# hp StorageWorks multi-site disaster tolerant solution

# implementation blueprint

# installation and configuration guide

## Notice



Edition 0503



## Format conventions

**Note** This is a Note.

**Caution** This is a Caution.

**Warning** This is a Warning.

| | |
|---|---|
| **`User Input`** | Specifies text to be typed exactly as shown, such as commands, path names, file names, and directory names. |
| *`variable`* | Indicates that you must supply a value. |
| `Screen text` | Denotes text displayed on the screen. |

# Contents

# Figures

# Introduction

This *Installation and Configuration Guide* describes one possible process for implementing the Multi-Site Disaster Tolerant (MSDT) Solution according to the solution's high-availability model configuration. The *MSDT Implementation Blueprint: Design Guide* describes this model configuration in detail; for reference, a summary is provided on page 10.

This guide presents the implementation process as a set of numbered tasks, and provides a suggested order for completing those tasks. Each numbered chapter provides supporting information, recommendations (where applicable), and instructions for completing a task. Each task includes a combination of design, installation, and configuration procedures, including:

- Creating low-level design diagrams, such as host-to-storage connection, CU/RCU mapping, and zone mapping diagrams
- Installing hardware and software components
- Configuring components specifically for the MSDT solution
- Testing the solution

| | |
|---|---|
| **Caution** | Before you begin the implementation process, it is vital that you understand both the process as a whole and the individual process tasks. Read this entire guide before you begin the process. |

When you fully understand the implementation process, you may decide to perform specific tasks or procedures in a different order.

## Audience

This guide is designed to be used by HP Consulting and Integration Engineers, HP Account Support Engineers, and HP Solution Architects.

## Required skills and knowledge

This guide assumes that the HP representative implementing the solution has an in-depth understanding of:

- Systems and network design and administration, including device mapping, device access, storage area networks (SANs), switches, DWDM, Fibre Channel (FC), etc.
- The MSDT Solution (detailed in the *MSDT Technical Blueprint* and *MSDT Implementation Blueprint: Design Guide*)
- Design and configuration of HP-UX servers and operating system
- Design and configuration of HP StorageWorks XP disk arrays
- HP StorageWorks RAID Manager XP (RAID Manager XP)
- HP StorageWorks Business Copy (BC)
- HP StorageWorks Continuous Access (CA) and Continuous Access Extension (CA Ext)
- HP cluster solutions, including HP ServiceGuard (SG), MetroCluster, Continental Clusters, and the XP-CA toolkit for MetroCluster
- HP StorageWorks LUN Configuration and Security Manager XP, HP StorageWorks LUN Configuration Manager XP, and HP StorageWorks Secure Manager XP

- HP Multi-Site DT Management Tools (detailed in the *MSDT Implementation Blueprint: Multi-Site DT Management Tools User Guide*)

# Additional resources

For more information on the MSDT Solution and components, refer to the following documents:

- *Multi-Site Disaster Tolerant Solution Business Blueprint*
- *Multi-Site Disaster Tolerant Solution Technical Blueprint*
- *Multi-Site Disaster Tolerant Solution Implementation Blueprint: Design Guide*
- *Designing Disaster Tolerant High Availability Clusters*: B7660-90013
- *Managing MC/ServiceGuard*: B3936-90065
- *HP ServiceGuard Quorum Server Version A.02.00 Release Notes*: B8467-90011
- *MetroCluster with Continuous Access XP Version A.04.20 Release Notes*: B8109-90013
- *Other applicable HP StorageWorks XP product documentation*

# Patch catalog and latest versions information pointer

http://wtec.cup.hp.com/~patches/catalog/

# Document assumptions

The terms “host” and “node” are used interchangeably.

The terms “application” and “package” are used interchangeably.

In the MSDT Solution, there are four device groups.

**device groups**

| | |
|---|---|
| CA_1 | All devices in the synchronous replication between Site 1 and Site 2. |
| BC_1 | All the devices used in the BC point-in-time copy created on Site 2. |
| CA_2 | All the devices used in the long-distance replication copy from Site 2 to Site 3. |
| BC_2 | All devices used in the creation of a point-in-time BC on Site 3. |

If the customer decides not to have a BC at Site 3, you would use only three device groups.

# Solution overview

In the MSDT Solution, two nearby sites—less than 100 km[1] apart—protect each other in case of an in-region disaster. Critical applications are mirrored synchronously, which provides high availability for online transaction processing. If a failure occurs at one site, the other site can take over application processing, with minimal interruption, at virtually the exact point where it was interrupted. A third site (Site 3) located well outside of the region offers protection should both primary sites (Sites 1 and 2) be lost. A point-in-time copy is made at Site 2 and then mirrored to Site 3 at customer-defined intervals. Copying is a scheduled, scripted process enabled by the MSDT Management Tools. The recovery point from Site 3 would be determined by the last time a copy cycle was completed.

The MSDT Solution solves the following business issues:

- Application continuance in case of a local or regional disaster while maximizing transaction-processing availability. Mirroring of Site 2 to Site 3 using a point-in-time copy allows placement of Site 3 out of the disaster region.
- High availability for 24 x 7 operations.
- ServiceGuard with MetroCluster and Continental Clusters provide a recovery time objective (RTO) of less than one hour. Synchronous capability between Sites 1 and 2 allows for recovery point objective (RPO) of virtually zero.
- Customizable management scripts allow specific integration with customer environment and applications.

[1] 100 km is the current tested limit although recent technology may allow for greater distances.

# Model configuration

In the MSDT Solution model configuration, data is distributed over three sites, utilizing one or more hosts per site.

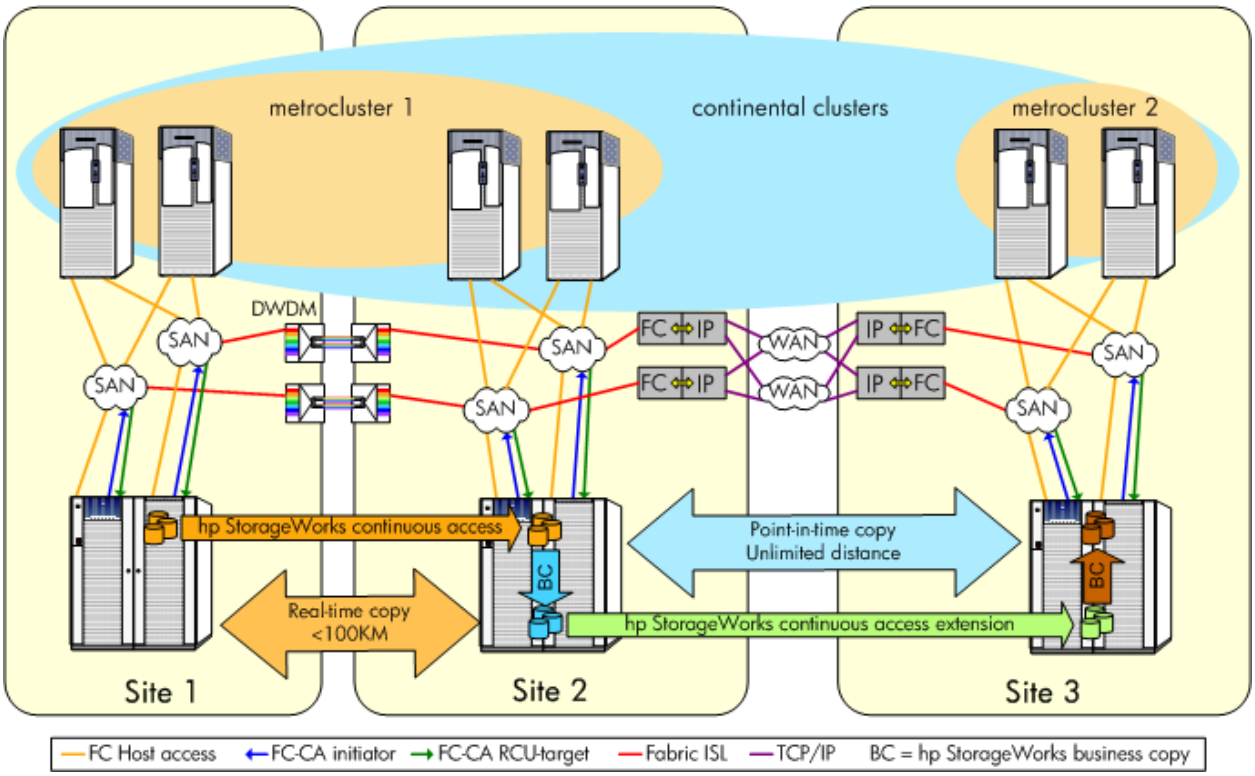



**Figure 1: Model configuration physical view**

The application is normally operational at Site 1. Site 2 is a fully operational data center with application hosts available. The hosts at Site 2 can be used for development work or any other application, including being a client for the application at Site 1. By means of a fully automated ServiceGuard, MetroCluster, and sync CA implementation, the application can automatically fail over between Sites 1 and 2 without data loss. The cluster configuration between Sites 1 and 2 is a full two-site MetroCluster implementation, managing both the application failover and the CA device pair.

Application data is synchronously replicated to Site 2, where the data is stored and forwarded in a constant cycle to Site 3. This process can continue uninterrupted when the application fails over from Site 1 to Site 2. The single- or multiple-node MetroCluster implemented at Site 3 allows startup and failover of the application to Site 3 after a manual decision was made to start the application. The MetroCluster software performs the takeover on the CA Ext devices during startup of the application. The Continental Clusters ensure that the application will only be active at one site at any time, and it provide a push-button failover option for the application. The BC copy at Site 3 ensures that there is always a valid, usable copy of the application data at Site 3.

# Implementation Process Overview

As stated in the "Introduction," this guide describes one possible process for implementing the MSDT Solution. Figure X provides a "big-picture" view of this process.

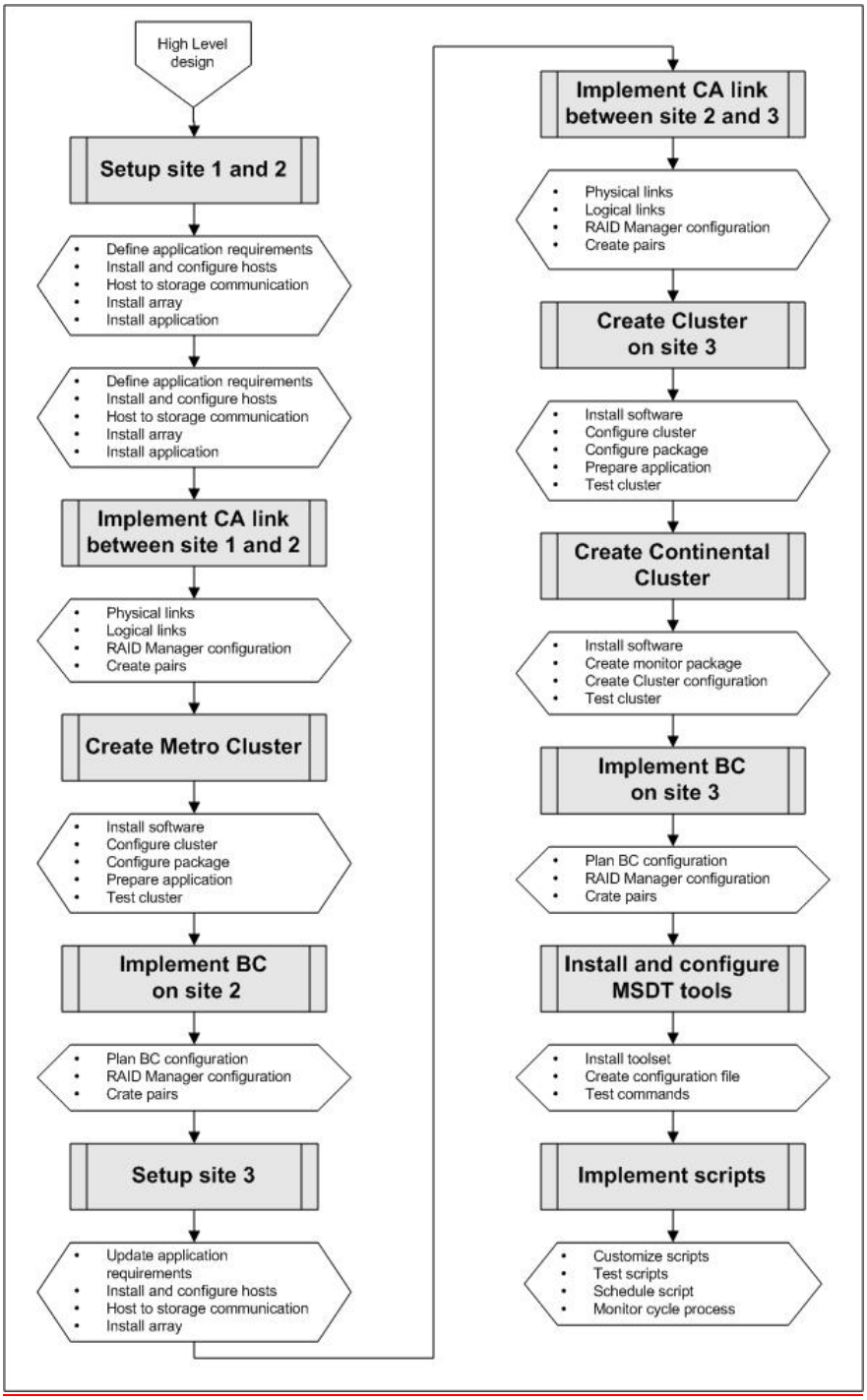



**Figure 2: Implementation process**

The goal of the process is to produce a customer-specific configuration diagram similar to Figure 3, and then implement that configuration.

