. on page 155 for all the PPRC copy pairs you require.

The procedure to create a task to *Establish PPRC copy pair* with the option to *Copy out-of-sync cylinders* is similar to creating the task outlined above. In the Select Copy Options wizard, click on *Copy out-of-sync cylinders* instead of *Do not copy volume* and name the task <VGname>EstSync.

7.1.6.2 Task to terminate PPRC pairs

Remember that secondaries must be terminated to make the volumes available to AIX.

Having created the tasks for establishing PPRC pairs, it is easy to create the tasks for terminating the pairs. For example, use the previously defined task that establishes pairs for *evg001* to create the task for terminating the pairs. Do this by modifying the task and saving it as a New task.

Here are the necessary steps.

1. At the Copy Services panel, select **Tasks**. (See Figure 109.)



Figure 109. Selecting task to run

2. Highlight the task `evg001EstDNCV`, then click **Properties** to reveal the Select task type wizard. (See Figure 110.)

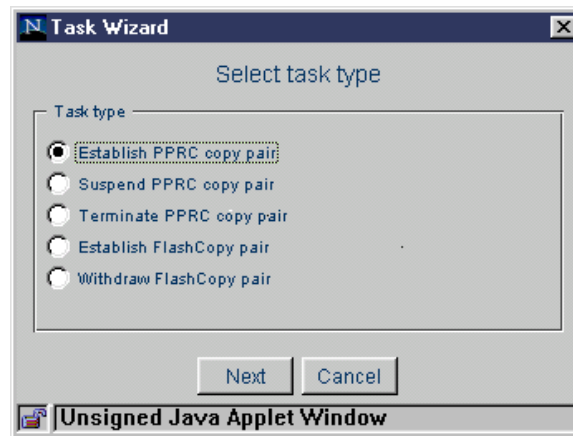


Figure 110. Select task type

3. Click **Terminate PPRC copy pair**, then click **Next**. (See Figure 111.)

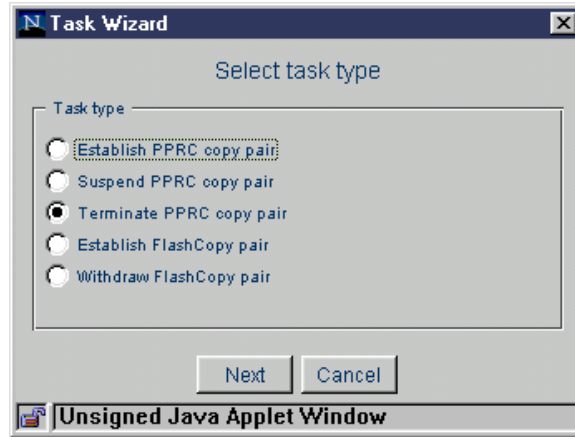


Figure 111. Terminate PPRC

This task must be scheduled with the secondary (target) storage server. Click **Schedule task with target storage server**, then click **Next**. (See Figure 112.)

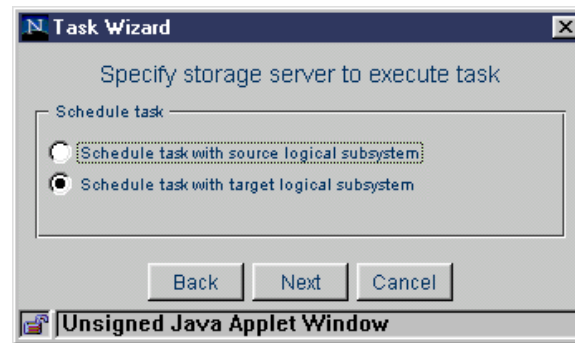


Figure 112. Execute task screen

Type in a Task Name and a Task Description, then click **New**. (See Figure 113.)

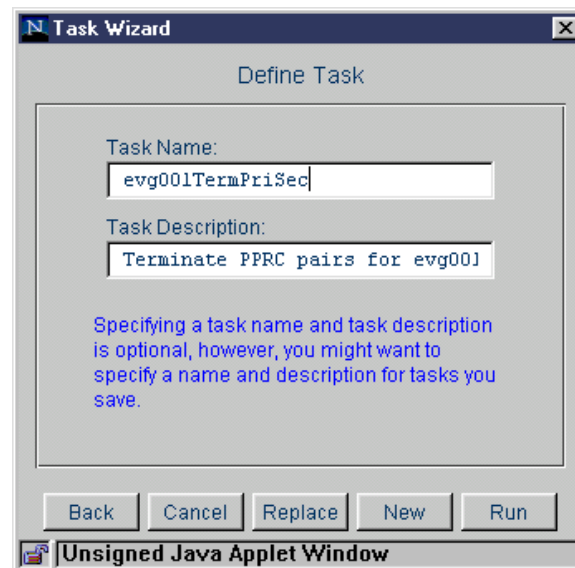


Figure 113. Define task screen

Now you have a task for terminating the pairs for evg001. (See Figure 114.)

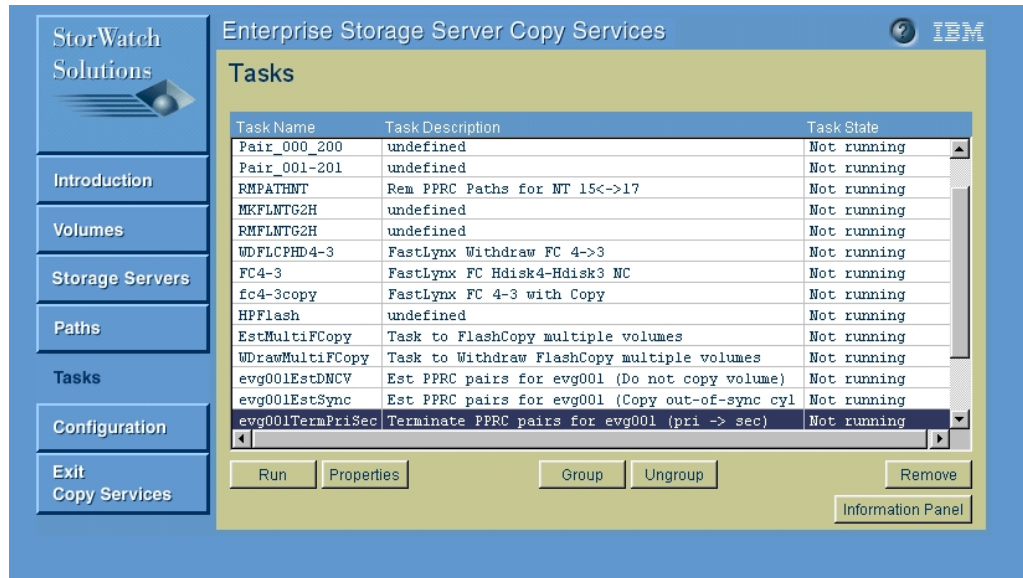


Figure 114. Task list screen

To create a task that terminates a pair that is defined for copying in the reverse direction, that is, from the secondary ESS to the primary ESS, you need to perform the following steps.

1. Ensure that you have at least one ESCON path defined in the reverse direction.
2. Establish a PPRC copy pair with the option to *Do not copy volume*.
3. Create a task to Terminate PPRC copy pair and save it with the name <VGname>TermSecPri, for example, evg001TermSecPri.
4. Execute evg001TermSecPri to terminate the pairs.

Figure 115 shows an example list of all the tasks needed for both volume groups, evg001 and evg002.

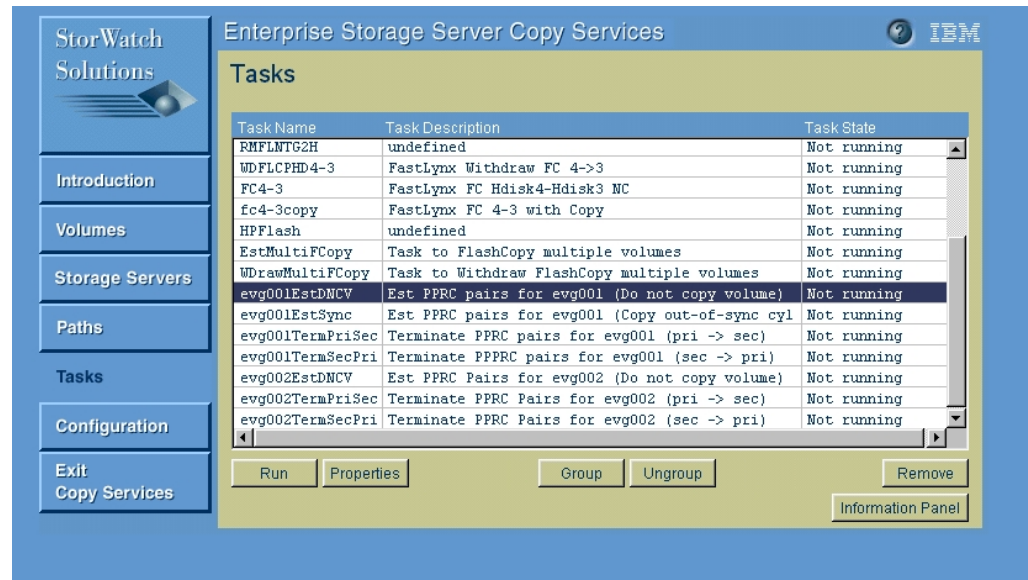


Figure 115. Selected task highlighted

7.1.7 Maintenance procedures

From time to time, you may want to add, change or delete logical storage within your production environment. Because management of storage resources in our specific example is external to the cluster configuration, you can make your changes with cluster services either running or stopped.

7.1.7.1 Adding a volume group

To add a volume group the cluster, you need to do the following things:

1. Review Section 7.1.4.2, "Tasks needed for cluster start-up, failover and re-integration".
2. On the source ESS, allocate the required storage and create the LUN(s) that will be used in the new volume group.
3. On the target ESS, allocate the storage and create LUN(s) of equal size to the corresponding source LUN(s) you created in step 2.
4. Stop Copy Services and restart it so that Copy Services is made aware of the new storage allocations.
5. Execute `rsPrimeServer.sh` on each host server to allow Copy Services to label the new LUNs with `hdisk` names.
6. On the primary host server, create the new volume group. Remember to choose one major number that is unique to both host servers and ensure that the volume group is set to not automatically activate at system restart.
7. Create the logical volumes with LV names that are unique to both host servers. Then create the filesystems on the "previously defined logical volumes". Remember to ensure that each filesystem is set to not automatically mount at system restart.
8. Rename the JFS log LV to a name that is unique to both host servers and change each filesystem's characteristics to point to it. For example:

```
# chfs -a log=/dev/elv1000 /u04
```
9. Confirm that all the new filesystems can be mounted, then unmount them.

10. Follow the procedure in Section 7.1.6, "Creating the Copy Services Tasks" and create the required Copy Services tasks.
11. Establish PPRC with each volume pair (hdisk pair) using the "Copy entire volume" option.
12. Use the Copy Services Specialist to monitor synchronization progress, then terminate each of the PPRC pairs once they are synchronized.
13. On the standby server, remove (`rmdev -dl`) the PPRC target hdisks if they are already configured.
14. Run configuration manager on the standby host server to reconfigure the target physical volumes (hdisks) with the new LVM data structures copied from their corresponding source physical volumes.
15. Confirm that the new physical volume definitions show PVIDs identical to those of the source physical volumes. If a PVID does not match its source, then a target may not have not been copied from its respective source or you have identified the incorrect disk.
16. Import the volume group to the standby server, for example:

```
# importvg -V60 -y evg005 hdisk28
```
17. Confirm that all the new filesystems can be mounted, then unmount them.
18. Deactivate (`varyoffvg`) the new volume group.
19. Establish the PPRC pairs with "Do not copy volume" by executing your newly created task `<VGname>EstDNCV`.
20. Add the filesystem names to the `FS_LIST` variable in the environment script `hacmpPPRC.vars`, or if using raw LVs, add the volume group name to the `VG_LIST` variable.
21. Copy `hacmpPPRC.vars` to the appropriate the directory primary host server and the standby host server.
22. On the primary host server, activate the volume group and mount the new filesystems.

7.1.7.2 Adding a filesystem to an existing volume group

Adding a filesystem is simple. There is no requirement to modify PPRC pairs, if adding the new filesystem does not involve allocating additional storage. The filesystem's data structures and data you add to the filesystem, will be copied to the target volumes as soon as you define it and start using it. Nevertheless, you do need to tell ODM and `/etc/filesystems` on the standby host server about the new filesystem.

Here are the steps you need to follow:

1. On the primary host server, create a logical volume with LV name that are unique to both host servers. Then create the filesystem on the "previously defined logical volume". Remember to ensure that the filesystem is set to not automatically mount at system restart.
2. Add the filesystem names to the `FS_LIST` variable in the environment script `hacmpPPRC.vars`.
3. Copy `hacmpPPRC.vars` to the appropriate the directory on primary host server and the standby host server.
4. Mount the new filesystem.

The standby host server now needs to know about the new filesystem. This can be done in the usual way by importing the volume group (`importvg -L`). If you cannot deactivate the volume group containing the new filesystem, then it can be imported to the standby server while on-line to the primary host server.

5. On the primary host server, unlock the volume group. For example:

```
# varyonvg -b -u evg005
```

6. On the standby server, import the volume group. For example:

```
# importvg -L evg005 hdisk28
```

7. On the primary host server, lock the volume group by reestablishing the SCSI reserves on the disks. For example:

```
# varyonvg evg005
```

Your volume group definitions now are synchronized on both host servers and the new filesystem is available to production.

