

Copyright © 2003-2007 EMC Corporation. All rights reserved.

Published August, 2007

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

All other trademarks used herein are the property of their respective owners.

# **Contents**

Preface		1X
Chapter 1	About CX300, CX300i, CX500, CX500i, and CX Systems	700
	Introducing CX300, CX300i, CX500, CX500i, and CX700	Storage
	Systems	1-2
	Fibre Channel	1-4
	Fibre Channel Storage Components	1-5
	Storage Component (Storage System Enclosures)	
	iSCSI	
	iSCSI Storage Components	
	Initiator Components	
	LAN Components	
	Target Components	
	Types of Storage-System Installations	
	About Switched Shared Storage and SANs (Storage Are	
	Networks)	
	Storage Groups	
Chapter 2	RAID Types and Trade-offs	
	Introducing RAID	2-2
	Disk Striping	
	Mirroring	
	RAID Groups and LUNs	

	RAID Types	2-4
	RAID 6 Group (Double Distributed Parity)	
	RAID 5 Group (Distributed Parity)	2-5
	RAID 3 Group (Single-Disk Parity)	2-7
	RAID 1 Mirrored Pair	2-8
	RAID 1/0 Group (Mirrored RAID 0 Group)	2-9
	RAID 0 Group (Nonredundant RAID Striping)	2-10
	Individual Disk Unit	2-11
	Hot Spare	2-11
	Proactive Sparing	2-13
	RAID Benefits and Trade-offs	2-14
	Performance	
	Storage Flexibility	
	Data Availability and Disk Space Usage	2-16
	Guidelines for RAID Groups	2-18
	Sample Applications for RAID Types	2-20
Chapter 3	Storage-System Management	
	Management Ports	
	Navisphere Management Suites	
	Access Logix Software	
	Navisphere Manager	
	Navisphere Host Agent and CLI	
	Navisphere Analyzer	3-7
	Navisphere Integrator	
	Using Navisphere Software	
	Navisphere Quality of Service Manager	
	Navisphere Management Worksheet	3-10
Chapter 4	Data Replication and	
•	Data Mobility Software	
	Data Replication and Data Mobility Software Overview	
	What Is SnapView?	4-4
	Clones Overview and Example	4-4
	Snapshots Overview and Example	4-6
	Clone and Snapshot Trade-offs	4-7
	SnapView Worksheets	
	What Is MirrorView/S?	4-12
	Sample MirrorView/S Configuration	4-13

	MirrorView/S Features and Benefits 4-14
	Provision for Disaster Recovery with Minimal Overhead 4-15
	CLARiiON MirrorView/S Environment 4-15
	Bidirectional Mirroring 4-15
	Integration with EMC SnapView Snapshot Software 4-16
	Integration with EMC SAN Copy Software 4-16
	MirrorView/S Planning Worksheet 4-17
	What Is MirrorView/A? 4-19
	Sample MirrorView / A Configuration 4-20
	MirrorView/A Features and Benefits 4-22
	Provision for Disaster Recovery with Minimal Overhead 4-22
	CLARiiON MirrorView/A Environment 4-22
	Bidirectional Mirroring 4-23
	Integration with EMC SnapView Snapshot Software 4-23
	Integration with EMC SAN Copy Software 4-23
	Replication Over Long Distances 4-24
	MirrorView/A Planning Worksheets 4-25
	What Is EMC SAN Copy? 4-27
	SAN Copy Example
	SAN Copy Features and Benefits 4-30
	Implementing SAN Copy Over Extended Distances 4-31
	SAN Copy Planning Worksheet 4-33
	What Is EMC SAN Copy/E? 4-35
	SAN Copy/E Example 4-37
	SAN Copy/E Features and Benefits 4-38
	Implementing SAN Copy/E over Extended Distances 4-40
	Performance Considerations
	SAN Copy/E Planning Worksheet 4-42
Chapter 5	Planning File Systems and LUNs
	Multiple Paths to LUNs 5-2
	Disk ÎDs 5-2
	Sample Shared Switched Installation 5-2
	Sample Unshared Direct Installation 5-5
	Planning Applications, LUNs, and Storage Groups 5-6
	Application and LUN Planning 5-7
	LUN and Storage Group Planning Worksheet 5-9
	LUN Details Worksheet 5-14

Chapter 6	Storage-System Hardware
	Storage-System Installation Types
	Systems
	Storage Hardware 6-3
	Storage Processor (SP)
	Disks
	Planning Your Hardware Components
	Components for Storage Systems 6-8
	Hardware Dimensions and Requirements 6-9
	Cabinets for Enclosures 6-10
	Fibre Channel Data Cable and Configuration Guidelines 6-11
	Fibre Channel Front-End Hardware Planning Worksheets 6-13
Chapter 7	iSCSI Configuration
	iSCSI Configuration Options
	iSNS
	CHAP
	Storage System iSCSI Data Port Worksheet
	iSCSI Initiator Port Worksheet
	iSNS Server Worksheet
	iSCSI CHAP Authentication Worksheets
	What Is CHAP?
	Initiator CHAP Worksheet
	Mutual CHAP Worksheet
Index	i-1

# **Figures**

1-1	CX700 Storage System	1-3
1-2	Nodes - Initiator and Target	1-4
1-3	Switch and Point-to-Point Connections	1-6
1-4	A Switch Zone	
1-5	Typical 16-Port Switch, Back View	1-7
1-6	CX700 Storage Processor Enclosure (SPE)	
1-7	Nodes - Initiator and Target	1-9
1-8	Switch and Point-to-Point Connections	1-10
1-9	Types of Storage-System Installation	1-12
1-10	Components of a SAN	1-13
1-11	Sample Shared Storage Configuration	1-15
1-12	Data Access Control with Shared Storage	
2-1	Multiple LUNs in a RAID Group	
2-2	RAID 6 Group	
2-3	RAID 5 Group	2-6
2-4	RAID 3 Group	2-7
2-5	RAID 1 Mirrored Pair	2-9
2-6	RAID 1/0 Group	2-10
2-7	How a Hot Spare Works	2-12
2-8	Disk Space Usage in the RAID Configuration	2-17
3-1	Storage System Serial Number Label (CX300i Enclosure)	3-3
3-2	Storage System Serial Number Label - CX700	
3-3	Storage Domains on the Internet	
4-1	Clones Example	4-5
4-2	Snapshot Example	4-7
4-3	Sample Remote Mirror Configuration	4-13
4-4	Sample Remote Mirror Configuration	4-20
4-5	Sample SAN Copy Sessions Between Storage Systems	4-29
4-6	Sample SAN Copy Over IP Configuration	
4-7	Sample SAN Copy/E Session Between Storage Systems	

#### Contents

4-8	Sample SAN Copy/E over IP Configuration	4-41
5-1	Sample Shared Switched Storage Configuration	
5-2	Unshared Direct Storage System	
6-1	Types of Fibre Channel Storage-System Installations	
6-2	CX700 Storage System	
6-3	Shared Storage Systems	
6-4	Disks and Their IDs	

#### **Preface**

As part of an effort to improve and enhance the performance and capabilities of its product line, EMC from time to time releases revisions of its hardware and software. Therefore, some functions described in this guide may not be supported by all revisions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes.

If a product does not function properly or does not function as described in this guide, please contact your EMC representative.

#### **About this Manual**

This planning guide introduces EMC® CLARiiON® CX300, CX300i, CX500, CX500i, and CX700 disk-array storage systems and offers useful background information and worksheets to help you plan.

#### Audience

#### Read this guide

- if you are considering purchase of a CX300, CX300i, CX500, CX500i, or CX700 disk-array storage system and want to understand its features; or
- before you plan the installation of a storage system.

You should be familiar with the host servers that will use the storage systems and with the operating systems on the servers. After reading this guide, you will be able to

- determine the best storage-system components for your installation,
- determine your site requirements, and
- configure storage systems correctly

#### Organization

This manual is organized as follows:

- Chapter 1 Provides background information about Fibre Channel and iSCSI features and explains the major types of storage.
- Chapter 2 Describes the RAID Groups and the different ways they store data.
- Chapter 3 Explains storage management software.
- Chapter 4 Describes the optional EMC storage-system based applications for replicating data or copying data within or between CLARiiON storage systems:
  - ◆ EMC SnapView<sup>™</sup> software for replicating data with a storage system using snapshots or clones.
  - ◆ EMC MirrorView<sup>TM</sup> /Synchronous and EMC MirrorView/Asynchronous remote mirroring software for replicating data between storage systems.
  - EMC SAN Copy ™ and EMC SAN Copy/E software for copying data directly from a logical unit on one storage system to destination logical units on supported remote systems without using host resources.
- Chapter 5 Helps you plan your storage-system file systems and LUNs.
- Chapter 6 Describes the hardware components of storage systems.
- Chapter 7 Explains how to configure CX300i and CX500i iSCSI storage systems.

# Conventions used in this guide

EMC uses the following conventions for notes and cautions.

A note presents information that is important, but not hazard-related.



#### CAUTION

A caution contains information essential to avoid data loss or damage to the system or equipment. The caution may apply to hardware or software.

#### Where to get help

EMC support, product, and licensing information can be obtained as follows.

**Product information** — For documentation, release notes, software updates, or for information about EMC products, licensing, and service, go to the EMC Powerlink website (registration required) at:

http://Powerlink.EMC.com

**Technical support** — For technical support, go to EMC WebSupport on Powerlink. To open a case on EMC WebSupport, you must be a WebSupport customer. Information about your site configuration and the circumstances under which the problem occurred is required.

#### Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Please send your opinion of this guide to:

techpub\_comments@EMC.com

Preface	
Troidee	

# About CX300, CX300i, CX500, CX500i, and CX700 Systems

This chapter introduces CLARiiON® CX300, CX300i, CX500, CX500i, and CX700 disk-array storage systems and storage area networks (SANs).

#### Major sections are

•	Introducing CX300, CX300i, CX500, CX500i, and CX700	
	Storage Systems	1-2
•	Fibre Channel	
•	iSCSI	1-9
<b>*</b>	Types of Storage-System Installations	1-12
	About Switched Shared Storage and SANs (Storage Area	
	Networks)	1-13

# Introducing CX300, CX300i, CX500, CX500i, and CX700 Storage Systems

EMC® CLARiiON CX300, CX300i, CX500, CX500i, and CX700 disk-array storage systems provide terabytes of disk storage capacity, high transfer rates, flexible configurations, and highly available data at low cost. Their hardware RAID features are provided by two storage processors (SPs).

Table 1-1 outlines each storage system enclosure and disk capacities.

Table 1-1 Storage System Enclosures and Disk Capacities

Storage System	Enclosure Type	Number of Disks per Storage System Enclosure	Maximum Number of 2-Gigabit Disk-Array Enclosures (DAEs) <sup>a</sup> and Disks
CX300 & CX300i	2-Gbit Disk Processor Enclosure (DPE2)	15 Fibre Channel disks	3 separate DAEs for a total of 60 disks
CX500 & CX500i	2-Gbit Disk Processor Enclosure (DPE2)	15 Fibre Channel disks	7 separate DAEs for a total of 120 disks
CX700	Storage Processor Enclosure (SPE)	An SPE does not include disks, and requires at least one 2-Gbit Disk-Array Enclosure (DAE) with Fibre Channel disks.	16 separate DAEs for a total of 240 disks

a. Each DAE enclosure has slots for 15 disks.

The CX-Series storage systems discussed in this manual support three types of DAEs:

- ◆ 2-gigabit point-to-point disk-array enclosure (DAE2P) with high-performance Fibre Channel disks.
- Original 2-gigabit disk-array enclosure (DAE2) with high-performance Fibre Channel disks.
- DAE2-ATA with economical ATA (Advanced Technology Attachment) disks.

Figure 1-1 shows a typical CX700 storage system.

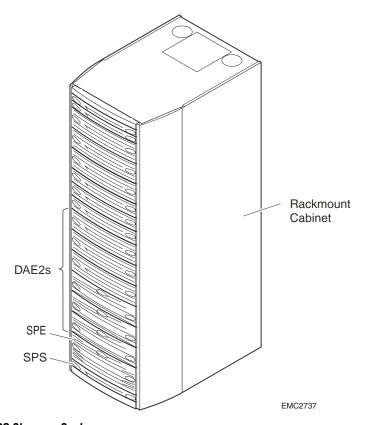


Figure 1-1 CX700 Storage System

A storage-system package includes storage management software, Fibre Channel interconnect hardware, and one or more storage systems. The host bus adapter driver hardware and software are available from outside vendors.

#### **Fibre Channel**

Fibre Channel is a high-performance serial protocol that allows transmission of both network and I/O channel data. It is a low level protocol, independent of data types, and supports such formats as SCSI and IP.

The Fibre Channel standard supports several physical topologies, including switch fabric point-to-point and arbitrated loop (FC-AL). The Fibre Channel storage systems described in this manual use switch fabric and FC-AL topologies.

A switch fabric is a set of point-to-point connections between nodes; each connection is made through one or more Fibre Channel switches. Each node may have its own unique address, but the path between nodes is governed by a switch. The nodes are connected by optical cable.

A Fibre Channel arbitrated loop is a circuit consisting of nodes. Each node has a unique address, called a Fibre Channel arbitrated loop address. The nodes connect with optical cables. An optical cable can transmit data over great distances for connections that span entire enterprises and can support remote disaster recovery systems.

Each connected device in a switched fabric or arbitrated loop is a server adapter (initiator) or a target (storage system). The switches are not considered nodes.

Figure 1-2 shows a node and initiator.

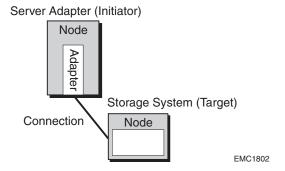


Figure 1-2 Nodes - Initiator and Target

#### **Fibre Channel Storage Components**

A Fibre Channel storage system has three main components:

- Server component (host bus adapter driver with adapter and software)
- ◆ Interconnect components (cables based on Fibre Channel standards, and switches)
- Storage component (storage system with storage processors, power supply, cooling hardware, and disks)

#### Server Component (Host Bus Adapter and Driver)

The host bus adapter is a printed-circuit board that slides into an I/O slot in the server's cabinet. It transfers data between server memory and one or more disk-array storage systems over Fibre Channel — as controlled by the support software (adapter driver).



#### **Interconnect Components**

The interconnect components include the optical cables between components and any Fibre Channel switches.

The maximum length of optical cable between a CX300-, CX500-, and CX700-Series storage system and a server or switch ranges from 150 to 500 meters (165 to 550 yards), depending on the type of cable and operating speed. With extenders, connections between servers, switches, and other devices can span up to 60 kilometers (36 miles) or more. This ability to span great distances is a major advantage of optical cable.

Details on cable lengths and rules appear later, in Chapter 6.

#### **Fibre Channel Switches**

A Fibre Channel switch, which is required for switched shared storage (a storage area network, SAN), connects all the nodes cabled to it using a fabric topology. A switch adds serviceability and scalability to any installation; it allows online insertion and removal of any device on the fabric and maintains integrity if any connected

device stops participating. A switch also provides server-to-storage-system access control and point-to-point connections.

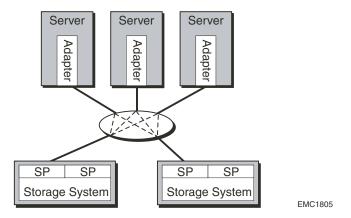


Figure 1-3 Switch and Point-to-Point Connections

You can cascade switches (connect one switch port to another switch) for additional port connections.

#### Switch Zoning

Switch zoning lets an administrator define paths between connected nodes based on the node's unique World Wide Name. Each zone includes a server adapter and/or one or more SPs. EMC recommends single-initiator zoning, which limits each zone to a single HBA port (initiator). The current connection limits are outlined in Table 1-2.

Table 1-2 Connection Limits

Storage System	Maximum HBA Ports per SP	Maximum HBA Ports per Storage System
CX300	32	128
CX500	64	256
CX700	64	512

In Figure 1-4, server 1 has access to one SP in storage systems 1 and 2; it has no access to any other SP.

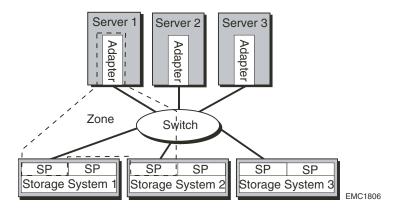


Figure 1-4 A Switch Zone

To illustrate switch zoning, Figure 1-4 shows just one HBA per server and one switch. Normally, such installations include multiple HBAs per server and two or more switches.

If you do not define a zone in a switch, all adapter ports connected to the switch can communicate with all SP ports connected to the switch. However, access to an SP does not necessarily provide access to the SP's storage. Access to storage is governed by the Storage Groups you create (explained in *Storage Groups* on page 1-14.

Fibre Channel switches are available with 8, 16, 32, or more ports. They are compact units that fit into a rackmount cabinet.

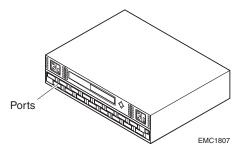


Figure 1-5 Typical 16-Port Switch, Back View

If your servers and storage systems will be far apart, you can place the switches closer to the servers or storage systems, as convenient. A switch is technically a repeater, not a node, in a Fibre Channel loop. However, it is bound by the same cabling distance rules as a node.

### **Storage Component (Storage System Enclosures)**

EMC CX300, CX500, and CX700 disk-array storage systems, with their storage processor (SP), power supplies, and cooling hardware form the storage component of a Fibre Channel system. The controlling unit is a CX300 or CX500 disk processor enclosure (DPE2) or a CX700 storage processor enclosure (SPE). A CX700 SPE, outside of its cabinet, is shown in Figure 1-6.

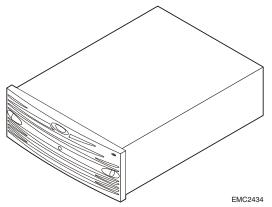


Figure 1-6 CX700 Storage Processor Enclosure (SPE)

Each CX300 (DPE2), CX500 (DPE2), and CX700 (SPE) has two SPs. The number of ports for each SP varies between storage systems. Table 1-3 lists the number of ports per SP for each system.

Table 1-3 Number of Storage System Ports per SP

Storage System	Front End Ports	Back End Ports
CX300	2	1
CX500	2	2
CX700	4	4

The front-end ports communicate with switches or servers and the back-end ports communicate with disks. Hardware details appear later in Chapter 6.

#### **iSCSI**

*Internet SCSI (iSCSI)* is a serial protocol that uses the internet protocol (IP) for transmission of both network and I/O data.

An iSCSI network (LAN) is a set of point-to-point connections between nodes; each connection is made directly or through one or more network components such as switches. Each node has its own unique IP address, while the LAN on which the storage system resides has a subnet mask address and a gateway address. Nodes are connected through a LAN by Ethernet CAT 6, CAT 5E, or CAT 5 copper cables. (We recommend CAT 6 cables for gigabit Ethernet connections.)

Each node in an iSCSI environment is either an initiator (in a server) or a target (in the storage system). Network switches are not nodes. Figure 1-2 shows an initiator node and a target node.

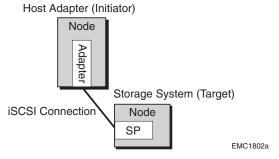


Figure 1-7 Nodes - Initiator and Target

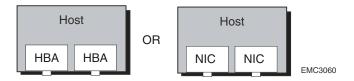
## iSCSI Storage Components

An iSCSI storage environment has three main components:

- Initiator components: servers containing host adapters, network interface cards (NICs), drivers, and software.
- LAN components: LAN cables and LAN devices such as switches and routers.
- Target components: storage systems with storage processors, power supply, cooling hardware, and disks.

#### **Initiator Components**

The network interface card (NIC) or host bus adapter (HBA) is a printed-circuit board that slides into an I/O slot in the server's cabinet. It transfers data between server memory and one or more storage systems over the LAN using the iSCSI protocol.



#### **LAN Components**

LAN components include the network cables (CAT 6 is recommended for gigabit Ethernet) and network components such as switches and routers.

#### **LAN Switches**

A LAN switch connects all the nodes cabled to it. A switch adds serviceability and scalability to any installation; it allows online insertion and removal of any device and maintains integrity if any connected device fails.

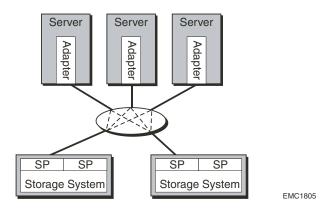


Figure 1-8 Switch and Point-to-Point Connections

#### **Target Components**

Target components are iSCSI storage systems, namely CX300i or CX500i disk-array storage systems, with storage processors (SPs), power supplies, and cooling hardware. The controlling unit is a CX300i/CX500i disk processor enclosure (DPE2). Each DPE2 has two SPs, each with two 1-gigabit iSCSI ports. The iSCSI front-end (data) ports communicate with switches or servers and the Fibre Channel back-end ports communicate with disks.

# **Types of Storage-System Installations**

You can use a storage system in any of several types of installation:

- Unshared direct with one server is the simplest and least costly.
- Shared-or-clustered direct, with a limit of two servers, lets two servers share storage resources with high availability.
- ◆ Shared switched, with two switch fabrics, lets two or more servers share the resources of several storage systems in a storage area network (SAN). Shared switched installations are available in high-availability versions (two HBAs per server) or with one HBA per server. Shared switched storage systems can have multiple paths to each SP, providing multipath I/O for dynamic load sharing and greater throughput.

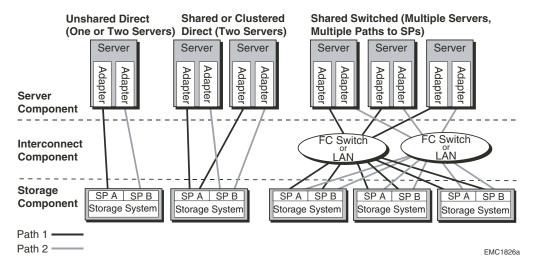


Figure 1-9 Types of Storage-System Installation

Storage systems for any shared installation require EMC Access Logix<sup>TM</sup> software to control server access to the storage-system LUNs.

The shared-or-clustered direct installation can be either shared (that is, use Access Logix to control LUN access) or clustered (without Access Logix, but with operating system cluster software controlling LUN access).

# About Switched Shared Storage and SANs (Storage Area Networks)

This section explains the features that let multiple servers share disk-array storage systems on a SAN (storage area network).

A SAN is one or more storage devices connected to servers through switches to provide a central location for disk storage. Centralizing disk storage among multiple servers has many advantages, including:

- highly available data
- flexible association between servers and storage capacity
- centralized management for fast, effective response to users' data storage needs
- easier file backup and recovery

An EMC SAN is based on shared storage; that is, the SAN requires EMC Access Logix software to provide flexible access control to storage-system LUNs. Within the SAN, a network connection to each SP in the storage system lets you configure and manage it.

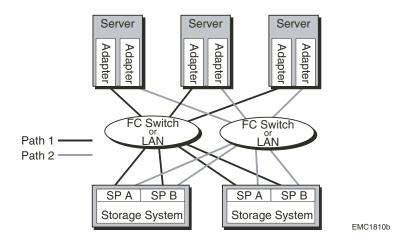


Figure 1-10 Components of a SAN

Fibre Channel switches can control data access to storage systems through the use of switch zoning, explained earlier on page 1-6. Ethernet switches in an iSCSI environment rely on authentication

methods such as CHAP to control access to storage processors. (See Chapter 7, *iSCSI Configuration*, for a description of host/target authentication).