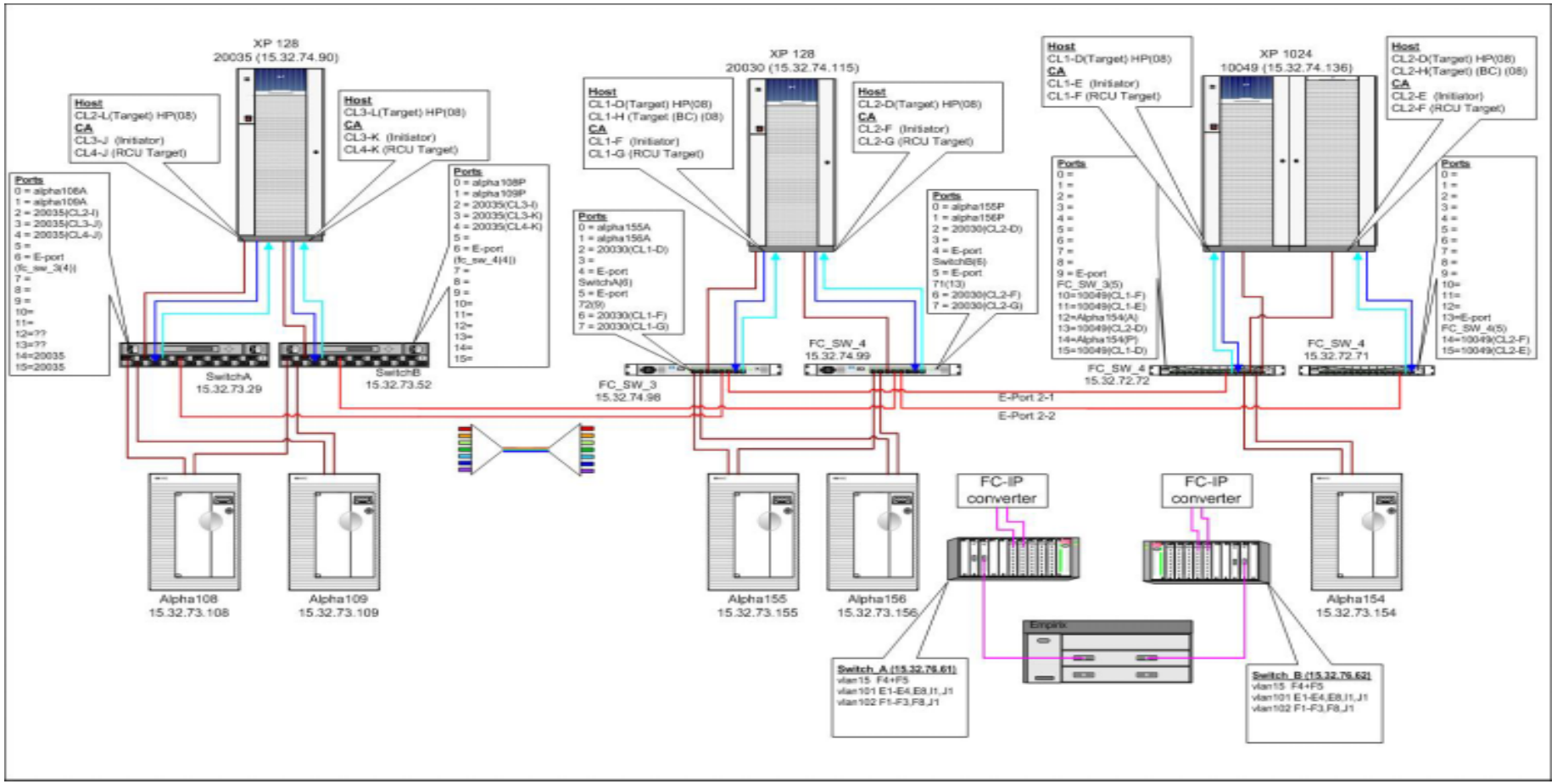



Figure 3: Detailed configuration diagram

# Before you start

Before starting the implementation process, you need the following deliverables from the high-level design (HLD) process discussed in the *MSDT Implementation Blueprint: Design Guide.*

**HLD deliverables**

| | |
|---|---|
| **Physical diagrams** | Basic physical diagrams of the customer-specific solution. Physical diagrams show:<br>• The general interconnections of the solution, including cable routes<br>• The various types of hardware used for this solution, including switches and converters<br>• Redundant local clusters and XP enclosures supporting CA and BC<br>• High-level software layers such as MetroCluster, Continental Clusters, BC, CA, and CA Ext |
| **Bills of materials (BOMs)** | Complete lists of components to be used in the solution, including suitable existing components and new components at each site |

The customer-specific solution documented in the HLD deliverables may differ from the solution's model configuration (page 10) in various ways. For example, a customer may not want redundant DWDM or WAN links, or may have only one host at Site 3. If the customer's configuration differs from the model configuration, you must adapt the information in this guide accordingly.

You must also complete the following tasks:

- HP Site Preparation personnel have assessed and prepared the physical sites, including confirming or establishing the FC connection between Sites 1 and 2 and the WAN connection between Sites 2 and 3
- The components identified in the BOMs for each site have been ordered and delivered to the appropriate site

Using the HLD BOMs, work with the customer to determine where the hardware will be located at each site, and then define cable lengths and racking requirements. Note that it is the customer's responsibility to meet the appropriate power supply and cooling requirements for each hardware component, according to the hardware product documentation.

# During implementation

During the implementation process, you will need:

- Product documentation for all solution components
- The *Multi-Site Disaster Tolerant Solution Implementation Blueprint: Administration Guide*
- Access to customer representatives who are familiar with the customer's current network configuration and disaster recovery business processes and procedures

# Network considerations

Three types of networks must be considered during various stages of the implantation process. These networks will not be covered in detail in this document, except for the private network required for CA. Any specific requirements for either the cluster solutions or end users access to applications must be addressed by the customer's network services.

## User access network

This is the network that application users use to access application information. It is part of the shared public network and must provide access for a number of users over a wide geographical area. Redundancy is the main factor to consider to ensure uninterrupted access to the application. Because MetroCluster utilizes the

same subnet across Sites 1 and 2, and utilizes the floating IP address concept in ServiceGuard, the application IP will remain the same no matter where in the MetroCluster configuration the application is running. The user will have access to the application as long as the external network infrastructure allows this.

## Heartbeat network

This network is for private heartbeat communication between the nodes in the cluster and is vital to the health of the cluster. Heartbeat networks should be dedicated to this purpose to prevent unnecessary timeouts due to large transfers by other applications. The user network interface can be utilized as a standby heartbeat network.

## Private IP network

The most important network for the MSDT Solution is the private network for CA over IP implementation between Site 2 and Site 3. This network should be dedicated for CA traffic only, with specific guaranteed service level agreements (SLA). The latency and packet loss characteristics of this network will determine the performance of the CA  copy, which impacts the time needed to cycle the data. This network should be re-evaluated on a regular basis to ensure that it scales with the growing demands of the application over time.

# 1 Set Up Sites 1 and 2

Use the information in this chapter to complete preliminary hardware and software setup at Sites 1 and 2 before connecting the sites. Many components may already be installed at the customer sites. If this is an extension of a previously installed and operational application environment, it is still recommended to work through all the steps to ensure all details of current the solution are recorded and taken in consideration for the entire solution.

## 1: Set up site 1

Perform the following tasks to complete the preliminary setup of Site 1. The objective of this step is to obtain a fully functional application environment and to document all physical device requirements for the application.

### 1a: Define application requirements

Evaluate the production application and physical requirements for the application, in particular the storage requirements. Most applications require a number of different storage areas with particular performance requirements. Mapping the application into a high-level "volume group" design enables you to identify every area and the specific requirements for that part of the application. The volume group is then broken into logical volumes each with specific needs. Striping filesystem and the number of devices in a file system, for example, play an important role. Lastly, the volume groups are also mapped to physical devices each with a specific size and emulation in mind. If this is a new installation, the designer could not map physical devices to logical devices within the array, and this would only be possible after the array is configured and the devices assigned to the host.

After the array is installed and configured and the host-to-array connections are in place, you must return to this section and complete the application mapping. If the application is already installed and running on this site, it is possible to immediately complete the mapping with all the logical device information. At this point, it is necessary to map the host devices to logical devices within the array, in the format of an array internal CU:LDEV number. The Control Unit (CU) number indicates the logical control table within the array in which this logical device is configured. One control table (CU) can manage 255 devices. The CU number is important when creating the logical paths (RCU) between the arrays, as you will need to create a logical path from every CU the application is using in the local array to a corresponding CU in the remote array.

A device mapping similar to the following example must be created for the device on Site 1.

| Application | Volume group | Logical Volume | Physical device | Size (GB) | Emulation | Site 1 Array 20035 CA_1 CU | Ldev |
|---|---|---|---|---|---|---|---|
| | | | | | | CL1-A and CI2-A | |
| Oracle1 | vg_archive | lv_arch1 | /dev/dsk/c0t1d0 | 14 | Open-E | 00 | 01 |
| | | | /dev/dsk/c0t1d1 | 14 | Open-E | 00 | 04 |
| | | lv_arch2 | /dev/dsk/c0t1d2 | 14 | Open-E | 00 | 08 |
| | | | /dev/dsk/c0t1d3 | 14 | Open-E | 00 | 12 |
| | vg_data1 | lv_data1 | /dev/dsk/c0t1d4 | 9 | Open-9 | 01 | 10 |
| | | | /dev/dsk/c0t1d5 | 9 | Open-9 | 01 | 12 |
| | vg_data2 | lv_data2 | /dev/dsk/c0t1d6 | 9 | Open-9 | 01 | 14 |
| | | | /dev/dsk/c0t1d7 | 9 | Open-9 | 01 | 18 |
| | vg_data3 | lv_data3 | /dev/dsk/c0t2d0 | 9 | Open-9 | 01 | 16 |
| | | | /dev/dsk/c0t2d1 | 9 | Open-9 | 01 | 18 |
| | vg_data4 | lv_data4 | /dev/dsk/c0t2d2 | 9 | Open-9 | 01 | 20 |
| | | | /dev/dsk/c0t2d3 | 9 | Open-9 | 01 | 22 |
| | vg_index | lv_index | /dev/dsk/c0t2d4 | 14 | Open-E | 00 | 30 |
| | | lv_index2 | /dev/dsk/c0t2d5 | 14 | Open-E | 00 | 31 |
| | vg_logA | lv_log1 | /dev/dsk/c0t2d6 | 3 | Open-3 | 02 | 32 |
| | | lv_log2 | /dev/dsk/c0t2d7 | 3 | Open-3 | 02 | 33 |
| | | lv_log3 | /dev/dsk/c0t3d0 | 3 | Open-3 | 02 | 34 |
| | | lv_log4 | /dev/dsk/c0t3d1 | 3 | Open-3 | 02 | 35 |
| | vg_log_b | lv_log1 | /dev/dsk/c0t3d2 | 3 | Open-3 | 02 | 62 |
| | | lv_log2 | /dev/dsk/c0t3d3 | 3 | Open-3 | 02 | 63 |
| | | lv_log3 | /dev/dsk/c0t3d4 | 3 | Open-3 | 02 | 64 |
| | | lv_log4 | /dev/dsk/c0t3d5 | 3 | Open-3 | 02 | 65 |
| | vg_admin | lv_error | /dev/dsk/c0t3d6 | 9 | Open-9 | 00 | 70 |
| | | lv_config | /dev/dsk/c0t3d7 | 9 | Open-9 | 00 | 71 |
| Command device (alpaha108) | | | | | | 00 | 9A |
| Command device (alpaha108) | | | | | | 01 | 8C |
| Command device (alpaha109) | | | | | | 00 | 9B |
| Command device (alpaha109) | | | | | | 01 | 8D |

**Table 1: Application to device mapping - Site 1**

From this table you can determine what size (emulation) devices are required for the application. This information is important for the configuration and setup of the array.