7.1.7.3 Deleting a filesystem

Deleting a filesystem is the reverse of adding a filesystem. There is no requirement to modify PPRC pairs. The filesystem's data structures will disappear from both the source and target volumes as soon as you delete the filesystem. Nevertheless, you do need to tell ODM and `/etc/filesystems` on the standby host server that the filesystem has been deleted.

Here are the steps you need to follow:

1. Delete the filesystem name from the `FS_LIST` variable in the environment script `hacmpPPRC.vars`.
2. Copy `hacmpPPRC.vars` to the appropriate the directory on the primary host server and the standby host server.
3. On the primary host server, unmount the filesystem.
4. Delete the filesystem (`rmfs -r`).

The standby host server now needs to know that the filesystem has been deleted. This can be done in the usual way by importing the volume group (`importvg -L`). If you cannot deactivate the volume group containing the new filesystem, then it can be imported to the standby server while on-line to the primary host server.

5. On the primary host server, unlock the volume group. For example:

```
# varyonvg -b -u evg005
```

6. On the standby server, import the volume group. For example:

```
# importvg -L evg005 hdisk28
```

7. On the primary host server, lock the volume group by reestablishing the SCSI reserves on the disks. For example:

```
# varyonvg evg005
```

Your volume group definitions now are synchronized on both host servers.

7.2 Implementing Microsoft Cluster Server with Copy Services

Microsoft Cluster Server for Windows NT (MSCS) that is supplied as a part of Windows NT Enterprise Edition is being used for high availability (HA), however, it cannot be easily used for disaster recovery (DR).

The difference between HA and DR can be understood as follows: HA is usually considered an automatic takeover of resources in case one server fails. DR is often viewed as a human-assisted start of processing on the recovery site after a responsible person declares a primary site failure.

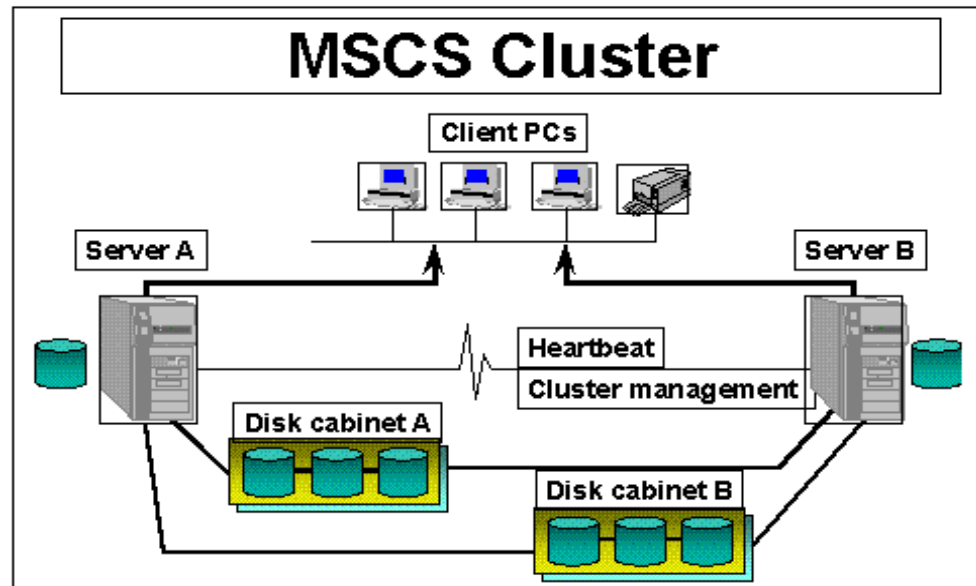


Figure 116. Microsoft Cluster Server

MSCS requires at least one disk accessible from both nodes (originally this meant the disk has to be on a SCSI bus shared by both servers) and places a *quorum* file containing the cluster configuration and resource list on the first partition of that shared disk. That partition is not recommended to be used for anything else — the data that will be accessible to cluster nodes has to be on another partition or another shared disk. See Figure 116.

On that shared disk, MSCS applies its *challenge/defense* protocol, which works as follows: SCSI-2 has reserve/release verbs with a semaphore on the disk controller. The owner of the disk controller gets a “lease” on the semaphore, which it can renew every three seconds. To preempt ownership, a challenger clears the semaphore with a SCSI bus reset, waits ten seconds (three seconds for renewal and two seconds for bus-settle time — twice, to give the current owner two chances to renew). If the semaphore is still clear, the challenger takes the lease from the former owner by issuing a reserve to acquire the semaphore. See Figure 117.

A SCSI Reserve command is issued to the SCSI target ID of the disk (note that you cannot reserve a LUN). ESS itself does not require the bus to be shared as it replicates the SCSI Reservation to all buses for which the SCSI ID of the volume is visible.

In a PPRC configuration, the primary ESS cannot propagate the SCSI Reserve to the secondary volumes that are offline.

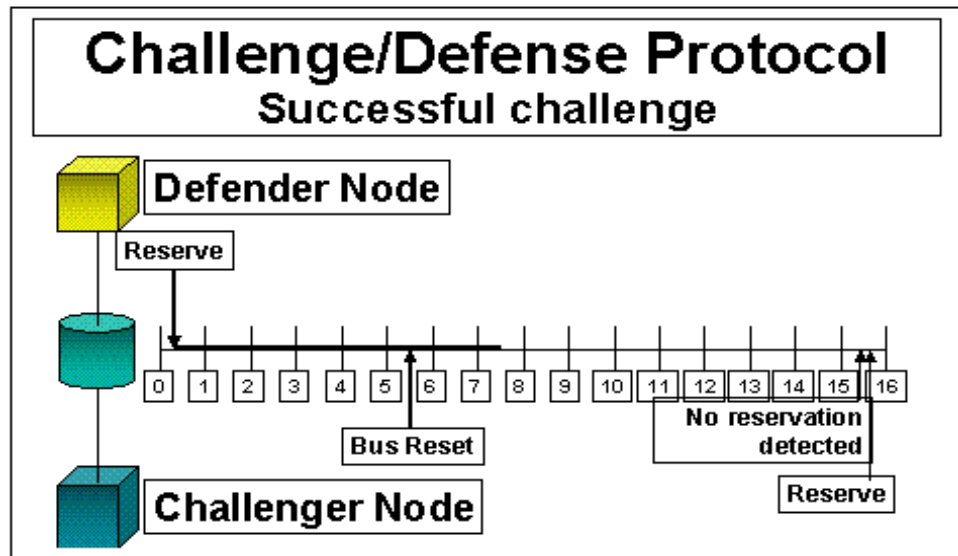


Figure 117. Challenge /defense protocol of MSCS

MSCS allows you to set up a two-node cluster only. That means any failure in the Windows NT server with cluster nodes located one at the primary and the other one at the secondary site would initiate a site failure. It will require you to use Fibre Channel for host connection to the ESS (native, or SAN Data Gateway) and link the two sites not only with ESCON but also with Fibre Channel.

We recommend that you set up a cluster on the primary site and a backup server or another pair of clustered servers on the secondary. With this, you can manually failover to the secondary site after declaring a site failure.

Appendix A. HACMP start and stop scripts

In this appendix, we provide code listings for HACMP start and stop scripts.

A.1 hacmpPPRC.vars

```
#!/usr/bin/ksh
#-----
# Name:      hacmpPPRC.vars
# Nodes:
# Path:      /usr/sbin/cluster/local
# Description:Profile to set up HACMP/PPRC dependencies and variables
# Written by: Andrew Beyer, IBM Australia
#-----
#
#####
#
# Notes:
# -----
# 1. Set file mode to 700 to avoid exposing the Copy Services username and
#    password.
#
#####
# Pre-requisites:
# -----
# Replace the values in:
# 1. FS_LIST with your filesystems; ordering is unimportant as we sort later.
#
# 2. VG_LIST with your Volume Groups containing raw LVs only. If there are no
#    raw LVs, then leave VG_LIST unchanged (null).
#
# 3. Copy Services Tasks must be created with the volume group name prefixed
#    to the specified task name. The overall length of the task name must be
#    no more than 16 characters, eg:
#    For a volume group named evg001, the task names are:
#    evg001EstSync, evg001EstDNCV evg001TermPriSec evg001TermSecPri
#
# 4. PRIMARY_NODE and STANDBY_NODE with the node names of your primary and
#    standby cluster nodes.
#
# 5. PRIMARY_HOST and STANDBY_HOST with the hostnames of your cluster nodes.
#
# 6. PRIMARY_ESS and STANDBY_ESS with the IP addresses of your primary and
#    secondary ESSs.
#
#-----
# ESS & cluster node variables
#-----

FS_LIST="/u01 /u02 /u03 /u04"

VG_LIST=""

PRIMARY_NODE=korca
PRIMARY_HOST=${PRIMARY_NODE}
```

```

PRIMARY_ESS=sfin6c1
STANDBY_NODE=fastlynx
STANDBY_HOST=${STANDBY_NODE}
STANDBY_ESS=sfin6c2

username=storwatch
password=specialist

#
-----
# Modify the following constants with care and only if necessary.
# These are needed to setup the environment & permit calls to HACMP event
# scripts work properly.
#-----

export PATH="/usr/sbin/cluster/utilities/cl_get_path all"
export VERBOSE_LOGGING=${VERBOSE_LOGGING:-high}
[ "$VERBOSE_LOGGING" = "high" ] && set -x
export EXPORT_FILESYSTEM=${EXPORT_FILESYSTEM:-''}

# Paths to Copy Services & HACMP utilities
#-----

# Copy Services TASK_LIST format:
# "CSsyncTask CSNoCopyTask CSTermPriSec CSTermSecPri"
# The volume group is prefixed to the task name later to make it unique.
TASK_LIST="EstSync EstDNCV TermPriSec TermSecPri"

CLI_CMDS="/usr/opt/ibm2105cli/" # Install directory for CLI scripts & code
HA_UTILS="/usr/sbin/cluster/utilities/"
CL_UTILS="/usr/sbin/cluster/events/utils/"
LOCAL="/usr/sbin/cluster/local/"

```

A.2 startESSapps

```

#!/usr/bin/ksh
#####
# NOTE: This script suitable for a two node cluster only where each node is
# connected to a separate ESS.
#
# Name:          startESSapps.sh
#
# Description:This script executes on the primary node when the primary node
#              has joined the cluster and on the standby node when the
#              primary node has failed. It manages the PPRC pairs, activates
#              volume groups, mounts filesystems and starts applications.
#
# Called by:    start_server
#
# Calls to:     get_local_nodename, cl_get_path, cl_fs2disk, cl_disk_available,
#              cl_activate_vgs, cl_sync_vgs, esscl_activate_fs,
#              rsTestConnection.sh, rsExecuteTask.sh, rsQueryComplete.sh,
#              rsQuery.sh
#
# Arguments:    none
#

```



```

# Written by: Andrew Beyer, IBM Australia
#
#####
# Notes:
# -----
# 1. This script will work only on hosts on which HACMP for AIX is installed.
# It calls utilities that are packaged only with HACMP.
#
# 2. Before this script will function properly with HACMP, the following
# prerequisites must be performed.
# All of the Volume Groups must be participating in PPRC pairs.
# Each pair needs to be terminated and their Targets must be imported as
# Volume Groups so that ODM and /etc/filesystems have been populated.
# If the Target disk volumes were previously known to AIX, then they must
# be removed and re-configured before the running importvg. If this is
# not done first, the PVIDs in ODM will differ from those in the VGDA and
# disk super blocks and importvg will occur improperly.
#
# 3. Because this script calls clvaryonvg, VG changes will be learned on
# failover. This means you can Add, Change, Delete logical volumes or
# filesystems at the source (primary site) and the updated definitions will
# be propagated to the target (standby site) at failover time.
#
# 4. This script does NOT handle SDD vpaths.
#
#####
# Pre-requisites:
# -----
# Replace the values in:
# 1. START_APP1, START_APP2, etc. with the names of the programs to start and
# stop your applications.
#
#####

# Import shell variables
. /usr/sbin/cluster/local/hacmpPPRC.vars

PROGRAMNAME=$(basename $0)
OUT=/tmp/hacmp.out

DATE()
{
    DATEVAL=$(date +"%h %d %H:%M:%S")
    print ${DATEVAL}
}

print "$(DATE) Start execution of ${PROGRAMNAME}" >> ${OUT}

# Place your application start calls here
#
# START_APP1=
# START_APP2=
# START_APP3=

rdupsort()
{
    #
    # Sorts a list and removes duplicates

```

```

# Do not join the lines containing the sed command
#
echo $* | sed -e 's/\ / \
/g' | sort -u
}

NODE()
{
#
# Explicitly determines on which node the script is running
#
if [[ ${NODENAME} = ${PRIMARY_NODE} && ${HOSTNAME} = ${PRIMARY_HOST} ]]
then
    print "primary"
elif [[ ${NODENAME} = ${STANDBY_NODE} && ${HOSTNAME} = ${STANDBY_HOST} ]]
then
    print "standby"
fi
}

Parse_CStasks()
{
    vg=${1}
    # Task names can be no longer than 16 characters
    # CSSyncTask CSNoCopyTask CSTermPriSec CSTermSecPri

    set -A TASK $(print ${TASK_LIST})

    CSSyncTask=${vg}${TASK[0]}
    CSNoCopyTask=${vg}${TASK[1]}
    CSTermPriSec=${vg}${TASK[2]}
    CSTermSecPri=${vg}${TASK[3]}

    print "${CSSyncTask} ${CSNoCopyTask} ${CSTermPriSec} ${CSTermSecPri}"
}

# Main Program
#
NODENAME=$((${HA_UTILS}get_local_nodename)
HOSTNAME=$(lsattr -El inet0 | awk '/hostname/ {print $2}')

#
# Determine if primary CS server is available, if not, try standby CS server
#
${CLI_CMDS}rsTestConnection.sh -s ${PRIMARY_ESS}
if [ $? -eq 0 ]
then
    CS_SERVER=${PRIMARY_ESS}
    print "Using Primary Copy Services server ${CS_SERVER}"
else
    #
    # Primary Copy Services server is down
    #
    print "Primary Copy Services server may be down, trying Secondary server..."
    ${CLI_CMDS}rsTestConnection.sh -s ${STANDBY_ESS}
    if [ $? -eq 0 ]
    then
        CS_SERVER=${STANDBY_ESS}
    fi
fi
}