Switch zoning cannot selectively control data access to LUNs in a storage system, because each SP appears as a single Fibre Channel device to the switch fabric. Switch zoning and restrictive authentication can prevent or allow communication with an SP, but not with specific disks or LUNs attached to an SP. For access control with LUNs, a different solution is required: Storage Groups.

#### **Storage Groups**

A Storage Group is one or more LUNs (logical units) within a storage system that is reserved for one or more servers and is inaccessible to other servers. Storage Groups are the central component of shared storage; storage systems that are unshared do not use Storage Groups.

When you configure shared storage, you specify servers and the Storage Group(s) each server can read from and/or write to. Access Logix software running in each storage system enforces the server-to-Storage Group permissions.

More than one server can access a Storage Group if all the servers run cluster software. The cluster software enforces orderly access to the shared Storage Group LUNs.

Figure 1-11 shows a simple shared storage configuration consisting of one storage system with two Storage Groups. One Storage Group serves a cluster of two servers running the same operating system, and the other Storage Group serves a UNIX® database server. Each server is configured with two independent paths to its data, including separate host bus adapters, switches, and SPs, so there is no single point of failure for access to its data.

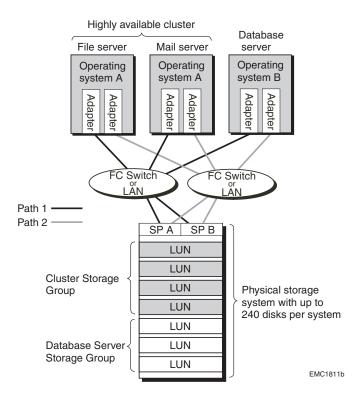


Figure 1-11 Sample Shared Storage Configuration

#### **Access Control with Shared Storage**

Access control permits or restricts a server's access to shared storage. Configuration access, the ability to configure storage systems, is governed by the storage system software security mechanism: a set of privileges based on username and password that is stored, encrypted, on one or more storage systems in a network.

Data access, the ability to read and write information to storage-system LUNs, is provided by Storage Groups. During storage-system configuration, the system administrator uses a management utility to associate a server with one or more LUNs. The associated LUNs compose a Storage Group.

Each server sees its Storage Group as if it were an entire storage system, and never sees the other LUNs on the storage system. Therefore, it cannot access or modify data on LUNs that are not part of its Storage Group. However, you can define a Storage Group as

sharable (accessible by more than one server), if, as shown above in Figure 1-11, the servers run cluster software.

Figure 1-12 shows access control through Storage Groups. Each server has exclusive read and write access to its designated Storage Group.

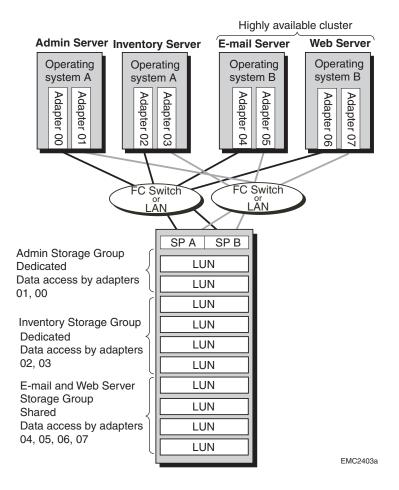


Figure 1-12 Data Access Control with Shared Storage

#### What Next?

For information about RAID types and RAID trade-offs, continue to the next chapter.

For information on the Navisphere  $^{\circledR}$  Manager Suite software, go to Chapter 3.

For information on the optional data replication software (SnapView $^{\text{TM}}$ , MirrorView $^{\text{TM}}$ /Synchronous, MirrorView/Asynchronous) or data mobility software (SAN Copy $^{\text{TM}}$ ), go to Chapter 4.

To plan LUNs and file systems, go to Chapter 5.

For details on the storage-system hardware, go to Chapter 6.

For worksheets and a description of iSCSI configurations, go to Chapter 7.

About CX300, CX300i, CX500, CX500i, and CX700 Systems

# RAID Types and Trade-offs

This chapter explains the RAID types you can choose for your storage-system LUNs. If you already know about RAID types and know which ones you want, you can skip this background information and go to the LUN planning chapter (Chapter 5).

## Major sections are

•	Introducing RAID	2-2
	RAID Types	
	RAID Benefits and Trade-offs	
	Guidelines for RAID Groups	
	Sample Applications for RAID Types	
•	2411p1011pp11e441e11510114112 1ype51111111111111	····· <b>–</b> –

# Introducing RAID

The storage system uses RAID (redundant array of independent disks) technology. RAID technology groups separate disks into one logical unit (LUN) to improve reliability and/or performance.

The storage system supports five RAID levels and two other disk configurations, the individual unit and the hot spare (global spare). You group the disks into one RAID Group by *binding* them using a storage-system management utility.

Four of the RAID levels use *disk striping* and two use *mirroring*.

#### **Disk Striping**

Using disk stripes, the storage-system hardware can read from and write to multiple disks simultaneously and independently. By allowing several read/write heads to work on the same task at once, disk striping can enhance performance. The amount of information read from or written to each disk makes up the stripe element size. The stripe size is the stripe element size multiplied by the number of disks in a group. For example, assume a stripe element size of 128 sectors (the default). If the group has five disks, you would multiply five by the stripe element size of 128 to yield a stripe size of 640 sectors.

The storage system uses disk striping with most RAID types.

## Mirroring

Mirroring maintains a copy of a logical disk *image* that provides continuous access if the original image becomes inaccessible. The system and user applications continue running on the good image without interruption.

You can create a mirror by binding disks as a RAID 1 mirrored pair or a RAID 1/0 Group (a mirrored RAID 0 Group); the hardware will then mirror the disks automatically.

With LUNs of any RAID type, you can also apply a variety of data replication and data mobility options discussed in Chapter 4, *Data Replication and Data Mobility Software*.

#### **RAID Groups and LUNs**

You can create multiple LUNs on one RAID Group and then allot each LUN to a different user, server, or application. For example, a five-disk RAID 5 Group that uses 73-Gbyte disks offers 292 Gbytes of space. You could bind three LUNs, with 60, 100, and 132 Gbytes of storage capacity, for temporary, mail, and customer files.

One disadvantage of multiple LUNs on a RAID Group is that I/O to each LUN may affect I/O to the others in the group; that is, if traffic to one LUN is very heavy, I/O performance with other LUNs may be degraded. The main advantage of multiple LUNs per RAID Group is the ability to divide the enormous amount of disk space provided by RAID Groups on newer, high-capacity disks.

Figure 2-1 shows three LUNs in one RAID Group.

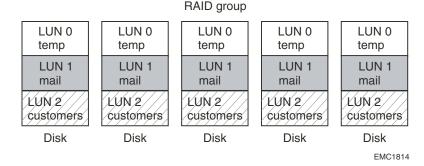


Figure 2-1 Multiple LUNs in a RAID Group

# **RAID Types**

Each RAID Group you create supports only one RAID type. You can choose from the following RAID types: RAID 6, RAID 5, RAID 3, RAID 1, RAID 0, RAID 1/0, individual disk unit, and hot spare.

RAID 6 Groups are available to CX300, CX300i, CX500, CX500i, and CX700 storage systems running FLARE OE 02.26 or higher.

You can choose additional types of data replication and mobility using the options discussed in Chapter 4, *Data Replication and Data Mobility Software*.

You cannot mix ATA (Advanced Technology Attachment) and Fibre Channel disk drives within a RAID Group.

### **RAID 6 Group (Double Distributed Parity)**

The RAID 6 Group is available on CX300, CX300i, CX500i, CX500i, and CX700 storage systems running FLARE Operating Environment (OE) version 02.26 or higher. A RAID 6 Group usually consists of 6 or 12 disks, but can have 4, 8, 10, 14, or 16 disks. Space equal to two disks is dedicated to parity. The remaining disk space is for data. A RAID 6 Group uses disk striping. With a RAID 6 Group, you can create up to 128 RAID 6 LUNs to allocate disk space to different users, servers, and applications.

The storage system writes two independent sets of parity information - row parity and diagonal parity - that lets the Group continue operating if one or two disks fail or if a hard media error occurs during a single-disk rebuild. When you replace the failed disks, the SP rebuilds, or with proactive sparing, continues rebuilding, the Group using the information stored on the working disks. Performance is degraded while the SP rebuilds the Group. This degradation is lessened with proactive sparing. During the rebuild, the storage system continues to function and give users access to all data, including data stored on the failed disks.

A RAID 6 Group distributes parity evenly across all drives so that parity drives are not a bottleneck for write operations. Figure 2-3 shows user and parity data with the default strip element size of 128 sectors (65,536 bytes) in a 6-disk RAID 6 Group. The stripe size is

comprised of all strip elements. Notice that the disk block addresses in the stripe proceed sequentially from the first disk to the second, third, fourth, fifth, and sixth, then back to the first, and so on.

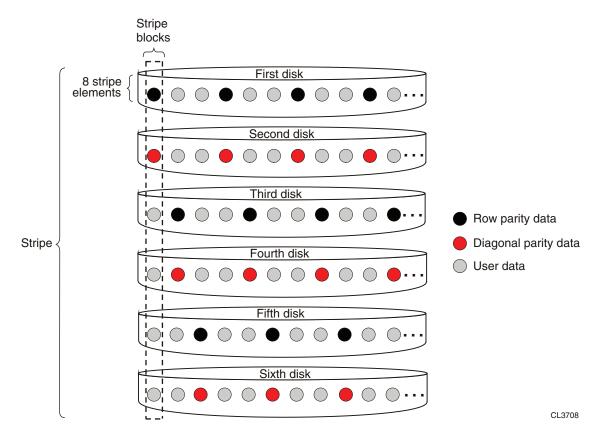


Figure 2-2 RAID 6 Group

A RAID 6 Group offers good read performance and good write performance. Write performance benefits greatly from storage-system caching.

# **RAID 5 Group (Distributed Parity)**

A RAID 5 Group usually consists of five disks (but can have three to sixteen). A RAID 5 Group uses disk striping. With a RAID 5 group,

you can create up to 128 RAID 5 LUNs to apportion disk space to different users, servers, and applications.

The storage system writes parity information that lets the Group continue operating if a disk fails. When you replace the failed disk, the SP rebuilds the group using the information stored on the working disks. Performance is degraded while the SP rebuilds the group. However, the storage system continues to function and gives users access to all data, including data stored on the failed disk.

Figure 2-2 shows user and parity data with the default stripe element size of 128 sectors (65,536 bytes) in a five-disk RAID 5 group. The stripe size comprises all stripe elements. Notice that the disk block addresses in the stripe proceed sequentially from the first disk to the second, third, and fourth, fifth, then back to the first, and so on.

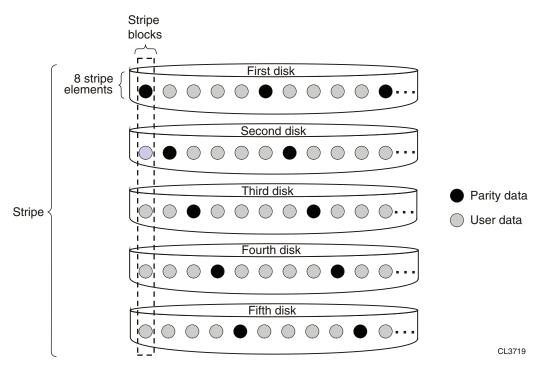


Figure 2-3 RAID 5 Group

RAID 5 Groups offer excellent read performance and good write performance. Write performance benefits greatly from storage-system caching.

## RAID 3 Group (Single-Disk Parity)

A RAID 3 Group consists of five or nine disks and uses disk striping. Limit concurrent access to RAID 3 LUNs for the best bandwidth performance. For example, a RAID 3 Group may have multiple LUNs, but the highest bandwidth is achieved with 1 to 4 threads of concurrent, large I/O.

The storage system writes parity information that lets the Group continue operating if a disk fails. When you replace the failed disk, the SP rebuilds the group using the information stored on the working disks. Performance is degraded while the SP rebuilds the group. However, the storage system continues to function and gives users access to all data, including data stored on the failed disk.

Figure 2-4 shows user and parity data with the default stripe element size of 128 sectors (65,536 bytes) in a RAID 3 Group. Notice that the disk block addresses in the stripe proceed sequentially from the first disk to the second, third, and fourth, then back to the first, and so on.

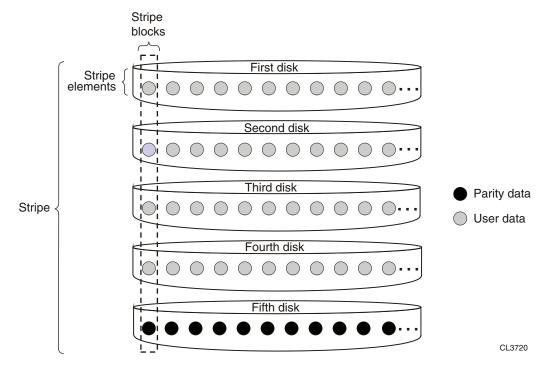


Figure 2-4 RAID 3 Group

RAID 3 differs from RAID 6 or RAID 5 in one major way. With a RAID 3 Group, the parity information is stored on one disk; with a RAID 6 or RAID 5 Group, it is stored on all disks. RAID 3 can perform sequential I/O better than RAID 6 or RAID 5, but does not handle random access as well. FLARE<sup>TM</sup> Operating Environment 2.07 improved RAID 3 processing in several ways:

- Use of the write cache is now equivalent to RAID 5
- I/O sizes to disk have been increased
- Maximum host transfer size has been increased.

RAID 3 is best thought of as a specialized RAID 5, where the expectation is either large or sequential I/O. However, with write cache active for RAID 3, RAID 3 is more flexible than in previous releases, and can handle some level of concurrency. A RAID 3 Group works well for applications that use I/Os of blocks 64 Kbytes and larger. By using both the read and write cache, a RAID 3 group can handle several concurrent streams of access.

RAID 3 Groups do not require any special buffer area. No fixed memory is required to use write cache with RAID 3. Simply allocate write cache as you would for RAID 5, and ensure that caching is turned on for the LUNs in the RAID 3 Groups. Access to RAID 3 LUNs is compatible with concurrent access to LUNs of other RAID types on the storage system.

#### **RAID 1 Mirrored Pair**

A RAID 1 Group consists of two disks that are mirrored automatically by the storage-system hardware. With a RAID 1 Group, you can create multiple RAID 1 LUNs to apportion disk space to different users, servers, and applications.

RAID 1 hardware mirroring within the storage system is not the same as software mirroring, remote mirroring, or hardware mirroring for other kinds of disks. Functionally, the difference is that you cannot manually stop mirroring on a RAID 1 mirrored pair, and then access one of the images independently. If you want to use one of the disks in such a mirror separately, you must unbind the mirror (losing all data on it), rebind the disk as the type you want, and software format the newly bound LUN.

With a storage system, RAID 1 hardware mirroring has the following advantages:

 automatic operation (you do not have to issue commands to initiate it)

- physical duplication of images
- a rebuild period that you can select during which the SP recreates the second image after a failure

With a RAID 1 mirrored pair, the storage system writes the same data to both disks, as follows.

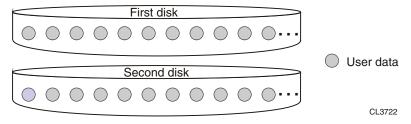


Figure 2-5 RAID 1 Mirrored Pair

# RAID 1/0 Group (Mirrored RAID 0 Group)

A RAID 1/0 Group consists of two, four, six, eight, ten, twelve, fourteen, or sixteen disks. These disks make up two mirror images, with each image including two to eight disks. The hardware automatically mirrors the disks. A RAID 1/0 Group uses disk striping. It combines the speed advantage of RAID 0 (Page 2-10) with the redundancy advantage of mirroring. With a RAID 1/0 Group, you can create up to 128 RAID 1/0 LUNs to apportion disk space to different users, servers, and applications.

Figure 2-6 shows the distribution of user data with the default stripe element size of 128 sectors (65,536 bytes) in a six-disk RAID 1/0 Group. Notice that the disk block addresses in the stripe proceed sequentially from the first mirrored disks (first and fourth disks) to the second mirrored disks (second and fifth disks), to the third

mirrored disks (third and sixth disks), and then from the first mirrored disks, and so on.

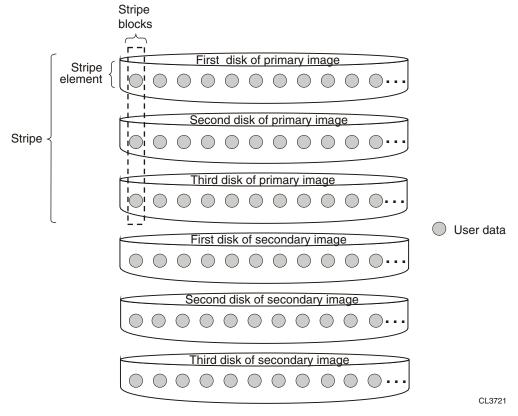


Figure 2-6 RAID 1/0 Group

A RAID 1/0 Group can survive the failure of multiple disks, providing that one disk in each image pair survives.

# **RAID 0 Group (Nonredundant RAID Striping)**



#### CAUTION

A RAID 0 Group provides no protection for your data. EMC does not recommend using a RAID 0 Group unless you have some way of protecting your data, such as software mirroring.

A RAID 0 Group consists of three to a maximum of sixteen disks. A RAID 0 Group uses disk striping, in which the hardware writes to or reads from multiple disks simultaneously. You can create up to 128 LUNs in any RAID 0 Group.

Unlike the other RAID levels, with RAID 0 the hardware does not maintain parity information on any disk; this type of group has no inherent data redundancy. As a result, if any failure (including an unrecoverable read error) occurs on a disk in the LUN, the information on the LUN is lost.

RAID 0 offers enhanced performance through simultaneous I/O to different disks. A desirable alternative to RAID 0 is RAID 1/0, which does protect your data.

#### Individual Disk Unit

An individual disk unit is a disk bound to be independent of any other disk in the cabinet. An individual unit has no inherent high availability, but you can make it highly available by using software mirroring with another individual unit.

## **Hot Spare**

A hot spare is a dedicated replacement disk on which users cannot store information. A hot spare is global: if any disk in a RAID 6 Group, RAID 5 Group, RAID 3 Group, RAID 1 mirrored pair, or RAID 1/0 Group fails, the SP automatically rebuilds the failed disk's structure on the hot spare. When the SP finishes rebuilding, the disk group functions as usual, using the hot spare instead of the failed disk. When you replace the failed disk, the SP copies the data from the former hot spare onto the replacement disk.

When the copy is done, the disk group consists of disks in the original slots, and the SP automatically frees the hot spare to serve as a hot spare again. A hot spare is most useful when you need the highest data availability. It eliminates the time and effort needed for someone to notice that a disk has failed, find a suitable replacement disk, and insert the disk.

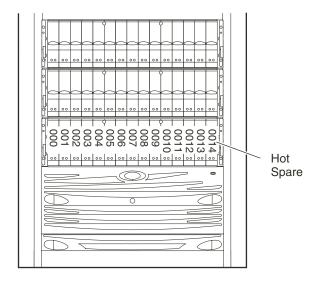
When you plan to use a hot spare, make sure the disk has the capacity to serve in any RAID Group in the storage-system chassis. A RAID Group cannot use a hot spare that is smaller than a failed disk in the group.

Your hot spare must match the drive type of the RAID group it serves; RAID groups made up of Fibre Channel disks require FC hot spares, and ATA RAID groups require ATA hot spares.

You can have one or more hot spares per storage-system chassis. You can make any disk in the chassis a hot spare, except a disk that stores the  $FLARE^{TM}$  operating environment (formerly called Base or Core Software) or the write cache vault; that is, a hot spare can be any disk except disk IDs 000-004.

If you use hot spares of different sizes, the storage system will automatically use the hot spare of the proper size in place of a failed disk.

An example of hot spare usage for a CX700 storage system follows.



- RAID 5 Groups consist of disk modules 0-4 and 5-9; mirrored pairs are modules 10-11 and 12-13, disk module 14 is a hot spare.
- 2. Disk module 3 fails
- 3. RAID 5 Group becomes modules 0, 1, 2, 14, and 4; now no hot spare is available.
- System operator replaces failed module 3 with a functional module.
- 5. Once again, RAID 5 Group consists of modules 0-4 and the hot spare is 14.

EMC2432

Figure 2-7 How a Hot Spare Works

# **Proactive Sparing**

Proactive sparing lets you create a proactive hot spare of a drive that is becoming prone to errors (a proactive candidate). The proactive sparing operation copies the contents of a drive to a hot spare before the drive fails. Subsequently, the drive can be removed from the storage system before it fails and the hot spare takes it place. The proactive sparing operation is initiated automatically or manually. When the FLARE storage-system software identifies certain types of frequencies of errors on a disk, it identifies the disk as a proactive candidate, and automatically begins the proactive sparing operation. FLARE copies the contents of the proactive candidate to the proactive spare. Additionally, you can manually start a copy of a proactive candidate to a proactive hot spare using Navisphere Manager. All data is copied from the proactive candidate to the proactive hot spare. When a proactive sparing coy completes, the proactive candidate is faulted. When you replace the faulted drive, the storage system equalizes.

Any available hot spare can be a proactive spare, but only one hot spare can be used for proactive sparing at a time. If the storage system has only one hot spare, it can be a proactive hot spare. The following table lists the number of concurrent proactive spare copies supported per storage system.

Unit Type	Number of supported proactive spares	
RAID 6, RAID 5, RAID 3	1	
RAID 1	1 per pair	
RAID 1/0	1 per mirrored pair	

Proactive sparing is not supported for RAID 0 or individual disk units. It is available on CX300, CX300i, CX500, CX500i, and CX600 storage systems running FLARE OE 02.26 or higher.

# **RAID Benefits and Trade-offs**

This section reviews RAID types and explains their benefits and trade-offs. You can bind disks into seven types of logical units (LUNs):

- RAID 6 Group (dual distributed parity)
- RAID 5 Group (distributed parity)
- RAID 3 Group (single-disk parity)
- RAID 1 mirrored pair
- ◆ RAID 1/0 Group (mirrored RAID 0 Group); a RAID 0 Group mirrored by the storage-system hardware
- RAID 0 Group (nonredundant RAID striping); no inherent high-availability features
- Individual unit; no inherent high-availability features
- Hot spare; serves only as an automatic replacement for any disk in a RAID type other than 0; does not store data during normal system operations

Plan the disk unit configurations carefully. After a disk has been bound into a LUN, you must have LUN migration support — available in Navisphere Release 16 or higher — to change the RAID type of that LUN without unbinding it. Unbinding a LUN involves losing all data on it.

Table 2-1 compares the read and write performance and relative cost per gigabyte (Gbyte) of the RAID types. Figures shown are theoretical maximums.