## 1b: Install and configure hosts

The process of designing the physical layout of the application hosts is a multi-step process that must be completed with input from the application design, array configuration, and array-to-host connection data available.

1. Design Server configuration

   In this step, the designer defines the exact layout of all the hardware required in the system. Special attention is required to ensure all system has equivalent hardware, and it is recommended to ensure that all hardware interfaces are configured similarly in all systems. Using this approach the configuration of the cluster software and later maintenance of the hosts systems is significantly simplified.

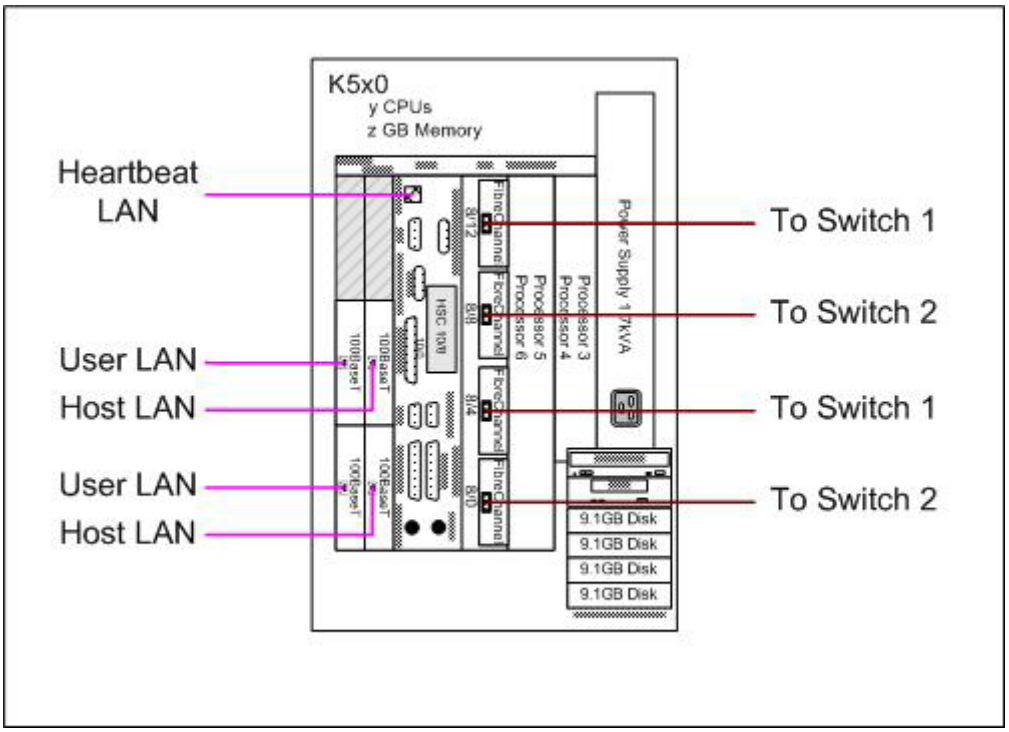



**Figure 4: Host configuration layout**

**Note:** Figure 4 is an example of a K-Class system layout diagram, and is not necessarily an exact replica of a system used in the solution.

Define the host name and IP addresses for each network interface. Identify a network interface as a primary LAN interface or a standby interface for cluster usages. Identify all heartbeat networks interfaces. Create a detailed diagram with all network connections and associated network switches. The following is an example of such a table.

| Network Card Layout For Host | | | | | |
|---|---|---|---|---|---|
| **Device** | **HW Path** | **Device File** | **IP Address** | **Sub net mask** | **Usage** |
| 100BaseT | 10/16/4 | /dev/lan1 | 135.33.52.5 | 255.255.248.0 | User accesses LAN |
| 100BaseT | 10/4/4 | /dev/lan2 | 140.55.24.5 | 255.255.248.0 | Hosts LAN (heartbeat) |
| 100BaseT | 10/16/8 | /dev/lan3 | Standby for 135 net | | Standby User LAN |
| 100BaseT | 10/4/8 | /dev/lan4 | Standby for 140 net | | Standby Host LAN |
| Onboard LAN | 10/12/6 | /dev/lan0 | 10.1.8.5 | 255.255.255.0 | Dedicated Heartbeat LAN |

**Table 2: Network map - Host 1 Site 1**

Design the operating system layout, and identify devices that will be the boot devices for the host. Ensure that this device is a high-availability device with some sort of hardware or software protection enabled. It is possible to boot from the XP array using devices with RAID protection or using stand-alone devices with software mirroring (HP MirrorDisk UX) protection. Design the operating system layout in such a way that future expansion is possible without any re-configurations.

| HP-UX Boot Disk layout | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Device** | **HW Path** | **Device file** | **Primary/ mirror** | **Volume group** | **Logical vol** | **Mount point** | **Size(MB)** |
| 9.1G Jbod | 10/0.6.0 | /dev/dsk/c6t6d0 | Primary | /dev/vg00 | lvol1 | /stand | 800 |
| | | | | | lvol2 | swap | 2000 |
| 9.1G Jbod | 10/0.5.0 | /dev/dsk/c6t6d0 | Primary | | lvol3 | / | 800 |
| | | | | | lvol4 | /home | 500 |
| 9.1G Jbod | 8/0.6.0 | /dev/dsk/c6t6d0 | Mirror | | lvol5 | /opt | 2000 |
| | | | | | lvol6 | /tmp | 2000 |
| 9.1G Jbod | 8/0.5.0 | /dev/dsk/c6t6d0 | Mirror | | lvol7 | /usr | 2000 |
| | | | | | lvol8 | /var | 3000 |

**Table 3: Operating system filesystem mapping**

**Note:** The default filesystem sizes used during the OS install are typically too small for productions systems to run without re-sizing the filesystem. Plan ahead and allow for application software and patch installations space. Table 3 shows suggested values used on the test system. Adjust them as required.

1. Identify a list of software components that must be installed on the system as well as all associated patches that might be required for every software package. The following is an example of the software packages installed on the test systems.

```
# swlist -l bundle
# Initializing...
# Contacting target "alpha108"...
# Target: alpha108:/
#
 HPUXBase64          B.11.11          HP-UX 64-bit Base OS
 HPUXBaseAux         B.11.11.0212         HP-UX Base OS Auxiliary
 BUNDLE11i           B.11.11.0102.2       Required Patch Bundle for HP-UX 11i, February 2001
 FEATURE11-11            B.11.11.0209.5       Feature Enablement Patches for HP-UX 11i, Sept 2002
 HWEnable11i     B.11.11.0212.4       Hardware Enablement Patches for HP-UX 11i, December 2002
 OnlineDiag          B.11.11.09.11        HPUX 11.11 Support Tools Bundle, Dec 2002
 B2491BA             B.11.11          MirrorDisk/UX
 B3935DA             A.11.14          MC / Service Guard
 B8109BA             A.04.20          MetroCluster with Continuous Access XP Toolkit
 T2346BA             A.04.00          HP Continental Clusters
 B5736DA             A.03.20.01           HA Monitors
 B7609BA             A.03.20.01           Event Monitoring Service
 B3835DA             C.02.00.02           HP Process Resource Manager
 B8324BA             B.01.04          HP Cluster Object Manager
 HPOVSAMHostAgt      03.00.00.0307        hp OpenView storage area manager hostagent
```

2. Install servers.

   Follow the hardware installation instructions to install all hardware in each server. Perform self-test and ensure that each hardware component passes the initial self-test. Ensure that all environmental and power requirements are met for each server.

3. Install the HP-UX operating system.

   At this point, the operating system can be installed. If it is required to boot the server from the array, then this step must wait until the host-to-storage connections are completed. Install the operating system according to the procedures documented in the operating installation manual. After the basic operating system installation, configure all network interfaces and set initial values for host. Install any additional software products in the installation sequence documented in each products' installation instructions. After all software installations are completed, install all recommended patches and latest patch bundles for the OS and application.

## 1c: Design and install host-to-storage communication

Host-to-storage connections are achieved by Fibre C-channel cabling from the host bus adapter (HBA) to the storage system port. In some cases, the host can be directly connected to the storage system, but in most large installations the use of switches between the host and array is recommended to ensure better connection rates to the storage system and to have better control for multiple systems to the array. For the latest supported switch information, reference the STREAMS documents for SAN configurations. To ensure high availability, HP recommends having two separate paths for every logical device. Route the paths in two separate routes from the server to the switch and from the switch to the storage. HP recommends that the customer not have the switches racked in the same cabinet. If you cannot avoid this, then the switches should at least have separate power supplies (most switches support redundant power supplies).

Record the ports used on the array for host connections as well as the port connections on the switches. The array connections must be from two separate power boundaries or clusters (CL1 and CL2 in the diagram). This information will help to debug installation issues during later configuration and testing.

In the test configuration, we utilized two switches providing two independent paths to the array for each of the hosts on the site. In most installations, the number of switches and paths to the array will be determined by the application requirements and performance requirements. At the end of the design phase of this step, produce a detailed host-to-storage diagram that will help the installation specialist to correctly install and connect all cabling. The following is an example of such a diagram.

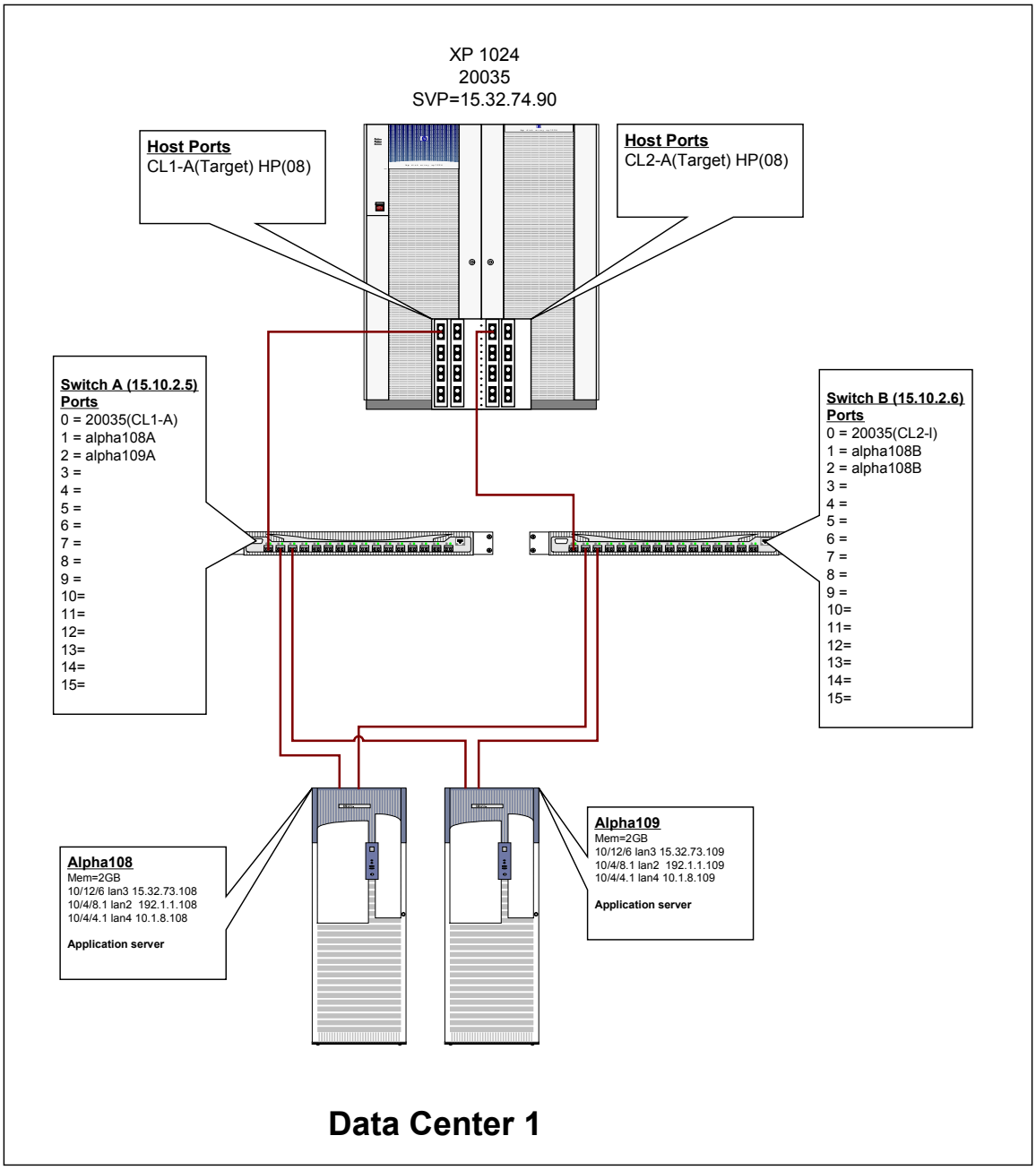



**Figure 5: Host-to-storage connections for Site 1**

## 1d: Install and configure arrays

If this is a new installation, it will be necessary to install and configure the array first before it is possible to access the devices and configure the application. If this is an existing installation, then the array is probably already installed and configured and the devices will be available to the host. In this case, this step is probably not necessary. The following steps must be performed.