```

```

    print "Using Standby Copy Services server ${CS_SERVER}"
else
    # retcode >0 (should be 40)
    # Standby Copy Services server is down
    #
    print "ERROR: Cannot contact a Copy Services server"
    print "Unable to continue without Copy Services. Exiting..."
    exit 1
fi
fi

#
# Create volume group list for the Copy Services tasks.
FS_LIST=$(rdupsort ${FS_LIST})
for FS in ${FS_LIST}
do
    VG=$((${CL_UTILS})cl_fs2disk -v ${FS})
    VG_LIST="${VG_LIST} ${VG}"
done
VG_LIST=$(rdupsort ${VG_LIST})

if [ $(NODE) = "primary" ]
then

    for vg in ${VG_LIST}
    do
        HDISK=$(lspv | awk -v VG=${vg} ' $0~VG {print $1}')
        # Pick one hdisk in the volume group, hence assumption that if one
        # volume (hdisk) is in particular state, then all others in the VG
        # are in the same state.
        HDISK=$(print ${HDISK} | cut -f1 -d" ")

        VOL=$(lscfg -pvl ${HDISK} | grep "Serial Number" | sed "s/.*[a-z]\.*/g")
        CONDITION=$((${CLI_CMDS})/rsQuery.sh -u ${username} -p ${password} -q ${VOL}
-s ${CS_SERVER} | awk -F "State=|status=|, " '/State/ {print $2 $4}')

        # Setup the Copy Services task names based on prefixed VG name
        Parse_CStasks ${vg}

        # The following case statement tests for the conditions listed
        #
        # sourcesuspendedvolume is source and suspended
        #     task: ${CSSyncTask}
        # simplexnot_suspendedvolume is simplex and not_suspended
        #     task: ${CSNoCopyTask}
        # targetsuspended | targetnot_suspended)
        #     volume is target, suspended or not_suspended
        #     tasks: ${CSTermSecPri}, ${CSNoCopyTask}
        case ${CONDITION} in

            sourcesuspended)# ${CSSyncTask}
            #
            # The primary node may be joining the cluster with standby node down
            # or the standby node entered the cluster first. In either case the
            # the PPRC pairs must be a suspended state for this condition to
            # execute.
            #
        esac
    done
fi

```

```

    ${CLI_CMDS}rsExecuteTask.sh -v -u ${username} -p ${password} -s
${CS_SERVER} ${CSSyncTask}
    if [ $? -gt 0 ]
    then
        # This condition does not prevent application restart.
        print "ERROR: Could not establish one or more PRI->SEC PPRC pairs"
        print "Manual intervention required to establish remote copying!"
        print "Continuing..."
    fi
;;

simplexnot_suspended)# ${CSNoCopyTask}
#
# Re-establish PRI-SEC PPRC pairs with "Do not copy Volume" option.
#
# This condition will occur if reintegrating the primary node (site)
# after a node failure or graceful stop with takeover (clstop -grsy)
#
    ${CLI_CMDS}rsExecuteTask.sh -v -u ${username} -p ${password} -s
${CS_SERVER} ${CSNoCopyTask} >/dev/null 2>&1

# CLI return codes: 0=success, 80=failed

if [ $? -gt 0 ]
then
    # This condition does not prevent application restart.
    print "Warning: Could not establish one or more PRI->SEC PPRC pairs"
    print "Manual intervention required to establish remote copying!"
    print "Continuing..."
fi
;;

targetsuspending | targetnot_suspended)# ${CSTermSecPri},
${CSNoCopyTask}

#
# We are reintegrating the primary node after a "site failure".
# This logic assumes that all of the Primary Volumes have been
# synchronized manually and and that the AIX Volume Groups have
# been recreated.
#
# If SECONDARY to PRIMARY pairs have not been terminated,
# terminate them here.
#
    ${CLI_CMDS}rsExecuteTask.sh -v -u ${username} -p ${password} -s
${STANDBY_ESS} ${CSTermSecPri} >/dev/null 2>&1
    rc=$?
    if [ ${rc} -gt 0 ]
    then

        print "ERROR: CLI return code=${rc}"
        print "Could not terminate one or more PPRC pairs"
        print "!!! MANUAL INTERVENTION IS REQUIRED !!!"
        print "Applications cannot be started! Exiting..."
        exit 1

    else

```

```

#
# Query task completion to include progress in hacmp.out
#
# Return code 80 will occur if the task is not running
# or has already completed.
${CLI_CMDS}rsQueryComplete.sh -u ${username} -p ${password} -s
${STANDBY_ESS} ${CSTermSecPri} >/dev/null 2>&1

#
# Re-establish PRI-SEC PPRC pairs with "Do not copy Volume" option.
#
${CLI_CMDS}rsExecuteTask.sh -v -u ${username} -p ${password} -s
${CS_SERVER} ${CSNoCopyTask} >/dev/null 2>&1

# CLI return codes: 0=success, 80=failed

rc=$?
if [ ${rc} -gt 0 ]
then

# This condition does not prevent application restart.
print "ERROR: CLI return code=${rc}"
print "Warning: Could not establish one or more PRI->SEC PPRC pairs"
print "Manual intervention required to establish remote copying!"
print "Continuing..."
fi
fi
;;

esac
done

fi # NODE = primary

if [ $(NODE) = "standby" ]
then
#
# Node (or site) failover has occurred
# Terminate Pairs to make the secondary volumes available to LVM
#
for vg in ${VG_LIST}
do
HDISK=$(lspv | awk -v VG=${vg} '$0~VG {print $1}')
# Pick one hdisk in the volumes group
HDISK=$(print ${HDISK} | cut -f1 -d" ")

VOL=$(lscfg -pvl ${HDISK}|grep "Serial Number"|sed "s/.*[a-z]\.*//g")
CONDITION=$( ${CLI_CMDS}/rsQuery.sh -u ${username} -p ${password} -q ${VOL}
-s ${CS_SERVER} | awk -F "State=|, " '/State/ {print $2}')

# Setup the Copy Services task names by prefixing VG name
Parse_CStasks ${vg}

if [ ${CONDITION} = "target" ]
then
${CLI_CMDS}rsExecuteTask.sh -v -u ${username} -p ${password} -s
${STANDBY_ESS} ${CSTermPriSec} >/dev/null 2>&1

```

```

# CLI return codes: 0=success, 80=failed

rc=$?
if [ ${rc} -gt 0 ]
then
    print "ERROR: CLI return code=${rc}"
    print "Could not terminate one or more PPRC pairs"
    print "!!! MANUAL INTERVENTION IS REQUIRED !!!"
    print "Applications cannot be started! Exiting..."
    exit 1
fi

#
# Query task completion to include progress in hacmp.out
#
# Return code 80 will occur if the task is not running
# or has already completed.
${CLI_CMDS}rsQueryComplete.sh -u ${username} -p ${password} -s
${STANDBY_ESS} ${CSTermPriSec} >/dev/null 2>&1
fi
done
fi

#
# From the list of filesystems, discover their volume groups and the disks
# containing the VGs.
#
STATUS=0
DISK_LIST=""
FS_LIST=$(rdupsort ${FS_LIST})
for FS in ${FS_LIST}
do
    VG=$((${CL_UTILS}cl_fs2disk -v ${FS})
    #
    # If the VG is not already varied on, append new VG to the previous
    # VG list, and add the new disks to the previous DISK list
    #
    if lsvg -o | grep -qx ${VG}
    then
        continue
    else
        VG_LIST="${VG_LIST} ${VG}"
        DISK=$((${CL_UTILS}cl_fs2disk -p ${FS})
        #
        # Append to the previous DISK list.
        #
        DISK_LIST="${DISK_LIST} ${DISK}"
    fi
done

#
# Determine if all disks in the list are available.
# Ghost disks are handled here!
#
if [ -n "${DISK_LIST}" ]
then
    DISK_LIST=$(rdupsort ${DISK_LIST})

```

```

    ${CL_UTILS}cl_disk_available "${DISK_LIST}"

    if [ $? -ne 0 ]
    then
        STATUS=1
    fi
fi

if [ ${STATUS} -eq 1 ]
then
    print "ERROR: One or more hdisks are unavailable"
    print "!!! Manual intervention is required !!!"
    print "Applications will not be started!"
    exit ${STATUS}
fi

#
# Activate all VGs in the list.
#
if [ -n "${VG_LIST}" ]
then
    VG_LIST=$(rdupsort ${VG_LIST})
    ${CL_UTILS}cl_activate_vgs -n "${VG_LIST}"

    if [ $? -ne 0 ]
    then
        STATUS=1
    fi
fi

#
# Check and mount all filesystems in the list.
#
if [ -n "${FS_LIST}" ]
then
    ${LOCAL}esscl_activate_fs "${FS_LIST}"

    if [ $? -ne 0 ]
    then
        STATUS=1
    fi
fi

#
# Export filesystems in /etc/exports
#
if [ -f /etc/xtab ]
then
    exportfs -a
fi

#
# Synchronize all VGs in background.
#
if [ -n "${VG_LIST}" ]
then
    ${CL_UTILS}cl_sync_vgs -b "${VG_LIST}"
fi

```

```

# Start applications
#$( ${START_APP1} )

print "$(DATE) End execution of ${PROGNAME}" >> ${OUT}

exit ${STATUS}

```

A.3 stopESSapps

```

#!/usr/bin/ksh
#####
# NOTE: This script suitable for a two node cluster only where each node is
# connected to a separate ESS.
#
# Name:          stopESSapps.sh
#
# Description:This script executes on an exiting cluster node if the cluster
#              resources exist on that node, and on the standby node when the
#              primary node is joining the cluster after a failover.
#              It stops applications, unmounts filesystems and deactivates
#              volume groups.
#
# Called by:    stop_server
#
# Calls to:    cl_get_path, cl_fs2disk, cl_deactivate_vgs,
#              esscl_deactivate_fs
#
# Arguments:   none
#
# Written by:  Andrew Beyer, IBM Australia
#
#####
# Pre-requisites:
# -----
# This script will work only on hosts on which HACMP for AIX is installed.
# It calls utilities that are packaged only with HACMP.
#
#####
# Notes:
#
#####
# Replace the values in:
# 1. STOP_APP1, STOP_APP2, etc. with the names of the programs to start and
#    stop your applications
#####

# Import shell variables
. /usr/sbin/cluster/local/hacmpPPRC.vars

PROGNAME=$(basename $0)
OUT=/tmp/hacmp.out

DATE()
{

```



```

DATEVAL=$(date +"%h %d %H:%M:%S")
print ${DATEVAL}
}

# Place your application stop calls here
#
#STOP_APP1=
#STOP_APP2=
#STOP_APP3=

STATUS=0
DISK_LIST=""

rdupsort ()
{
#
# Sorts a list in reverse order and removes duplicates
# Do not join the lines containing the sed command
#
echo $* | sed -e 's/\ / \
/g' | sort -ru
}

# Main Program

# Stop applications
# ${STOP_APP1}

#
# From the list of filesystems, discover their volume groups.
#
FS_LIST=$(rdupsort ${FS_LIST})
for FS in ${FS_LIST}
do
    VG=$((${CL_UTILS}cl_fs2disk -v ${FS})
    VG_LIST="${VG_LIST} ${VG}"
done

#
# Unmount all filesystems in the list.
#
if [ -n "${FS_LIST}" ]
then
    ${LOCAL}esscl_deactivate_fs "${FS_LIST}"

    if [ $? -ne 0 ]
    then
        STATUS=1
    fi
fi

#
# Deactivate all VGs in the list.
#
if [ -n "${VG_LIST}" ]
then
    VG_LIST=$(rdupsort ${VG_LIST})
    ${CL_UTILS}cl_deactivate_vgs "${VG_LIST}"

```

```
if [ $? -ne 0 ]
then
    STATUS=1
fi
fi

print "$(DATE) End execution of ${PROGNAME}" >> ${OUT}

exit ${STATUS}
```

Appendix B. Understanding logical subsystems

Up until now the ESS Specialist has virtually hidden the requirement of knowing or having to do anything about logical subsystems (LSSs) for Open Systems. Now, with Copy Services and Flash Copy specifically, it becomes critical that the person configuring the ESS knows what the boundaries are on an LSS for Open Systems. Using Flash Copy, all source/target volumes copied must be copied (from/to) within the same LSS. The following note will define what an LSS is for Open Systems.

Figure 118 defines one ESS containing up to 32 LSSs; this is only true for an ESS with both ESCON and Open Systems ports. ESCON can have up to 16 LSS and Open Systems can have up to 16 LSSs. An OPEN SYSTEM LSS can only have up to 256 LUNs created in a single LSS.