Table 2-1 Performance, Availability, and Cost of RAID Types (Individual Unit = 1.0)

Disk Configuration	Relative Read Performance Without Cache	Relative Write Performance Without Cache	Relative Cost Per Gbyte
RAID 6	Up to 5x with six disks (for small I/O requests that are 2 to 8 KB)	Up to 1x with six disks (for small I/O requests that are 2 to 8 KB)	1.5
RAID 5 Group with five disks	Up to 5x with five disks (for small I/O requests that are 2 to 8 Kbytes)	Up to 1.25x with five disks (for small I/O requests that are 2 to 8 Kbytes)	1.25
RAID 3 Group with five disks	Up to 4x (for large I/O requests)	Up to 4x (for large I/O requests)	1.25
RAID 1 mirrored pair	Up to 2x	Up to 1x	2
RAID 1/0 Group with 10 disks	Up to 10x	Up to 5x	2
Individual unit	1x	1x	1

Notes: These performance numbers are not based on storage-system caching. With caching, the performance numbers for RAID 5 writes improve significantly. Performance multipliers vary with the load on the server and storage system.

#### Performance

RAID 6 and RAID 5, with individual access, provide high read throughput by allowing simultaneous reads from each disk in the group. RAID 5 write performance is excellent when the storage system uses write caching.

RAID 3, with parallel access, provides high throughput for sequential requests. Large block sizes (more than 64 Kbytes) are most efficient. RAID 3 attempts to write full stripes to the disk, and avoid parity update operations.

Generally, the performance of a RAID 3 Group increases as the size of the I/O request increases. Read performance increases incrementally with read requests up to 1 Mbyte. Write performance increases incrementally for sequential write requests that are greater than 256 Kbytes. For applications issuing large or sequential I/O requests to CLARiiON ATA drives, a RAID 3 LUN provides significantly better write performance than a RAID 6 or RAID 5 LUN.

RAID 1 read performance will be higher than that of an individual disk while write performance remains approximately the same as that of an individual disk.

A RAID 0 Group (nonredundant RAID striping) or RAID 1/0 Group (mirrored RAID 0 Group) can have as many I/O operations occurring simultaneously as there are disks in the group. In general, the performance of RAID 1/0 equals the number of disk pairs times the RAID 1 performance number. If you want high throughput for a specific LUN, use a RAID 1/0 or RAID 0 Group. A RAID 1/0 Group requires at least two disks; a RAID 0 Group, at least three disks.

If you create multiple LUNs on a RAID Group, the LUNs share the RAID Group disks, and the I/O demands of each LUN affect the I/O service time to the other LUNs.

#### **Storage Flexibility**

You can create up to 128 LUNs in each RAID Group. This adds flexibility, particularly with large disks, since it lets you apportion LUNs of various sizes to different servers, applications, and users.

## Data Availability and Disk Space Usage

If data availability is critical and you cannot afford to wait hours to replace a disk, rebind it, make it accessible to the operating system, and load its information from backup, then use a redundant RAID Group: RAID 5, RAID 3, RAID 1 mirrored pair, or RAID 1/0. If data availability is not critical, or disk space usage is critical, bind an individual unit.

A RAID 1 mirrored pair or RAID 1/0 Group provides very high data availability. They are more expensive than RAID 6, RAID 5, or RAID 3 Groups, since only 50 percent of the total disk capacity is available for user data, as shown in Figure 2-8 on page 2-17.

A RAID 6, RAID 5, or RAID 3 Group provides high data availability, but requires more disks than a mirrored pair. In a RAID 6 Group, the disk space available for user data is the total capacity of two less than the number of disks in the RAID Group. In a RAID 5 or RAID 3 Group, the disk space available for user data is the total capacity of one less than the number of disks in the RAID Group. For example, in a five-disk RAID 5 Group, the capacity of four disks is available for user data, which is 80% of the group's total disk capacity. So RAID 6, RAID 5, and RAID 3 Groups use disk space much more efficiently than a mirrored pair. A RAID 6, RAID 5, or RAID 3 Group is usually

more suitable than a RAID 1 mirrored pair for applications where high data availability, good performance, and efficient disk space usage are all of relatively equal importance.

Figure 2-8 illustrates RAID type disk space usage.

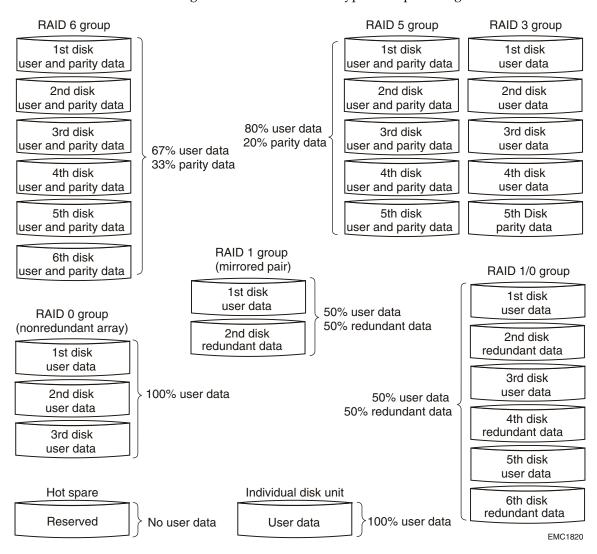


Figure 2-8 Disk Space Usage in the RAID Configuration

A RAID 0 Group (nonredundant RAID striping) provides all its disk space for user files, but does not provide any high availability

features. For high availability, you can use a RAID 1/0 Group instead.

A RAID 1/0 Group provides the best combination of performance and availability, at the highest cost per Gbyte of disk space.

An individual unit, like a RAID 0 Group, provides no high-availability features. All its disk space is available for user data, as shown in Figure 2-8 above.

# **Guidelines for RAID Groups**

To decide when to use a RAID 5 Group, RAID 3 Group, mirror (that is, a RAID 1 mirrored pair or RAID 1/0 Group, a RAID 0 Group, individual disk unit, or hot spare), you need to weigh these factors:

- Importance of data availability
- Importance of performance
- Amount of data stored
- Cost of disk space

The following guidelines will help you decide on RAID types.

# Use a RAID 5 Group (distributed parity) for applications where

- Data availability is very important.
- Large volumes of data will be stored.
- Multitask applications use I/O transfers of different sizes.
- Excellent read and good write performance is needed (write performance is very good with write caching).
- You want the flexibility of multiple LUNs per RAID Group.

# Use a RAID 3 Group (single-disk parity) for applications where

- Data availability is very important.
- Large volumes of data will be stored.

Use RAID 3 with ATA drives for high bandwidth applications (those with sequential access or large I/O request sizes). Use RAID 3 with Fibre Channel drives where you expect similar access patterns, random access is extremely unlikely, and you require the highest possible bandwidth performance.

#### Use a RAID 1 mirrored pair for applications where

- Data availability is very important.
- Speed of write access is important and write activity is heavy.

# Use a RAID 1/0 Group (mirrored nonredundant array) for applications where

- Data availability is critically important.
- Overall performance is very important.

# Use a RAID 0 Group (nonredundant RAID striping) for applications where

- High availability is not important.
- You can afford to lose access to all data stored on a LUN if a single disk fails.
- Overall performance is very important.

#### Use an individual unit for applications where

- High availability is not important.
- Speed of write access is somewhat important.

#### Use a hot spare where

- In any RAID 5, RAID 3, RAID 1/0 or RAID 1 Group, high availability is so important that you want to regain data redundancy quickly without human intervention if any disk in the group fails.
- Minimizing the degraded performance caused by disk failure in a RAID 5 or RAID 3 Group is important.

# Sample Applications for RAID Types

This section describes some types of applications in which you would want to use a RAID 6 Group, RAID 5 Group, RAID 3 Group, RAID 1 mirrored pair, RAID 0 Group (nonredundant array), RAID 1/0 Group, or individual unit.

**RAID 6 (Dual Distributed Parity) or RAID 5 Group (Distributed Parity)** — A RAID 6 or RAID 5 Group is useful as a database repository or a database server that uses a normal or low percentage of write operations (writes are 33 percent or less of all I/O operations). Use a RAID 6 or RAID 5 Group where multitasking applications perform I/O transfers of different sizes. Write caching can significantly enhance the write performance of a RAID 6 or RAID 5 Group. For higher availability, use a RAID 6 Group instead of a RAID 5 Group.

For example, a RAID 6 or RAID 5 Group is suitable for multitasking applications that require a large history database with a high read rate, such as a database of legal cases, medical records, or census information. A RAID 6 or RAID 5 Group also works well with transaction processing applications, such as an airline reservations system, where users typically read the information about several available flights before making a reservation, which requires a write operation. You could also use a RAID 6 or RAID 5 Group in a retail environment, such as a supermarket, to hold the price information accessed by the point-of-sale terminals. Even though the price information may be updated daily, requiring many write operations, it is read many more times during the day.

RAID 3 Group (Single-Disk Parity)— A RAID 3 Group is ideal for high-bandwidth reads or writes, that is, applications that perform either logically sequential I/O or use large I/O sizes (stripe size or larger). Using standard read and write cache, several applications can read and write from a RAID 3 Group. Random access in a RAID 3 Group is not optimal, so the ideal applications for RAID 3 are backup to disk (particularly with ATA drives), real-time data capture, and storage of extremely large files.

You might use a RAID 3 Group for a single-task application that does large I/O transfers, like a weather tracking system, geologic charting application, medical imaging system, or video storage application.

**RAID 1 mirrored pair** — A RAID 1 mirrored pair is useful for logging or record-keeping applications because it requires fewer

disks than a RAID 0 Group (nonredundant array) and provides high availability and fast write access. Or you could use it to store daily updates to a database that resides on a RAID 6 or RAID 5 Group. Unlike a RAID 1/0 Group, a RAID 1 Group is not expandable to more than two disks.

**RAID 0 Group (nonredundant RAID striping)** — Use a RAID 0 Group where the best overall performance is important. A RAID 0 Group (like a RAID 5 Group) requires a minimum of three disks. A RAID 0 Group is useful for applications using short-term data to which you need quick access.

RAID 1/0 Group (mirrored RAID 0 Group) — A RAID 1/0 Group provides the best balance of performance and availability. You can use it very effectively for any of the RAID 6 or RAID 5 applications. A RAID 1/0 Group requires a minimum of two disks.

**Individual unit** — An individual unit is useful for print spooling, user file exchange areas, or other such applications, where high availability is not important or where the information stored is easily restorable from backup.

The performance of an individual unit is slightly less than a standard disk not in an storage system. The slight degradation results from SP overhead.

**Hot spare** — A hot spare provides no data storage but enhances the availability of each RAID 6, RAID 5, RAID 3, RAID 1, and RAID 1/0 Group in a storage system. Use a hot spare where you must regain high availability quickly without human intervention if any disk in such a RAID Group fails. A hot spare also minimizes the period of degraded performance after a RAID 5 or RAID 3 disk fails. A hot spare can also be used as a proactive hot spare in

#### What Next?

For information on the Navisphere Manager Suite software, go to Chapter 3.

For information on the optional data replication software (SnapView<sup>TM</sup>, MirrorView<sup>TM</sup>, MirrorView/Asynchronous) or data mobility software (SAN Copy<sup>TM</sup>), go to Chapter 4.

To plan LUNs and file systems, go to Chapter 5.

For details on the storage-system hardware, go to Chapter 6.

For worksheets and a description of iSCSI configurations, go to Chapter 7.

RAID Types and Trade-offs		
	1	

# Storage-System Management

This chapter describes the storage-system management ports, asks you to fill out a management port worksheet, and describes applications you can use to manage storage systems from servers.

# Major sections are

•	Management Ports	3-2
	Navisphere Management Suites	
•	Using Navisphere Software	3-8

# **Management Ports**

The storage system has two management ports, one per SP. Each port must be configured with network and security characteristics that you must define and record in the Management Port Information Worksheet. The term SP means storage processor; LAN means local area network.

Record network information for the management interface. Your network administrator should provide most of this information.

#### Table 3-1 Management Port Worksheet

Storage-System Serial Number	er:			
Management Ports - Network	Information:			
IP Address	SP A:		SP B:	·
Subnet Mask	SP A:		SP B:	·
Gateway	SP A:		SP B:	·
Management Ports - Login In	formation			
Username:		Pass	sword:	
Roles: Monitor N	Manager	Administrator	Security Administra	tor

# Storage-System Serial Number

The serial number of the storage system. On CX300 and CX500 systems, a label with the hardware serial number is located on the

rear of the chassis, on the far left, between power supply B and storage processor B.

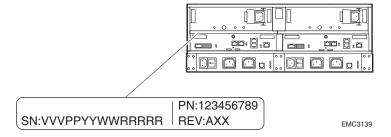


Figure 3-1 Storage System Serial Number Label (CX300i Enclosure)

On CX700-Series systems, the serial number label is behind the front bezel, as shown in Figure 3-2.

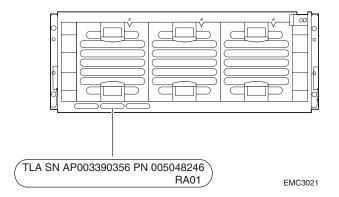


Figure 3-2 Storage System Serial Number Label - CX700

#### **IP Address**

The static network IP (Internet Protocol) address (for example, 128.222.78.10) for communication with a 10/100 management port of a storage processor (SP A or SP B). There is one management port per SP.

#### **Subnet Mask**

The subnet mask for the LAN to which the storage system is connected for management; for example, 255.255.255.0.

#### **Gateway**

The gateway address for the LAN to which the storage-system management port is connected.

Use the *Management Ports - Login Information* section to record the username and password to login to each management port. Your network administrator may provide these. If not, then you need to create one or both.

#### Username

A username for the management interface. It must start with a letter and may contain 1 to 32 letters and numbers. The name may not contain punctuation, spaces, or special characters. You can use uppercase and lowercase characters. Usernames are case-sensitive. For example, **ABrown** is a different username from **abrown**.

#### **Password**

A password to connect to the management interface. It may contain 1 to 32 characters, consisting of uppercase and lowercase letters and numbers. As with the username, passwords are case-sensitive. For example, **Azure23** differs from **azure23**. The password is valid only for the username you specified.

#### **User roles**

Four basic user roles are available - monitor, manager, administrator, and security administrator. All users, except a security administrator, can monitor the status of a storage system. Users with the manager role can also configure a storage system. Administrators can maintain user accounts, as well as configure a storage system. Security administrators can *only* manage security and domain settings.

# **Navisphere Management Suites**

You can choose one of three different Navisphere management suites: Navisphere Workgroup, Navisphere Departmental, or Navisphere Enterprise. Each suite provides the following software:

- Access Logix
- Navisphere Manager
- Navisphere Host Agents and host-based command line interpreter (CLI) for any operating system
- Navisphere Analyzer, provided with Navisphere Enterprise and an optional product you can purchase for Navisphere Workgroup and Navisphere Departmental.

The maximum number of servers that you can connect to a storage system varies with the management suite as follows:

Management Suite	Maximum Number of Servers Connected to a Storage System	
Workgroup	2	
Departmental	7	
Enterprise	Maximum number defined for the specific storage-system type	

# **Access Logix Software**

Access Logix lets you restrict server access to LUNs through the use of Storage Groups. A server can access only the LUNs in the Storage Group that you connect to the server. When you create a Storage Group, you select the LUNs it contains. If multiple servers are connected to a storage system, you should connect them to different Storage Groups, unless they are part of a cluster.

# Navisphere Manager

Navisphere Manager (called Manager) lets you manage multiple storage systems on multiple servers simultaneously. It includes an Event Monitor that checks storage systems for fault conditions and can notify you and/or customer service if any fault condition occurs.

With Manager you can create RAID groups, LUNs, and Storage Groups, manipulate caches, examine storage-system status and logged events, transfer control from one SP to another, and examine events recorded in storage-system event logs.

Manager features a user interface (UI) with extensive online help. UIs and online help for each of the data replication, data mobility, and Analyzer software integrate seamlessly into the Manager UI and online help.

Navisphere Manager provides the following security functions and benefits:

- Authentication
- Authorization
- ◆ Privacy
- ◆ Audit

#### Authentication

Manager uses password-based authentication that is implemented by the storage management server on each storage system in the domain. You assign a username and password when you create either global or local user accounts. Global accounts apply to all storage systems in a domain and local user accounts apply to a specific storage system. A global user account lets you manage user accounts from a single location. When you create or change a global user account or add a new storage system to a domain, Manager automatically distributes the global account information to all storage systems in the domain.

#### **Authorization**

Manager bases authentication on the role associated with the authenticated user. Four user roles are available - monitor, manager, administrator, security administrator. All users, except a security administrator, can manage the status of a storage system. Users with the manager role can also configure a storage system. Administrators can maintain user accounts, as well as configure a storage system. Security administrators can *only* manage security and domain settings.

The security administrator role is available for CX300, CX300i, CX500, CX500i, and CX700 storage systems running FLARE OE version 02.26 or higher.

#### Privacy

Manager encrypts all data that passes between the browser and the storage-system management server, as well as the data that passes between storage management servers. This encryption protects the transferred data whether it is on local networks behind corporate firewalls or on the Internet.

#### Audit

Manager maintains an SP event log that contains a time-stamped record for each event. This record includes information such as an event code and event description. Manager also adds time-stamped audit records to the SP event log each time a user logs in or enters a request. These records include information about the request and requestor.

#### Navisphere Host Agent and CLI

Navisphere Host Agents are provided for different operating systems. Each server connected to a storage system must run the Host Agent for its operating system. The Host Agent registers the host HBAs with the storage system and provides the storage system with information on how the server maps devices to LUNs (for example, the Windows drive letter for a LUN).

The CLI provides commands that implement Navisphere Manager UI functions in addition to commands that implement the functions of the UIs for the optional data replication and data mobility software. A major benefit offered by the CLI is the ability to write command scripts to automate management operations.

# Navisphere Analyzer

Navisphere Analyzer lets you measure, compare, and chart the performance of SPs, LUNs, and disks.

# Navisphere Integrator

Navisphere Integrator provides an interface between Navisphere products and HP OpenView<sup>®</sup>, CA Unicenter<sup>®</sup>, and Tivoli<sup>®</sup> network management software.

# **Using Navisphere Software**

With Manager you can assign storage systems on an intranet or internet to a storage domain. For any installation, you can create one of more domains, provided that each storage system is in only one domain. Each storage domain must have at least one member with installed Manager.

All CX-Series storage systems now ship from the factory with Manager installed on them.

Each storage system in the domain is accessible from any other in the domain. Using an internet browser, you point at a storage system that has Manager installed. The security software then prompts you to log in. After logging in, depending on the privileges of your account, you can monitor, manage, and/or define user accounts for any storage system in the domain. You cannot view storage systems outside the domain from within the domain.

You can run the internet browser on any supported station (often a PC or laptop) with a network controller and browser software. At least one storage system — ideally at least two for higher availability — in a domain must have Manager installed.

Figure 3-3 shows an internet that connects nine storage systems. It shows two domains, a U.S. Division domain with five storage systems (four systems on SANs) and a European Division with four storage systems. The 13 servers that use the storage may be connected to the same or a different network, but the intranet shown is the one used to manage the storage systems.

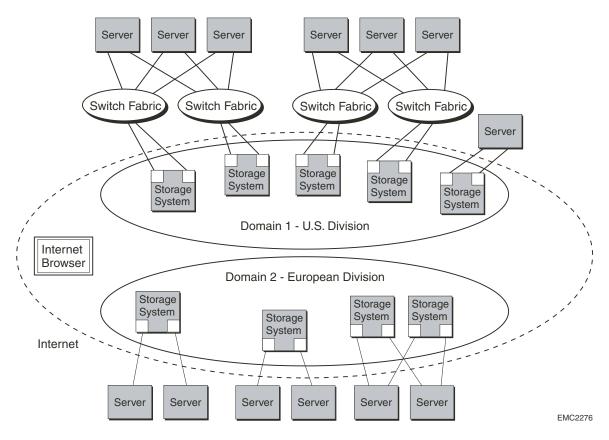


Figure 3-3 Storage Domains on the Internet

# **Navisphere Quality of Service Manager**

The optional Navisphere Quality of Service Manager (NQM or Navisphere QoS Manager) lets you allocate storage-system performance resources on an application-by-application basis. You can use NQM to solve performance conflicts in consolidated environments where multiple applications share the same storage system. Within storage-system capacity, NQM lets you meet specific performance targets for applications, and create thresholds to prevent applications from monopolizing storage-system performance.

# Navisphere Management Worksheet

The following worksheet will help you plan your Navisphere storage management configuration. A sample worksheet follows the blank worksheet.

## Storage Management Worksheet for Navisphere

Server Name	Operating System
Navisphere Workgroup	Navisphere Enterprise   Navisphere Analyzer
Storage System Name	Domain Name
Server Name	Operating System
Navisphere Workgroup  Navisphere Departmental  Navisphere Quality of Service Manager	Navisphere Enterprise   Navisphere Analyzer
Storage System Name	Domain Name
Server Name	Operating System
Navisphere Workgroup ☐ Navisphere Departmental ☐ Navisphere Quality of Service Manager ☐	Navisphere Enterprise   Navisphere Analyzer
Storage System Name	Domain Name

## Sample Storage Management Worksheet for Navisphere

Server Name <u>cpcnyc1</u>	Operating System Windows Server 2003
Navisphere Workgroup ☐ Navisphere Department ☐ Navisphere Quality of Service Manager ☐	Navisphere Enterprise   Navisphere Analyze
Storage System Name <u>hwdeng1</u>	Domain Name <u>nychwd</u>
Storage System Name <u>hwdeng2</u>	Domain Name <u>nychwd</u>
Storage System Name	Domain Name
Storage System Name	Domain Name
Server Name cpcnc2	Operating System Windows Server 2003
Navisphere Workgroup ☐ Navisphere Departmenta ☐ Navisphere Quality of Service Manager ☐	Navisphere Enterprise   Navisphere Analyzer
Storage System Name <u>hwdeng1</u>	Domain Name <u>nychwd</u>
Storage System Name <u>hwdeng2</u>	Domain Name <u>nychwd</u>
Storage System Name	Domain Name
Storage System Name	Domain Name
Server Name cpcnyc4	Operating System Windows 2000
Navisphere Workgroup ☐ Navisphere Departmental ☐ Navisphere Quality of Service Manage ☐	Navisphere Enterprise ☐ Navisphere Analyzer ☐
Storage System Name <u>hwdqa2</u>	Domain Name <u>nychwd</u>
Storage System Name	Domain Name
Storage System Name	Domain Name
Storage System Name	Domain Name

#### What Next?

For information on the optional data replication software (SnapView $^{\text{TM}}$ , MirrorView $^{\text{TM}}$ , MirrorView/Asynchronous) or data mobility software (SAN Copy $^{\text{TM}}$ ), go to Chapter 4.

To plan LUNs and file systems, go to Chapter 5.

For details on the storage-system hardware, go to Chapter 6.

For worksheets and a description of iSCSI configurations, go to Chapter 7.

Storage-System Management	
- storage system management	

# Data Replication and Data Mobility Software

This chapter introduces EMC CLARiiON data replication software (SnapView<sup>TM</sup>, MirrorView<sup>TM</sup>/Synchronous, and MirrorView<sup>TM</sup>/Asynchronous software) and the EMC data mobility software (SAN Copy<sup>TM</sup> and SAN Copy<sup>TM</sup>/E software) for CX-Series storage systems. These software applications allow you to create copies of logical units (LUNs).

Your storage system may not support all of the software options discussed in this chapter. Refer to the ELab Interoperability Navigator on EMC Powerlink for an up-to-date listing of platforms and supported applications.