1. Install hardware.

The first step in the process is to determine the number and type of devices and fiber interfaces required in the array. The number of physical devices will determine the configuration of the array. Physical devices can be installed in sets of four or eight and then formatted in one of the supported emulation modes. The logical devices created after the format process are then assigned to control units (CUs) in the array. Ensure that the emulation mode used (and therefore the size of the devices) is supported on the array type at each site.

Part of the installation process is to enable all software licenses that were purchased with the array. Most of the time these licenses are installed at the factory, but it still requires a step to ensure the necessary licenses are installed and enabled on the array. Ensure that the appropriate CA licenses are enabled on the array.

The products to install depends on the arrays used in the customer configuration.

| XP128 or XP1024 arrays | XP48 or XP512 arrays |
|---|---|
| HP StorageWorks LUN Configuration and Security Manager XP | HP StorageWorks LUN Configuration Manager XP<br>*and*<br>HP StorageWorks Secure Manager XP |

2. Format array.

   If this is a new installation, format the array. This process initializes all the logical devices and prepares them for host access. Allocate sufficient time for this process as it typically a lengthy process.

3. Port map.

   The next step in the preparation of the array is to allocate the logical devices to the host ports for host access. First set the host port to the correct mode setting for the host that will allocate the devices. In the case of HP-UX, the mode setting is 08. Ensure that the port is set as a “Target” port and not as an “Initiator” or RCU Target” port. Then allocate the devices to the port providing a Target and LUN number for each device. Ensure that each device is mapped to at least two ports for high availability. Remember to also map some command devices to the ports for RAID Manager access from the host.

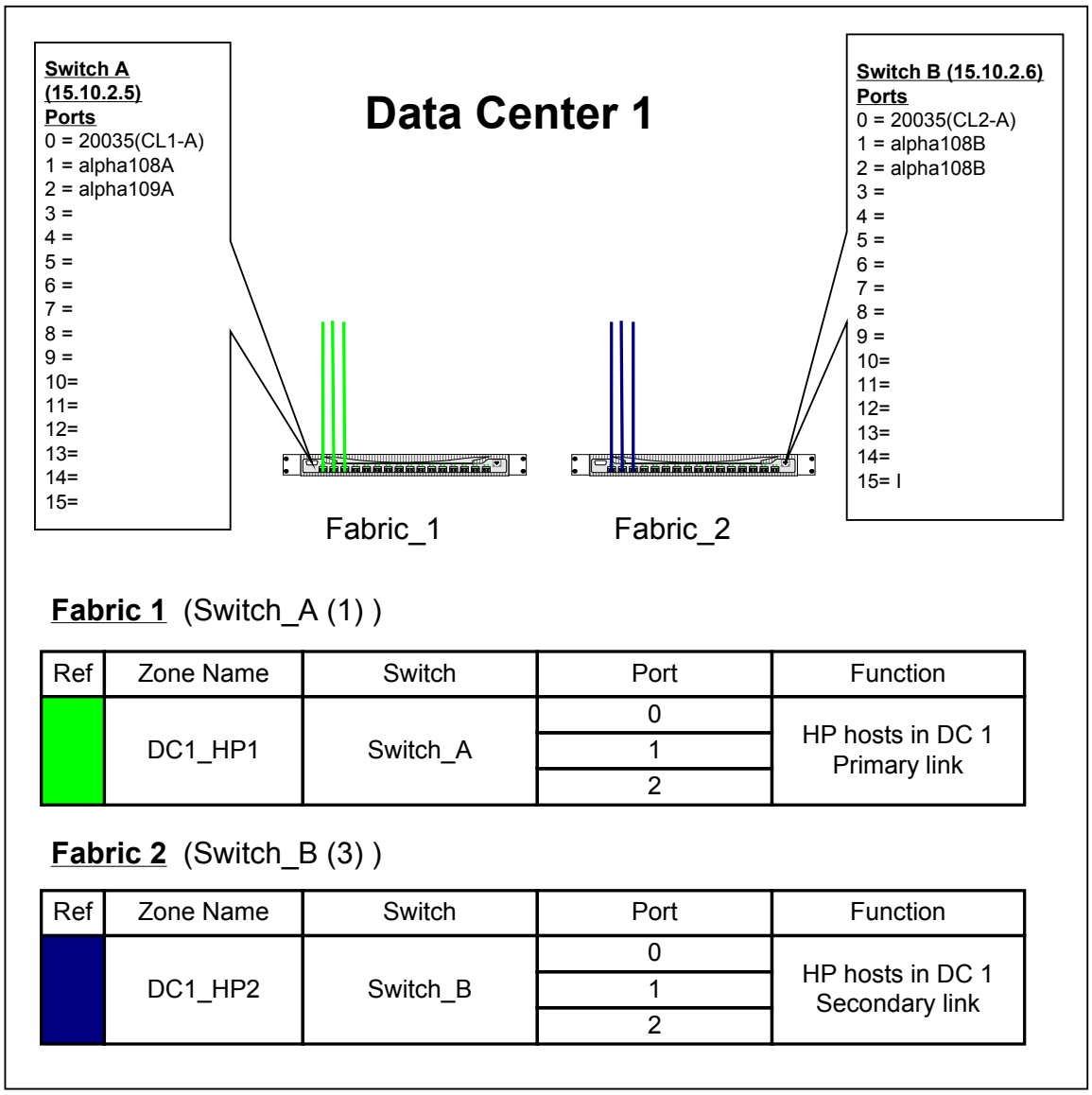
4. Connect according to task 3 (host-to-storage connections).

   Complete the host-to-storage connections as designed in step 1c by connecting the array to the switches and the switches to the hosts. Ensure that all cabling is operational and that the port status shows active ports.

5. Enable security.

   At this point it is necessary to enable security features on either the array or the switches (or both) to ensure appropriate access to all devices. On the arrays, this can be achieved by enabling security on the host ports and configuring only the appropriate host ports World Wide Number (WWN) to have accesses to the port.

   It is required to zone the ports on the switches to ensure that the host has only access to the storage arrays host’s ports and not to the CA ports (to be configured later). At the end of this step, the following zone setting should be implemented on the switches.




**Fabric 1** (Switch_A (1) )

| Ref | Zone Name | Switch | Port | Function |
|---|---|---|---|---|
| | DC1_HP1 | Switch_A | 0 | HP hosts in DC 1 Primary link |
| | | | 1 | |
| | | | 2 | |

**Fabric 2** (Switch_B (3) )

| Ref | Zone Name | Switch | Port | Function |
|---|---|---|---|---|
| | DC1_HP2 | Switch_B | 0 | HP hosts in DC 1 Secondary link |
| | | | 1 | |
| | | | 2 | |

**Figure 6: Switch zone mapping — Site 1**

## 1e: Install the application

With all the physical installations completed, verify that devices are available to the host and visible on all required hosts. Now is the time to install and configure the application software. This step requires some input from the application vendor and assistance from the system administrator and data base administrators.

There are two possible approaches to take when installing the application:

- Each host has a local set of basic application software installed, and the application data is shared between the hosts using mirroring. This approach is useful when upgrading application software since the application has some old copies of the software to operate on while updating one system at a time with new software. You can move the application from one system to the next if there is a software problem on one system. A drawback of this approach is that the administrator must manage so many individual copies of the application software.
- There is only one set of application software that is mirrored among all systems and application software and data move together from system to system. This approach has the advantage of having

only one set of software to maintain, but could lead to extended troubleshooting during failover because of application software errors.

Either of the two approaches can be used in the multi-site solution as long as all administrators understand the impact of each approach. Ensure that applications utilize the correct devices as determined in Table 1. Update the table with logical device numbers. Ensure that all devices used by the application are recorded in the table with the correct size and logical device numbers. This information will be used when creating the CA groups between the sites.

For details on the installation and configuration of the application, reference the application-specific documentation.

# Results

At the end of this task you will have the application up and running on Site 1. The array and switches are installed and access to the array is verified. The physical device's space required by the application is verified and documented.

A single production system is fully operational with no mirror protection.

# 2: Set up site 2

Perform the following tasks to complete the preliminary setup of Site 2. The setup and configuration for Site 2 is very similar to that of Site 1 except for the application design and installation.

## 2a: Define application requirements

At the end of step 1 you have a table of all the application requirements for Site 1. Because this will be a mirror copy of the application at Site 1, you only need to build on the table with physical device information for the application at Site 2. Plan the layout of the array and host-to-storage connections in the same way as Site 1 with the objective to have access to exactly the same number and size of devices as Site 1. If the array on Site 2 is not yet installed, you cannot add the CU:Ldev numbers to the table at this stage. After the array is installed and the host-to-array connections are completed, update the table created in step 1a to resemble Table 4.

**Application and CA/BC mappings for Multi Site Solution**

| Application | Volume group | Logical Volume | Physical device | Size (GB) | Emulation | Site 1 Array 20035 CA_1 CU | Site 1 Array 20035 CA_1 Ldev | Site 2 Array 20049 CA_1/BC_1 CU | Site 2 Array 20049 CA_1/BC_1 Ldev |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | CL1-A and Cl2-A | | CL1-A and Cl2-A | |
| Oracle1 | vg_archive | lv_arch1 | /dev/dsk/c0t1d0 | 14 | Open-E | 00 | 01 | 00 | 20 |
| | | | /dev/dsk/c0t1d1 | 14 | Open-E | 00 | 04 | 00 | 21 |
| | | lv_arch2 | /dev/dsk/c0t1d2 | 14 | Open-E | 00 | 08 | 00 | 22 |
| | | | /dev/dsk/c0t1d3 | 14 | Open-E | 00 | 12 | 00 | 23 |
| | vg_data1 | lv_data1 | /dev/dsk/c0t1d4 | 9 | Open-9 | 01 | 10 | 01 | 35 |
| | | | /dev/dsk/c0t1d5 | 9 | Open-9 | 01 | 12 | 01 | 36 |
| | vg_data2 | lv_data2 | /dev/dsk/c0t1d6 | 9 | Open-9 | 01 | 14 | 01 | 37 |
| | | | /dev/dsk/c0t1d7 | 9 | Open-9 | 01 | 18 | 01 | 38 |
| | vg_data3 | lv_data3 | /dev/dsk/c0t2d0 | 9 | Open-9 | 01 | 16 | 01 | 39 |
| | | | /dev/dsk/c0t2d1 | 9 | Open-9 | 01 | 18 | 01 | 3A |
| | vg_data4 | lv_data4 | /dev/dsk/c0t2d2 | 9 | Open-9 | 01 | 20 | 01 | 3B |
| | | | /dev/dsk/c0t2d3 | 9 | Open-9 | 01 | 22 | 01 | 3C |
| | vg_index | lv_index | /dev/dsk/c0t2d4 | 14 | Open-E | 00 | 30 | 00 | 43 |
| | | lv_index2 | /dev/dsk/c0t2d5 | 14 | Open-E | 00 | 31 | 00 | 44 |
| | vg_logA | lv_log1 | /dev/dsk/c0t2d6 | 3 | Open-3 | 02 | 32 | 02 | 53 |
| | | lv_log2 | /dev/dsk/c0t2d7 | 3 | Open-3 | 02 | 33 | 02 | 56 |
| | | lv_log3 | /dev/dsk/c0t3d0 | 3 | Open-3 | 02 | 34 | 02 | 57 |
| | | lv_log4 | /dev/dsk/c0t3d1 | 3 | Open-3 | 02 | 35 | 02 | 58 |
| | vg_log_b | lv_log1 | /dev/dsk/c0t3d2 | 3 | Open-3 | 02 | 62 | 02 | 59 |
| | | lv_log2 | /dev/dsk/c0t3d3 | 3 | Open-3 | 02 | 63 | 02 | 5A |
| | | lv_log3 | /dev/dsk/c0t3d4 | 3 | Open-3 | 02 | 64 | 02 | 5B |
| | | lv_log4 | /dev/dsk/c0t3d5 | 3 | Open-3 | 02 | 65 | 02 | 5C |
| | vg_admin | lv_error | /dev/dsk/c0t3d6 | 9 | Open-9 | 00 | 70 | 00 | 8A |
| | | lv_config | /dev/dsk/c0t3d7 | 9 | Open-9 | 00 | 71 | 00 | 8B |
| Command device (alpaha108) | | | | | | 00 | 9A | | |
| Command device (alpaha108) | | | | | | 01 | 8C | | |
| Command device (alpaha109) | | | | | | 00 | 9B | | |
| Command device (alpaha109) | | | | | | 01 | 8D | | |
| Command device (alpaha155) | | | | | | | | 00 | 9A |
| Command device (alpaha155) | | | | | | | | 01 | 8A |
| Command device (alpaha156) | | | | | | | | 00 | 9B |
| Command device (alpaha156) | | | | | | | | 01 | 8B |