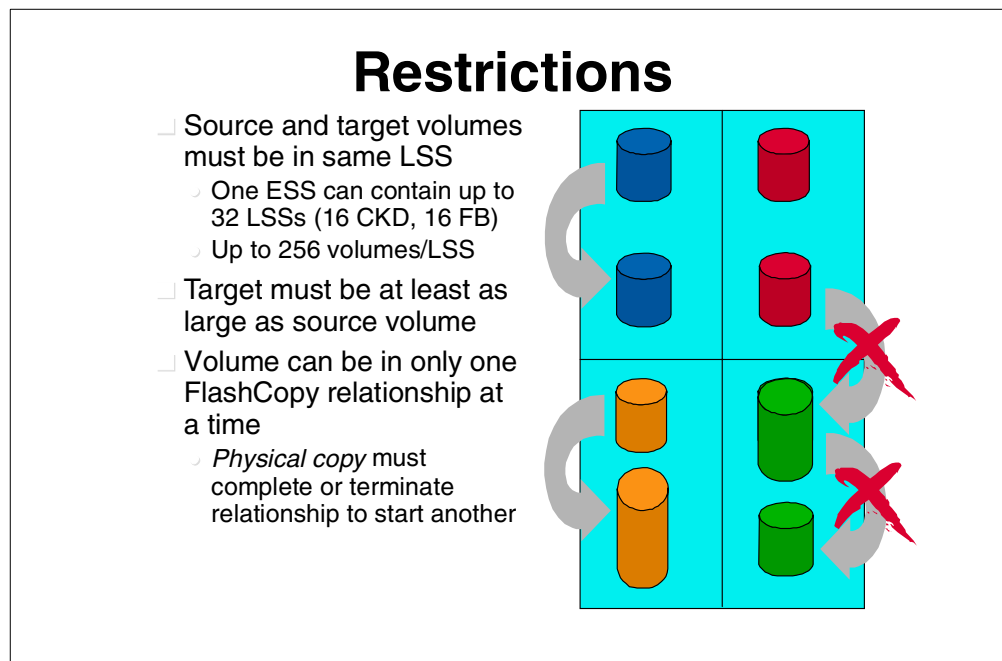


Figure 118. Restrictions regarding source and target volumes

An ESS can have 0, 8, or 16 LSSs *defined* for Open Systems storage. Zero is for an ESS that is all ESCON and will have no storage defined for Fixed Block.

Note: In all further text, we will refer to ESS definitions that only include SCSI ports, not a box that includes ESCON ports.

An ESS that has 8 LSSs defined, but only has 4 eight-packs in the 'A' box, could end up with only 4 LSSs until more eight-packs are added. An ESS that has 16 eight-packs and is defined for 16 LSSs, could have 8, 9, 10, or any number up to 16 LSSs. If you define an ESS for 16 LSSs, the code will configure the box as if it were defined for 8 LSSs until an LSS gets more than 192 LUNs (192 = 75% of 256). At this point, additional eight-packs added to this LSS, or any unconfigured eight-packs that are currently in this LSS, will go into a new LSS number. This change *ONLY* affects this LSS, none of the other LSSs get changed.

Figure 119 shows the LSS boundaries for an ESS with a full configuration that is defined with 8 LSSs.

Note: LSS numbering is usually in HEX. The first LSS for Open Systems = LSS 16. We will only use both Decimal and Hex numbering for all LSS definitions.

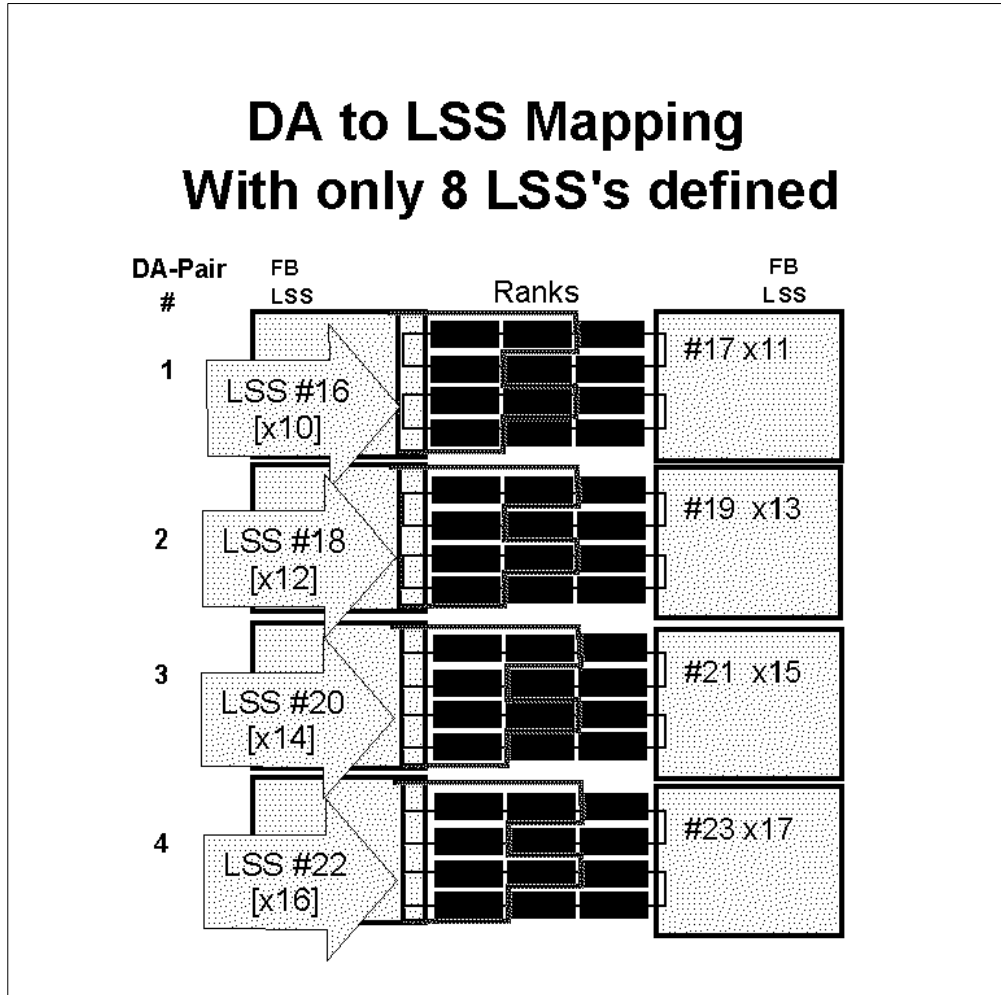


Figure 119. DA to LSS mapping

Figure 120 shows a fully configured ESS base box configured with 16 eight-packs or 128 disks, and also configured with only 8 LSSs for Fixed Block storage.

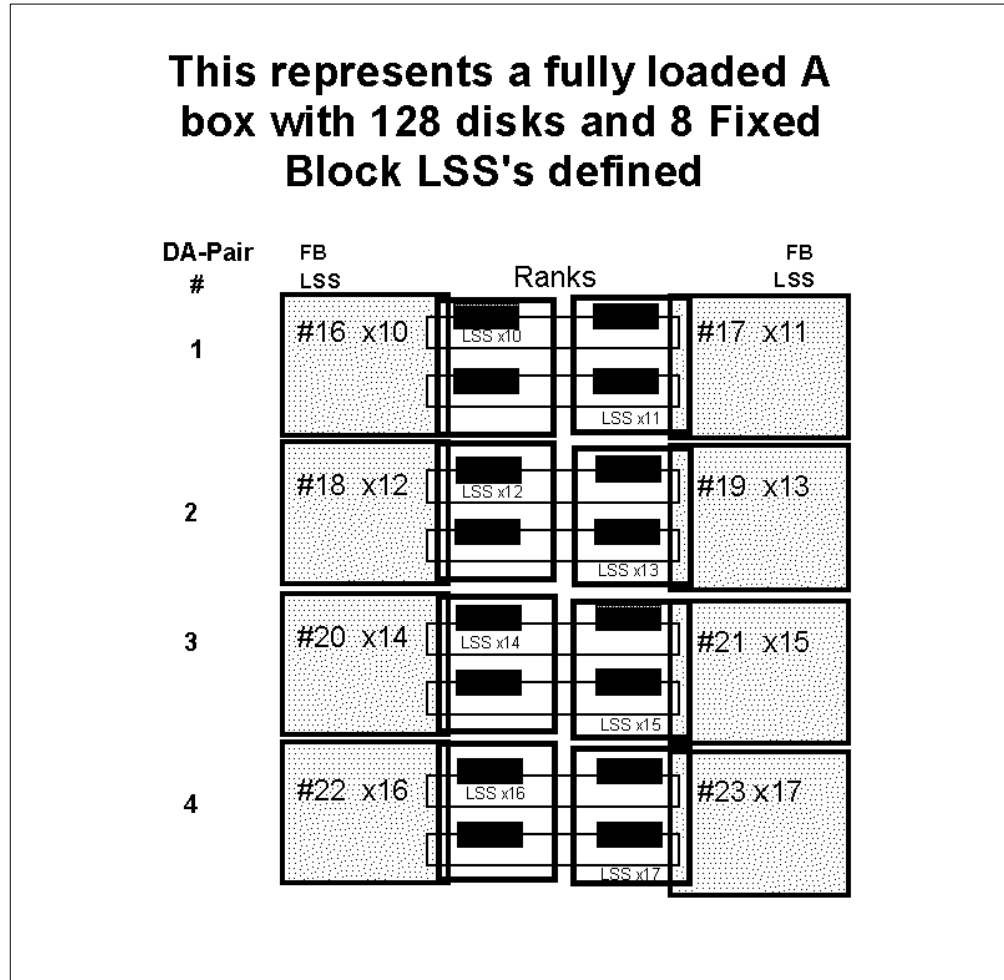


Figure 120. Fully configured ESS A box showing 8 LSS mappings

Figure 121 shows a fully loaded A box with 16 eight-packs and 16 fixed block LSSs defined. Note the "dotted" lines that split the ranks. Until a rank has more than 192 LUNs created and the other rank is *NOT* configured, both ranks will stay in the same LSS. If you have already started configurations on both ranks and eventually get to the point where you have 192 LUNs, only additional ranks added to the loop would acquire a new LSS number. The eight-packs do not change their LSS designation if they have been configured, they can only "change" if they are in an unconfigured state. Once ranks are configured, they do not change LSS numbers, unless the rank is erased and configuration of that rank starts over.

For example: On LSS #16 (x10), if both ranks were set up and 180 LUNs were created upon initial installation, and over the next 6 months additional LUNs get created in LSS #16, such that now there are 192 LUNs created on LSS #16, then there will be no change. Both ranks will still be LSS #16. Any additional ranks added to this LSS group would become LSS #24 (x18). LSS boundaries are at a rank definition. No rank will be split to allow for a change in LSS definition.

This represents a fully loaded A box with 16 eight packs (128 disks) and 16 Fixed Block LSS's defined

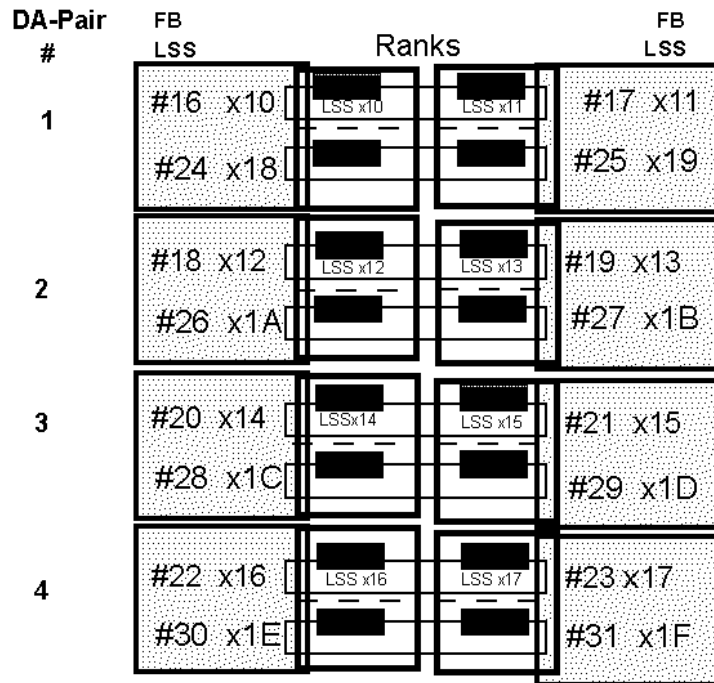


Figure 121. Fully configured ESS A box showing 16 LSS mappings

In Figure 122, regardless of definitions for 8 LSSs or for 16 LSSs, you end up with 4 ranks (numbers #20 [x14], #21 [x15], #22 [x16], and #23 [x17]) that comprise a single LSS. This means that the source and target of a Flash Copy must be within a single rank on these 4 LSSs.