This chapter refers to the EMC MirrorView/Synchronous software as MirrorView/S, and EMC MirrorView/Asynchronous software as MirrorView/A.

Major sections are

# Data Replication and Mobility Software Summary◆ Data Replication and Data Mobility Software Overview4-3SnapView◆ What Is SnapView?4-4◆ Clones Overview and Example4-4◆ Snapshots Overview and Example4-6◆ Clone and Snapshot Trade-offs4-7◆ SnapView Worksheets4-9

MirrorView/Synchronous	
What Is MirrorView/S?	4-12
• Sample MirrorView/S Configuration	
<ul> <li>MirrorView/S Features and Benefits</li> </ul>	
♦ MirrorView/S Planning Worksheet	4-17
MirrorView/Asynchronous	
<ul> <li>What Is MirrorView/A?</li> </ul>	4-19
• Sample MirrorView / A Configuration	
MirrorView/A Features and Benefits	
MirrorView/A Planning Worksheets	
SAN Copy	
What Is EMC SAN Copy?	4-27
<ul> <li>Implementing SAN Copy Over Extended Distances.</li> </ul>	
SAN Copy Features and Benefits	
SAN Copy Planning Worksheet	
SAN Copy/E	
What Is EMC SAN Copy/E?	4-35
<ul> <li>Implementing SAN Copy/E over Extended Distance</li> </ul>	
• SAN Copy/E Features and Benefits	
SAN Copy/E Planning Worksheet	

# Data Replication and Data Mobility Software Overview

Table 4-1 summarizes the common use, supported storage systems, and requirements for the data replication and data mobility software explained in this chapter.

Table 4-1 Data Replication and Data software Summary

	Data Replication		Data Mobility	
	SnapView	MirrorView/S	MirrorView/A	SAN Copy
Common Usage	Local backup and recovery     Application testing	Disaster recovery     Remote backup -     limited distance	Disaster recovery     Remote backup - unlimited distance	Data distribution     Data migration     Remote backup - unlimited distance
Supported Storage Systems	CLARiiON only	CLARiiON only	CLARiiON only	CLARiiON, Symmetrix and third party storage systems.
Requirements	Software must be installed on the CLARiiON storage system.	Software must be installed on every CLARiiON storage system with LUNs in a remote mirror.	Software must be installed on every CLARiiON storage system with LUNs in a remote mirror.	Software must be installed on at least one CLARiiON storage system participating in the SAN Copy session. <sup>a</sup>

For a full copy, SAN Copy must be installed on either the storage system with the source logical unit or the storage system with the target logical unit. For an incremental copy, SAN Copy must be installed on the storage system with the source logical unit.

# What Is SnapView?

SnapView is a storage-system-based software application that allows you to create a copy of a LUN by using either clones or snapshots.

A clone, also referred to as a Business Continuance Volume (BCV), is an actual copy of a LUN and takes time to create, depending on the size of the source LUN. A snapshot is a virtual point-in-time copy of a LUN and takes only seconds to create.

SnapView has the following important benefits:

- It allows full access to production data with modest impact on performance and without the risk of damaging the original data.
- For decision support or revision testing, it provides a coherent, readable and writable copy of real production data.
- For backup, it practically eliminates the time that production data spends offline or in hot backup mode, and it offloads the backup overhead from the production server to another server.

Depending on your application needs, you can create clones, snapshots, or snapshots of clones. For a detailed overview of clones or snapshots, refer to the *Clones Overview and Example* (see page 4-4) or to the *Snapshots Overview and Example* (see page 4-6). For a comparison of using clones, snapshots, and snapshots of clones, refer to Table 4-2, *A Comparison of Clones and Snapshots*, on page 4-8.

# Clones Overview and Example

A clone is a complete copy of a source LUN. You specify a source LUN when you create a Clone Group. The copy of the source LUN begins when you add a clone LUN to the Clone Group. The software assigns each clone a clone ID. This ID remains with the clone until you remove the clone from its group.

While the clone is part of the Clone Group and unfractured (not accessible to a secondary server), any server write requests made to the source LUN are simultaneously copied to the clone. Once the clone contains the desired data, you can fracture the clone. Fracturing the clone breaks it from its source LUN and, once activated, makes it available to a secondary server.

Clone private LUNs record information that identifies data chunks on the source LUN and clone LUN that have been modified after you fractured the clone. A modified data chunk is a chunk of data that a server changes by writing to the clone or source LUN. A log in the clone private LUN records this information, but no actual data is written to the clone private LUN. This log reduces the time it takes to synchronize or reverse synchronize a clone and its source LUN since the software copies only modified (changed) chunks.

Writes made to the source LUN from the production server are copied to the fractured clone only when you manually perform a synchronization, which unfractures the clone and updates the contents on the clone with its source.

Figure 4-1 shows an example of how a fractured clone works. Note, as the production server writes to the source LUN, and the secondary server writes to the clone, the clone private LUN tracks areas on the source and clone that have changed since the clone was fractured.

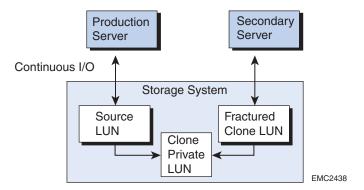


Figure 4-1 Clones Example

#### **Snapshots Overview and Example**

A snapshot is a virtual LUN that allows a secondary server to view a point-in-time copy of a source LUN. You determine the point in time when you start a SnapView session. The session keeps track of the source LUN's data at a particular point in time.

During a session, the production server is still able to write to the source LUN and modify data. When this happens, the software stores a copy of the original point-in-time data on a reserved LUN in the SP's LUN pool. This operation is referred to as copy-on-first-write because it occurs only when a data chunk is first modified on the source LUN.

As the session continues and additional I/O modifies other data chunks on the source LUN, the amount of data stored in the reserved LUN pool grows. If needed, you can increase the size of the reserved LUN pool by adding more LUNs to the LUN pool.

**Important** An adequate number of reserved LUNs is essential since SnapView terminates sessions if the reserved LUN runs out of space and no additional LUNs are in the SP's LUN pool.

From a secondary server, you can view a session's point-in-time data by activating (mapping) a snapshot to the session. You can activate only one snapshot at a time to a session. If another point-in-time view is desired, you can deactivate (unmap) a snapshot from a session and activate it to another session of the same source LUN.

Though a snapshot appears as a conventional LUN to other servers, its data does not reside on a disk like a conventional LUN. A snapshot is a composite of the unchanged data chunks on the source LUN and data chunks that have changed on the reserved LUN. The data chunks on the source LUN are those that have not been modified since you started the session. The data chunks in the reserved LUN pool are copies of the original source LUN data chunks that have been modified since you started the session.

Figure 4-2 shows an example of how snapshots work. The snapshot is a composite of the source LUN and the reserved LUN in the SP's LUN pool.

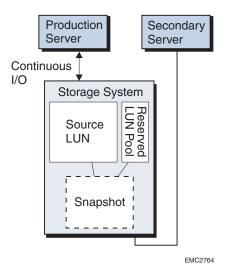


Figure 4-2 Snapshot Example

SnapView also allows you to instantly restore a session's point-in-time data back to the source LUN, if the source LUN were to become corrupt or if a session's point-in-time data is desired for the source. You can do this by using SnapView's rollback feature.

# Clone and Snapshot Trade-offs

The following table describes the benefits and trade-offs of using clones, snapshots, and snapshots of clones.

Table 4-2 A Comparison of Clones and Snapshots

	Clones	Snapshots	Snapshots of Clones
Benefits	Provides immediacy in replacing the contents of the source LUN with the contents of the clone LUN, should the source become corrupted.  Makes backup operation nondisruptive. Provides enhanced protection against critical data loss because it is an actual LUN. Provides immediacy in redirecting servers from the source LUN to the clone LUN, should the source become corrupted.	Provides immediacy in replacing the contents of the source LUN with the contents of the session, should the source LUN become corrupted.  Makes backup operation nondisruptive. Provides a quick and instant copy because it is a virtual LUN.	Provides immediacy in replacing the contents of the source LUN with the contents of the session, should the source LUN become corrupted.  Makes backup operation nondisruptive. Provides an extra level of protection against critical data loss if both the source LUN and clone LUN become corrupted.
Creation Time	Minutes to hours. The creation time depends on the size of the source LUN. Subsequent synchronizations are incremental.	Instantaneous	Instantaneous
Disk Space Used	Uses the same amount of disk space as the source LUN.	Uses reserved LUN pool space, which is usually 10% to 20% of the source LUN size per session, but will vary depending on how much data has changed on the source LUN.	Uses reserved LUN pool space (for the snapshot) and full disk space (for the clone), which usually totals 100% of the source LUN size for clones and 10% to 20% of the source LUN size per session, but will vary depending on how much data has changed on the source LUN.
Data Recovery Time After Source LUN Failure/ Corruption	Instantaneous after initializing a reverse synchronization.	Instantaneous after initializing a rollback operation.	Restore time from network or backup tape if fractured clone LUN received server I/O.
Performance Impact on the Storage System	There is no performance impact when a clone LUN is in a fractured state.  For the initial synchronization of the clone LUN, performance decreases. Subsequent synchronizations or reverse synchronizations have minimal impact on performance. Impact is also determined by the synchronization rate, which is set when the clone LUN is added to the Clone Group.	A slight performance decrease due to the copy-on-first-write.	Combination of both clone LUNs and snapshot LUNs.

## **SnapView Worksheets**

The following three worksheets will help you plan your clones and snapshots. The first worksheet is for clone planning; the second is for snapshot planning; the third is for reserved LUN pool planning for snapshots.

## **Clone Planning Worksheet**

The following worksheet will help you plan for clones. Note that the size of the clone private LUNs (one for each SP) is fixed and it is quite modest, so the worksheet does not include it.

### **Clone Planning Worksheet**

			LUN	
Production Host Name	Source LUN ID	Storage Group ID	Size (Gbytes)	Application, File System, or Database Name

### **Snapshot Planning Worksheet**

For each SnapView session, you must complete a snapshot session worksheet. Note that you must assign the LUN and snapshot to different Storage Groups. One Storage Group must include the production host and source LUN; another Storage Group must include the second host and the snapshot.

### **Snapshot Session Worksheet**

			LUN		Reserve				
Production Host Name	Source LUN ID	Storage Group ID	Size (MB)	Application, File System, or Database Name	LUN ID or Name	Size (MB)	SP (Both LUN and Cache)	Time of Day To Copy	Session Name

Since snapshots use the reserved LUN pool, you need the following information for system setup so you can bind LUNs for the reserved LUN pool. The number of reserved LUNs should always be twice the number of source LUNs. Note that MirrorView/A and incremental SAN Copy also use the reserved LUN pool. One application can use a reserved LUN as long as another application is not using it.

### Reserved LUN Pool Setup Worksheeta

SP (A or B) Only for Navisphere Manager 6.22 or lower <sup>b</sup>	Source LUN Size	RAID Type for Reserved LUN Pool	RAID Group ID of Parent RAID Group	Reserved LUN Size <sup>c</sup> (Gbytes)	Reserved LUN ID (complete after binding)

a. This worksheet is also provided in the MirrorView/A and incremental SAN Copy sections.

b. For Navisphere 6.24 or higher you no longer assign LUNs in the reserved LUN pool to an SP, so the first column of the worksheet - SP (A or B) - is not needed. We recommend that the reserved LUN size is 20% of the source LUN size.

c. EMC recommends 20% of the source LUN size.

## What Is MirrorView/S?

MirrorView/S is a software application that maintains a copy image of a logical unit (LUN) at separate locations in order to provide for disaster recovery; that is, to let one image continue if a serious accident or natural disaster disables the other.

The production image (the one mirrored) is called the primary image; the copy image is called the secondary image. MirrorView/S supports up to two remote images, but since you operate on one image at a time, the examples in this manual show a single image. Each image resides on a separate storage system. The primary image receives I/O from a server called the production server; the secondary image is maintained by a separate storage system that can be a stand-alone storage system or connected to a secondary server. The same client that is managing the storage system, which can promote the secondary image if the primary image becomes inaccessible, manages both storage systems.

MirrorView/S supports MirrorView/S Consistency Groups, which this manual refers to as *Consistency Groups*. A Consistency Group is a set of synchronous mirrors whose secondary images need to be kept consistent with each other in order to be useful. That is, the data on the set of secondary images must have existed on the set of primary images previously. This allows an application to use the secondary images if the primary storage system fails.

The primary images of mirrors in a MirrorView/S Consistency Group must reside on a single storage system, and the secondary images of the mirrors must reside on a single (but different) storage system. This contrasts with volumes in a Symmetrix  $^{\circledR}$  Consistency Group, which can reside on multiple storage systems.

# Sample MirrorView/S Configuration

Figure 4-3 shows a sample remote mirror configuration with two sites and a primary and secondary image that includes the database of four LUNs.

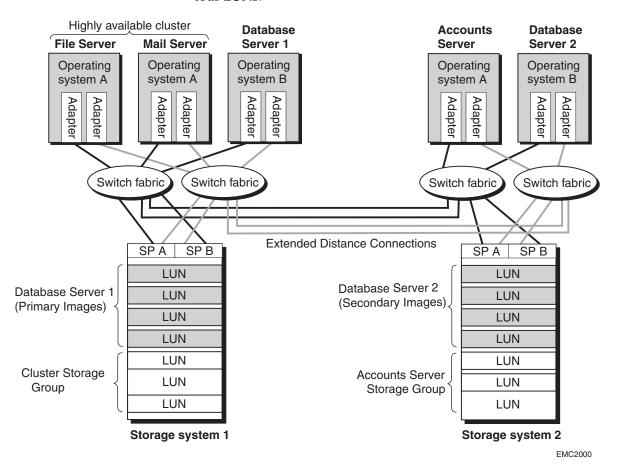


Figure 4-3 Sample Remote Mirror Configuration

In Figure 4-3, Database Server 1, the production server, executes customer applications. These applications access data on Storage System 1, in the Database Server Storage Group. Storage System 2 mirrors the data on the Database Server Storage Group. The mirroring is synchronous, so that Storage System 2 always contains

all data modifications that are acknowledged by Storage System 1 to the production server.

We recommend that you attach a secondary server to Storage System 2, so that if a complete site failure occurs where Storage System 1 and Database Server 1 are located, you can completely fail over to the secondary site, and thus minimize the outage window. The server at the standby site is not required; but because it is recommended, this example includes it in the overall configuration.

Each server has a path to each SP through each fabric — to each storage system. If a failure occurs in a path, software installed on the server (for example, EMC PowerPath®) can switch to the path through the other SP and continue accessing the data, transparent to the applications on the server.

The production server sends a write request to an SP in Storage System 1, which then writes data to the local LUN. The change to the primary LUN is recorded, and at an interval that you define, all changes are copied to the secondary storage system.

If a failure occurs in Storage System 1, an administrator can use the client that is managing the storage system to promote the image on Storage System 2 to the role of primary image.

The mirrored data is inaccessible until the secondary image is promoted to a primary image.

Then the appropriate applications can start on any connected server (here, Database Server 2) with full access to the data. The mirror can be accessible in minutes, although the time needed for applications to recover will vary.

If you want to create a second secondary image, the configuration requires three storage systems.

# MirrorView/S Features and Benefits

MirrorView/S mirroring has the following features:

- Provision for disaster recovery with minimal overhead
- ◆ CLARiiON<sup>®</sup> environment
- Bidirectional mirroring

- ◆ Integration with EMC SnapView<sup>™</sup> LUN copy software
- ◆ Integration with EMC SAN Copy<sup>™</sup> software

## Provision for Disaster Recovery with Minimal Overhead

Provision for disaster recovery is the major benefit of MirrorView/S mirroring. Destruction of the data at the primary site would cripple or ruin many organizations. After a disaster, MirrorView/S lets data processing operations resume with minimal overhead. MirrorView/S enables a quicker recovery by creating and maintaining a copy of the data on another storage system.

MirrorView/S is transparent to servers and their applications. Server applications do not know that a LUN is mirrored and the effect on performance is minimal.

MirrorView/S uses synchronous writes, which means that server writes are acknowledged only after all secondary storage systems commit the data.

MirrorView/S is not server-based; therefore it uses no server I/O or CPU resources. The additional processing for mirroring is performed on the storage system.

## **CLARiiON MirrorView/S Environment**

MirrorView/S operates in a highly available environment, leveraging the dual-SP design of CLARiiON systems. If one SP fails, MirrorView/S running on the other SP will control and maintain the mirrored LUNs. If the server is able to fail over I/O to the remaining SP, then mirroring will continue as normal. The high-availability features of RAID protect against disk failure, and mirrors are resilient to an SP failure in the primary or secondary storage system.

## **Bidirectional Mirroring**

A single storage system may be primary (that is, hold the primary image) for some mirrors and secondary (that is, hold the secondary image) for others. This enables bidirectional mirroring.

A storage system can never hold more than one image of a single mirror.

### Integration with EMC SnapView Snapshot Software

SnapView software lets you create a snapshot of an active LUN at any point in time; however, do this only when the mirror is not synchronizing the secondary image. Since the secondary image is not viewable to any servers, but you can use SnapView in conjunction with MirrorView/S to create a snapshot of a secondary image on a secondary storage system to perform data verification and run parallel processes, for example backup.

You can create a snapshot of a mirrored LUN. You can also create a clone of a mirrored LUN on CX500 or CX700 storage systems running FLARE OE version 02.26 or higher.

Before starting a SnapView session, make sure that the secondary image is in the synchronized or consistent state. Starting a SnapView session of a secondary LUN when MirrorView/S is synchronizing the secondary will not give consistent data. Also note that data cached on the server, as well as data written to the primary, but waiting to be transferred to the secondary on the next update, will not be included in the session of the secondary image.

For more information about SnapView, see the Navisphere Manager help on the Powerlink website.

## Integration with EMC SAN Copy Software

SAN Copy™ software lets you create an intra- or inter-storage system copy of a LUN at any point in time; however, do this only when the image state is either Synchronized or Consistent. The copy is a consistent image that can serve for other application purposes while I/O continues to the source LUN. The MirrorView/S secondary image is not viewable to any servers, but you can use SAN Copy to create a copy of the secondary image on a secondary storage system to perform data verification and run parallel processes.

Related to the process of making the MirrorView/S secondary unavailable for server I/O, you cannot run SAN Copy full copy sessions on MirrorView/S secondary images. You can, however, run SAN Copy incremental sessions on MirrorView/S secondary images.

For more information about SAN Copy, see the Navisphere Manager help on the Powerlink website.

## MirrorView/S Planning Worksheet

The following worksheets help you plan your remote mirrors. The first worksheet is for remote mirror planning, and the second is for the optional write intent log planning. You must assign each primary image LUN to a Storage Group (as with any normal LUN), but you must not assign a secondary image LUN to a Storage Group.

#### **Remote Mirrors Worksheet**

Production Host Name	Primary LUN ID, Size, and File System Name	Storage Group Number or Name	Use Write Intent Log - Y/N <sup>a</sup>	SP (A/B)	Remote Mirror Name	Secondary Host Name	Secondary Image LUN ID

a. By default, newly created remote mirrors do not use the write intent log, so you do not need to allocate space for the write intent log. However, before you use it, you must allocate space for it. If you want to use the write intent log, answer Y, and, after completing the worksheet, fill out the Write Intent Log Worksheet on the next page.

#### **Data Replication and Data Mobility Software**

If you decide to use the write intent log feature for even one remote mirror, then you must bind two LUNs of at least 128 Mbytes for the write intent log.

## Optional Write Intent Log Worksheet

	LUN Name or ID	LUN Size (128 Mbytes minimum)
First LUN		
Second LUN		

# What Is MirrorView/A?

MirrorView/A lets you periodically update a remote copy of production data. It is a software application that keeps a point-in-time copy of a logical unit (LUN) and periodically replicates the copy to a separate location in order to provide disaster recovery, that is, to let one image continue to be active if a serious accident or natural disaster disables the other. It provides data replication over long distances (hundreds to thousands of miles).

The production image (the one mirrored) is called the primary image; the copy image is called the secondary image. MirrorView/A supports one remote image per primary. The primary image receives I/O from a server called the production server; a separate storage system maintains the secondary image. This storage system can optionally have a failover/standby computer connected to it or can be connected to its own computer system. Both storage systems can be in different domains, which you manage with the User Interface (UI). The client that is managing the storage system can promote the secondary image if the primary image becomes inaccessible. After initial synchronization, the remote site always has a consistent point-in-time copy of the primary data.

As a storage system-based mirroring product, MirrorView/A does not flush server buffers before replicating the primary data. Therefore, the copy is a crash-consistent image of the primary data. You must verify data integrity of the secondary image before using it for disaster recovery. (The verification process varies by application type.)

MirrorView/A supports MirrorView/A Consistency Groups, which this manual refers to as *Consistency Groups*. A Consistency Group is a set of asynchronous mirrors whose secondary images need to be kept consistent with each other in order to be useful. That is, the data on the set of secondary images must have existed on the set of primary images previously. This allows an application to use the secondary images if the primary storage system fails.

The primary images of mirrors in a MirrorView/A Consistency Group must reside on a single storage system, and the secondary images of the mirrors must reside on a single (but different) storage system. This contrasts with volumes in a Symmetrix<sup>®</sup> Consistency Group, which can reside on multiple storage systems.

# Sample MirrorView/A Configuration

Figure 4-4 shows a sample remote mirror configuration with two sites and a primary and secondary image that includes the database of four LUNs.

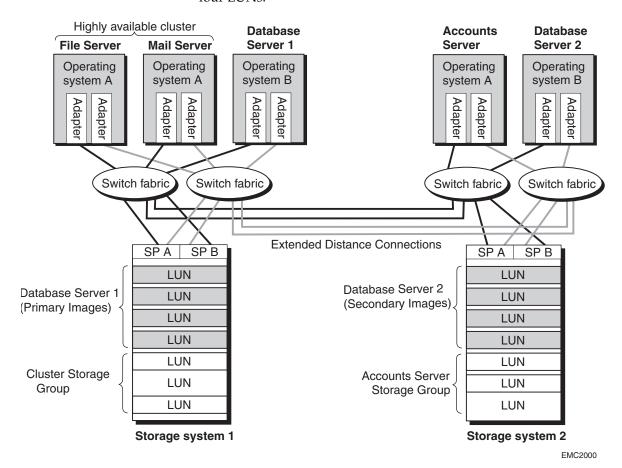


Figure 4-4 Sample Remote Mirror Configuration

In Figure 4-4, Database Server 1, the production server, executes customer applications. These applications access data on Storage System 1, in the Database Server Storage Group. Storage System 2 mirrors the data on the Database Server Storage Group.

Storage Groups are different than Consistency Groups, which are described on page 4-19. For more information about Storage Groups, see the Navisphere Manager help on the Powerlink website.

We recommend that you attach a secondary server to Storage System 2, so that if a complete site failure occurs where Storage System 1 and Database Server 1 are located, you can completely fail over to the secondary site, and thus minimize the outage window. The server at the standby site is not required; but because it is recommended, this example includes it in the overall configuration.

Each server has a path to each SP through each fabric — to each storage system. If a failure occurs in a path, software installed on the server (for example, EMC PowerPath®) can switch to the path through the other SP and continue accessing the data, transparent to the applications on the server.