**Table 4: Application to device mapping - Site 1**

## 2b: Install and configure hosts

Using the same design and installation methods as in step 1b, plan and install the hosts on Site 2. Pay close attention to the network interface design. Because the hosts on Site 2 will be part of the metro cluster between Sites 1 and 2, it is necessary that these hosts are in the same network subnet as the hosts in Site 1.

Install the operating system with the same file system sizes as the systems in Site 1. Install all the additional applications and software on the host systems in the same sequence as that of Site 1, and ensure all patches are installed on the systems. If the host hardware is not exactly the same as that in Site1, it might be necessary to install some host/hardware-specific patches in addition to the patches installed on Site 1.

### 2c: Design and install host-to-storage communication

The host-to-storage communication must be designed in the same way as Site 1. Different model switches can be on this site as long as the switches are still supported by the SAN streams documents as well as the extender/converters used later in the installation processes. More or less host-to-storage connections can be used if required, but HP recommends using the same design processes in each site to ensure similar performance characteristic on each site.

### 2d: Install and configure arrays

Using the same step as in 1d design and install the array on Site 2. It will be necessary to format the array similar to the array on Site 1 ensuring the correct number and size devices are available for the CA link. Keep in mind that this array will have both CA and BC devices for the application, and therefore will have to have double the configuration of that in Site 1.

Ensure that all array software licenses are installed and active on the array. Because this array will require using both synchronous and asynchronous replication, it is necessary to have both the CA and CA extension licenses installed on this array.

### 2e: Install the application

The only installation required on Site 2 for the application is to install the basic application software on this site (if this is not shared between sites). The remaining application configuration will be completed after the CA link between Sites 1 and 2 is established and the data is copied from Site 1 to 2.

Update Table 4 ensuring that there are sufficient devices available for application on Site 2 and that all hosts on site have access to these devices.

## Results

At the end of this step you will have a fully configured host, with access to the necessary devices on the array at Site 2. The application installation will be completed after the mirror process is completed.

# 2 Implement the CA_1 link between Sites 1 and 2

In this step you design and install the physical links between Sites 1 and 2, then create a logical communication path between the control units (CUs) of the arrays. After the connection between the arrays is established, you will create the RAID Manager XP instances, and then finally create the device pairs between the arrays, which starts the replication processes.

The process of creating the physical and logical links between the arrays does not influence the application, and can be performed while the application is running on Site 1. When creating the CA pair between Sites 1 and 2, the synchronous replication can affect the application performance on Site 1 but can still be done with the application up and running in Site 1. The application does not need to be stopped during this step of the configuration.

## 1: Design and install physical/logical links between sites

### 1a: Identify "Initiators" and "RCU Targets"

The physical communication between the arrays requires a physical path from an "Initiator" port on one array to an "RCU Target" port on the other array. "Initiator" ports are sender or talker ports and initiate communication while the "RCU Target" port is a receiver or listener port. The communication path from an "Initiator" to a "RCU Target" provides one-way communication from one array to another. You must communicate in both directions and to fail over applications automatically between arrays, it is necessary to create an additional "Initiator" to "RCU Target" path in the opposite direction. The ability of a port to act as an "Initiator" or "RCU Target" is determined by the processor controlling the ports. On some of the older arrays, a processor manages two adjacent ports and therefore both the ports managed by the processor have to be set to the same function.

Carefully plan the "Initiator" to "RCU Target" configurations, paying special attention to high availability and any performance requirements. It is possible to share the same switches with host and CA communication as long as the host and CA ports are zoned in separate zones. When using switches, the switch combines the "Initiator" and "RCU Target" communication into one inter switch link (ISL) that must be extended between sites. Ensure that the ISL has sufficient bandwidth available for the CA traffic, and does not overload the ISL with multiple "Initiators." Although the ISL cannot be zoned only for CA traffic, it is not advisable and not supported to have host access over the ISL between sites. Allowing a host on Site 1 to access storage on Site 2 can become very problematic during site failures, and allowing hosts to perform large transactions over the ISL used for CA communications can negatively impact the CA communication.

The following diagram is an example of the physical path specifications between the two arrays.

20035

| Local Array | Local Port Initiator | Remote Port RCU Target | Remote Array |
|---|---|---|---|
| 20035 | CL1-C | CL1-D | 20049 |
| | CL2-C | CL2-D | |

10049

| Local Array | Local Port Initiator | Remote Port RCU Target | Remote Array |
|---|---|---|---|
| 20049 | CL1-C | CL1-D | 20035 |
| | CL2-C | CL2-D | |

**Figure 7: Physical path definition between two arrays**

The physical connections between the arrays depend on the configuration decided upon during the HLD process. The following sample diagram shows a configuration using only one dense wavelength division multiplexing (DWDM) system per site, sharing switches with the host systems. Bi-directional CA requires an initiator-to-RCU target link from Site 1 to Site 2, and an initiator-to-RCU target link from Site 2 to Site 1. To ensure high availability, HP recommends having two links in each direction, with the initiator for link 1 from Site 1 to 2 and the RCU for link 1 from Site 2 to 1 located in the same power boundary (cluster) on the array.

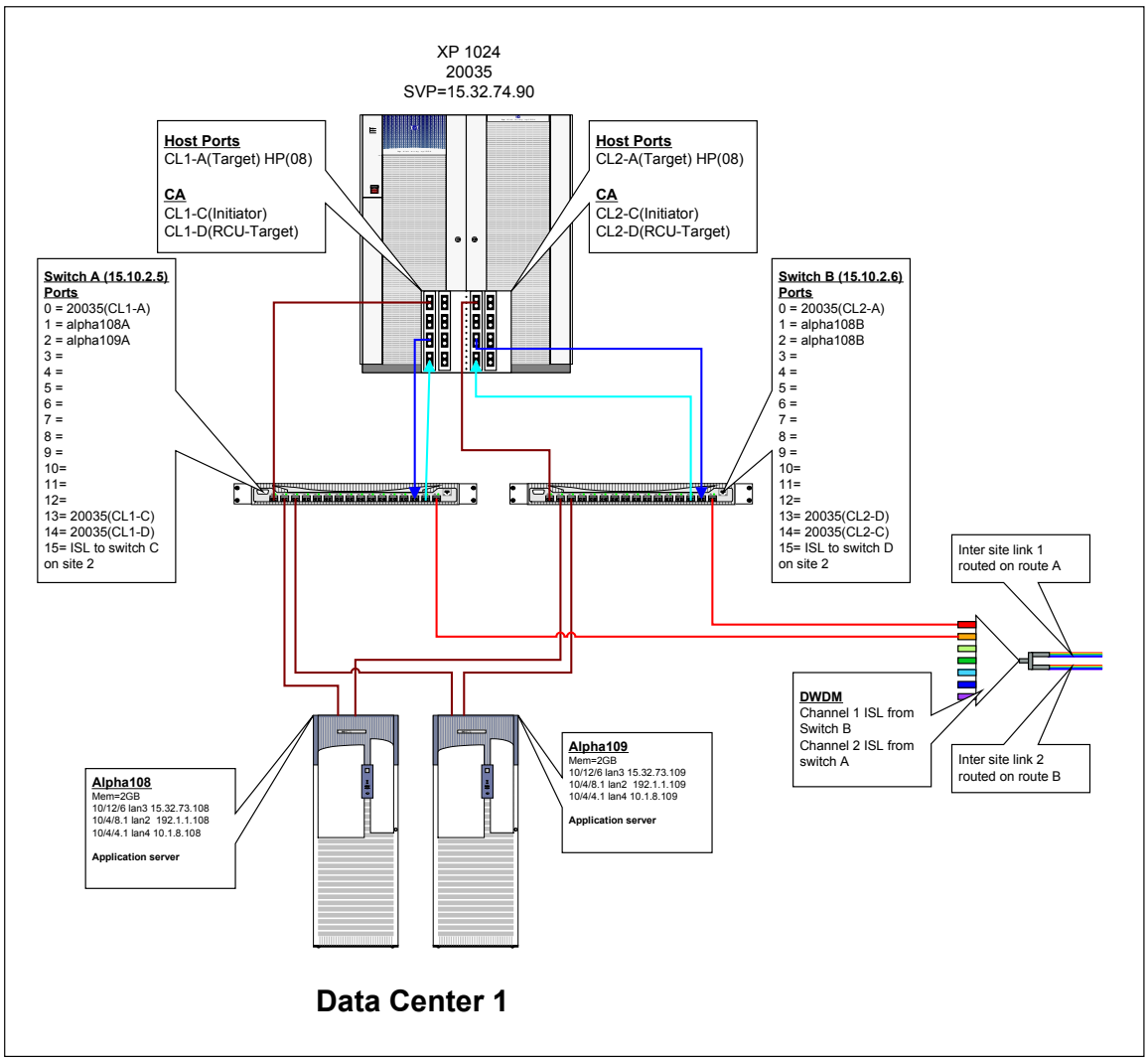



**Figure 8: "Initiator" and "RCU Target" ports on Site 1**

## 1b: Define intersite links

The physical link between Sites 1 and 2 can be implemented with DWDM technology or by using a standard ISL. The HLD physical diagrams indicate whether the customer has chosen DWDM or direct ISL between switches. With DWDM you can implement multiple CA links on the same physical fiber sharing the infrastructure and optimizing the utilization of the physical link. The link can even be shared with other applications using different protocol (like TCP/IP). A direct ISL link requires a dedicated Fibre Channel (FC) link for each ISL implemented.

The model configuration assumes FC-based replication with the use of fiber switches. To improve performance, consider using extended fabric buffer management on the switch.

After the physical link between the sites is installed, it is possible to connect the DWDM system to the physical link and ensure that the communication between DWDM systems is operational. The next step is to connect the switches to the DWDM extender with each switch on its own channel. It is important to document the channel used on the DWDM since this will determine the connection for the remote switch.

Before connecting the switches to the DWDM extender, merge the two switch configurations and zoning information since the connections will create one central fabric between the switches. It is recommended to perform the connection with the application down since the merging of the zones might disrupt host-to-storage communications.

The following diagram shows the interconnections between Sites 1 and 2 using one DWDM extender per site with two independent communication links between the sites. For redundancy these two links must be routed by different routs. The configuration creates two independent fabric configurations between the switches, with each fabric providing an “Initiator” to “RCU Target” communication channel in both directions.