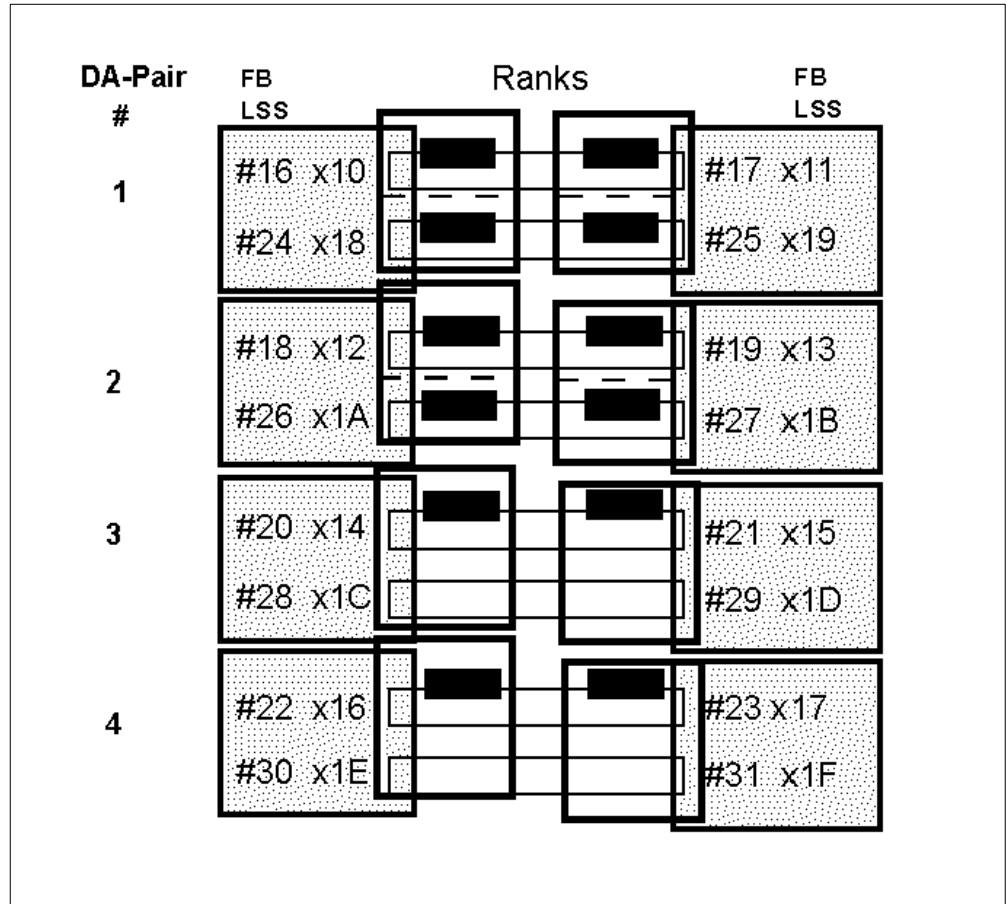


Figure 122. Partially configured ESS

Appendix C. Operating system considerations

In this appendix we discuss the preparations that are needed on an AIX system before FlashCopy can be used as part of an integrated backup solution.

C.1 Definition of a "typical" backup solution

There are different understandings about what a "typical" backup solution is. The scenario that we tested is described in detail together with the sample scripts that we used. While these scripts worked in our environment, there is no guarantee that they will work in other situations. They do, however, provide a basis on which you can develop your own backup scripts.

One RS6000 Server running the AIX operating system was attached to an IBM 2105 ESS with a storage pool (LUNs) assigned to it. This Server is known as the production Server, on which the production applications would be run. The LUNs of this Server are the source LUNs with respect to FlashCopy. The target LUNs are located on the same ESS but are assigned to a different RS6000 server running the AIX operating system which is called the backup server. We recommend that all I/O to the FlashCopy source volumes is quiesced and that all buffers are flushed before the copy is made. In the case of a file system this would mean issuing the `sync` command. Databases usually have their own command to ensure that all pending writes are flushed to disk.

C.2 Preparation after the first FlashCopy

After all the production volumes have been FlashCopied, you have to make the FlashCopy target volumes available to the backup server. Ensure that they are all varied online, and that no volume group is created over them. As this is the first FlashCopy, there is no information about the Volume Group structure in the ODM of the backup server. This means that you have to import the volume groups to the backup server. This is done by using the procedure described below. Since the FlashCopy copies the entire LUN, it will also copy the physical volume identifier (PVID) of the source volume. The PVID is a unique identifier which AIX uses to identify its physical disks.

1. Identify the PVID of at least one hdisk from each volume group on the production server that you intend to FlashCopy.

To determine the pvid of an hdisk in volume group `samplevg`, issue the command `lsvg -p samplevg`. This will give the output shown in Table 3.

Table 3. Output `lsvg` command issued against `samplevg`

PV_NAME	PV STATE	TOTAL PPs	FREE PPs	FREE DISTRIBUTION
hdisk26	active	953	2	00..00..00..00..02
hdisks27	active	953	0	00..00..00..00..00
hdisk28	active	953	0	00..00..00..00..00
hdisk29	active	953	0	00..00..00..00..00
hdisk30	active	953	0	00..00..00..00..01
hdisk31	active	953	0	00..00..00..00..00
hdisk32	active	953	0	00..00..00..00..00
hdisk33	active	953	0	00..00..00..00..00

- Query one of the hdisks by issuing the `lspv` command. For example, if we wish to use `hdisk26`, the command that is issued would be `lspv hdisk26`. The output from this command is shown in Figure 123.

```

PHYSICAL VOLUME:   hdisk26                VOLUME GROUP:   samplevg
PV IDENTIFIER:    000c3c7d5c13c183      VG IDENTIFIER   000c3c7d5c13dd48
PV STATE:        active
STALE PARTITIONS: 0                      ALLOCATABLE:   yes
PP SIZE:         8 megabyte(s)          LOGICAL VOLUMES: 1
TOTAL PPs:       953 (7624 megabytes)   VG DESCRIPTORS: 1
FREE PPs:        2 (16 megabytes)
USED PPs:        951 (7608 megabytes)
FREE DISTRIBUTION: 00..00..00..00..02
USED DISTRIBUTION: 191..191..190..190..189
    
```

Figure 123. Output of `lspv hdisk26`

The PV Identifier is the required PVID. Alternatively, the command `lspv | grep hdisk26` can be issued, which will produce the following output:

```
hdisk26      000c3c7d5c13c183      samplevg
```

The second column is the PVID. You now have to determine on the backup server which `hdisk` corresponds to this PVID. To do this you issue the command:

```
lspv | grep 000c3c7d5c13c183
```

The output of this command will have the following format:

```
hdisk16      000c3c7d5c13c183      none
```

The "none" in the 3-rd column indicates that this `hdisk` does not belong to any known volume group on the backup server.

- Import the volume group on the backup server. This is done by issuing the command: `importvg -y samplevg hdisk16`

All of these steps should be repeated for every volume group from the production server.

This procedure should be used the first time that you FlashCopy the LUNs from a volume group on one server and mount them on a second server.

C.2.1 Further FlashCopy invocations

Before invoking FlashCopy make sure that all volume groups on the target volumes of the backup server are varied off. On the testing that we carried out, we found that it was not necessary to remove the hdisk definitions from the backup server.

C.2.2 Invoke FlashCopy

Run the AIX configuration manager (`cfmgmr`) on the backup server, since there may be new target LUNs added, which have not yet been configured. If you did not export the volume groups (`exportvg`) on the backup server before FlashCopy invocation, it may not be necessary to export and import them back, but just a simple `varyonvg` might be sufficient.

A `varyonvg` is much faster than an `export/importvg`. In the example above, an `importvg -y samplevg hdisk26` will take about 32 seconds. A `varyonvg` will take only 0.5 sec. The time will vary depending on how many hdisks are in one volume group and also on the number of logical volumes per volume group.

C.2.2.1 When is an Importvg required?

An `importvg` is required when:

1. You do not have an entry of the volume group in the ODM. This means you either FlashCopy the corresponding LUNs of a volume group for the first time or you did an explicitly `exportvg` before invoking FlashCopy.
2. You added or removed some hdisks, logical volumes and/or file systems to (from) the volume group on the production server before invoking FlashCopy - simply speaking you changed the volume group structure

To determine if an `importvg` is necessary, or just a `varyonvg` is sufficient, we have to see if there is an entry in the ODM for the volume group. This is done by issuing the command:

```
lsvg | grep vgroup
```

This will answer the question “is there an entry for the volume group named `vgroup` in the ODM?”

The second question that needs answering is “Is there any change of the volume group structure?”. This can be obtained from the Volume group Descriptor Area (VGDA). Any time you change the structure of a volume group, the VGDA gets changed. This is reflected in the changed time-stamp of the VGDA.

Note

Simply adding/changing/deleting a file from a filesystem of the volume-group does NOT require an `export/importvg`, as this does not change the VGDA.

You can query the time-stamp of a VGDA from a hdisk by issuing:

```
# lquerypv -h /dev/hdiskX 11000 10
```

In this command, `hdiskX` is a member hdisk of a volume group.

This will produce an output similar to the following:

```
00011000 39E4E87D 1FAC826C 000C3C7D 5C13DD48 |9..}...1..<}\\..H|
```

The second column is the time-stamp.

You can query the time-stamp of the ODM copy of the VGDA for a particular volume group by issuing:

```
odmget -q "attribute like timestamp and name like vgroupname" CuAt
```

In this command, `vgroupname` is the volume-group name.

Here is a sample output for the `samplevg` command:

```
# odmget -q "attribute like timestamp and name like samplevg" CuAt
CuAt:
name = "samplevg"
attribute = "timestamp"
value = "39e4e87d1fac826c"
type = "R"
generic = "DU"
rep = "s"
nls_index = 0
```

Here, the first 8 digits in the field named `value` are the time-stamp. As you can see, both values are the same and a `varyonvg` will be sufficient.

C.2.3 Backup scripts

In this section we describe the primary and secondary scripts.

C.2.3.1 Secondary script

The script `primary.sh` expects a list of the volume groups which should be backed up via FlashCopy. It should be installed on the production server. It produces an output file (whose name you have to specify) with the following format:

- `vgroupname1 pvid1`
- `vgroupname2 pvid2`
-

So the script actually carries out the steps described in Appendix C.2, "Preparation after the first FlashCopy" on page 183

The format of the input file should be:

- `vgroupname1`
- `vgroupname2`
-

Note that empty lines are not allowed, and that this script should always be executed before the script `secondary.sh`.

```

#####
# #
# primary.sh #
# #
# Script for supporting a FlashCopy Backup solution (AIX 4.3.X) #
# #
# To be run on the Primary (Productional) Sever #
# #
# Run this script prior to running secondary.sh !!! #
# #
# Author: Vladimir Atanaskovik #
# #
# (c) by IBM #
# #
# Initial Coding: 10/19/2000 Vladimir Atanaskovik #
#####
#####

# Enter here the File containing the Volume Groups
FileVolGrups=
# Enter here the Output File
# Default /tmp/primary_list.lst
OutputFile=/tmp/primary_list.lst

if [ -z "$FileVolGrups" ]
then
echo ERROR! Logical Volume Group Input File not specified!
exit -1
fi

if [ -f $FileVolGrups ]
then
:
else
echo ERROR! Could not find specified Logical Volume Group Input File!
exit -1
fi

if [ -z "$OutputFile" ]
then
echo ERROR! Output File not specified!
exit -1
fi

if [ -f $OutputFile ]
then
rm $OutputFile
fi

VolGrups=$(cat $FileVolGrups)

if [ -z "$VolGrups" ]
then

```

```

    echo ERROR! No volume Groups specified!
    exit -1
fi

echo $0 running .....

for Ix in $VolGrups
do
    PhVol=$(lsvg -p $Ix|awk 'NR==3 {print $1}')
    PVID=$(lspv $PhVol|grep IDENTIFIER|awk '{print $3}')
    echo $Ix $PVID >> $OutputFile
done

```

C.2.3.2 Secondary script

The script `secondary.sh` is located on the backup server. Prior to running this script, execute `primary.sh`. You need to enter the name (or IP address) of the production server, and the name (with full path) of the output file produced with `primary.sh`. Since this script will (`rcp`) remote copy the output file from the `primary.sh` script, make sure that the following things have been taken care of:

- The user who executes the `secondary.sh` script also has an account with the same name on the production server.
- Edit the `/etc/hosts.equiv` and `.rhosts` file (in the home-directory of the user) both on the production server, so that an `rcp` from the backup server is possible.

The script `secondary.sh` does all steps described in the sections: "Preparing for the first FlashCopy" and Further FlashCopy invocations"

```

#####
#####
#
#
# secondarysh
#
#
# Script for supporting a FlashCopy Backup solution (AIX 4.3.X)
#
#           #
# To be run on the Backup Sever           #
#
#
# Previous to this script run primary.sh on the primary (productional) server
!!!      #
#
#
# Author: Vladimir Atanaskovik (vladimir@de.ibm.com)
#
#
# (c) by IBM
#
#
# Initial Coding: 10/19/2000 Vladimir Atanaskovik
#

```

```

#           10/30/2000  VGDA Time Stamp Query and comparison added
#
#           10/31/2000  Minor Changes
#
#####
#####

#!ksh

# Specify the Primary host
PriHost=
# Specify the path and name of the output file created with primary.sh on the
primary server
# Default /tmp/primary_list.lst
PriFile=/tmp/primary_list.lst
# Specify local file name
LocalFile=./LocalFile.lst
# Specify Temp File to be used
TempFile="$_Secondary.tmp"
# Specify 1-use vpaths 0-do not use vpaths
# Use Vpath only if IBM Subsystem device Driver is installed
let UseVpath=0

if [ -z "$PriHost" ]
then
echo ERROR! Primary Host not specified!
exit -1
fi

if [ -z "$PriFile" ]
then
echo ERROR! Primary host vg/pvid list file not specified!
exit -1
fi

if [ -z "$LocalFile" ]
then
echo ERROR! Local File for primary host vg/pvid list not specified!
exit -1
fi

if [ -z "$TempFile" ]
then
echo ERROR! Temp File not specified!
exit -1
fi

echo $0 running .....

rcp "$PriHost":"$PriFile" $LocalFile

RetVal=$?

if [ $RetVal -ne 0 ]
then
echo ERROR! Could not rcp Config File!