The production server sends a write request to an SP in Storage System 1, which then writes data to the local LUN. The change to the primary LUN is recorded, and at an interval that you define, all changes are copied to the secondary storage system.

If a failure occurs in Storage System 1, an administrator can use the client that is managing the storage system to promote the image on Storage System 2 to the role of primary image.

The mirrored data is inaccessible until the secondary image is promoted to a primary image.

Then the appropriate applications can start on any connected server (here, Database Server 2) with full access to the data. The mirror can be accessible in minutes, although the time needed for applications to recover will vary.

If you want to create a second secondary image, the configuration requires three storage systems.

## MirrorView/A Features and Benefits

MirrorView/A mirroring has the following features:

- Provision for disaster recovery with minimal overhead
- ◆ CLARiiON<sup>®</sup> environment
- Bidirectional mirroring
- ◆ Integration with EMC SnapView<sup>™</sup> LUN copy software
- Replication over long distances

You can also script application integration using the Navisphere CLI commands. (See the *EMC MirrorView/Asynchronous Command Line Interface (CLI) Reference*.)

## Provision for Disaster Recovery with Minimal Overhead

Provision for disaster recovery is the major benefit of MirrorView/A mirroring. Destruction of the data at the primary site would cripple or ruin many organizations. After a disaster, MirrorView/A lets data processing operations resume with minimal overhead. MirrorView/A enables a quicker recovery by creating and maintaining a copy of the data on another storage system.

MirrorView/A is transparent to servers and their applications. Server applications do not know that a LUN is mirrored and the effect on performance is minimal.

With MirrorView/A, secondary systems are periodically updated.

MirrorView/A is not server-based; therefore it uses no server I/O or CPU resources. The processing for mirroring is performed on the storage system.

### **CLARiiON MirrorView/A Environment**

MirrorView/A operates in a highly available environment, leveraging the dual-SP design of CLARiiON systems. If one SP fails, MirrorView/A running on the other SP will control and maintain the mirrored LUNs. If the server is able to fail over I/O to the remaining SP, then periodic updates will continue. The high-availability features of RAID protect against disk failure, and mirrors are resilient to an SP failure in the primary or secondary storage system.

### **Bidirectional Mirroring**

A single storage system may be primary (that is, hold the primary image) for some mirrors and secondary (that is, hold the secondary image) for others. This enables bidirectional mirroring.

A storage system can never hold more than one image of a single mirror.

## Integration with EMC SnapView Snapshot Software

SnapView software lets you create a snapshot of an active LUN at any point in time; however, do this only when the mirror is not updating to the secondary image. Since the secondary image is not viewable to any servers, you can use SnapView in conjunction with MirrorView/A to create a snapshot of a secondary image on a secondary storage system to perform data verification and run parallel processes, for example backup.

You can create a snapshot of a mirrored LUN. You can also create a clone of a mirrored LUN on CX500, or CX700 storage systems running FLARE OE version 02.26 or higher

Before starting a SnapView session, make sure that the secondary image is in the synchronized or consistent state. Starting a SnapView session of a secondary LUN when MirrorView/A is updating the secondary will not give consistent data. Also note that data cached on the production server, as well as data written to the primary image, but waiting to be transferred to the secondary on the next update, will not be included in the session of the secondary image.

For more information about SnapView, see the Navisphere Manager help on the Powerlink website.

## Integration with EMC SAN Copy Software

SAN Copy™ software lets you create an intra- or inter-storage system copy of a LUN at any point in time; however, do this only when the image state is either Synchronized or Consistent and the mirror is not updating to the secondary image. The copy is a consistent image that can serve for other application purposes while I/O continues to the source LUN. The MirrorView/A secondary image is not viewable to any servers, but you can use SAN Copy to create a copy of the

secondary image on a secondary storage system to perform data verification and run parallel processes.

Related to the process of making the MirrorView/A secondary unavailable for server I/O, you cannot run SAN Copy full copy sessions on MirrorView/A secondary images. You can, however, run SAN Copy incremental sessions on MirrorView/A secondary images.

For more information about SAN Copy, see the Navisphere Manager help on the Powerlink website.

## **Replication Over Long Distances**

MirrorView/A uses FC-to-IP devices to provide replication over long distances (hundreds to thousands of miles).

# MirrorView/A Planning Worksheets

The following worksheets will help you plan your remote asynchronous mirrors. The first is for mirror planning, and the second is for reserved LUN pool planning.

You must assign each primary image LUN to a Storage Group (as with any normal LUN), but you must not assign a secondary image LUN to a Storage Group. If primary image LUNs are in a Consistency Group, all secondary image LUNs must be on the same remote storage system.

### **Remote Asynchronous Mirrors Worksheet**

Production Host Name	Primary LUN ID, Size, and File System Name	Storage Group Number or Name	SP (A/B)	Remote Mirror Name	Secondary Host Name	Secondary Image LUN ID	Group Name (if applicable)

Since MirrorView/A uses the reserve LUN pool, you need the following information for system setup so you can bind LUNs for the reserved LUN pool. The number of reserved LUNs should always be twice the number of source LUNs. A MirrorView/A primary image LUN is a source LUN. Note that SnapView and incremental SAN Copy also use the reserved LUN pool. One application can use a reserved LUN as long as another application is not using it..

### Reserved LUN Pool Setup Worksheeta

SP (A or B) Only for Navisphere Manager 6.22 or lower <sup>b</sup>	Source LUN Size	RAID Type for Reserved LUN Pool	RAID Group ID of Parent RAID Group	Reserved LUN Size <sup>c</sup> (Gbytes)	Reserved LUN ID (complete after binding)

- a. This worksheet is also provided in the MirrorView/A and incremental SAN Copy sections.
- b. For Navisphere 6.24 or higher you no longer assign LUNs in the reserved LUN pool to an SP, so the first column of the worksheet SP (A or B) is not needed. We recommend that the reserved LUN size is 20% of the source LUN size.
- EMC recommends 20% of the source LUN size.

# What Is EMC SAN Copy?

EMC SAN Copy is storage-system-based software for copying data directly from a logical unit on one storage system to destination logical units on supported remote systems without using host resources. SAN Copy connects through a storage area network (SAN), and also supports protocols that let you use the IP WAN to send data over extended distances (see *Implementing SAN Copy Over Extended Distances* on page 4-31).

SAN Copy runs in the SPs of a supported storage system (called a SAN Copy storage system), and not on the host servers connected to the storage systems. As a result, the host processing resources are free for production applications while the SPs copy data.

SAN Copy can copy data between logical units as follows:

- Within a CLARiiON storage system
- On CLARiiON storage systems
- On CLARiiON and Symmetrix storage systems
- On CLARiiON and qualified non-EMC storage systems

SAN Copy can use any CLARiiON SP ports to copy data, provided the port is not being used for MirrorView connections. Multiple sessions can share the same port. You choose which ports SAN Copy sessions use through switch zoning.

You can use SAN Copy to create full or incremental copies of a source logical unit. An incremental session copies only the data that has changed since the last copy session. This can significantly reduce the time needed to copy the data, thereby allowing the copy operation to be performed more frequently and more effectively. Unlike full copy sessions, the source logical unit for an incremental session can remain online during the copy process. Regardless of the type of copy session — full or incremental — the participating logical units must meet certain requirements.

For full and incremental sessions, the source and destination logical units must be

- the same size, or the destination logical unit must be larger.
- for use by the same operating system.
- owned by the same SP, for a copy within one storage system.

In addition, for full sessions:

- The source and destination logical units must be inactive (offline and/or unmounted) during the copy process.
- No writes to the source logical unit (unless you use replication software), and no application I/O with the destination logical unit can occur during the copy session.

In addition, for incremental sessions, the source logical unit

- *cannot* be a SnapView snapshot.
- must reside on the SAN Copy storage system.
- can remain online during the copy process.

You can use SAN Copy for the following tasks:

- One-time data migration
- Routine copying of data for application testing
- ◆ Content distribution
- Scheduled updates to multiple databases

To manage SAN Copy sessions, you can use either the UI-based Navisphere Manager, or the SAN Copy Command Line Interface.

## **SAN Copy Example**

Figure 4-5 shows two logical units being copied to two different storage systems. One SAN Copy session is copying LUN 15 in the CX700 storage system to volume 32 in the Symmetrix storage system. The second SAN Copy session is copying LUN 6 in the CX500 storage system to LUN 14 in the CX700 storage system. In this example, SAN Copy is required only on the CX700.

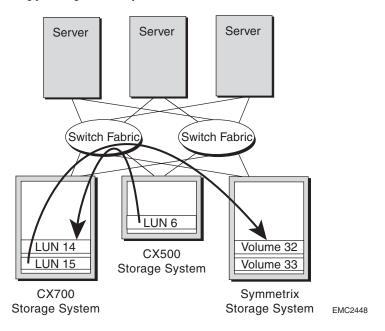


Figure 4-5 Sample SAN Copy Sessions Between Storage Systems

### **SAN Copy Features and Benefits**

SAN Copy adds value to customer systems by offering the following features:

- Storage-system-based data mover application that offloads the host, thereby improving host performance
- Software that you can use in conjunction with replication software, allowing I/O with the source logical unit to continue during the copy process
- Simultaneous sessions that copy data to multiple CLARiiON and Symmetrix storage systems
- Easy-to-use, web-based application for configuring and managing SAN Copy

### Storage-System-Based Data Mobility Application

SAN Copy is a storage-system based data mobility application that uses a SAN to copy data between storage systems. Since SAN Copy runs on the storage systems, the need to move data to and from the attached hosts is eliminated, freeing host resources for users and applications. Since the host is not involved in the copy process and the data migration takes place on the SAN, the copy process is much faster than a LAN-based, host-involved copy process.

## **Use with Replication Software**

If the storage system containing the source logical unit has device replication software, such as EMC SnapView or EMC TimeFinder<sup>TM</sup>, SAN Copy can use a snapshot, clone, or Symmetrix BCV (business continuation volume) as its source logical unit — allowing I/O with the source logical units for the snapshot, clone, or BCV source logical unit to continue during the copy process.

## Simultaneous Sessions Between Multiple Storage Systems

SAN Copy lets you have more than one session active at the same time. The number of supported concurrent active sessions and the number of logical units per session depends on the storage-system type.

## **Implementing SAN Copy Over Extended Distances**

SAN Copy supports protocols that let you use the IP WAN to send LUN data over extended distances. To do this, you need to extend your switch fabric by using Fibre Channel-to-IP converters that allow Fibre Channel devices to communicate over TCP/IP networks.

When using SAN Copy over IP, a read operation will be faster than a write operation.

#### **Performance Considerations**

To get the best performance when implementing SAN Copy over IP, it is important to consider the following performance factors:

- Network performance
- SAN Copy configuration
- Data compression

#### **Network Performance**

The quality of the network is a critical factor in determining SAN Copy performance, especially with respect to latency (delay in the network) and bandwidth (frequency of the transferred bits).

## **SAN Copy Configuration**

In order to address the network performance issues, SAN Copy lets you configure the buffers the SAN Copy operations use.

## **Data Compressibility**

Data compression allows more data to be sent a given amount of time, effectively multiplying the effective bandwidth. The amount of data compression depends on many factors, such as data type, compression algorithm of the equipment, and line speed.

Figure 4-6 shows a sample SAN Copy over IP configuration. In this example, LUN data is being transferred over the IP storage network using two FC-IP converters. Depending on the direction of the data transfer, the FC-to-IP converter either converts Fibre Channel data to IP data or IP data to Fibre Channel data.

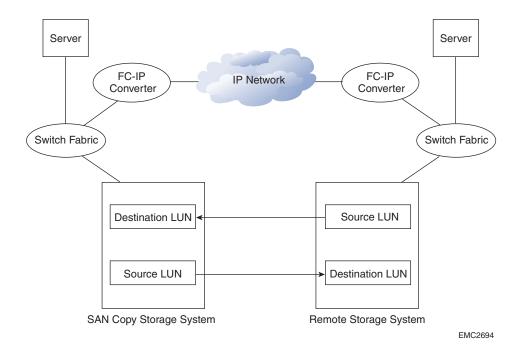


Figure 4-6 Sample SAN Copy Over IP Configuration

SAN COPY FIGHTIII	ng Worksheet
	The following worksheets will help you plan SAN Copy sessions. The first worksheet is for sessions between storage systems; the second is for sessions within the same storage system; and the third is for reserved LUN pool planning.
	One of the storage systems in the session must be a SAN Copy storage system; that is, it must have SAN Copy installed on it.
	SAN Copy Worksheet for a Session Between Storage Systems
	Session Name:
	SAN Copy Storage System Name:
	Source Storage-System Name: Source Logical Unit Name:
	Destination Storage-System Name: Destination Logical Unit Names:
	For a SAN Copy session between logical units in the same storage system, the same SP must own the logical units.  SAN Copy Worksheet for A Session Within Storage System
	Session Name:
	Source Storage-System Name:
	Source Logical Unit Name:

Incremental SAN Copy uses the reserve LUN pool, so if you want to use this type of SAN Copy, you need the following information for system setup so you can bind LUNs for the reserved LUN pool. The number of reserved LUNs should always be twice the number of source LUNs. Note that MirrorView/A and SnapView also use the reserved LUN pool. One application can use a reserved LUN as long as another application is not using it..

### Reserved LUN Pool Setup Worksheeta

SP (A or B) Only for Navisphere Manager 6.22 or lower <sup>b</sup>	Source LUN Size	RAID Type for Reserved LUN Pool	RAID Group ID of Parent RAID Group	Reserved LUN Size <sup>c</sup> (Gbytes)	Reserved LUN ID (complete after binding)

- This worksheet is also provided in the MirrorView/A and incremental SAN Copy sections.
- b. For Navisphere 6.24 or higher you no longer assign LUNs in the reserved LUN pool to an SP, so the first column of the worksheet SP (A or B) is not needed. We recommend that the reserved LUN size is 20% of the source LUN size.
- c. EMC recommends 20% of the source LUN size.

# What Is EMC SAN Copy/E?

Like EMC SAN Copy, EMC SAN Copy/E is storage-system-based software for copying data directly from a logical unit on one storage system to destination logical units on supported remote systems without using host resources. SAN Copy/E connects through a storage area network (SAN), and also supports protocols that let you use the IP WAN to send data over extended distances (see *Implementing SAN Copy/E over Extended Distances* on page 4-40). SAN Copy/E runs only on a CX300 or AX-Series storage system and can only copy data to CX-Series storage systems that are running San Copy.

SAN Copy/E can use any CLARiiON SP ports to copy data, provided the port is not being used for MirrorView connections. Multiple sessions can share the same port. You choose which ports SAN Copy sessions use through switch zoning.

You can use SAN Copy/E to create full or incremental copies of a source logical unit. A full copy session copies all the data on the source logical unit to the destination logical units. An incremental session copies only the data that has changed since the last copy session. This can significantly reduce the time needed to copy the data, thereby allowing the copy operation to be performed more frequently and more effectively. Unlike full copy sessions, the source logical unit for an incremental session can remain online during the copy process. Regardless of the type of copy session — full or incremental — the participating logical units must meet certain requirements.

For full and incremental sessions

- the source logical unit must reside on the storage system running SAN Copy/E software.
- the destination logical units must reside on CX-Series storage systems running SAN Copy software.
- the source and destination logical units must be the same size, or the destination logical unit must be larger.
- the source and destination logical units must for use by the same operating system.

In addition, for full sessions:

- The source and destination logical units must be inactive (offline and/or unmounted) during the copy process.
- No writes to the source logical unit (unless you use replication software), and no application I/O with the destination logical unit can occur during the copy session.

In addition, for incremental sessions, the source logical unit

- cannot be a SnapView snapshot.
- can remain online during the copy process.

You can use SAN Copy/E for the following tasks:

- One-time data migration
- Routine copying of data for application testing
- Content distribution
- Scheduled updates to multiple databases

To manage SAN Copy/E sessions, you can use either the UI-based Navisphere Manager, or the SAN Copy Command Line Interface.

## **SAN Copy/E Example**

In Figure 4-7, the SAN Copy storage system is a dual-SP AX100 storage system running SAN Copy/E, and the remote storage system is a CX700 storage system running SAN Copy. In this example, two copy sessions are running. One session is copying data from LUN 1 to LUN 3. The other session is copying data from LUN 2 to LUN 4

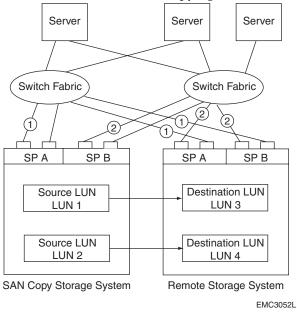


Figure 4-7 Sample SAN Copy/E Session Between Storage Systems

### **SAN Copy/E Features and Benefits**

The SAN Copy/E adds value to customer systems by offering the following features:

- Incremental copy sessions from AX-Series or CX300 storage systems to SAN Copy systems located in the data center.
- Storage-system-based data mover application that uses a storage area network (SAN) rather than host resources to copy data resulting in a faster copy process.
- Easy-to-use, web-based application for configuring and managing SAN Copy/E.
- Software that you can use in conjunction with replication software, allowing I/O with the source logical unit to continue during the copy process

### Incremental Copying from Remote Systems to Data Center

SAN Copy/E lets you use a LUN that resides on an AX-Series or CX300 storage system as the source of an incremental copy session. The incremental copy feature lets you copy only the data that has changed since the last copy session. This significantly reduces the time needed to copy the data, thereby allowing the copy operation to be performed more frequently and more effectively. Unlike full copy sessions, the source logical unit can remain online during the copy process.

To use the incremental copy feature

- The source logical unit *cannot* be a SnapView snapshot.
- You must configure the reserved LUN pool (see the Navisphere Manager help on the Powerlink website).

## Storage-System-Based Data-Mover Application

SAN Copy/E is a storage-system-based data-mover application that uses a SAN (storage area network) to copy data between storage systems. Since SAN Copy/E runs on the storage systems, this eliminates the need to move data to and from the attached hosts and reserves host processing resources for users and applications. Since the host is not involved in the copy process, and the data migration takes place on the SAN, the copy process is much faster than the LAN-based, host-involved copy process.

### **Easy-to-Use Web-Based Application**

SAN Copy/E is an integrated part of Navisphere Manager, the easy-to-use, web-based management tool for CLARiiON storage systems. Manager lets you select the source and destination logical units for the copy session, create copy sessions and then manage these sessions.

### **Use with Replication Software**

If the storage system containing the source logical unit has device replication technology, such as EMC SnapView  $^{\text{TM}}$  software, SAN Copy/E can use a snapshot (full copy sessions only), or clone as its source logical unit — allowing I/O with the source logical units for the snapshot or clone to continue during the copy process.

AX-Series storage systems do not support SnapView clones.

You must fracture a SnapView clone before using it as a source LUN in a copy session. For more details, refer to the SnapView documentation and online help.

# Implementing SAN Copy/E over Extended Distances

SAN Copy/E supports protocols that let you use the IP WAN to send LUN data over extended distances. In order to do this, you need to extend your switch fabric by using Fibre Channel to IP converters that allow Fibre Channel devices to communicate over TCP/IP networks.

When using SAN Copy/E over IP, a read operation will be faster than a write operation.

### **Performance Considerations**

To get the best performance when implementing SAN Copy/E over IP, it is important to consider the following performance factors:

- Network performance
- ◆ SAN Copy/E configuration
- Data compression

#### **Network Performance**

The quality of the network is a critical factor in determining SAN Copy/E performance, especially with respect to latency (delay in the network) and bandwidth (frequency of the transferred bits). SAN Copy/E lets you set the link bandwidth for your network configuration.

## **SAN Copy/E Configuration**

In order to address the network performance issues, SAN Copy/E lets you configure the buffers used by SAN Copy/E operations.

### **Data Compression**

Data compression allows more data to be sent in the same amount of time, effectively multiplying the effective bandwidth. The amount of data compression depends on many factors, such as data type, compression algorithm of the equipment, and line speed.

Figure 4-8 shows a sample SAN Copy/E over IP configuration. In this example, LUN data is being transferred over the IP storage network using two FC-IP converters. Depending on the direction of the data transfer, the converter either converts Fibre Channel data to IP data or IP data to Fibre Channel data.

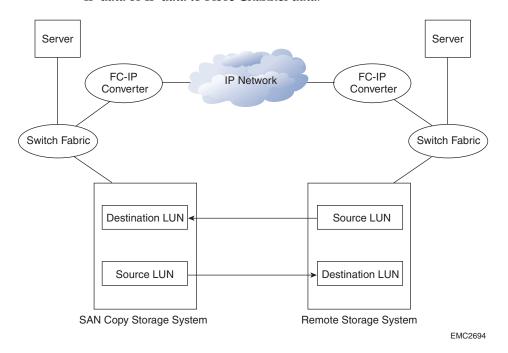


Figure 4-8 Sample SAN Copy/E over IP Configuration

### **SAN Copy/E Planning Worksheet**

The following two worksheets will help you plan SAN Copy/E sessions. The first worksheet is for sessions between storage systems; and the second is for reserved LUN pool planning.

The source storage system must be running SAN Copy/E software; the destination storage systems must be CX-Series storage systems running SAN Copy.

### SAN Copy/E Worksheet for a Session Between Storage Systems

Incremental SAN Copy/E uses the reserve LUN pool, so if you want to use this type of copy session, you need the following information for system setup so you can bind LUNs for the reserved LUN pool. The number of reserved LUNs should always be twice the number of source LUNs. Note that MirrorView/A and SnapView also use the reserved LUN pool. One application can use a reserved LUN as long as another application is not using it..

### Reserved LUN Pool Setup Worksheeta

SP (A or B) Only for Navisphere Manager 6.22 or lower <sup>b</sup>	Source LUN Size	RAID Type for Reserved LUN Pool	RAID Group ID of Parent RAID Group	Reserved LUN Size <sup>c</sup> (Gbytes)	Reserved LUN ID (complete after binding)
<u> </u>					

- a. This worksheet is also provided in the MirrorView/A and incremental SAN Copy sections.
- b. For Navisphere 6.24 or higher you no longer assign LUNs in the reserved LUN pool to an SP, so the first column of the worksheet SP (A or B) is not needed. We recommend that the reserved LUN size is 20% of the source LUN size.
- EMC recommends 20% of the source LUN size.

#### What Next?

To plan LUNs and file systems, go to Chapter 5.

For details on the storage-system hardware, go to Chapter 6.

For worksheets and a description of iSCSI configurations, go to Chapter 7.

Data Replication and Data Mobility Software	

# Planning File Systems and LUNs

This chapter shows sample installations with shared switched and unshared direct storage, and then provides worksheets for planning your own storage installation.

# Major sections are

<b>♦</b>	Multiple Paths to LUNs	5-2
	Disk IDs	
•	Sample Shared Switched Installation	5-2
	Sample Unshared Direct Installation	
	Planning Applications, LUNs, and Storage Groups	

# Multiple Paths to LUNs

A shared storage-system installation includes two or more servers and one or more storage systems, each with two SPs and the Access Logix option. Often shared storage installations include two or more switches or routers.

In properly configured shared storage (switched or direct), each server has at least two paths to each LUN in the storage system. The Access Logix software detects all paths and, using optional PowerPath® failover software, can automatically switch to the other path, without disrupting applications, if a device (such as a host bus adapter or cable) fails. And with two adapters and two or more ports per SP zoned to them, PowerPath can send I/O to each available path in a user-selectable sequence (multipath I/O) for load sharing and greater throughput.