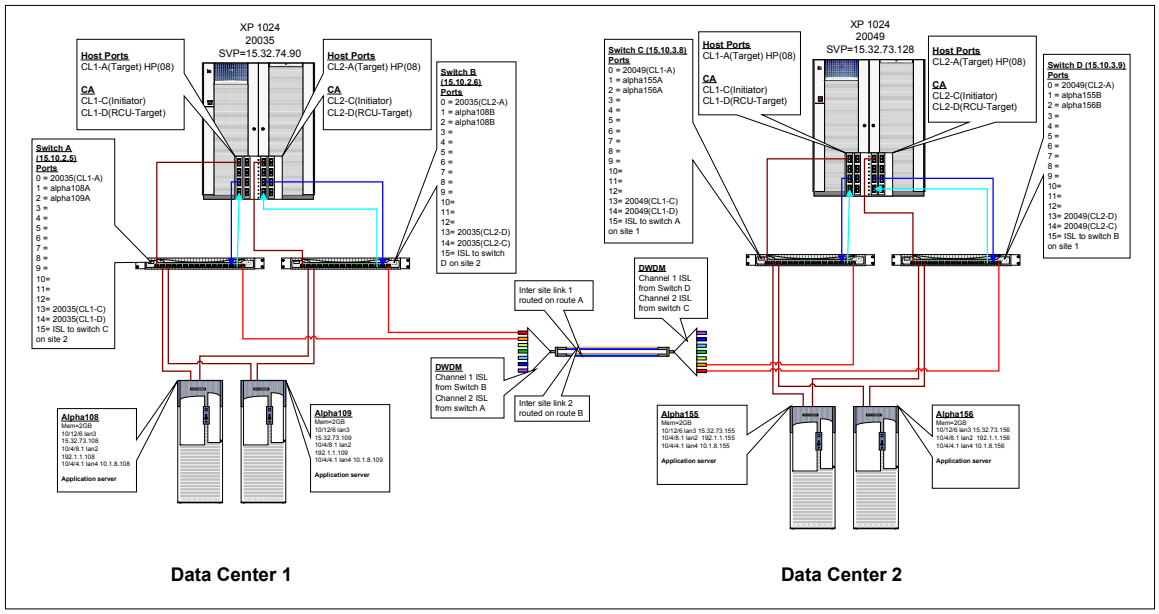



**Figure 9: Physical connections between Sites 1 and 2**

## 1c: Design switch zoning

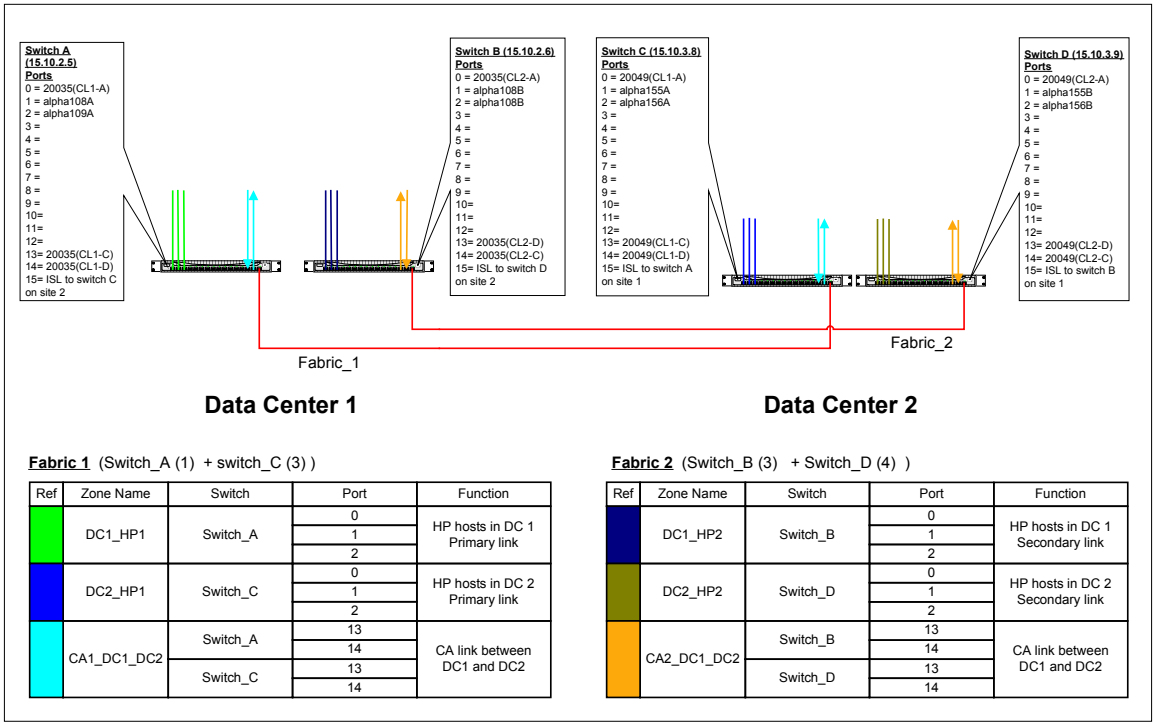
When connecting the switches on Site 1 with those on Site 2 a single fabric will be established between the sites with shared configuration and zoning information. To create the shared fabric, it is required that the configuration setting on all switches is similar and that each switch has its own unique domain ID number. This can be performed by telnet into the switch and changing the configuration settings.

**Note:** Changing the domain ID number on the switch may invalidate any zone configuration on a switch.

The best way to ensure that the configurations are the same is to download the configuration from each switch and compare all parameters. Change all the parameters to a common set, and then upload the new configuration. To complete this operation, the switches must be disabled. It is possible to continue operations through one set of switches while upgrading the other set by using alternative paths set on the host. However, to reduce any risk of interrupting production systems, HP recommends that you schedule downtime for this operation.

**Note:** To change configuration settings, the switch must be offline, which suspends all hosts-to-storage operations.

To merge the zone information, it will be necessary to delete all zone configurations on one of the switches, then complete the interconnection of the switches and redefine the zones deleted earlier. The new configuration will now be shared between both switches. You must define switch port zoning to ensure that the host systems do not see any devices at the other site. The use of either hard port zoning or worldwide name zoning is allowed. ISL links cannot be zoned only for CA access because ISL ports are open to be used by any port that is in a zone covering multiple switches. The following diagram shows the completed zone configuration for the two sites. Note that there are two independent fabrics available between the sites.




**Fabric 1** (Switch_A (1) + switch_C (3) )

| Ref | Zone Name | Switch | Port | Function |
|---|---|---|---|---|
| | DC1_HP1 | Switch_A | 0<br>1<br>2 | HP hosts in DC 1 Primary link |
| | DC2_HP1 | Switch_C | 0<br>1<br>2 | HP hosts in DC 2 Primary link |
| | CA1_DC1_DC2 | Switch_A | 13<br>14 | CA link between DC1 and DC2 |
| | | Switch_C | 13<br>14 | |

**Fabric 2** (Switch_B (3) + Switch_D (4) )

| Ref | Zone Name | Switch | Port | Function |
|---|---|---|---|---|
| | DC1_HP2 | Switch_B | 0<br>1<br>2 | HP hosts in DC 1 Secondary link |
| | DC2_HP2 | Switch_D | 0<br>1<br>2 | HP hosts in DC 2 Secondary link |
| | CA2_DC1_DC2 | Switch_B | 13<br>14 | CA link between DC1 and DC2 |
| | | Switch_D | 13<br>14 | |

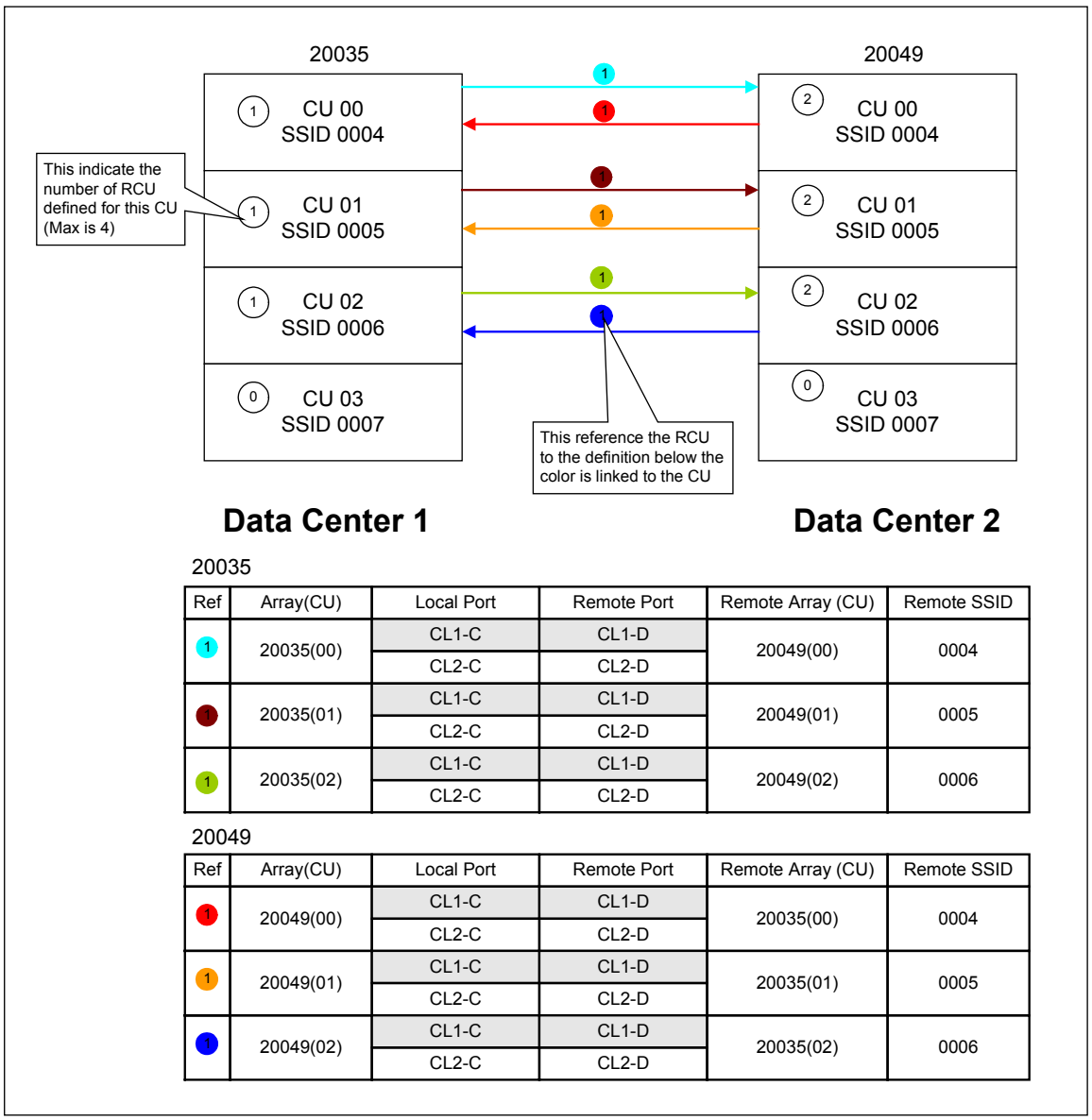
**Figure 10: Sample zone mapping between Sites 1 and 2**

## 1d: Design CU/RCU mapping

On the XP array, you format physical devices (array group) with certain emulation types and logical sizes, and then allocate these logical devices to a control unit (one CU can address 255 logical devices). You then assign a logical device from the CU to a host port with a specific target and LUN number. Use devices from several CUs rather than only one unit since most of the devices in one CU are normally located on the same physical drive. When using CA to replicate data between arrays, it is necessary to create a logical path between arrays called a remote control unit (RCU). An RCU must be configured between a source CU (initiator) and a target CU to create a CA device pair using devices controlled by the specific CUs. For full failover capabilities, configure a reverse RCU from the target array to the source array. A CU can only support four RCU definitions. A RCU definition can define one or more physical paths between the arrays as communication paths for the device pairs. All I/O will be shared between these paths and failure of one path does not affect the pair status, only the mirror performance.

**Note:** To ensure optimal performance, HP recommends that alternative paths between the arrays have the same performance characteristics. The use of different technologies to create the links between arrays may negatively affect the CA performance.