```

```

    exit -1
fi

VolGr=$(cat $LocalFile|cut -d ' ' -f1)

if [ -z "$VolGr" ]
then
    echo ERROR! Invalid Format of Input File!
    exit -1
fi

PhVolId=$(cat $LocalFile|cut -d ' ' -f2)

if [ -z "$PhVolId" ]
then
    echo ERROR! Invalid Format of Input File!
    exit -1
fi

lspv > $TempFile

let Jxvg=1

for Ixpv in $PhVolId
do
    Hdisk=$(grep $Ixpv $TempFile|awk 'NR==1{print $1}')

    if [ -z "$Hdisk" ]
    then
        echo ERROR! Could not find hdisk with PVID $Ixpv!
        rm $TempFile
        exit -1
    fi

    Vgname=$(echo $VolGr | cut -d ' ' -f $Jxvg)

    NeedImport=0

    ### Is there already an ODM Entry for the Volume Group

    DoesExist=$(lsvg|grep $Vgname)

    if [ -n "$DoesExist" ]

    ### Yes, ODM Entry for the volume group does exist
    then

    # Check if the volume group is already varied on
    IsVaryOn=$(lsvg -o|grep $Vgname)

    if [ -n "$IsVaryOn" ]
    then
        echo ERROR! Volume Group $Vgname already Varied-On! Possible Data
        Corruption!
    fi
fi

```



```

    echo Vary Off $Vgname and re-establish FlashCopy!
    rm $TempFile
    exit -1
fi

# Check VGDA Timeststamp
# on the hdisk

HdVGDA=$(lquerypv -h /dev/$Hdsk 11000 10|awk '{print $2}'|sed -e
'y/ABCDEF/abcdef/')

# and in the ODM
OdmVGDA=$(odmget -q "attribute like timestamp and name like $Vgname"
CuAt|grep value|awk '{print $3}'|cut -c 2-9)

## Does VGDA Timeststamp on the hdisk match VGDA Timeststamp in the ODM
if [ "$HdVGDA" != "$OdmVGDA" ]

## No - then importvg
then
    exportvg $Vgname
    # update TempFile since Vgname has been deleted
    lspv > $TempFile
    NeedImport=1

## Yes - then just varyon
else
    varyonvg $Vgname
    RetValue=$?
    if [ $RetValue -ne 0 ]
    then
        echo ERROR! Could Not Vary-On Volume Group $Vgname
        rm $TempFile
        exit -1
    fi

fi

### No, ODM Entry for the volume group does not exist
else
    NeedImport=1
fi

if [ NeedImport -eq 1 ]
then

    TmpVlgrp=$(grep $Ixp $TempFile|awk 'NR==1{print $3}')

    if [ "$TmpVlgrp" != "None" ]
    then
        echo ERROR! Can Not Import Volume Group $Vgname
        echo Hdisk $Hdsk already assigned to another volume group!
        rm $TempFile
        exit -1
    fi

fi

```

```
importvg -y $Vgname $Hdsk
RetVal=$?

if [ $RetVal -ne 0 ]
then
    echo ERROR! Could Not Import Volume Group $Vgname
    rm $TempFile
    exit -1
fi

if [ $UseVpath -eq 1 ]
then
    hd2vp $Vgname
fi

fi

let Jxvg=$Jxvg+1
done

rm $TempFile
exit 0
```

Appendix D. Special notices

This publication is intended to help IBMers, Business Partners, and customers who are involved with storage subsystems to specify, install, and use ESS Copy services functions in UNIX and NT environments. The information in this publication is not intended as the specification of any programming interfaces that are provided with the ESS. See the PUBLICATIONS section of the IBM Programming Announcement for the IBM2105 ESS for more information about what publications are considered to be product documentation.

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

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

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

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


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

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

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

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

IBM ®

Redbooks
Redbooks Logo 

AIX	Netfinity
AS/400	OS/390
AT	OS/400
CT	RACF
CUA	RAMAC
Current	RS/6000
DFSMS/MVS	S/390
DFSMSdss	S/390 Parallel Enterprise Server
Enterprise Storage Server	Seascape
Enterprise Systems Connection Architecture	SP
ESCON	StorWatch
FICON	System/370
	System/390
MVS/DFP	S/390 Parallel Enterprise Server
MVS/ESA	Wave
	Wizard

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere., The Power To Manage., Anything. Anywhere., TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli A/S.

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

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

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Appendix E. Related publications

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

E.1 IBM Redbooks

For information on ordering these publications see “How to get IBM Redbooks” on page 197.

- *IBM Enterprise Storage Server*, SG24-5465
- *Implementing the Enterprise Storage Server in Your Environment*, SG24-5420
- *Implementing ESS Copy Services in a System/390 Environment*, SG24-5680

E.2 IBM Redbooks collections

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at ibm.com/redbooks for information about all the CD-ROMs offered, updates and formats.

CD-ROM Title	Collection Kit Number
IBM System/390 Redbooks Collection	SK2T-2177
IBM Networking Redbooks Collection	SK2T-6022
IBM Transaction Processing and Data Management Redbooks Collection	SK2T-8038
IBM Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
IBM AS/400 Redbooks Collection	SK2T-2849
IBM Netfinity Hardware and Software Redbooks Collection	SK2T-8046
IBM RS/6000 Redbooks Collection	SK2T-8043
IBM Application Development Redbooks Collection	SK2T-8037
IBM Enterprise Storage and Systems Management Solutions	SK3T-3694

E.3 Other resources

These publications are also relevant as further information sources:

- *ESA/390 Principles of Operation*, SA22-7201
- *IBM Enterprise Storage Server Users Guide 2105 Models E10 and E20*, SC26-7295
- *IBM Enterprise Storage Server Introduction and Planning Guide, Models E10 and E20*, GC26-7294
- *Enterprise Storage Server Configuration Planner*, SC26-7353. Available in softcopy only.
- *IBM Enterprise Storage Server System/390 Command Reference*, SC26-7298.
- *3990/9390 Storage Control Planning, Installation, and Storage Administration Guide*, GA32-0100
- *3990/9390 Operations & Recovery Guide*, GA32-0253
- System Overview 9672 Generation 4, 9674 Coupling Facility C05, GA22-7154.

- *ESS Performance White Paper* ,
<http://www.storage.ibm.com/hardsoft/products/ess/whitepaper.htm>
- *DFSMS/MVS Software Support for IBM Enterprise Storage Server*, SC26-7318. Available in softcopy only
- *OS/390 V2R8.0 MVS System Commands*, GC28-1781

E.4 Referenced Web sites

These Web sites are also relevant as further information sources:

- <http://www.redbooks.ibm.com>
IBM Redbooks
- <http://www.storage.ibm.com/hardsoft/products/ess/supserver.htm>
IBM ESS List of Supported Servers
- <http://www.ibm.com/storage/ess>
IBM Enterprise Storage Server
- <http://www.elink.ibm.com/pbl/pbl>
IBM Publications

How to get IBM Redbooks

This section explains how both customers and IBM employees can find out about IBM Redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** ibm.com/redbooks

Search for, view, download, or order hardcopy/CD-ROM Redbooks from the Redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this Redbooks site.

Redpieces are Redbooks in progress; not all Redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

Send orders by e-mail including information from the IBM Redbooks fax order form to:

	e-mail address
In United States or Canada	pubscan@us.ibm.com
Outside North America	Contact information is in the "How to Order" section at this site: http://www.elink.ibm.ibm.com/pbl/pbl

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the "How to Order" section at this site: http://www.elink.ibm.ibm.com/pbl/pbl

- **Fax Orders**

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	Fax phone number is in the "How to Order" section at this site: http://www.elink.ibm.ibm.com/pbl/pbl

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the Redbooks Web site.

IBM Intranet for Employees

IBM employees may register for information on workshops, residencies, and Redbooks by accessing the IBM Intranet Web site at <http://w3.itso.ibm.com/> and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access MyNews at <http://w3.ibm.com/> for redbook, residency, and workshop announcements.

Glossary

Glossary

This glossary contains a list of terms used within this redbook.

A

allegiance. The ESA/390 term for a relationship that is created between a device and one or more channel paths during the processing of certain condition.

allocated storage. On the ESS, this is the space that you have allocated to volumes, but not yet assigned.

application system. A system made up of one or more host systems that perform the main set of functions for an establishment. This is the system that updates the primary DASD volumes that are being copied by a copy services function.

AOM. Asynchronous operations manager.

APAR. Authorized program analysis report.

array. An arrangement of related disk drive modules that you have assigned to a group.

assigned storage. On the ESS, this is the space that you have allocated to volumes, and assigned to a port.

asynchronous operation. A type of operation in which the remote copy XRC function copies updates to the secondary volume of an XRC pair at some time after the primary volume is updated. Contrast with synchronous operation.

ATTIME. A keyword for requesting deletion or suspension at a specific target time.

availability. The degree to which a system or resource is capable of performing its normal function.

B

bay. Physical space on an ESS rack. A bay contains SCSI, ESCON or Fibre Channel interface cards and SSA device interface cards.

backup. The process of creating a copy of data to ensure against accidental loss.

C

cache. A random access electronic storage in selected storage controls used to retain frequently used data for faster access by the channel.

cache fast write. A form of fast write where the subsystem writes the data directly to cache, where it is available for later destaging.

CCA. Channel connection address.

CCW. Channel command word.

CEC. Central electronics complex.

channel. (1) A path along which signals can be sent; for example, data channel and output channel. (2) A functional unit, controlled by the processor, that handles the transfer of data between processor storage and local peripheral equipment.

channel connection address (CCA). The input/output (I/O) address that uniquely identifies an I/O device to the channel during an I/O operation.

channel interface. The circuitry in a storage control that attaches storage paths to a host channel.

channel path. The ESA/390 term for the interconnection between a channel and its associated controllers.

channel subsystem. The ESA/390 term for the part of host computer that manages I/O communication between the program and any attached controllers.

CKD. Count key data. An ES/390 architecture term for a device that specifies the format of and access mechanism for the logical data units on the device. The logical data unit is a track that can contain one or more records, each consisting of a count field, a key field (optional), and a data field (optional).

CLIST. TSO command list.

cluster. See storage cluster.

cluster processor complex (CPC). The unit within a cluster that provides the management function for the storage server. It consists of cluster processors, cluster memory, and related logic.

concurrent copy. A copy services function that produces a backup copy and allows concurrent access to data during the copy.

concurrent maintenance. The ability to service a unit while it is operational.

consistency group time. The time, expressed as a primary application system time-of-day (TOD) value, to which XRC secondary volumes have been updated. This term was previously referred to as "consistency time".

consistent copy. A copy of data entity (for example a logical volume) that contains the contents of the entire data entity from a single instant in time.

contingent allegiance. ESA/390 term for a relationship that is created in a controller between a

device and a channel path when unit-check status is accepted by the channel. The allegiance causes the controller to guarantee access; the controller does not present the busy status to the device. This enables the controller to retrieve sense data that is associated with the unit-check status, on the channel path with which the allegiance is associated.

control unit address (CUA). The high order bits of the storage control address, used to identify the storage control to the host system.

Note: The control unit address bits are set to zeros for ESCON attachments.

CUA. Control unit address.

D

daisy chain. A method of device interconnection for determining interrupt priority by connecting the interrupt sources serially.

DA. Device adapter.

DASD. Direct access storage device. See disk drive module.

data availability. The degree to which data is available when needed. For better data availability when you attach multiple hosts that share the same data storage, configure the data paths so that data transfer rates are balanced among the hosts.

data sharing. The ability of homogenous or divergent host systems to concurrently utilize information that they store on one or more storage devices. The storage facility allows configured storage to be accessible to any attached host systems, or to all. To use this capability, you need to design the host program to support data that it is sharing.

DDM. Disk drive module

data compression. A technique or algorithm that you use to encode data such that you can store the encoded result in less space than the original data. This algorithm allows you to recover the original data from the encoded result through a reverse technique or reverse algorithm.

data field. The third (optional) field of a CKD record. You determine the field length by the data length that is specified in the count field. The data field contains data that the program writes.

data record. A subsystem stores data records on a track by following the track-descriptor record. The subsystem numbers the data records consecutively, starting with 1. A track can store a maximum of 255 data records. Each data record consists of a count field, a key field (optional), and a data field (optional).

DASD-Fast Write. A function of a storage controller that allows caching of active write data without

exposure of data loss by journaling of the active write data in NVS.

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

data in transit. The update data on application system DASD volumes that is being sent to the recovery system for writing to DASD volumes on the recovery system.

data mover. See system data mover.

dedicated storage. Storage within a storage facility that is configured such that a single host system has exclusive access to the storage.

demote. The action of removing a logical data unit from cache memory. A subsystem demotes a data unit in order to make room for other logical data units in the cache. It could also demote a data unit because the logical data unit is not valid. A subsystem must destage logical data units with active write units before they are demoted.

destage. (1) The process of reading data from cache. (2) The action of storing a logical data unit in cache memory with active write data to the storage device. As a result, the logical data unit changes from cached active write data to cached read data.

device. The ESA/390 term for a disk drive.

device address. The ESA/390 term for the field of an ESCON device-level frame that selects a specific device on a control-unit image. The one or two leftmost digits are the address of the channel to which the device is attached. The two rightmost digits represent the unit address.

device adapter. A physical sub unit of a storage controller that provides the ability to attach to one or more interfaces used to communicate with the associated storage devices.

device ID. An 8-bit identifier that uniquely identifies a physical I/O device.

device interface card. A physical sub unit of a storage cluster that provides the communication with the attached DDMs.

device number. ESA/390 term for a four-hexadecimal-character identifier, for example 13A0, that you associate with a device to facilitate communication between the program and the host operator. The device number that you associate with a subchannel.

device sparing. Refers to when a subsystem automatically copies data from a failing DDM to a spare DDM. The subsystem maintains data access during the process.

Device Support Facilities program (ICKDSF). A program used to initialize DASD at installation and perform media maintenance.

DFDSS. Data Facility Data Set Services.