Unshared storage, if the server has two adapters, provides two paths to each LUN. With two adapters, PowerPath performs the same function as with shared systems: automatically switches to the other path if a device (such as host bus adapter or cable) fails.

# **Disk IDs**

Disk IDs have the form b e d, where

- *b* is the back-end loop (also referred to as a back-end bus) number (0, or 1)
- *e* is the enclosure number, set on the enclosure rear panel (0 for a DPE2 and the first DAE with an SPE)
- d is the disk position in the enclosure (left is 0, right is 14)

# Sample Shared Switched Installation

Figure 5-1 and the table that follows it show a sample shared storage system connected to three servers: two servers in a cluster and one server running a database management program. Note that each server has a complete independent connection to SP A and SP B.

Figure 5-1 shows a single-cabinet storage system. You can configure a storage system with multiple cabinets.

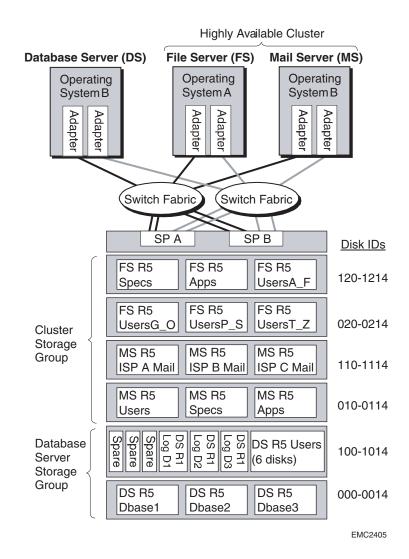


Figure 5-1 Sample Shared Switched Storage Configuration

The storage-system disk IDs, RAID types, storage types, capacities, and drive letters for the configuration in Figure 5-1 are as follows. The capacities shown reflect 73-Gbyte disks.

					Clustered S	System LUN	ls	
Database Server LUNs (DS) - SP A 1,440 Gbytes on Seven LUNs			Mail Server LUNs (MS) - SP A 1,460 Gbytes on Seven LUNs			File Server LUNs (FS) - SP B 1,460 Gbytes on Seven LUNs		•
Disk IDs	RAID Type	Storage Type, Capacity, Drive	Disk IDs	RAID Type	Storage Type, Capacity, Drive			Storage Type, Capacity, Drive
000-004	RAID 5	Dbase1, 292 Mbytes, drive G	010-014	RAID 5	Users, 292 Mbytes, drive N	020-024	RAID 5	UsersG_O, 292 Mbytes, drive T
005-009	RAID 5	Dbase2, 292 Mbytes, drive H	015-019	RAID 5	Specs, 292 Mbytes, drive O	025-029	RAID 5	UsersP_S, 292 Mbytes, drive U
0010-0014	RAID 5	Dbase3, 292 Mbytes, drive I	0110- 0114	RAID 5	Apps, 292 Mbytes, drive P	0210- 0214	RAID 5	UsersT_Z, 292 Mbytes, drive V
103-104	RAID 1	Log file,Dbase1, 73 Mbytes, drive J	110-114	RAID 5	ISP A Mail, 292 Mbytes, drive Q	120-124	RAID 5	Specs, 292 Mbytes, drive U
105-106	RAID 1	Log file, Dbase2, 73 Mbytes, drive K	115-119	RAID 5	ISP B Mail, 292 Mbytes, drive R	125-129	RAID 5	Apps, 292 Mbytes drive V
107-108	RAID 1	Log file, Dbase3, 73 Mbytes drive L	1110- 1114	RAID 5	ISP C Mail, 292 Mbytes, drive S	1210- 1214	RAID 5	UsersA_F, 292 Mbytes, drive W
109-1014	RAID 5	Users, 345 Mbytes (6 disks), drive M			•			•

Note: The available storage is 50% for a RAID 1 LUN and 80% for a 5-disk RAID 5 LUN.

# Sample Unshared Direct Installation

This section shows the disks and LUNs in an unshared direct storage-system installation.

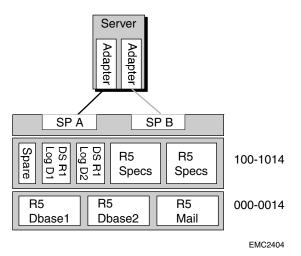


Figure 5-2 Unshared Direct Storage System

The example below describes a storage system with two back-end loops. Each disk holds 73 Gbytes, so the storage system provides the server with 504 Gbytes of disk storage, all highly available. The storage-system disk IDs and LUNs are as follows:

0 )				
LUNs - SP A and SP B - 1,440 Gbytes on Seven LUNs				
Disk IDs	RAID Type	Storage Type, Capacity, Drive		
000-004	RAID 5	Dbase1, 292 Mbytes, drive G		
005-009	RAID 5	Dbase2, 292 Mbytes, drive H		
0010-0014	RAID 5	Mail, 292 Mbytes, drive I		
101-102	RAID 1	Log file, Dbase1, 73 Mbytes, drive J		
103-104	RAID 1	Log file, Dbase2, 73 Mbytes, drive K		
105-109	RAID 1	Users, 292 Mbytes, drive L		
1010-1014	RAID 5	Specs, 292 Mbytes, drive M		
100 – Hot spare (automatically replaces a failed disk in any LUN)				

Note: The available storage is 50% for a RAID 1 LUN and 80% for a 5-disk RAID 5 LUN.

# Planning Applications, LUNs, and Storage Groups

This section will help you plan your storage use — the applications to run, the LUNs that will hold them, and, for shared storage, the Storage Groups that will belong to each server.

Unshared storage systems do not use Storage Groups. For unshared storage, skip the Storage Group entry on the LUN and Storage Group worksheet.

#### The worksheets include

- Application and LUN planning worksheet lets you outline your storage needs.
- LUN and Storage Group planning worksheet lets you decide on the disks to compose the LUNs and the LUNs to compose the Storage Groups for each server.
- ◆ LUN details worksheet lets you plan each LUN in detail.

Make as many copies of each blank worksheet as you need. You will need this information later when you configure the storage system(s).

Sample worksheets appear later in this chapter.

# **Application and LUN Planning**

Use the following worksheet to list the applications you will run, and the RAID type and size of LUN to hold them. For each application that will run, write the application name, file system (if any), RAID type, LUN ID (ascending integers, starting with 0), disk space required, and finally the name of the servers and operating systems that will use the LUN.

### **Application and LUN Planning Worksheet**

Application	File System, Partition, or Drive	RAID Type of LUN	LUN ID (hex)	Disk Space Provided (Gbytes)	Server Hostname and Operating System

# Sample Application and LUN Planning Worksheet

Application	File system, Partition, or Drive	RAID Type of LUN	LUN	Disk Space Provided (Gbytes)	Server Hostname and Operating System
Dbase1	S	RAID 5	0	292 GB	Server1, Windows
Dbase2	T	RAID 5	1	292 GB	Server2,Windows
Log file for Dbase1	U	RAID 1	2	73 GB	Server1,Windows
Log file for Dbase2	V	RAID 1	3	73 GB	Server2,Windows
Users	W	RAID 5	4	365 GB	Server1, Windows

### Completing the Application and LUN Planning Worksheet

**Application**. Enter the application name or type.

**File system, partition**, or **drive**. Write the planned **file system**, **partition**, or **drive** name here.

With a Windows operating system, the LUNs are identified by drive letter only. The letter does not help you identify the disk configuration (such as RAID 5). We suggest that later, when you use the operating system to create a partition on a LUN, you use the disk administrator software to assign a volume label that describes the RAID type. For example, for drive T, assign the volume ID RAID5\_T. The volume label will then identify the drive letter.

**RAID type of LUN**. This is the RAID Group type you want for this file system, partition, or drive. The features of RAID types are explained in Chapter 2. You can create one or more LUNs on the RAID Group.

**LUN ID.** The LUN ID is a hexadecimal number assigned when you bind the disks into a LUN. By default, the ID of the first LUN bound is 0, the second 1, and so on. Each LUN ID must be unique within the storage system, regardless of its Storage Group or RAID Group.

The maximum number of LUNs supported on one host bus adapter depends on the operating system.

**Disk space required (Gbytes)**. Consider the largest amount of disk space this application will need, and then add a factor for growth.

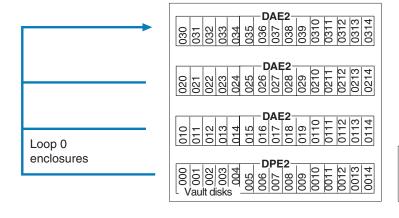
If sometime in the future, you find that you need more space for this application, you can expand the capacity of the LUN by adding other LUNs to it. For information on expanding LUNs, see the Navisphere Manager help on the Powerlink website.

**Server hostname and operating system**. Enter the server hostname (or, if you don't know the name, a short description that identifies the server) and the operating system name, if you know it.

# **LUN and Storage Group Planning Worksheet**

Use the appropriate worksheet to select the disks that will make up the LUNs and Storage Groups in each storage system. A storage system is any group of enclosures connected to a DPE2 (CX300-Series or CX500-Series storage system) or an SPE (CX700 storage system). A CX300/CX300i storage system can include up to three DAE enclosures for a total of 60 disks; a CX500/CX500i storage system can include up to seven DAE enclosures for a total of 120 disks; a CX700 storage system can include up to 16 DAE enclosures for a total of 240 disks. Unshared storage systems do not use Storage Groups. For unshared storage, skip the Storage Group entry.

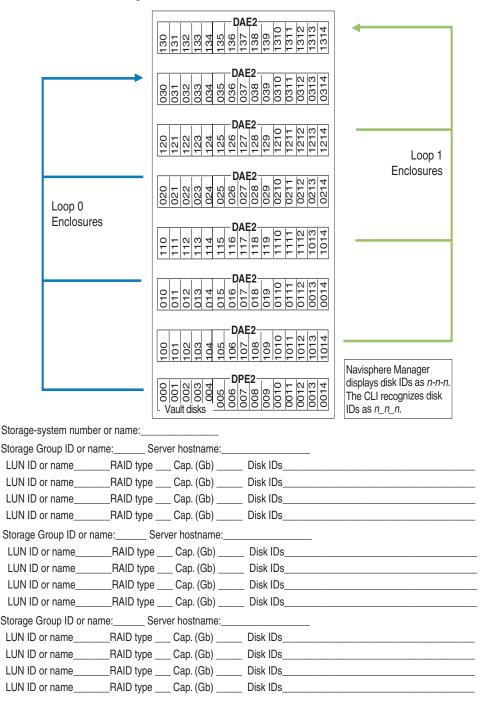
### LUN and Storage Group Worksheet for CX300/CX300i



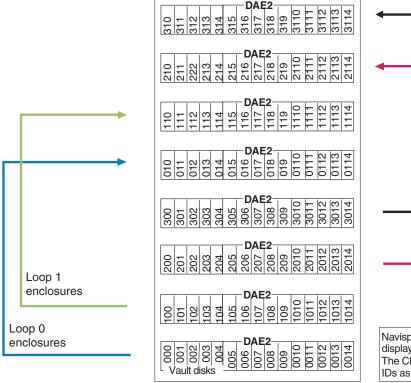
Navisphere Manager displays disk IDs as *n-n-n*. The CLI recognizes disk IDs as *n\_n\_n*.

Storage-system numbe	er or name:			
Storage Group ID or na	ame: Serv	/er hostname:_		
LUN ID or name	RAID type	_ Cap. (Gb)	Disk IDs	 
LUN ID or name	RAID type	_ Cap. (Gb)	Disk IDs	 
LUN ID or name	RAID type	_ Cap. (Gb)	Disk IDs	 
LUN ID or name	RAID type	_ Cap. (Gb)	Disk IDs	 
Storage Group ID or n	ame: Ser			
LUN ID or name	RAID type _	Cap. (Gb)	Disk IDs	 
LUN ID or name	RAID type _	Cap. (Gb)	Disk IDs	 
LUN ID or name	RAID type _	Cap. (Gb)	Disk IDs	 
LUN ID or name	RAID type	Cap. (Gb)	Disk IDs	

### LUN and Storage Group Worksheet for CX500CX500i



### LUN and Storage Group Worksheet for CX700 - Single Cabinet

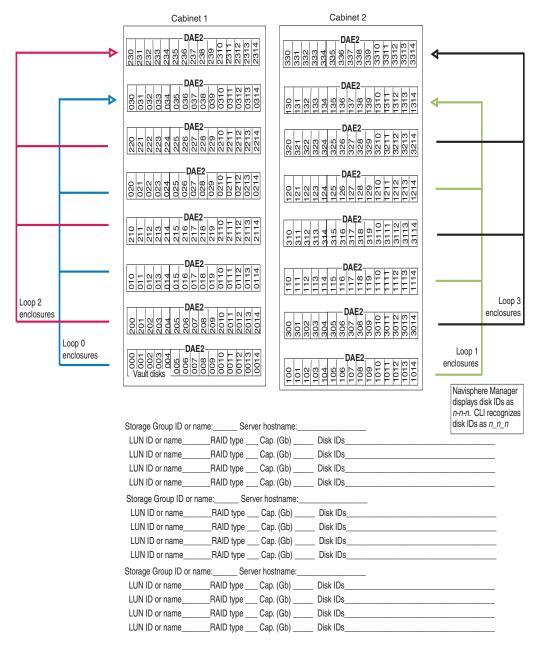


Loop 3 enclosures Loop 2 enclosures

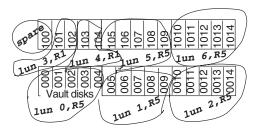
Navisphere Manager displays disk IDs as *n-n-n*. The CLI recognizes disk IDs as *n\_n\_n*.

Storage-system number or name:\_\_\_ Storage Group ID or name: Server hostname: LUN ID or name \_RAID type \_\_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs \_RAID type \_\_\_\_ Cap. (Gb) \_\_\_\_\_ Disk IDs LUN ID or name LUN ID or name \_RAID type \_\_\_ Cap. (Gb) \_\_\_ Disk IDs LUN ID or name RAID type \_\_\_ Cap. (Gb) \_\_\_ Disk IDs Storage Group ID or name:\_\_\_\_\_ Server hostname: RAID type \_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs\_\_ LUN ID or name LUN ID or name \_RAID type \_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs\_\_\_ LUN ID or name \_RAID type \_\_\_\_ Cap. (Gb) \_\_\_ Disk IDs RAID type \_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs\_\_\_ LUN ID or name\_ Storage Group ID or name:\_\_\_\_\_ Server hostname:\_\_\_ LUN ID or name\_ \_RAID type \_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs\_\_\_ RAID type Cap. (Gb) Disk IDs LUN ID or name LUN ID or name \_\_\_RAID type \_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs\_\_\_\_ LUN ID or name\_\_\_\_ RAID type \_\_\_ Cap. (Gb) \_\_\_\_ Disk IDs\_\_\_

### LUN and Storage Group Worksheet for CX700 - Dual Cabinets



Part of a sample LUN and Storage Group worksheet for one cabinet follows.



Storage-system number or name:ss1
Storage Group ID or name:o Server hostname:server1 Dedicated XShared □
LUN ID or nameoRAID type _5_ Cap. (Gb) _292_ Disk IDs000, 001, 002, 003, 004
LUN ID or name1RAID type _5_ Cap. (Gb) _292_ Disk IDs005, 006, 007, 008, 009
LUN ID or name2RAID type _5_ Cap. (Gb) _292_ Disk IDs0010, 0011, 0012, 0013, 0015
LUN ID or name3RAID type _1_ Cap. (Gb)73_ Disk IDs101, 102
Storage Group ID or name:o Server hostname:server1 Dedicated Shared
LUN ID or name4RAID type _1_ Cap. (Gb)73_ Disk IDs103,104
LUN ID or name5RAID type _5_ Cap. (Gb) _292_ Disk IDs105, 106, 107, 108, 109
LUN ID or name6RAID type Cap. (Gb) _292_ Disk IDs1010, 1011, 1012, 1013, 1014
LUN ID or name RAID type Cap. (Gb) Disk IDs

# Disk Configuration Rules and Recommendations

The following rules and recommendations apply to CX300, CX300i, CX500, CX500i, and CX700 systems:

- ◆ You cannot use any of the disks 000 through 004 (enclosure 0, loop 0, disks 0-4) as a hot spare in a CX-Series system.
- ◆ The hardware reserves several gigabytes on each of disks 000 through 004 for the cache vault and internal tables. To conserve disk space, you should avoid binding any other disk into a RAID Group that includes any of these disks. Any disk you include in a RAID Group with a cache disk 000-004 is bound to match the lower unreserved capacity, resulting in lost storage of several gigabytes per disk.
- ◆ Each disk in the RAID Group should have the same capacity. All disks in a Group are bound to match the smallest capacity disk, and you could waste disk space. The first five drives (000-004) should always be the same size.

- You cannot mix ATA (Advanced Technology Attachment) and Fibre Channel disk drives within a RAID Group.
- Hot spares for Fibre Channel drives must be Fibre Channel drives; ATA drives require ATA hot spares.
- If a storage system uses disks of different capacities and/or speeds (for example, 146 Gbytes and 73 Gbytes, or 10K and 15K rpm) within any enclosure, then EMC recommends that you place them in a logical order, such as the following:
  - Place drives with the highest capacity in the first (leftmost) slots, followed by drives of lower capacities.
  - Within any specific capacity, place drives of the highest speed first, followed by drives of lower speed.
- You should always use disks of the same speed and capacity in any RAID Group.
- Do not use ATA drives to store boot images of an operating system. You must boot host operating systems from a Fibre Channel drive.

### Completing the LUN and Storage Group Planning Worksheet

As shown, draw circles around the disks that will compose each LUN, and within each circle specify the RAID type (for example, RAID 5) and LUN ID. This is information you will use to bind the disks into LUNs. For disk IDs, use the form shown on page 5-10.

Next, complete as many of the Storage System sections as needed for all the Storage Groups in the SAN (or as needed for all the LUNs with unshared storage). Copy the (blank) worksheet as needed.

For shared storage, if a Storage Group will be dedicated (not accessible by another server in a cluster), mark the **Dedicated** box at the end of its line; if the Storage Group will be accessible to one or more other servers in a cluster, write the hostnames of all servers and mark the **Shared** box.

For unshared storage, ignore the **Dedicated/Shared** boxes.

### **LUN Details Worksheet**

Use the LUN Details Worksheet to plan the individual LUNs. Blank and sample completed LUN worksheets follow.

Complete as many blank worksheets as needed for all LUNs in storage systems. For unshared storage, skip the Storage Group entries.

LUN Details W	orksheet
Storage system (complete this section one	ce for each storage system)
Storage-system number or name:Storage-system installation type  Unshared Direct  Shared-or-Clustered Direct	Shared Switched
SP information: SP A: IP address or hostname:Port SP B: IP address or hostname	rt ALPA ID: Memory (Mbytes)
LUN ID: SP owner: 0	□A □B SP loop (0 or 1):
RAID Group ID:Size,GB: LUN size,GB:Disk IDs:_ RAID type: □RAID 5 □RAID 3 □RAID 1/0 □Individual disk  Caching: □Read and write □Write □Read □None	□RAID 1 mirrored pair □RAID 0 □Hot spare
Servers that can access this LUN's Storage Group:	
Operating system information: Device name:	_File system, partition, or drive:
RAID Group ID:Size,GB: LUN size,GB:Disk IDs:  RAID type: □RAID 5 □RAID 3 □RAID 1/0 □Individual disk	□RAID 1 mirrored pair □RAID 0
Caching: □Read and write □Write □Read □None	□Hot spare
Servers that can access this LUN's Storage Group:	
Operating system information: Device name:	_File system, partition, or drive:
LUN ID: SP owner: □A IRAID Group ID:Size,GB: LUN size,GB:Disk IDs:	
RAID type: □RAID 5 □RAID 3 □Individual disk  Caching: □Read and write □Write □Read □None	□RAID 1 mirrored pair □RAID 0 □Hot spare
Servers that can access this LUN's Storage Group:	
Operating system information: Device name:	_riie system, partition, or drive:

### Completing the LUN Details Worksheet

Complete the header portion of the worksheet for each storage system as described next. Copy the blank worksheet as needed.

### **Storage-System Entries**

**Storage-system installation type:** specify Unshared Direct, Shared-or-Clustered Direct, or Shared Switched.

SP information: IP address or hostname. The IP address is required for communication with the SP. You don't need to complete it now, but you will need it when the storage system is installed so that you can set up communication with the SP.

**Port ALPA ID.** For shared storage (switches), the SP ALPA ID is ignored for all except HP/UX servers. For unshared storage, each SP connected to a server must have a unique SP Port ALPA ID; for example, for two storage systems, use IDs 0, 1, 2, and 3. The SP Port ALPA ID, like the IP address, is generally set at installation.

**Memory (Mbytes)**. Each CX300/CX300i has 2 Gbytes of memory (one per storage processor). Each CX500/CX500i has 4 Gbytes of memory, and each CX700 has 8 Gbytes of memory.

**Caching.** You can use SP memory for read/write caching. You can use different cache settings for different times of day. For example, for user I/O during the day, use more write cache; for sequential batch jobs at night, use more read cache. You enable caching for specific LUNs — allowing you to tailor your cache resources according to priority. If you choose caching, check the box and continue to the next cache item.

**Read cache percentage.** If you want a read cache, generally it should be about 20% of the cache memory. For large configurations, assign at least 200MB to the read cache.

**Write cache percentage.** Generally, the write cache should be whatever is left of total available cache memory after allocating read cache.

**Cache page size.** This applies to both read and write caches. It can be 2, 4, 8, or 16 Kbytes.

As a general guideline, we suggest 8 Kbytes. The ideal cache page size depends on the operating system and application.

### **RAID Group/LUN Entries**

Complete a RAID Group/LUN entry for each LUN and hot spare.

**LUN ID**. The LUN ID is a hexadecimal number assigned when you bind the disks into a LUN. By default, the ID of the first LUN bound is 0, the second 1, and so on. Each LUN ID must be unique within the storage system, regardless of its Storage Group or RAID Group.

The maximum number of LUNs supported on one host bus adapter depends on the operating system.

**SP** owner. Specify the SP that will own the LUN: SP A or SP B. You can let the management program automatically select the SP to balance the workload between SPs; to do so, leave this entry blank.

**SP** back-end loops (buses). Each CX300/CX300i SP supports one physical loop that is paired with its counterpart SP to create a redundant loop (loop 0).

Each CX500/CX500i SP supports two physical loops that are paired with its counterpart SP to create two redundant loops (loop 0 and 1).

Each CX700 SP supports four physical loops that are paired with its counterpart SP to create four redundant loops (loop 0, 1, 2, and 3).

Ideally, you will place the same amount of load on each loop. This may mean placing two or three heavily used LUNs on one loop, and six or eight lightly used LUNs on the other loop. The loop designation appears in the disk ID (form b-e-d, where b is the back-end loop number, e is the enclosure number, and d is the disk position in the enclosure. For example, 0-1-3 indicates the fourth disk on loop 0, in enclosure 1 (numbering 0, 1, 2, 3...from the left) in the storage system.

**RAID Group ID**. This ID is a hexadecimal number assigned when you create the RAID Group. By default, the number of the first RAID Group in a storage system is 0, the second 1, and so on, up to the number of disk drives in the system.