To replicate the entire application, you must configure RCU links for every CU the application is using. Using the application to device mapping (Table 4), determine all the CUs in each array that require a logical path between them. In the mapping of the devices groups (table 4) you took great care to ensure that you paired devices from the same CU group throughout the configuration. This may not always be possible, but HP recommends staying within the same CU for each device pair and using the BC devices to switch to another CU within a site. In the following sample, an RCU link exists between CU 0 on array 20035 and CU 0 on array 20049. A reverse RCU link exists between CU 0 on 20049 and CU 0 on array 20035. Similar RCU links exist for CU 01 and 02. At this stage you only utilize 1 RCU per CU on each array.




20035

| Ref | Array(CU) | Local Port | Remote Port | Remote Array (CU) | Remote SSID |
|---|---|---|---|---|---|
| 1 | 20035(00) | CL1-C | CL1-D | 20049(00) | 0004 |
| | | CL2-C | CL2-D | | |
| 1 | 20035(01) | CL1-C | CL1-D | 20049(01) | 0005 |
| | | CL2-C | CL2-D | | |
| 1 | 20035(02) | CL1-C | CL1-D | 20049(02) | 0006 |
| | | CL2-C | CL2-D | | |

20049

| Ref | Array(CU) | Local Port | Remote Port | Remote Array (CU) | Remote SSID |
|---|---|---|---|---|---|
| 1 | 20049(00) | CL1-C | CL1-D | 20035(00) | 0004 |
| | | CL2-C | CL2-D | | |
| 1 | 20049(01) | CL1-C | CL1-D | 20035(01) | 0005 |
| | | CL2-C | CL2-D | | |
| 1 | 20049(02) | CL1-C | CL1-D | 20035(02) | 0006 |
| | | CL2-C | CL2-D | | |

**Figure 11: Sample CU/RCU mapping**

**Note:** Each CU has an SSID number defined on the array. This number is required when creating the RCU.

## 1e: Configure RCUs

With the physical paths between the arrays configured and installed and the logical paths designed, it is time to create the RCUs between the arrays, which is done using the Command View array console software or the internal laptop of the array (SVP). Ensure that all ports are configured as Initiators and RCU Targets as defined, using the port configuration menus. Create the RCU specifying the remote array serial numbers and CU numbers as well as the SSID of the remote CU. Specify all paths to the remote array as well as the minimum paths that are required to keep the link operational. HP recommends that the minimum path is set to 1 to ensure that the CA link remains operational during link failures until there is no more link available between the arrays. Check the RCU status to ensure that all paths show operational status. Repeat the process

for each CU on the source array as well as for all CUs on the target array, creating logical paths in both directions.

# Results

At the end of this task, you have a physical and logical path configured between the array at Site 1 and the array at Site 2. At this stage, you are ready to start the replication process between the arrays.

# 2: Create the CA_1 pair

At this stage it is assumed that the physical connections between Sites 1 and 2 exist and that all RCUs have been configured. Creating the pair relationship between devices can be performed from either the array's service processor (SVP) or from the host using RAID Manager (RAID Manager XP). Since the RAID Manager XP configuration is required for the cluster configuration and for management of the solution, HP recommends using RAID Manager XP to configure the device pairs.

## 2a: Install RAID Manager

If the RAID Manager XP software was not installed during the host installation, install it at this point. Follow the instruction in the "readme" file distributed with the software for installation procedures. Ensure that the latest version of RAID Manager XP is installed and it is at least version 01.11.00 or later. The RAID Manager XP software must be installed on all hosts in the configuration on both Sites 1 and 2.

## 2b: Design RAID Manager instance mapping

When creating device pairs using RAID Manager XP, it is required to configure at least two instances of RAID Manager XP, one instance to manage the primary devices (P-vol) and one instance to manage the secondary devices (S-vol). In the case of CA pairs the two instances will be located on two different hosts, one at Site 1 and one at Site 2. RAID Manager XP can automatically attempt communications with another instance if the first instance fails to respond. In the example, there are two hosts on Site 1 and two on Site 2, each able to run a RAID Manager XP instance. The two hosts on Site 1 have an instance that can manage the primary devices of the pair and attempt to communicate with either one of the instances on Site 2.

When mapping RAID Manager XP instances, try to keep them in sequence; for example, 0 = Application, 1 = CA_1, 2 = BC_1, and so on. RAID Manager XP runs on any host attached to the array. RAID Manager XP does not need to see all the devices, just the command device for the array.

The following diagram shows the RAID Manager XP instances required to create and manage the CA_1 device pairs between Sites 1 and 2.

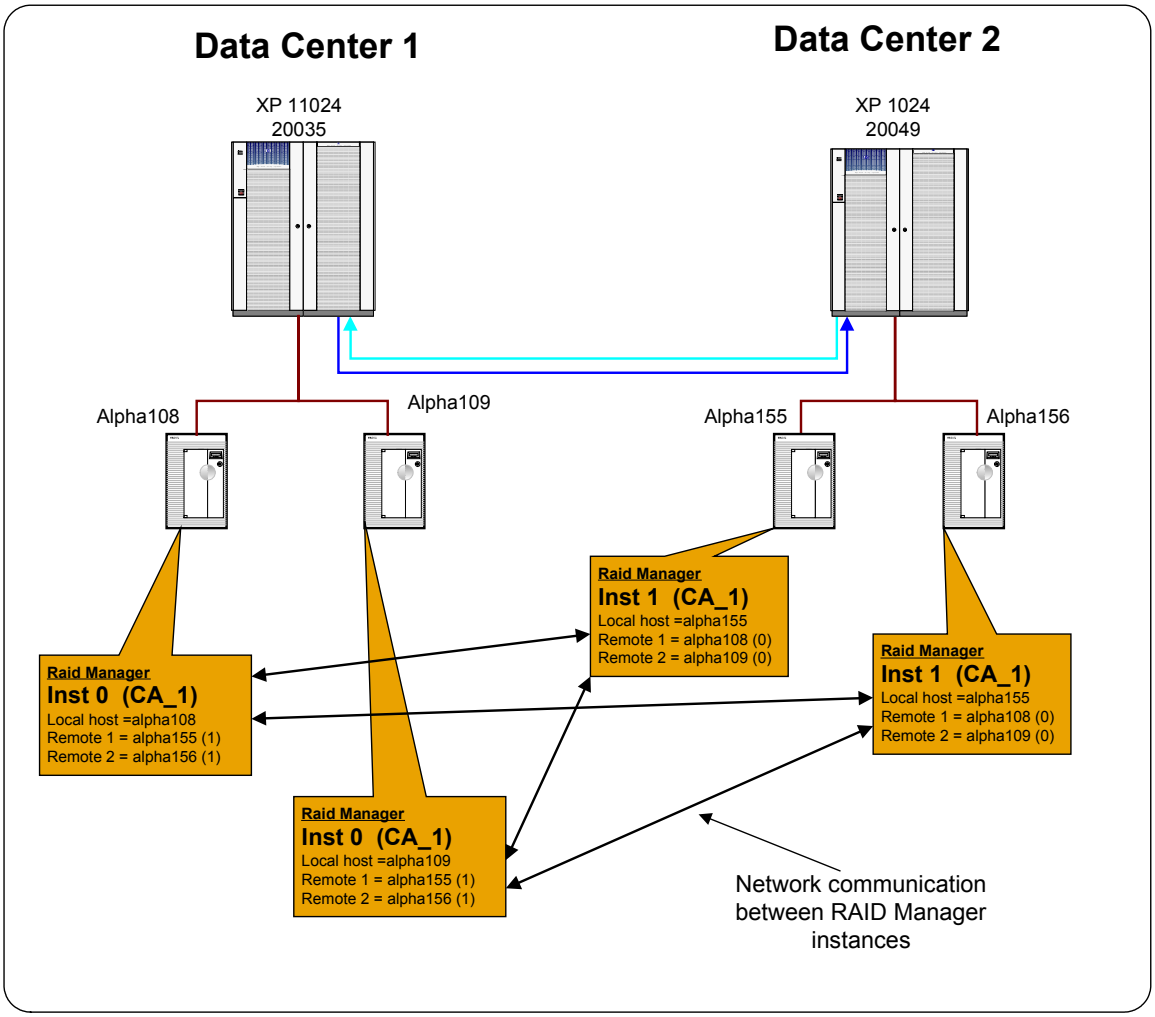



**Figure 12: RAID Manager instance for CA_1 pair**

## 2c: Identify command devices

Using the "ioscan" command, find the command devices allocated to the specific host (see Table 4 for the command device CU:ldev numbers). A command device is recognized by the "-CM" suffix to the device description. Find the associated device file for each command device.

## 2d: Create RAID Manager XP configuration files

To maintain consistency between all the devices used in the application, HP recommends grouping all devices for the application in one device group in the RAID Manager XP configuration file. An RAID Manager XP instance must be created on all hosts that can manage the device pairs. RAID Manager XP requires two instances to communicate with each other to manage a single group. Each instance defines the local devices that the instance can manage. The group and device names within the configuration file define the relationships between the devices. When executing the RAID Manager XP "**paircreate**" command, the direction and characteristics of the device group will be defined.

On each host in the configuration, create a temporary RAID Manager XP instance by defining only the instance variables and commands devices in the configuration file. RAID Manager XP configuration files are located in the **/etc** directory and is named **horcmxx.conf** (**xx** indicates the instance number). The following is an example of the temporary configuration file.

```
#/************************ For Local instance **************************/
HORCM_MON
#ip_address    service    poll(10ms)  timeout(10ms)
alpha108     horcm0     1000       3000

#/************************ For local command devices *******************/
HORCM_CMD
#dev_name       dev_name       dev_name
/dev/rdsk/c4t15d7
```

After creating the configuration file, be sure to add a entry to the **/etc/services** file to reserve a port number for each RAID Manager XP instance. The following is an example of the entries in the **/etc/services** file. Ensure that the entries are the same on all hosts in the configuration using RAID Manager XP.

```
horcm0              51000/udp   #RAIDManager instance 0
horcm1              51001/udp   #RAIDManager instance 1
```

Start the RAID Manager XP instance on the local host using the "**horcmstart.sh xx**" command, where **xx** is the instance number to start. After a successful startup of the RAID Manager XP instance, set the "**HORCMINST=xx**" environment variable to ensure the next RAID Manager XP commands connect to the correct RAID Manager XP instance.

Find the internal target and LUN numbers for each of the devices used by the application by using the "**raidscan -p <port_number> -fx**" command. Complete the configuration file by adding the device group name, device name, and port:Target:Lun number for each device. Define any remote instances that can manage the remote site of the device group and their services alias. The following is an example of a completed device group configuration file.

```
#ip_address    service    poll(10ms)  timeout(10ms)
alpha108     horcm0     1000       3000

#/************************ For local command devices *******************/
HORCM_CMD
#dev_name       dev_name       dev_name
/dev/rdsk/c4t15d7

#/************************** For device groups ***************************/
HORCM_DEV
#dev_group    dev_name    port#   TargetID  LU#   MU#
ca1_oracle      ora_data1  CL1-A          1     1        0
ca1_oracle      ora_data2  CL1-A          1     2        0
ca1_oracle     ora_index1      CL1-A          1     3        0
ca1_oracle     ora_index1      CL1-A          1     4        0
```

```
ca1_oracle    system1     CL1-A              1    5       0

#/*************** For remote instance definition ************************/
HORCM_INST
#dev_group   ip_address   service
ca1_oracle         alpha155    horcm1
ca1_oracle         alpha156    horcm1
```

Stop and restart the instance using the "**horcmshutdown.sh xx**" and "**horcmstart.sh xx**" commands.

## 2e: Validate the RAID Manager configuration files

RM only reports the status of the devices defined within the configuration file of the RAID Manager XP instance used, so you must ensure that all instance configuration files address the correct number of devices as well as the correct logical devices. The configuration file uses the target:lun number of a device to refer to a logical device in the array. It is very easy to mistype a value, therefore, it is important to verify the configuration file after any changes are made to the file. Use the **pairdisplay** command to check the group definition and to ensure that all the correct logical devices are used in the configuration.

## 2f: Create the CA_1 pair

After validating the RAID Manager XP configuration file, create the device pair and start the initial copy of data from Site 1 to Site 2. Connect to one of the hosts on Site 1. Set the **HORCMINST** environment variable to the local instance number (this should be 0 for this example). Using the **paircreate** command, create the device group. When creating the device group, it is necessary to specify the fence level of the device group. For the CA_1 link, the fence level should be one of the synchronous fence levels, either **DATA** or **NEVER**. The fence level determines if the access to the primary device will be blocked, if the remote I/O fails. With fence level **DATA,** local I/O will always be blocked (fenced) if the I/O cannot be written to the remote site. This will ensure full data integrity between the sites, but will stop the application if the link between sites fails. With fence level **NEVER,** the array will not block I/O to the primary devices if the link fails between the sites. This can result in inconsistent data between the sites. HP recommends using fence level **DATA** to ensure consistency between Sites 1 and 2.