DFSMSdss. A functional component of DFSMS/MVS used to copy, dump, move, and restore data sets and volumes.

director. See storage director and ESCON Director.

disaster recovery. Recovery after a disaster, such as a fire, that destroys or otherwise disables a system. Disaster recovery techniques typically involve restoring data to a second (recovery) system, then using the recovery system in place of the destroyed or disabled application system. See also recovery, backup, and recovery system.

disk drive module. The primary nonvolatile storage medium that you use for any host data that is stored within a subsystem. Number and type of storage devices within a storage facility may vary.

drawer. A unit that contains multiple DDMs, and provides power, cooling, and related interconnection logic to make the DDMs accessible to attached host systems.

DRAIN. A keyword for requesting deletion or suspension when all existing record updates from the storage control cache have been cleared.

drawer. A unit that contains multiple DDMs, and provides power, cooling, and related interconnection logic to make the DDMs accessible to attached host systems.

dump. A capture of valuable storage information at the time of an error.

dual copy. A high availability function made possible by the nonvolatile storage in cached IBM storage controls. Dual copy maintains two functionally identical copies of designated DASD volumes in the logical storage subsystem, and automatically updates both copies every time a write operation is issued to the dual copy logical volume.

duplex pair. A volume comprised of two physical devices within the same or different storage subsystems that are defined as a pair by a dual copy, PPRC, or XRC operation, and are in neither suspended nor pending state. The operation records the same data onto each volume.

E

ECSA. Extended common service area.

EMIF. ESCON Multiple Image Facility. An ESA/390 function that allows LPARs to share an ESCON channel path by providing each LPAR with its own channel-subsystem image.

environmental data. Data that the storage control must report to the host; the data can be service information message (SIM) sense data, logging mode

sense data, an error condition that prevents completion of an asynchronous operation, or a statistical counter overflow. The storage control reports the appropriate condition as unit check status to the host during a channel initiated selection. Sense byte 2, bit 3 (environmental data present) is set to 1.

Environmental Record Editing and Printing (EREP) program. The program that formats and prepares reports from the data contained in the error recording data set (ERDS).

EREP. Environmental Record Editing and Printing Program.

ERP. Error recovery procedure.

ESCD. ESCON Director.

ESCM. ESCON Manager.

ESCON. Enterprise Systems Connection Architecture. An ESA/390 computer peripheral interface. The I/O interface utilizes ESA/390 logical protocols over a serial interface that configures attached units to a communication fabric.

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

extended remote copy (XRC). A hardware- and software-based remote copy service option that provides an asynchronous volume copy across storage subsystems for disaster recovery, device migration, and workload migration.

ESCON Manager (ESCM). A licensed program that provides host control and intersystem communication capability for ESCON Director connectivity operations.

F

failover. The routing of all transactions to a second controller when the first controller fails. Also see cluster.

fast write. A write operation at cache speed that does not require immediate transfer of data to a DDM. The subsystem writes the data directly to cache, to nonvolatile storage, or to both. The data is then available for destaging. Fast write reduces the time an application must wait for the I/O operation to complete.

FBA. Fixed block address. An architecture for logical devices that specifies the format of and access mechanisms for the logical data units on the device. The logical data unit is a block. All blocks on the device are the same size (fixed size); the subsystem can access them independently.

FC-AL. Fibre Channel - Arbitrated Loop. An implementation of the fibre channel standard that uses a ring topology for the communication fabric.

FCS. See fibre channel standard.

fibre channel standard. An ANSI standard for a computer peripheral interface. The I/O interface defines a protocol for communication over a serial interface that configures attached units to a communication fabric. The protocol has two layers. The IP layer defines basic interconnection protocols. The upper layer supports one or more logical protocols (for example FCP for SCSI command protocols, SBCON for ESA/390 command protocols). **fiber optic cable.** A fiber, or bundle of fibers, in a structure built to meet optic, mechanical, and environmental specifications.

fixed utility volume. A simplex volume assigned by the storage administrator to a logical storage subsystem to serve as working storage for XRC functions on that storage subsystem.

FlashCopy. A point-in-time copy services function that can quickly copy data from a source location to a target location.

floating utility volume. Any volume of a pool of simplex volumes assigned by the storage administrator to a logical storage subsystem to serve as dynamic storage for XRC functions on that storage subsystem

G

GB. Gigabyte.

gigabyte. 1 073 741 824 bytes.

group. A group consist of eight DDMs. Each DDM group is a raid array.

GTF. Generalized trace facility.

H

HA. Home address, host adapter.

hard drive. A storage medium within a storage server used to maintain information that the storage server requires.

HDA. Head and disk assembly. The portion of an HDD associated with the medium and the read/write head.

HDD. Head and disk drive.

home address. A nine-byte field at the beginning of a track that contains information that identifies the physical track and its association with a cylinder.

host adapter. A physical sub unit of a storage controller that provides the ability to attach to one or more host I/O interfaces.

I

ICKDSF. See Device Support Facilities program.

identifier (ID). A sequence of bits or characters that identifies a program, device, storage control, or system.

IML. Initial microcode load.

initial microcode load (IML). The act of loading microcode.

I/O device. An addressable input/output unit, such as a direct access storage device, magnetic tape device, or printer.

I/O interface. An interface that you define in order to allow a host to perform read and write operations with its associated peripheral devices.

implicit allegiance. ESA/390 term for a relationship that a controller creates between a device and a channel path, when the device accepts a read or write operation. The controller guarantees access to the channel program over the set of channel paths that it associates with the allegiance.

Internet Protocol (IP). A protocol used to route data from its source to its destination in an Internet environment.

invalidate. The action of removing a logical data unit from cache memory because it cannot support continued access to the logical data unit on the device. This removal may be the result of a failure within the storage controller or a storage device that is associated with the device.

IPL. Initial program load.

ITSO. International Technical Support Organization.

J

JCL. Job control language.

Job control language (JCL). A problem-oriented language used to identify the job or describe its requirements to an operating system.

journal. A checkpoint data set that contains work to be done. For XRC, the work to be done consists of all changed records from the primary volumes. Changed records are collected and formed into a "consistency group", and then the group of updates is applied to the secondary volumes.

K

KB. Kilobyte.

key field. The second (optional) field of a CKD record. The key length is specified in the count field. The key length determines the field length. The program writes the data in the key field. The subsystem uses this data to identify or locate a given record.

keyword. A symptom that describes one aspect of a program failure.

kilobyte (KB). 1 024 bytes.

km. Kilometer.

L

LAN. See local area network.

least recently used. The algorithm used to identify and make available the cache space that contains the least-recently used data.

licensed internal code (LIC).

(1) Microcode that IBM does not sell as part of a machine, but licenses to the customer. LIC is implemented in a part of storage that is not addressable by user programs. Some IBM products use it to implement functions as an alternative to hard-wired circuitry.

(2) LIC is implemented in a part of storage that is not addressable by user programs. Some IBM products use it to implement functions as an alternative to hard-wired circuitry.

link address. On an ESCON interface, the portion of a source, or destination address in a frame that ESCON uses to route a frame through an ESCON director. ESCON associates the link address with a specific switch port that is on the ESCON director. Equivalently, it associates the link address with the channel-subsystem, or controller-link-level functions that are attached to the switch port.

link-level facility. ESCON term for the hardware and logical functions of a controller or channel subsystem that allows communication over an ESCON write interface and an ESCON read interface.

local area network (LAN). A computer network located on a user's premises within a limited geographical area.

logical address. On an ESCON interface, the portion of a source or destination address in a frame used to select a specific channel-subsystem or control-unit image.

logical data unit. A unit of storage which is accessible on a given device.

logical device. The functions of a logical subsystem with which the host communicates when performing I/O operations to a single addressable-unit over an I/O interface. The same device may be accessible over more than one I/O interface.

logical disk drive. See logical volume.

logical subsystem. The logical functions of a storage controller that allow one or more host I/O interfaces to access a set of devices. The controller aggregates the devices according to the addressing mechanisms of the associated I/O interfaces. One or more logical subsystems exist on a storage controller. In general, the controller associates a given set of devices with only one logical subsystem.

logical unit. The SCSI term for a logical disk drive.

logical unit number. The SCSI term for the field in an identifying message that is used to select a logical unit on a given target.

logical partition (LPAR). The ESA/390 term for a set of functions that create the programming environment that is defined by the ESA/390 architecture. ESA/390 architecture uses this term when more than one LPAR is established on a processor. An LPAR is conceptually similar to a virtual machine environment except that the LPAR is a function of the processor. Also the LPAR does not depend on an operating system to create the virtual machine environment.

logical volume. The storage medium associated with a logical disk drive. A logical volume typically resides on one or more storage devices. A logical volume is referred to on an AIX platform as an hdisk, an AIX term for storage space. A host system sees a logical volume as a physical volume.

LSS. See logical subsystem.

LUN. See logical unit number.

least-recently used (LRU). A policy for a caching algorithm which chooses to remove the item from cache which has the longest elapsed time since its last access.

M

MB. Megabyte.

megabyte (MB). 1 048 576 bytes.

metadata. Internal control information used by microcode. It is stored in reserved area within disk array. The usable capacity of the array take care of the metadata.

million instructions per second (MIPS). A general measure of computing performance and, by implication, the amount of work a larger computer can do. The term is used by IBM and other computer manufacturers . For large servers or mainframes, it is

also a way to measure the cost of computing: the more MIPS delivered for the money, the better the value.

MTBF. Mean time between failures. A projection of the time that an individual unit remains functional. The time is based on averaging the performance, or projected performance, of a population of statistically independent units. The units operate under a set of conditions or assumptions.

Multiple Virtual Storage (MVS). One of a family of IBM operating systems for the System/370 or System/390 processor, such as MVS/ESA.

MVS. Multiple Virtual Storage.

N

nondisruptive. The attribute of an action or activity that does not result in the loss of any existing capability or resource, from the customer's perspective.

nonvolatile storage (NVS). Random access electronic storage with a backup battery power source, used to retain data during a power failure. Nonvolatile storage, accessible from all cached IBM storage clusters, stores data during DASD fast write, dual copy, and remote copy operations.

NVS. Nonvolatile storage.

O

open system. A system whose characteristics comply with standards made available throughout the industry, and therefore can be connected to other systems that comply with the same standards.

operating system. Software that controls the execution of programs. An operating system may provide services such as resource allocation, scheduling, input/output control, and data management.

orphan data. Data that occurs between the last, safe backup for a recovery system and the time when the application system experiences a disaster. This data is lost when either the application system becomes available for use or when the recovery system is used in place of the application system.

P

path group. The ESA/390 term for a set of channel paths that are defined to a controller as being associated with a single LPAR. The channel paths are in a group state and are on-line to the host.

path-group identifier. The ESA/390 term for the identifier that uniquely identifies a given LPAR. The path-group identifier is used in communication between the LPAR program and a device to associate the path-group identifier with one or more channel paths. This identifier defines these paths to the control unit as being associated with the same LPAR.

partitioned data set extended (PDSE). A system-managed, page-formatted data set on direct access storage.

P/DAS. PPRC dynamic address switching.

PDSE. Partitioned data set extended.

peer-to-peer remote copy (PPRC). A hardware based remote copy option that provides a synchronous volume copy across storage subsystems for disaster recovery, device migration, and workload migration.

pending. The initial state of a defined volume pair, before it becomes a duplex pair. During this state, the contents of the primary volume are copied to the secondary volume.

pinned data. Data that is held in a cached storage control, because of a permanent error condition, until it can be destaged to DASD or until it is explicitly discarded by a host command. Pinned data exists only when using fast write, dual copy, or remote copy functions.

port. (1) An access point for data entry or exit. (2) A receptacle on a device to which a cable for another device is attached.

PPRC. Peer-to-peer remote copy.

PPRC dynamic address switching (P/DAS). A software function that provides the ability to dynamically redirect all application I/O from one PPRC volume to another PPRC volume.

predictable write. A write operation that can cache without knowledge of the existing formatting on the medium. All writes on FBA DASD devices are predictable. On CKD DASD devices, a write is predictable if it does a format write for the first record on the track.

primary device. One device of a dual copy or remote copy volume pair. All channel commands to the copy logical volume are directed to the primary device. The data on the primary device is duplicated on the secondary device. See also secondary device.

PTF. Program temporary fix.

R

RACF. Resource access control facility.

rack. A unit that houses the components of a storage subsystem, such as controllers, disk drives, and power.

random access. A mode of accessing data on a medium in a manner that requires the storage device to access nonconsecutive storage locations on the medium.

read hit. When data requested by the read operation is in the cache.

read miss. When data requested by the read operation is not in the cache.

recovery. The process of rebuilding data after it has been damaged or destroyed. In the case of remote copy, this involves applying data from secondary volume copies.

recovery system. A system that is used in place of a primary application system that is no longer available for use. Data from the application system must be available for use on the recovery system. This is usually accomplished through backup and recovery techniques, or through various DASD copying techniques, such as remote copy.

remote copy. A storage-based disaster recovery and workload migration function that can copy data in real time to a remote location. Two options of remote copy are available. See peer-to-peer remote copy and extended remote copy.

reserved allegiance. ESA/390 term for a relationship that is created in a controller between a device and a channel path, when a Sense Reserve command is completed by the device. The allegiance causes the control unit to guarantee access (busy status is not presented) to the device. Access is over the set of channel paths that are associated with the allegiance; access is for one or more channel programs, until the allegiance ends.

restore. Synonym for recover.

resynchronization. A track image copy from the primary volume to the secondary volume of only the tracks which have changed since the volume was last in duplex mode.

RVA. RAMAC Virtual Array Storage Subsystem.

S

SAID. System adapter identification.

SAM. Sequential access method.