**Size** (RAID Group size). Enter the user-available capacity in gigabytes (Gbytes) of the whole RAID Group. You can determine the capacity as follows:

RAID 6 disk-size \* (number-of-disks - 2)
RAID 5 or RAID 3 Group: disk-size \* (number-of-disks - 1)
RAID 1/0 or RAID 1 Group: (disk-size \* number-of-disks) / 2
RAID 0 Group: disk-size \* number-of-disks

Individual unit: disk-size

Usable disk space differs from the drive size reported by the manufacturer. To calculate usable bytes, you must account for

- private/reserved space (approximately 35 Mbytes on most disk drives, 6.3 Gbytes on drives 0-4)
- 8 bytes per 520-byte block used by the system
- the difference between a drive manufacturer's definition of a gigabyte (1,000,000,000) and the storage system's "true" definition (1024 x 1024 x 1024, or 1,073,741,824).

Typical conversions are shown below.

Listed Capacity	Available Gbytes			
	Drive 0-4	Other		
73 Gbytes	60.39 Gbytes	66.61 Gbytes		
146 Gbytes	127.43 Gbytes	133.65 Gbytes		
250 Gbytes (ATA)	NA	230.13 Gbytes		
320 Gbytes (ATA)	NA	297.01 Gbytes		

### For example,

- A five-disk RAID 5 or RAID 3 Group of 73-Gbyte disks holds 241.57 Gbytes if it includes disks 0-4; 266.42 Gbytes if it does not include the first five disks;
- An eight-disk RAID 1/0 Group of 73-Gbyte disks holds 241.57 with vault disks or 266.42 Gbytes without disks 0-4;
- A RAID 1 mirrored pair of 73-Gbyte disks holds 66.61 Gbytes (60.39 if the pair includes disks 0-4); and
- An individual disk of an 73-Gbyte disk also holds 60.39 or 66.61 Gbytes.

**LUN size**. Enter the user-available capacity in gigabytes (Gbytes) of the LUN. You can make this the same size or smaller than the RAID Group. You might make a LUN smaller than the RAID Group size if you wanted a RAID 5 Group with a large capacity and wanted to place many smaller capacity LUNs on it; for example, to specify a LUN for each user.

Multiple LUNs per disk group share the same total performance capability of the disk drives, so plan carefully.

If you want multiple LUNs per RAID Group, then use a RAID Group/LUN series of entries for each LUN.

**Disk IDs**. Enter the IDs of all disks that will make up the LUN or hot spare. These are the same disk IDs you specified on the previous worksheet. For example, for a RAID 5 Group in DAE on loop 0 (disks 10 through 14), enter **0010**, **0011**, **0012**, **0013**, and **0014**.

**RAID** type. Copy the RAID type from the previous worksheet; for example, RAID 5 or hot spare. For a hot spare (not strictly speaking a LUN at all), skip the rest of this LUN entry and continue to the next LUN entry (if any).

Caching. If you want to use caching (entry on page 5-16), you can specify the type of caching you want — read and write, read, or write for this LUN. Generally, write caching improves performance far more than read caching. Read caching is nonetheless crucial for good sequential read performance, as seen in backups and table scans. The ability to specify caching on a LUN basis provides additional flexibility, since you can use caching for only the units that will benefit from it. Read and write caching is recommended for any type RAID Group, particularly RAID 5.

Table 5-1 lists the amount of storage processor memory available for caching in systems running Navisphere 6.17.

Table 5-1 Available Cache Memory per Storage Processor

Storage System	Maximum Cache	Maximum Write Cache	Recommended Read Cache	Recommended Write Cache
CX700	3194 Mbytes	3072 Mbytes	250 Mbytes	2944 Mbytes
CX500	1497 Mbytes	1497 Mbytes	100 Mbytes	1397 Mbytes
CX500i	1497 Mbytes	1497 Mbytes	100 Mbytes	1397 Mbytes
CX300	561 Mbytes	561 Mbytes	100 Mbytes	461 Mbytes
CX300i	562 Mbytes	562 Mbytes	100 Mbytes	462 Mbytes

**Servers that can access this LUN's Storage Group**. For shared switched storage or shared-or-clustered direct storage, enter the name of each server (copied from the LUN and Storage Group worksheet). For unshared direct storage, this entry does not apply.

**Operating system information: Device name**. Enter the operating system device name, if this is important and if you know it. Depending on your operating system, you may not be able to complete this field now.

**File system, partition, or drive**. Write the name of the file system, partition, or drive letter you will create on this LUN. This is the same name you wrote on the application worksheet.

On the following line, write any pertinent notes; for example, the file system mount- or graft-point directory pathname (from the root directory). If any of this storage system's LUNs will be shared with another server, and the other server is the primary owner of this LUN, write **secondary**. (As mentioned earlier, if the storage system will be used by two servers, we suggest you complete one of these worksheets for each server.)

LUN Details Worksheet
Storage system (complete this section once for each storage system)  Storage-system number or name:ss1 Storage-system installation type  If Unshared Direct
SP information: SP A: IP address or hostname: _ss1spa_Port ALPA ID: _0 Memory(Mbytes): _1 GB_ SP B: IP address or hostname: _ss1spb_Port ALPA ID: _1 Memory(Mbytes): _1 GB_ Caching Read cache: 100 MB Write cache: 461 MB Cache page size (Kbytes): _8_
LUN ID:o SP owner: □A □B SP loop (0 or 1):_o
RAID Group ID:_o_Size,GB:_241.5_ LUN size,GB:_241.5_Disk IDs:_000,001,002,003,004  RAID type: ARAID 5
Servers that can access this LUN's Storage Group:server1
Operating system information: Device name:File system, partition, or drive:
LUN ID: _1 _ SP owner: □A □B SP loop (0 or 1):  RAID Group ID: _1 _ Size, GB: _266.4 LUN size, GB: _266.4 Disk IDs: 005,006,007,008,009  RAID type: ⑤ RAID 5 □ RAID 3 □ RAID 1 mirrored pair □ RAID 0
□RAID 1/0 □Individual disk □Hot spare
Caching: □Read and write □Write □Read □None
Servers that can access this LUN's Storage Group:server1
Operating system information: Device name:File system, partition, or drive:
LUN ID:2 SP owner: □A □B SP loop (0 or 1):1_
RAID Group ID: 2 Size, GB: 66.6 LUN size, GB: 66.6 Disk IDs: 0101, 1011, 1012, 1013, 1014
RAID type: □RAID 5 □RAID 3 □RAID 1 mirrored pair □RAID 0 □RAID 1/0 □Individual disk □Hot spare  Caching: □Read and write □Read □None
Servers that can access this LUN's Storage Group:server1
Operating system information: Device name:File system, partition, or drive:

### **Planning File Systems and LUNs**

### What Next?

If you have completed the worksheets to your satisfaction, you are ready to learn about the hardware needed for these systems as explained in Chapter 6.

For worksheets and a description of iSCSI configurations, go to Chapter 7.

# Storage-System Hardware

This chapter describes the storage-system hardware components.

# Major sections are

•	Storage-System Installation Types	6-2
	Hardware for CX300, CX300i, CX500, CX500i, and CX700	
	Storage Systems	6-3
•	Planning Your Hardware Components	
•	Hardware Dimensions and Requirements	6-9
•	Cabinets for Enclosures	6-10
•	Fibre Channel Data Cable and Configuration Guidelines	6-11
	Fibre Channel Front-End Hardware Planning Worksheets	

# **Storage-System Installation Types**

The storage systems attach to the server and the interconnect components as described in Chapter 1. The installation types are

- Unshared direct, with one server, is the simplest and least costly;
- Shared-or-clustered direct, with a limit of two servers, lets two servers share storage resources with high availability; and
- Shared switched, which has two switch fabrics, lets two or more servers share the resources of several storage systems in a storage area network (SAN).

Table 6-1 illustrates the storage-system installation types.

CX300, CX300i, CX500, CX500i, and CX700 storage systems require at least one network connection for management.

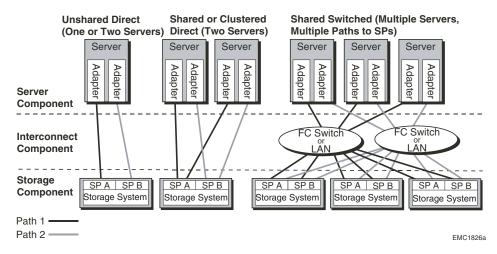


Figure 6-1 Types of Fibre Channel Storage-System Installations

# Hardware for CX300, CX300i, CX500i, CX500i, and CX700 Storage Systems

The main hardware component for a CX300, CX300i, CX500, or CX500i storage system is a disk processor enclosure (DPE2); for a CX700 storage system, it is a storage processor enclosure (SPE). Each storage system has two storage processors (SPs).

Table 1-1 on page 1-2 outlines each storage-system enclosure and disk capacity.

### **Storage Hardware**

A CX300, CX300i, CX500, and CX500i DPE2 or a CX700 SPE enclosure is a sheet-metal housing with a front door, midplane, and slots for the two storage processors (SPs), two power supplies, and fan modules. A CX300, CX300i, CX500, or CX500i DPE2 also has space for 15 disks. All components are replaceable units (FRUs) that can be replaced under power.

A separate standby power supply (SPS) is required to support write caching. Two are required for the CX500 and CX700; only one SPS is required for a CX300/CX300i, but it is more highly available with two.

Figure 6-2 shows SPE, DAE2s, SPSs, and cabinet components of a CX700 storage system. DAE2Ps and components of CX300-Series and CX500-Series storage systems are similar.

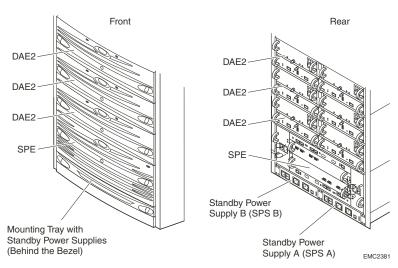


Figure 6-2 CX700 Storage System

### **Storage Processor (SP)**

The SP provides the intelligence of the storage system. Using its own proprietary software (called FLARE<sup>TM</sup> Operating Environment), the SP processes the data written to or read from the disk modules, and monitors the modules themselves. A CX700 SP consists of a printed-circuit board with memory modules, a board that provides SAN features (called a personality board), and status lights.

Each SP has front-end ports to communicate with switches or servers and back-end ports to communicate with disks. A CX300/CX300i DPE2 SP has two front-end ports and one back-end port; a CX500/CX500i DPE2 SP has two front-end ports and two back-end ports; a CX700 SPE SP has four front-end ports and four back-end ports.

The two iSCSI front-end ports on CX300i and CX500i SPs communicate either through a gigabit Ethernet LAN to clients or directly to an iSCSI HBA or a gigabit network interface card (NIC) in a client system.

For high availability, a storage system comes with two SPs. The second SP provides a second route to a storage system and also lets the storage system use write caching (described in *Storage-System Caching* on page 6-6) for enhanced write performance.

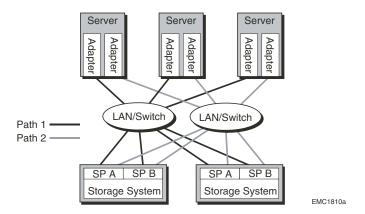


Figure 6-3 Shared Storage Systems

### **Storage-System Caching**

Storage-system caching improves read and write performance for several types of RAID Groups. Write caching, particularly, helps write performance — an inherent problem for RAID types that require writing to multiple disks. Read and write caching improve performance in two ways:

- For a read request If a read request seeks information that's already in the read or write cache, the storage system can deliver it immediately, much faster than a disk access can.
- ◆ For a write request the storage system writes updated information to SP write-cache memory instead of to disk, allowing the server to continue as if the write had actually completed. The write to disk from cache occurs later, at the most expedient time. If the request modifies information that's in the cache waiting to be written to disk, the storage system updates the information in the cache before writing it to disk; this requires just one disk access instead of two.

A standby power supply (SPS) protects data in the cache from power loss. If line power fails, the SPS provides power to let the storage system write cache contents to the vault disks. The vault disks are standard disk modules that store user data but have space reserved outside operating system control. When power returns, the storage system reads the cache information from the vault disks, and then writes it to the file systems on the disks. This design ensures that all write-cached information reaches its destination.

During normal operation, no I/O occurs with the vault; therefore, a disk's role as a vault disk has no effect on its performance.

# **SP Management Network Connection**

Each SP has an Ethernet connection through which the EMC Navisphere management software lets you configure and reconfigure the LUNs and Storage Groups in the storage system. Each SP connects to a network; this lets you reconfigure your system, if needed, should one SP fail.

### **SP iSCSI Network Connection**

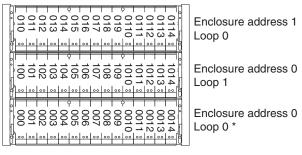
Each SP in an iSCSI storage system has iSCSI gigabit Ethernet connections through which the target and initiator communicate. Each iSCSI port connects to a gigabit Ethernet port on a network device such as a switch or router, or directly to an initiator iSCSI port on an HBA or NIC.

### Disks

The disks — available in differing capacities — fit into slots in the DAE enclosure. The CX-Series storage systems discussed in this manual support three types of DAEs:

- 2-gigabit point-to-point disk-array enclosure (DAE2P) with high-performance Fibre Channel disks.
- Original 2-gigabit disk-array enclosure (DAE2) with high-performance Fibre Channel disks
- ◆ DAE2-ATA version with economical ATA (Advanced Technology Attachment) disks.

Each disk has a unique ID that you use when binding or monitoring its operation. The ID is derived from the Fibre Channel loop number, enclosure address and disk slot in the enclosure. (Chapter 5 describes disk IDs in detail.)



<sup>\*</sup> Disks in DPE2 for CX300/CX300i or CX500/CX500i and in Fibre Channel DAE2/2PDAE for CX700

EMC2786

Figure 6-4 Disks and Their IDs

Disks with IDs 000 through 004 must be Fibre Channel disks. For a CX300/CX300i or a CX500/CX500i storage system, the slots for these disks are in the DPE2. For a CX700 storage system, the slots for these disks must be in a standard DAE2 or DAE2P and *not* a DAE2-ATA.

Enclosure 0 on loop 0 in a CX300, CX300i, CX500, CX500i, or CX700 storage system must contain disks with IDs 000 through 004. The remaining disk slots can be empty.

# **Planning Your Hardware Components**

This section helps you plan the hardware components — adapters, cables, and storage systems and site requirements — for each server in your installation.

For shared switched storage or shared-or-clustered direct storage, you must use high-availability options: two SPs per storage system and at least two HBAs per server. For shared switched storage, two switch fabrics are required.

For unshared direct storage, a server may have one or two HBAs.

### **Components for Storage Systems**

### **Components for Shared Switched Storage**

The minimum hardware configuration required for shared switched storage is two servers, each with two host bus adapters, two Fibre Channel switch fabrics with one switch per fabric, and two SPs per storage system. SPS units (standby power supplies) are also required for highest availability. You can use more servers (up to 15 are allowed), more switches per fabric, and more storage systems (up to four are allowed).

# Components for Shared-or-Clustered Direct Storage

The minimum hardware required for shared-or-clustered direct storage is two servers, each with two host bus adapters, and one storage system with two SPs. You can use more storage systems (up to four are allowed).

# Components for Unshared Direct Storage

The minimum hardware required for unshared direct storage is one server with two host bus adapters and one storage system with two SPs. You can choose more storage systems (up to four are allowed).

# **Hardware Dimensions and Requirements**

The DPE2, SPE and SPS units, DAE2s, and DAE2Ps require the following amounts of space and power in a standard 19-inch cabinet.

Table 6-1 CX300, CX300i, CX500, CX500i, and CX700 Hardware Dimensions and Requirements

Device	Dimensions	Vertical Size	Weight	Power <sup>a</sup>	VA Rating
DPE2	13.34 cm (5.25 in) x 45.01 cm (17.72 in) x 60.33 cm (23.75 in)	3 U (1.75 in per 1 U)	44.5 kg (98.3 lb) with SPs,15 disks, and mounting rails.	100 V-240 V ac At 100 V ac, 5.9 A. At 200 V ac, 2.9 A.	590 max (fully configured) <sup>b</sup>
SPE	17.78 cm (7.0 in) x 44.45 cm (17.5 in) x 67.11 cm (26.42 in)	4 U (1.75 in per 1 U)	32.9 kg (72.6 lb) with rails	100 V-240 V ac At 100 V ac, 5.2 A max At 200 V ac, 2.6 A max	520 max
DAE2 (FC)	13.34 cm (5.25 in) x 45.01 cm (17.72 in) x 60.33 cm (23.75 in)	3 U (1.75 in per 1 U)	43.2 kg (95.3 lb) with LCCs, 15 disks, and mounting rails	100 V-240 V ac At 100 V ac, 4.0 A. At 200 V ac, 1.6 A.	400 max (fully configured) <sup>b</sup>
DAE2-ATA				100 V-240 V ac At 100 V ac, 3.0 A. At 200 V ac, 1.4 A.	300 max (fully configured) <sup>a</sup>
DAE2P	13.34 cm (5.25 in) x 45.01 cm (17.72 in) x 35.6 cm (14 in)		30.91 kg (68 lb) with LCCs, 15 disks, and mounting rails	100 V-240 V ac At 100 V ac, 4.4 A. At 200 V ac, 2.2 A.	440 max (fully configured) <sup>a</sup>
SPS Tray	4.01 cm(1.58 in) x 20.96 cm (8.25 in) x 60.33 cm (23.75 in)	1 U (1.75 in per 1 U)	25.9 kg (57.15 lb) with tray (two SPS units at 25 lb each)	100 V-240 V ac At 100 V ac, 0.2 A max At 200 V ac, 0.1 A max	60 max

a. Platform-specific power ratings are published in the Hardware Reference manual for each storage system, and include dc power ratings where appropriate.

The total of the device VA ratings cannot exceed 4800 VA per cabinet.

b. A fully configured DPE2 or DAE includes 15 disk drives.

# **Cabinets for Enclosures**

19-inch-wide cabinets, pre-wired with ac power strips and ready for installation, are available in the following dimensions to accept storage-system enclosures.

Dimension	40U-C Cabinet	40U Cabinet*	
Height (internal, usable for devices)	40U (179 cm; 70 inches) from floor pan to cabinet top (fan installed)		
Height (overall)	190 cm (75 inches)		
Width (usable for devices)	NEMA 19-inch standard; rail holes 18.312 inches (45.78 cm) apart center-to-center		
Width (overall)	60 cm (24 inches)		
Depth (front to back rail)	60 cm (24 inches)		
Depth (overall)	98.425 cm (39.37 inches); 103.75 cm (41.5 inches) with optional front door)	91 cm (36 inches); 101 cm (40 inches) with optional front door	
Weight (empty)	390 lb (177 kg) maximum without front door 440 lb (200 kg) maximum with front door	164.7 kg (363 pounds)	
Maximum total device weight supported	945 kg (2100 pounds)		

<sup>\*</sup>The 40U-C cabinet replaced 40U models in March, 2004.

Both cabinets accept combinations of:

- ◆ SPEs at 4 U,
- ◆ DPE2s at 3 U
- DAE2/DAE2Ps at 3 U,
- SPS units at 1 U,
- Switches at 1 or 2 U

The cabinets require 200-240 volts ac at 50/60 Hz, and include two to four power strips with compatible outlets. Plug options are L6–30P and IEC 309 30 A.

Filler panels of various sizes are available.

# Fibre Channel Data Cable and Configuration Guidelines

CX300, CX500, and CX700 storage systems require optical cable between servers, switches, and SPs. The cabling between SPE or DPE2 and DAE2/DAE2P enclosures is copper.

EMC standard multimode optical configurations use 50 micron cable. EMC does not recommend mixing 62.5 micron and 50 micron optical cable in the same link. In certain situations, however, you can add a 50 micron adapter cable to the end of an already installed 62.5 micron cable plant. Contact your customer service representative for details.

Generally, you should minimize the number of cable connections, since each connection degrades the signal slightly and shortens the maximum distance the signal can carry.

When you plan and set up a storage system, we recommend that you assign each DAE enclosure to the next logically available back-end (BE) bus and address, and that you balance the number of enclosures on each bus. (Refer to your device-specific hardware reference for details on setting enclosure addresses.) For example, a system that supports two buses must have an enclosure 0 on bus 0; a second DAE should start a second bus as enclosure 0 on bus 1. If the same two-BE system has three enclosures on bus 0 and two enclosures on bus 1, add a new DAE as enclosure 2 on bus 1. Always optimize your system by using every available bus, and spreading the number of enclosures as evenly as possible across the buses.

Wherever possible, you should further optimize your system by configuring DAE2Ps on separate buses (loops) from DAE2s to achieve the maximum advantage of their point-to-point fault and data isolation.

Table 6-2 Fibre Channel Data Cable Types and Sizes

Length	Typical Use
1 m (3.3 ft) copper	In a single cabinet, connects the SPE or DPE2 or DAE2 to a DAE2.
5 m (16.5 ft), 10 m (33 ft) equalized copper	Cabinet to cabinet, connects the SPE or DPE2 or DAE2 to a DAE2.
3 m (10 ft) optical, LC-LC	In a single cabinet, connects the SPE or DPE2 to switches.
2 meter (6.6 feet), 5 meter (16.5 feet), 8 meter (26.4 feet) copper, HSSDC-HSSDC2	Connect SPE, DPE2, or DAE2 to DAE2P
2 meter (6.6 feet), 5 meter (16.5 feet), 8 meter (26.4 feet) copper, HSSDC2-HSSDC2	Connect DAE2P to DAE2P

### **Storage-System Hardware**

Length	Typical Use
10 m (33 ft) optical, LC-LC 30 m (100 ft) optical LC-LC 50 m (164 ft) optical, LC-LC 100 m (328 ft) optical, LC-LC 300 m (1000 ft, 0.2 mi) optical, LC-LC	In a room, connects the SPE or DPE2 to switches or HBAs. (Adapter must support optical cable.)
10 m (33 ft) optical, LC-SC 30 m (100 ft) optical LC-SC 50 m (164 ft) optical, LC-SC 100 m (328 ft) optical, LC-SC 300 m (1000 ft,.2 mi) optical, LC-SC	LC-SC cables are used on the EMC 1-Gbit Fibre Channel products. You can use LC-SC cables to connect an HBA to a switch, a switch to the SPE, or an HBA to the SPE.

The minimum bend radius for optical cables is 3 cm (1.2 in).

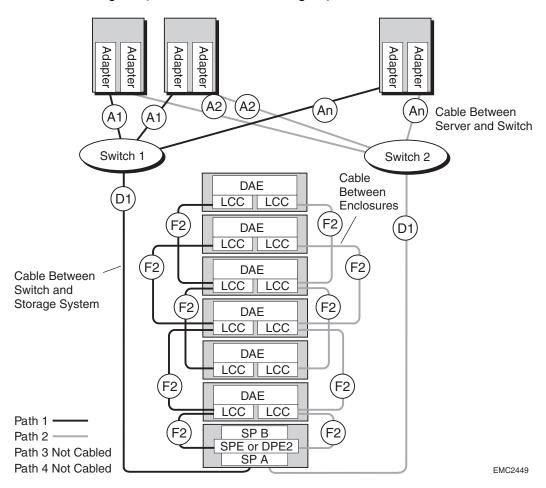
For the 50-micron cable operating at 2.125 Gbit/s the recommended maximum distance is 300 meters and at 1.0625 Gbit/s, the recommended maximum distance is 500 meters. For the 62.5-micron cable operating at 2.125 Gbit/s the recommended maximum distance is 150 meters and at 1.0625 Gbit/s, the recommended maximum distance is 300 meters. All of these maximum lengths assume no more than two connections/splices in the link.

Component planning diagrams and worksheets follow.

# Fibre Channel Front-End Hardware Planning Worksheets

Following are worksheets to note hardware components you want. Some installation types do not have switches or multiple servers.

### Cable Planning Template for a Shared Storage System

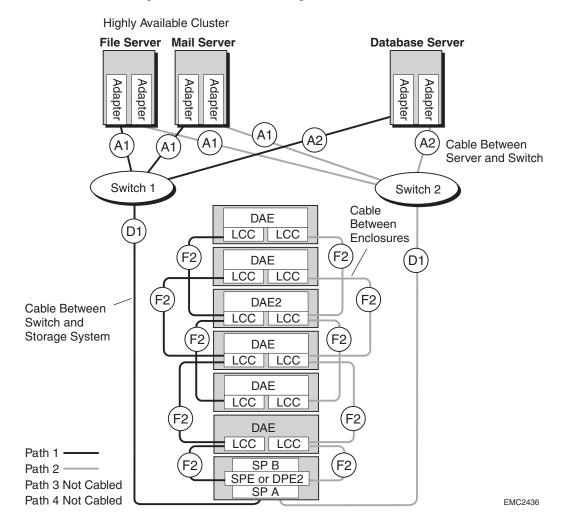


**Note:** The cable identifiers apply to all storage systems, and are described on the following worksheet.

# **Hardware Component Worksheet**

Number of servers: Adapters in servers: Switches: 32-port: 24-port	t:16-port:	8-port:
SPEs or DPE2s:Rackmount cabinets:		
DAEs:		
Cables between server and switch port- Cable A Cable A <sub>1,</sub> Optical: Number: Cable A <sub>2,</sub> Optical: Number: Cable A <sub>n</sub> ,Optical: Number:	Length	m or ft
Cables between switch ports and storage-system SP ports - Cable D Cable D <sub>1</sub> , CX300/500, up to 2/SP; CX700, up to 4/SP, Optical:Number: Cable D <sub>m</sub> CX300/500, up to 2/SP; CX700, up to 4/SP, Optical:Number:		
Cables between enclosures - Cable F Cable F <sub>2</sub> (Copper, 1.0 m): Number (2 per DAE):		

# Sample Cable Diagram for a Shared Storage System



# Sample Hardware Component Worksheet

Servers:3_ Adapters in servers:6_ Switches: 32-port: 24-port: 16-port:2_ 8-port:
SPEs or DPE2s: 1Rackmount cabinets:_1_
DAEs: <b>6</b> LCCs: <b>12</b>
Cables between server and switch - Cable A  Cable A <sub>1</sub> , Optical: Number:4
Cable D <sub>1</sub> CX300/500, up to 2/SP; CX700, up to 4/SP, Optical:Number:2 Length33m or ft Cable D <sub>m</sub> CX300/500, up to 2/SP; CX700, up to 4/SP, Optical: Number: Lengthm or ft
Cables between enclosures Cable F <sub>2</sub> (Copper, 1.0 m): Number (2 per DAE): <b>12</b> _

# **What Next?**

For worksheets and a description of iSCSI configurations, go to Chapter 7.

If you have completed the worksheets to your satisfaction, then you are ready to order some of this equipment.

# iSCSI Configuration

This chapter provides worksheets for configuring iSCSI initiator servers and iSCSI target storage systems in IP networks using the iSCSI protocol. It contains the following major sections:

•	iSCSI Configuration Options	7-2
	Storage System iSCSI Data Port Worksheet	
	iSCSI Initiator Port Worksheet	
•	iSNS Server Worksheet	7-5
٠	iSCSI CHAP Authentication Worksheets	7-6

CX300i and CX500i storage processors (targets) connect to iSCSI NICs or HBAs (initiators) through internet SCSI (iSCSI) interfaces. Initiators connect to targets either directly (initiator iSCSI port connected directly to target iSCSI port) or through a LAN. Each initiator contains either Network Interface Cards (NICs or iSCSI host bus adapters (HBAs), which are the physical iSCSI interfaces.

For the storage system, you will record each SP's iSCSI port network settings which include the IP address, subnet mask, and gateway address. For the initiators, you will record each initiator's iSCSI port network settings.

# iSCSI Configuration Options

CX-Series iSCSI storage systems support the following configuration options.

### **iSNS**

You can use the internet Storage Naming Service (iSNS) on Windows platforms that are part of an iSCSI network configuration. (Microsoft requires iSNS for WHCL certification.) iSNS requires an iSNS server on the storage network and a client storage system; other hosts connected to the storage system are also iSNS clients.

The iSNS server acts as the repository where hosts and storage devices register their lists of initiators and targets. Each individual initiator or target registers its components with the iSNS server; the server then answers the initiator's queries for devices with the list of registered targets. Discovery domains implemented on the server restrict what targets an initiator sees.

### **CHAP**

In addition, you can configure CHAP (Challenge Handshake Authentication Protocol) on the storage system. For CHAP authentication, you specify a username and password (called a secret) that any initiator must use to connect to the target storage system's iSCSI ports; the target storage system authenticates the initiators. You can also configure *mutual* CHAP authentication where initiators authenticate target iSCSI storage system(s).

For configuration rules, refer to the *EMC CLARiiON Open Systems Configuration Guide*, available on EMC Powerlink.

# Storage System iSCSI Data Port Worksheet

Prepare the following storage system iSCSI data port worksheet.

### Table 7-1 Storage System iSCSI Data Port Worksheet

iSCSI Front-End Data Ports - SP A:			
IP Address	Port 0:	Port 1:	
Subnet Mask	Port 0:	Port 1:	
Gateway	Port 0:	Port 1:	
iSCSI Front-End Data Ports - SP B:			
IP Address	Port 0:	Port 1:	
Subnet Mask	Port 0:	Port 1:	
Gateway	Port 0:	Port 1:	

#### IP Address of iSCSI Data Port

The static network IP (Internet Protocol) address (for example, 128.222.78.10) for communication with a 1-gigabit iSCSI data port of a storage processor (SP A or SP B). There are *two* 1-gigabit iSCSI data ports per SP.

#### **Subnet Mask for LAN**

The subnetwork mask for the LAN subnet to which the 1-gigabit iSCSI data port is connected.

## **Default Gateway for LAN**

The default gateway address for the LAN to which the 1-gigabit iSCSI data port is connected. The gateway specification is optional. Leave this field blank if your storage network is contained within a single subnet.

# iSCSI Initiator Port Worksheet

Use the following worksheet to record network information for as many as eight initiator ports on iSCSI HBAs or NICs. (Descriptions of each field are on the next page.)

Table 7-2 iSCSI Initiator Port Worksheet

Network configuration information for as many as eight iSCSI initiators:			
Server Name/Port Identifier	1:	2:	
IP Address			
Subnet Mask			
Gateway			
Server Name/Port Identifier	3:	4:	
IP Address			
Subnet Mask			
Gateway			
Server Name/Port Identifier	5:	6:	
IP Address			
Subnet Mask			
Gateway			
Server Name/Port Identifier	7:	8:	
IP Address			
Subnet Mask			
Gateway			

#### Port Identifier

Record an identifier for the port.

#### **IP Address for iSCSI Initiator Port**

The network IP (Internet Protocol) address of the initiator port.

#### **Subnet Mask for iSCSI Initiator Port**

The subnet mask for the LAN to which the initiator port is connected.

## **Default Gateway for iSCSI Initiator Port**

The default gateway address for the LAN to which the initiator port is connected. This parameter is not applicable with a direct connection.

# iSNS Server Worksheet

The storage system actively communicates with only one iSNS server at a time (the *primary* server). You can add additional servers to act as backup servers. If you elect to use iSNS on your storage network, enter the public IP address and name of your network's iSNS server. (Note that the public IP address is not related to the server's iSCSI port addresses.)

#### Table 7-3 iSNS Server IP Addresses

	IP Address/	Server Name
Primary iSNS Server:		
Secondary iSNS Server:		

# **iSCSI CHAP Authentication Worksheets**

If the storage system is on a private LAN, you can elect not to configure CHAP authentication. If the storage system is on a public LAN, we strongly recommend that you set CHAP security. If you do not set CHAP security for the storage system, any host connected to the LAN can read from and write to the storage system.



#### **CAUTION**

Whenever you record any security information such as the CHAP usernames and passwords, it is imperative that you store the data in a secure location.

#### What Is CHAP?

Challenge Handshake Authentication Protocol (CHAP) is an *optional* iSCSI authentication method where the target authenticates iSCSI initiators. CHAP consists of initiator CHAP and mutual CHAP, depending on which way the authentication occurs. For initiator CHAP, the target authenticates the initiator. Mutual CHAP can be configured *in addition to* initiator CHAP. For mutual CHAP, the initiator authenticates the target.

#### Initiator CHAP

When an initiator sends a request to an iSCSI target, it sends its initiator CHAP username and secret. The storage system compares the username and secret with a database of CHAP user accounts to authenticate the initiator. If you set up initiator CHAP authentication, you will enter the settings on the target and on all initiators. The initiator CHAP data is entered on the storage system and on initiators that will connect to the storage system.

#### Mutual CHAP

In addition to initiator CHAP, you can *optionally* set up the initiators to authenticate targets. This is mutual CHAP. For mutual CHAP, you create a target username and secret for the storage system. The storage system sends the username and password to the initiator and the initiator authenticates the storage system. The mutual CHAP data is entered on the storage system and on all initiators that will connect to the storage system.

If you are planning to set up *optional* CHAP authentication security on the storage system, prepare the following storage system CHAP security worksheets for initiator CHAP and mutual CHAP.

### **Initiator CHAP Worksheet**

If you want the storage system to authenticate initiators, fill out the Initiator CHAP Worksheet for all initiator accounts. Initiator CHAP must be set up and enabled for iSCSI security to work. The easiest way to configure initiator CHAP is to create a friendly username and secret for all initiators.

#### Table 7-4 Initiator CHAP Worksheet

Initiator name:		
Use Initiator Name as System CHAP User Name □		
Allow any Initiator to Login with this Username and Secret $\Box$		
CHAP Username		_
CHAP Secret:	Specified in hex 🖵	
Initiator name:		_
Use Initiator Name as System CHAP User Name □		
Allow any Initiator to Login with this Username and Secret $\Box$		
CHAP Username		_
CHAP Secret:	Specified in hex 🗆	
Initiator name:		_
Use Initiator Name as System CHAP User Name □		
Allow any Initiator to Login with this Username and Secret $\Box$		
CHAP Username		_
CHAP Secret:	Specified in hex 🖵	
Initiator name:		_
Use Initiator Name as System CHAP User Name □		
Allow any Initiator to Login with this Username and Secret $\Box$		
CHAP Username		_
CHAP Secret:	Specified in hex □	

#### Table 7-4 Initiator CHAP Worksheet (continued)

Initiator name:		
Use Initiator Name as System CHAP User Name □		-
Allow any Initiator to Login with this Username and Secret □		
•		
CHAP UsernameCHAP Secret:	Specified in hex □	
Initiator name:	<u>'</u>	
Use Initiator Name as System CHAP User Name □		-
Allow any Initiator to Login with this Username and Secret		
,		
CHAP Secret:	Specified in hex 🖵	
CHAP Secret:	·	
Initiator name:		-
Use Initiator Name as System CHAP User Name ☐		
Allow any Initiator to Login with this Username and Secret □		
CHAP Username		
CHAP Secret:	Specified in hex 🖵	
Initiator name:		_
Use Initiator Name as System CHAP User Name □		
Allow any Initiator to Login with this Username and Secret □		
CHAP Username		
CHAP Secret:	Specified in hex □	

### **Initiator Name**

The initiator name is a predefined name that usually begins with **iqn** (for iSCSI Qualified Name). You can determine the name by running the SANSurfer utility (shipped with QLogic HBAs) or the Microsoft iSCSI Initiator software for gigabit Ethernet NICs. The Microsoft iSCSI Initiator software is available on the Microsoft website.

### Use Initiator Name as System CHAP Username

If you want to use the initiator name as the CHAP username, check the **Use Initiator Name as System CHAP User Name** box. If you want to create a new username, leave the box blank.

#### **CHAP Username**

The CHAP username is a name that you assign for one user or for all users that connect to the storage system iSCSI port. You can use the same name as the initiator name or you can create a new name. If you use the same name as the initiator name, you are guaranteed to have a unique name. If you create a new friendly username, that name might not be unique unless you are on a private network and create a unique name for the network.

### **CHAP Secret (Password)**

For ASCII, the valid range is 12 to 16 printable characters. For hexadecimal (hex), the valid range is 24 to 32 hexadecimal characters, displayed in pairs. If you want to specify in hex, mark the box.

Allow Any Initiator to Login with this Username and Secret To allow any initiator to log in to the storage system, check the Allow any Initiator to Login with this Username and Secret box. If all initiators are trusted (for example, in a department or organization), you might want to check this for convenience. You can always grant or restrict initiator access to this storage system through the virtual disks.

#### **Mutual CHAP Worksheet**

Mutual (bidirectional) CHAP management provides an optional extra level of security that lets initiators verify the identity of the target storage system. If you want the initiators to authenticate targets, fill out the Mutual CHAP Worksheet for all mutual CHAP accounts.

#### Table 7-5 Mutual CHAP Worksheet

Target CHAP Username:		-
Target CHAP Secret (Passwor	d):	Specified in hex 🗆
<u>(1)</u>	2	y efforts will be nullified if you leave any leets in public places. Store them in a secure

10001 0 10 11		
iSCSI Configuration		

# Index

A	cache, storage-system 6-5
ac power requirements, storage system 6-10	size 5-16
Access Logix option 1-13	capacity
ALPA ID 5-16	disk 5-13
application	RAID Groups and LUNs 5-18
planning	cascading switches 1-5, 1-10
LUN 5-7	clone
MirrorView 4-17, 4-25	compared to snapshot 4-7
SAN Copy 4-33, 4-42	definition of
SnapView 4-9	ID 4-4
sample, for RAID groups 2-20	overview
array, see disk-array storage system	components, storage system 6-3
ATA (Advanced Technology Attachment), see	configurations
DAE2 (2-gigabit disk-array enclosure)	LUN and file system planning 5-6
ATA disks, configuration rules 5-13	RAID, planning 5-6
Till aloko, comigaration rates 5 15	see also installation types
<b>n</b>	Consistency Group
В	definition 4-12
backup	copy-on-first-write, definition of 4-6
benefit with SnapView 4-4	17
benefits	D
clones and snapshots 4-9	
SAN Copy 4-30, 4-38	DAE2 (2-gigabit disk-array enclosure)
bidirectional mirroring 4-15, 4-23	ATA 1-2, 6-7
business continuance volumes (BCVs), see clone	dimensions and weight 6-9
· · · · · ·	standard 1-2, 6-7
С	types of 1-2, 6-7
	data mobility application 4-30
cabinets for rackmount storage systems 6-10	see also SAN Copy
cabling	data sheets, hardware 6-10
guidelines	decision support
Fibre Channel 6-11	benefit with SnapView 4-4
types and sizes	device name, operating system 5-20
Fibre Channel 6-11	1 0,

disk	footprint
capacity 5-13	SPE 6-9
configuration rules 5-13	storage system 6-10
LUNs, planning 5-6	FRUs (replaceable units) 6-3
mirror, defined 2-2	
module IDs 5-4	G
RAID types	global spare, see hot spare
compared 2-14	giobai spare, see not spare
guidelines 2-18	••
selecting for RAID Groups 5-13	Н
shared storage systems, examples 5-2	hardware
speeds 5-13	data sheets 6-10
unit number on worksheet 5-9, 5-17	mirroring 2-2
disk-array storage system	planning 6-11
application and LUN planning worksheet	planning worksheets
5-7	Fibre Channel 6-13
cabinets 6-10	storage 6-3
components 6-3	trade-offs 6-8
dimensions 6-9	height, storage system 6-10
hardware requirements shared/unshared	high availability
6-8	hardware trade-offs 6-8
sample shared installation 5-2	shared switched installation 1-9
site requirements 6-9, 6-10	host, see server
SP 6-5	hot spare
unshared 5-5  DPE2 (Dick Processor Englosure 2 Chit)	defined 2-11
DPE2 (Disk Processor Enclosure, 2-Gbit)	proactive sparing 2-13
dimensions and weight 6-9	restrictions 5-14
statistics 6-9	sample applications 2-21
dual paths to LUNs 5-2	when to use 2-19
E	I
<del>-</del>	•
enclosure address (EA) 6-7	image (mirror) 2-2 defined 2-2
F	individual access array, see RAID 5 Group
fabric, switch, introduced 1-4	individual disk unit
Fibre Channel	defined 2-11
adapter 1-5	sample applications 2-21
components 1-5	when to use 2-19
defined 1-4	installation
switch, description 1-5	LUN and file system planning 5-6
see also switch	sample 5-2
file system	types 1-9
name 5-20	sample unshared direct 5-5
planning 5-6	sample, shared switched 5-2
1 0 -	integration with EMC SAN Copy software 4-16,
	4-23

integration with EMC SnapView Snapshot	RAID 1 mirrored pair 2-8
software 4-16, 4-23	RAID 1/0 Group 2-9
interconnect components, cables, hubs, switches	RAID types
1-5	compared 2-7, 2-14
IP address of SP 5-16	LUN (logical unit), Chapter 2
iSCSI 7-1	shared storage examples 5-2
components 1-9	unshared storage example 5-5
defined 1-9	worksheet 5-15
iSCSI CHAP authentication worksheets 7-6	
iSCSI initiator CHAP worksheet 7-7	M
iSCSI initiator port worksheet 7-4	
iSCSI mutual CHAP worksheet 7-9	management port worksheet 3-2
iSCSI port worksheet 7-3	memory, SP, specifying on worksheet 5-16
switch, description 1-10	mirrored pair, see RAID 1 mirrored pair
see also switch	mirrored RAID 0 Group, see RAID 1/0 Group
iSCSI CHAP authentication worksheets 7-6	mirroring
iSCSI disk-array storage system	defined 2-2
installation types 1-9	remote, see MirrorView software or remote
iSCSI initiator CHAP worksheet 7-7	mirroring
iSCSI initiator port worksheet 7-4	MirrorView
iSCSI interconnect components, cables, hubs,	overview 4-12
switches 1-10	planning worksheet 4-17
iSCSI mutual CHAP worksheet 7-9	MirrorView/A
iSCSI port worksheet 7-3	overview 4-19
iSCSI storage	planning worksheets 4-25
types	multipath I/O 1-12
defined 1-9	multiple paths to LUNs 5-2
iSCSl	
adapter 1-10	N
•	Navisphere 3-6
L	node, defined 1-4
LAN	node, LAN, defined 1-9
	nonredundant array, see RAID 0 Group
introduced 1-9	J
logical volume, see file system	0
LUN (logical unit)	
capacity 5-18	operating system
configurations, planning 5-6	device name for disk unit 5-20
disk mirror, defined 2-2	optical cable, types and sizes, Fibre Channel 6-11
hot spare 2-11	
in RAID Group 2-3	P
individual disk 2-11	page size, storage-system cache 5-16
number on worksheet 5-8, 5-17	parallel access array, see RAID 3 Group
paths to 5-2	paths to LUNs 5-2
planning file systems 5-6	performance, RAID Group 2-15
planning worksheet 5-7, 5-14	physical disk unit, see LUN (logical unit)
RAID 0 Group 2-10	r, s distribution and set 2011 (10 Steam distri

physical volume, see LUN (logical unit)	type guidelines 2-18
planning	types, compared 2-14
hardware 6-11	RAID Groups and LUNs 2-3
LUNs and file systems 5-6	read cache 6-6
see also worksheet	remote mirroring, promotion 4-14, 4-21
port, SP ALPA ID 5-16	replication over long distances 4-24
power requirements, storage system 6-9	revision testing, benefit with SnapView 4-4
primary image 4-12, 4-19	Ο' 1
proactive sparing	c
defined 2-13	\$
promotion	SAN Copy
of secondary image 4-14, 4-21	benefit 4-30, 4-38
	benefits 4-30, 4-38
D	example of 4-29, 4-37
R	planning worksheets 4-33, 4-42
rackmount storage system hardware components	simultaneous sessions 4-30
6-3	using with replication software 4-30
RAID 0 Group	secondary image
defined 2-10	promotion 4-14, 4-21
sample applications 2-21	server
when to use 2-19	cabling guidelines
RAID 1 mirrored pair	Fibre Channel 6-11
defined 2-8	connection to storage system 1-4
sample applications 2-20	see also cabling
when to use 2-19	shared storage systems
RAID 1/0 Group	application and LUN planning 5-7
defined 2-9	components 6-8
sample applications 2-21	examples 5-2
when to use 2-19	hardware planning worksheets 6-13
RAID 3 Group	requirements 6-8
defined 2-7	shared-or-clustered direct installation type,
sample applications 2-20	defined 1-9
when to use 2-18	shared switched installation type
RAID 5 Group	defined 1-9
defined 2-5	sample 5-2
sample applications 2-20	site requirements, storage systems 6-9, 6-10
when to use 2-18	size
RAID 6 Group	RAID group 5-18
defined 2-4	storage-system cache 5-16
sample applications 2-20	storage-system components 6-9
RAID Group	snapshot
application guidelines 2-20	compared to clone 4-7
capacity 5-18	overview 4-6
performance trade-offs 2-15	SnapView
planning worksheet 5-6	planning worksheet 4-9
sample applications 2-20	software mirroring 2-2
shared storage examples 5-2	O .

SP (storage processor)	unshared storage systems
about 6-5	application and LUN planning 5-7
IP address 5-16	components 6-8
port ALPA ID 5-16	defined 1-9
storage-system cache size 5-16	hardware planning worksheets 6-13
SPE (Storage Processor Enclosure)	requirements 6-8
components 6-3	sample 5-5
dimensions and weight 6-9	
speed, disk 5-13	V
storage	•
system components 6-3	vault disks 2-12
types	volume name 5-20
see also disk-array storage system,	
installation types	W
storage system, see disk-array storage system	weight, storage system 6-9, 6-10
stripe	worksheet
with RAID 1/0, RAID 0 2-9	hardware
with RAID 5 2-6	Fibre Channel 6-13
switch	iSCSI CHAP authentication worksheets 7-6
cascading 1-5, 1-10	iSCSI initiator CHAP worksheet 7-7
description 1-5, 1-10	iSCSI initiator port worksheet 7-4
fabric, introduced 1-4	iSCSI mutual CHAP worksheet 7-9
introduced 1-4	iSCSI port worksheet 7-3
sample shared storage installation 5-2	LUN 5-7, 5-14
zone 1-6	management port worksheet 3-2
	MirrorView 4-17
T	MirrorView/A 4-25
trade-offs	SAN Copy 4-33, 4-42
between snapshots and clones 4-7	SnapView 4-9
-	write cache 6-6
hardware, shared storage 6-8	
RAID types 2-14	Z
U	zone, switch 1-6
unshared direct installation type, sample 5-5	

Index	