SCSI. Small Computer System Interface. An ANSI standard for a logical interface to computer peripherals and for a computer peripheral interface. The interface utilizes a SCSI logical protocol over an I/O interface that configures attached targets and initiators in a multi-drop bus topology.

SCSI ID. A unique identifier assigned to a SCSI device that is used in protocols on the SCSI interface to identify or select the device. The number of data bits

on the SCSI bus determines the number of available SCSI IDs. A wide interface has 16 bits, with 16 possible IDs. A SCSI device is either an initiator or a target.

Seascape architecture. A storage system architecture developed by IBM for open system servers and S/390 host systems. It provides storage solutions that integrate software, storage management, and technology for disk, tape, and optical storage.

secondary device. One of the devices in a dual copy or remote copy logical volume pair that contains a duplicate of the data on the primary device. Unlike the primary device, the secondary device may only accept a limited subset of channel commands.

sequential access. A mode of accessing data on a medium in a manner that requires the storage device to access consecutive storage locations on the medium.

server. A type of host that provides certain services to other hosts that are referred to as clients.

service information message (SIM). A message, generated by a storage subsystem, that is the result of error event collection and analysis. A SIM indicates that some service action is required.

sidefile. A storage area used to maintain copies of tracks within a concurrent copy domain. A concurrent copy operation maintains a sidefile in storage control cache and another in processor storage.

SIM. Service information message.

simplex state. A volume is in the simplex state if it is not part of a dual copy or a remote copy volume pair. Ending a volume pair returns the two devices to the simplex state. In this case, there is no longer any capability for either automatic updates of the secondary device or for logging changes, as would be the case in a suspended state.

SMF. System Management Facilities.

SMS. Storage Management Subsystem.

SRM. System resources manager.

SnapShot copy. A point-in-time copy services function that can quickly copy data from a source location to a target location.

spare. A disk drive that is used to receive data from a device that has experienced a failure that requires disruptive service. A spare can be pre-designated to allow automatic dynamic sparing. Any data on a disk drive that you use as a spare is destroyed by the dynamic sparing copy process.

SSA. Serial Storage Architecture. An IBM standard for a computer peripheral interface. The interface uses a SCSI logical protocol over a serial interface that configures attached targets and initiators in a ring topology.

SSID. Subsystem identifier.

stacked status. An ESA/390 term used when the control unit is holding for the channel; the channel responded with the stack-status control the last time the control unit attempted to present the status.

stage. The process of reading data into cache from a disk drive module.

storage cluster. A power and service region that runs channel commands and controls the storage devices. Each storage cluster contains both channel and device interfaces. Storage clusters also perform the DASD control functions.

storage control. The component in a storage subsystem that handles interaction between processor channel and storage devices, runs channel commands, and controls storage devices.

STORAGE_CONTROL_DEFAULT. A specification used by several XRC commands and messages to refer to the timeout value specified in the maintenance panel of the associated storage control.

storage device. A physical unit which provides a mechanism to store data on a given medium such that it can be subsequently retrieved. Also see disk drive module.

storage director. In an IBM storage control, a logical entity consisting of one or more physical storage paths in the same storage cluster. See also storage path.

storage facility. (1) A physical unit which consists of a storage controller integrated with one or more storage devices to provide storage capability to a host computer. (2) A storage server and its attached storage devices.

Storage Management Subsystem (SMS). A component of MVS/DFP that is used to automate and centralize the management of storage by providing the storage administrator with control over data class, storage class, management class, storage group, aggregate group and automatic class selection routine definitions.

storage server. A unit that manages attached storage devices and provides access to the storage or storage related functions for one or more attached hosts.

storage path. The hardware within the IBM storage control that transfers data between the DASD and a channel. See also storage director.

storage subsystem. A storage control and its attached storage devices.

string. A series of connected DASD units sharing the same A-unit (or head of string).

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

subchannel. A logical function of a channel subsystem associated with the management of a single device.

subsystem. See DASD subsystem or storage subsystem.

subsystem identifier (SSID). A user-assigned number that identifies a DASD subsystem. This number is set by the service representative at the time of installation and is included in the vital product data.

suspended state. When only one of the devices in a dual copy or remote copy volume pair is being updated because of either a permanent error condition or an authorized user command. All writes to the remaining functional device are logged. This allows for automatic resynchronization of both volumes when the volume pair is reset to the active duplex state.

synchronization. An initial volume copy. This is a track image copy of each primary track on the volume to the secondary volume.

synchronous operation. A type of operation in which the remote copy PPRC function copies updates to the secondary volume of a PPRC pair at the same time that the primary volume is updated. Contrast with asynchronous operation.

system data mover. A system that interacts with storage controls that have attached XRC primary volumes. The system data mover copies updates made to the XRC primary volumes to a set of XRC-managed secondary volumes.

system-managed data set. A data set that has been assigned a storage class.

T

TCP/IP. Transmission Control Protocol/Internet Protocol.

TOD. Time of day.

Time Sharing Option (TSO). A System/370 operating system option that provides interactive time sharing from remote terminals.

timeout. The time in seconds that the storage control remains in a "long busy" condition before physical sessions are ended.

timestamp. The affixed value of the system time-of-day clock at a common point of reference for all write I/O operations directed to active XRC primary volumes. The UTC format is yyyy.ddd hh:mm:ss.thmiju.

track. A unit of storage on a CKD device that can be formatted to contain a number of data records. Also see home address, track-descriptor record, and data record.

track-descriptor record. A special record on a track that follows the home address. The control program uses it to maintain certain information about the track. The record has a count field with a key length of zero, a data length of 8, and a record number of 0. This record is sometimes referred to as R0.

TSO. Time Sharing Option.

U

Ultra-SCSI. An enhanced small computer system interface.

unit address. The ESA/390 term for the address associated with a device on a given controller. On ESCON interfaces, the unit address is the same as the device address. On OEMI interfaces, the unit address specifies a controller and device pair on the interface.

Universal Time, Coordinated. Replaces Greenwich Mean Time (GMT) as a global time reference. The format is yyyy.ddd hh:mm:ss.thmiju.

utility volume. A volume that is available to be used by the extended remote copy function to perform data mover I/O for a primary site storage control's XRC-related data.

UTC. Universal Time, Coordinated.

V

vital product data (VPD). Nonvolatile data that is stored in various locations in the DASD subsystem. It includes configuration data, machine serial number, and machine features.

volume. An ESA/390 term for the information recorded on a single unit of recording medium. Indirectly, it can refer to the unit of recording medium itself. On a non-removable medium storage device, the terms may also refer, indirectly, to the storage device that you associate with the volume. When you store multiple volumes on a single storage medium transparently to the program, you may refer to the volumes as logical volumes.

vital product data (VPD). Information that uniquely defines the system, hardware, software, and microcode elements of a processing system.

VSAM. Virtual storage access method.

VTOC. Volume table of contents.

W

workload migration. The process of moving an application's data from one set of DASD to another for

the purpose of balancing performance needs, moving to new hardware, or temporarily relocating data.

write hit. A write operation where the data requested is in the cache.

write miss. A write operation where the data requested is not in the cache.

write penalty. The term that describes the classical RAID write operation performance impact.

write update. A write operation that updates a direct access volume.

X

XDF. Extended distance feature (of ESCON).

XRC. Extended remote copy.

XRC planned-outage-capable. A storage subsystem with an LIC level that supports a software bitmap but not a hardware bitmap.

Index

Numerics

2103-H07 4
2109 4

A

Adaptec 8
AIX 7, 27, 131
 Configuration Manager 28
Application Server 138
asynchronous 4

B

backup 183

C

cache 1, 3, 4
CLI 5, 21
 rsList2105s.sh 35
Cluster 131
Cluster Complex 3
Cluster Processor 3
Command Line Interface 21, 24
Concurrent Copy 3
Copy Services 4, 5, 134
 backup server 11
 Command Line Interface 132
 Menu 12
 primary server 10
 server 13
 Specialist 152
 Task Wizard 153
 terminate PPRC pairs 155

D

Data path Optimizer 21
Device Adapters 2
disaster 131
Disaster Recovery 136
disk resources 137

E

e-business 3
Emulex 9
ESCON 2, 137, 177
ESS 10, 131, 177
 Copy Services 131
 serial interface 11
 service terminal 11
ESS overview 2

F

FC-AL 2, 4
FCP 4

Fibre Channel 135
filesystem 138
FlashCopy 4, 5, 19, 183, 185
 background copy 22
 bitmap 19, 20
 cfgmgr 28
 chdev 29, 30, 33
 exportvg 29
 Fibre Channel LUN 21
 fsck 28, 31
 importvg 28
 JFS log logical volume 31
 Long method 29
 mklv 30
 mkvg 30
 physical volume map 30
 point-in-time copy 5, 19
 recreatevg 34
 rmdev 29
 SCSI 21
 SCSI target ID 21
 Short method using recreatevg 33
 source volume 20, 31
 T0 (time-zero) copy 5, 19
 target volume 20, 29, 31, 35
 varyoffvg 29
 varyonvg 28

H

HACMP 131, 134
 Application Server 141
 Cascading 135
 cluster 136
 Concurrent 135
 config_too_long 147
 node_up event 147
 Recovery 136
 resource group 140
 resources 140
 Rotating 135
 start script 141
 stop script 142
 two-node cluster 135, 139
HACMP cluster
 Failover 144
 Re-integration 144
 startup 143
HACMP for AIX 131
HACMP/ES 135
HAGEO 136
hdisk 134
Host Adapters 2
hostname 13
HP 10
HP UX 10

I

Importvg 185
IP address takeover 138
IPAT 138

J

Java Runtime for AIX 132
JNI 9

L

Logical Subsystem 20
Logical Volume Manager 28
lspv 184
LSS 20, 134, 177
LUN 5, 179
LVM 6, 28

M

management 3
McData 4
memory 3
multiple LUNs per SCSI ID 151

N

node failover 139
Nonvolatile Storage 3
NT
 Disk Administrator 26
NUMA-Q 10

O

ODM 183
Operating system 183

P

Physical Volume ID 135
Physical Volume Identifier 28
PPRC 1, 3, 5
 pairs 136
PPRC pairs 152
PVID 135, 183, 184
PVID - Physical Volume Identifier 28

Q

QLogic 8

R

RAID 1
recreatevg 31, 32
 FlashCopy procedure 33
resynchronization 144
RS/6000 131
 SP 131
RS6000 7

S

SAN 3
scalability 3
script 188
SCSI 2, 135
SCSI reserve 151
Seascape 1, 3
serial interface 11
service terminal 11
shared filesystems 147
short-wave 4
Solaris 9
SPARC 9
Specialist 177
static routes 140
StorWatch 3
subnet 139
subnets 148
Subsystem Device Driver 21
Sun 9
SUN Solaris 35
supercomputers 1
supported servers 7
switch 4
Symbios 8
synchronous 4

T

T0 5
target 183
Task names, length of 143
TCP/IP 21
TCP/IP address 13

U

UDP broadcast packet 140, 148

V

VGDA - Volume Group Descriptor Area 28
VGID - Volume Group Identifier 28
Volume Group 183
volume group 138
Volume Group Descriptor Area 28
Volume Group Identifier 28

W

Windows NT 24

X

XRC 1, 3

IBM Redbooks review

Your feedback is valued by the Redbook authors. In particular we are interested in situations where a Redbook "made the difference" in a task or problem you encountered. Using one of the following methods, **please review the Redbook, addressing value, subject matter, structure, depth and quality as appropriate.**

- Use the online **Contact us** review redbook form found at ibm.com/redbooks
- Fax this form to: USA International Access Code + 1 845 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Document Number	SG24-5757-00
Redbook Title	Implementing ESS Copy Services on UNIX and Windows NT/2000
Review	
What other subjects would you like to see IBM Redbooks address?	
Please rate your overall satisfaction:	<input type="radio"/> Very Good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor
Please identify yourself as belonging to one of the following groups:	<input type="radio"/> Customer <input type="radio"/> Business Partner <input type="radio"/> Solution Developer <input type="radio"/> IBM, Lotus or Tivoli Employee <input type="radio"/> None of the above
Your email address: The data you provide here may be used to provide you with information from IBM or our business partners about our products, services or activities.	<input type="checkbox"/> Please do not use the information collected here for future marketing or promotional contacts or other communications beyond the scope of this transaction.
Questions about IBM's privacy policy?	The following link explains how we protect your personal information. ibm.com/privacy/yourprivacy/

Implementing ESS Copy Services on UNIX and Windows NT/2000

(0.2" spine)
0.17" x 0.473"
90 x 249 pages



Redbooks

Implementing ESS Copy Services on UNIX and Windows NT/2000

How to exploit the advanced functionality of the ESS

This IBM Redbook will help you to install, tailor, and configure the new Copy Services functions of the IBM Enterprise Storage Server (ESS) on the UNIX, Windows NT, and Windows 2000 platforms.

FlashCopy your data and eliminate downtime for backup

The Copy Services functions include Peer-to-Peer Remote Copy (PPRC), FlashCopy, Extended Remote Copy (XRC), and Concurrent Copy (CC). It should be noted that the latter two Copy Services functions, XRC and CC, are not available on UNIX and NT platforms. They are only available on System/390.

PPRC, the way to protect data against a site failure

This redbook provides a broad understanding of these functions, describes the prerequisites and corequisites, and then shows you how to implement each of the functions into your environment to ensure efficient usage and to maximize the benefits that these functions provide. This redbook also shows how to automate site failover using HACMP.

**INTERNATIONAL
TECHNICAL
SUPPORT
ORGANIZATION**

**BUILDING TECHNICAL
INFORMATION BASED ON
PRACTICAL EXPERIENCE**

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

**For more information:
ibm.com/redbooks**

SG24-5757-00

ISBN 0738418838