## 2g: Wait for pair status

After the **paircreate** command completes, the devices will start to copy all data from the local site (Site 1) to the remote site (Site 2). This is known as the initial copy process. Depending on the performance of the links between the sites and the total disk space used by the application, this process can take between a number of hours to a day to complete. During initial copy all data blocks on the source device are copied to the target device, even if the block is not actively used by the application or is empty. This process is required to ensure that the devices are exact replica of each other. Using the **pairdisplay** command, wait for the pair status to reach "**PAIR**" status.

# Results

At the end of this task you have a replica of the production data available on the secondary site (Site 2).

# 3 Create the MetroCluster for Sites 1 and 2

In this task, you create the metro cluster between Sites 1 and 2, and test failover in the metro cluster configuration

## 1: Install Cluster software

If the cluster software was not installed during the host installation, you must install the software at this point. If software was installed previously, verify that the software was installed on all hosts.

### 1a: Install ServiceGuard

The first step is to install the base ServiceGuard software on each host in the configuration that will be part of the metro cluster. Follow the installation instructions in the product documentation. While the data is being synchronized between Sites 1 and 2, you can perform the cluster installation on the hosts. When the device pair reaches "pair" status, the cluster configuration can be completed and failover and failback of the application be tested. Be sure you install all required patches for the base cluster software on all hosts as well as any hardware specific patches for cluster usage.

### 1b: Install MetroCluster and MetroCluster XP-CA toolkit

Install the MeroCluster and XP-CA toolkit on all hosts in the configuration. Be sure you install all the latest MetroCluster patches on each host.

### 1c: Install quorum server software

The metro cluster configuration requires either an arbitrator node or a quorum server node to ensure that there is always more than 50% of a cluster available after a site failure. It is recommended to use a third site for the arbitrator node. Since the arbitrator node must be part of the cluster, it is not advisable to use an arbitrator in this configuration. The use of a quorum server allows the user flexibility in the type of host to use for quorum serves, and this node does not have to be part of the cluster configuration. The ideal hosts for quorum services are the host on Site 3. Since you have not installed this host yet, you will either have to find another host or configure the cluster without a quorum server. You can also wait with the cluster configuration until the third site host is available. Follow the installation instruction to install quorum services on specific hosts.

## 2: Configure the cluster

The following task is the main step in the process of configuring a metro cluster with XP-CA toolkit. For detailed instructions, refer to the appropriate installation manuals.

### 2a: Create the cluster

The first step in configuring the cluster is to create a basic cluster configuration file identifying all nodes and hardware components in the cluster. Use the **`cmquerycl`** command to create a template configuration file with all the appropriate node information added to the file. Edit the file created by the **`cmquerycl`** command, and modify the cluster name and package configuration. Check the network interface definitions, and ensure that the correct network is used for the heartbeat. After all modifications are completed check the configuration file using the **`cmcheckconf`** command. After a successful check, the configuration can be applied using the **`cmapplycnf`** command. All the cluster commands can be executed from any node in the cluster although HP suggests you select one node and the configuration node.

### 2b: Start the cluster

When the **cmapplyconf** command successfully delivers the cluster configuration to all nodes in the cluster, the cluster can be started using the **cmruncl** command. Use the **cmviewcl** command to check the cluster status, and ensure all nodes and network interfaces are available. Edit the **/etc/rc.config.d/cmcluster** file, and enable automatic startup of the cluster services on each of the nodes. This will ensure that the node automatically rejoins the cluster after a failure or reboot.

## 3: Configure cluster package

In the cluster configuration, the application and all application resources are grouped together in what is called a package. The package can then be moved from one node to the other either manually or automatically. Package configuration exists out of two sections; a package configuration file and a package script file.

### 3a: Create package configuration file

Using the **cmcreatepkg** command, create the package configuration file. The file is normally located in the **/etc/cmcluster/<package_name>/<package_name>.ascii** file. This file defines the package name, nodes that can run the application, failover policies, resources, and startup values. After editing the default package configuration file, use the **cmapplyconf** command to add the package to the cluster configuration. Any changes to this file must be applied to the cluster using the **cmapplyconf** command.

### 3b: Create package script file

Use the **cmcreatepkg** command to create the default package script file for the package. Edit this file and add all the volume groups and logical volumes to the file that the application must use. Customize the resources and application startup/shutdown commands. This is the script that gets executed as soon as the application starts on the node to ensure that all resources are available for the application. The script by itself does not check the CA mirror status, and requires the CA mirror configuration file to enable mirror checking. If the package is started without the CA toolkit configuration file, the status of the mirror pairs will not be checked and the application may not start. For more detail on the package startup script, see the cluster documentation. This file is normally located in the package directory **/etc/cmcluster/<package_name>** and is called **<package_name>.sh.** This file must be manually copied to all nodes in the cluster. Any changes must be done to all nodes in the cluster.

### 3c: Create CA-XP toolkit configuration file

The CA-XP toolkit configuration file indicates to the package script that some hardware replication is in place and forces the script to first check the mirror status before starting the application. The configuration file is located in the package directory and is called **<packge_name>_caxp.env.** The configuration file has a number of parameters that determine application startup condition. Reference the toolkit documentation for a detailed description of each parameter. The file also contains the RAID Manager XP instance number and device group name that manage the application data.

## 4: Prepare application and package for cluster

In this step you ensure that all application resources are available on all of the hosts.

### 4a: Distribute volume group information

During the application installation only one host has been configured with all volume group information defined in the OS. All the other hosts have access to the devices but do not have the volume group information configured. To enable the application to run on any of the nodes in the cluster, it is necessary to distribute the volume group information to all nodes. Most volume groups are self-contained. This means that the volume group ID is written in the reserve areas of the devices although custom-naming conventions are

not maintained on the device. Use the **vgexport -s -p -m <mapfile.map> <vg_name>** command on the primary host to create a map file for each volume group. The map file contains the volume group id (provided by the **-s** option) and a map of custom names to logical segments on the devices. The **-p** option prevents the volume group from being permanently exported from the system. Copy the map file to all hosts in the configuration. To be able to read volume group ID from the local devices on each host, it is necessary to suspend the CA link with read/write access to all devices. Use the RAID Manager XP **pairsplit** command with anr
**-rw** option to enable all the devices for read/write access and to suspend mirroring of the devices. Create the device files for the volume groups using the **mknod** command, then use the **vgimport** command to import the volume group information on each node.

**Note:** Do not activate the volume groups on more than one system at any time. Doing this will cause file system corruption.

When the volume group definitions are imported to all nodes, ensure that the volume group is disabled on all nodes and reactivate the CA device mirroring using the **pairresync** command.

### 4b: Enable volume group for cluster awareness

One of the services the cluster software provides is to ensure that a volume group can only be active on one system at any time. To enable this function the volume group must be configured as cluster aware. Use the **vgchange -c y <volume_group>** command to enable the volume group as cluster aware. To execute the command the application must be down and the volume group de-activated using the **vgchange -a n** command.

## 5: Test the cluster

At this point the application is ready to be tested in the cluster environment. Use the **cmrunpkg** command to start the application on primary node. Check the startup log file to ensure that all operations was successful, and ensure that the CA-XP toolkit was enabled and did check the device status. The package startup log file is located in the package directory with filename **<package_name>.log**. After a successful startup on the primary node, stop the application using the **cmhaltpkg** command, and restart the application on next node in the cluster. Repeat this process for all nodes in the cluster, ensuring that the application starts successfully and all resources are available after startup. Any cluster failure test can be performed to see if the application reacts automatically to failure conditions in the cluster.

## Results

At the end of this step, you have successfully installed and configured a two-node Sync CA metro cluster.

# 4 Implement the BC_1 BusinessCopy at Site 2

At this stage you have a two-node metro cluster configuration with sync CA replication between Sites 1 and 2. To create a point-in-time copy of data that can be replicated to Site 3, you must implement a business copy (BC) at Site2. The following step describes the process of creating the business copy at Site 2.

## 1: Plan BC_1 configuration

### 1a: Define application requirements

In the CA_1 pair, you paired each of the devices used by the application on the primary Site with a device on the secondary site. You now have to pair each one of these secondary site devices with another device on this site and create a business copy pair with these devices. The first step in this process is to identify all the devices that will be used in the BC device pair, and add them to the table created with the CA_1 pair (table 4). Identify devices of the same emulation and size and record their CU:Ldev numbers in the table. Keep in mind that using devices from the same CU will simplify the CA_2 configuration. After this step, the updated table will look like the following.

**Application and CA/BC mappings for Multi Site Solution**

| Application | Volume group | Logical Volume | Physical device | Size (GB) | Emulation | Site 1 Array 20035 CA_1 CU | Site 1 Array 20035 CA_1 Ldev | Site 2 Array 20049 CA_1/BC_1 CU | Site 2 Array 20049 CA_1/BC_1 Ldev | Site 2 Array 20049 BC_1/CA_2 CU | Site 2 Array 20049 BC_1/CA_2 Ldev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | CL1-A and CI2-A | | CL1-A and CI2-A | | CL1-B | |
| Oracle1 | vg_archive | lv_arch1 | /dev/dsk/c0t1d0 | 14 | Open-E | 00 | 01 | 00 | 20 | 00 | A0 |
| | | | /dev/dsk/c0t1d1 | 14 | Open-E | 00 | 04 | 00 | 21 | 00 | A1 |
| | | lv_arch2 | /dev/dsk/c0t1d2 | 14 | Open-E | 00 | 08 | 00 | 22 | 00 | A2 |
| | | | /dev/dsk/c0t1d3 | 14 | Open-E | 00 | 12 | 00 | 23 | 00 | A3 |
| | vg_data1 | lv_data1 | /dev/dsk/c0t1d4 | 9 | Open-9 | 01 | 10 | 01 | 35 | 01 | B0 |
| | | | /dev/dsk/c0t1d5 | 9 | Open-9 | 01 | 12 | 01 | 36 | 01 | B1 |
| | vg_data2 | lv_data2 | /dev/dsk/c0t1d6 | 9 | Open-9 | 01 | 14 | 01 | 37 | 01 | B2 |
| | | | /dev/dsk/c0t1d7 | 9 | Open-9 | 01 | 18 | 01 | 38 | 01 | B3 |
| | vg_data3 | lv_data3 | /dev/dsk/c0t2d0 | 9 | Open-9 | 01 | 16 | 01 | 39 | 01 | B4 |
| | | | /dev/dsk/c0t2d1 | 9 | Open-9 | 01 | 18 | 01 | 3A | 01 | B5 |
| | vg_data4 | lv_data4 | /dev/dsk/c0t2d2 | 9 | Open-9 | 01 | 20 | 01 | 3B | 01 | B6 |
| | | | /dev/dsk/c0t2d3 | 9 | Open-9 | 01 | 22 | 01 | 3C | 01 | B7 |
| | vg_index | lv_index | /dev/dsk/c0t2d4 | 14 | Open-E | 00 | 30 | 00 | 43 | 00 | A9 |
| | | lv_index2 | /dev/dsk/c0t2d5 | 14 | Open-E | 00 | 31 | 00 | 44 | 00 | AA |
| | vg_logA | lv_log1 | /dev/dsk/c0t2d6 | 3 | Open-3 | 02 | 32 | 02 | 53 | 02 | C0 |
| | | lv_log2 | /dev/dsk/c0t2d7 | 3 | Open-3 | 02 | 33 | 02 | 56 | 02 | C1 |
| | | lv_log3 | /dev/dsk/c0t3d0 | 3 | Open-3 | 02 | 34 | 02 | 57 | 02 | C2 |
| | | lv_log4 | /dev/dsk/c0t3d1 | 3 | Open-3 | 02 | 35 | 02 | 58 | 02 | C3 |
| | vg_log_b | lv_log1 | /dev/dsk/c0t3d2 | 3 | Open-3 | 02 | 62 | 02 | 59 | 02 | C4 |
| | | lv_log2 | /dev/dsk/c0t3d3 | 3 | Open-3 | 02 | 63 | 02 | 5A | 02 | C5 |
| | | lv_log3 | /dev/dsk/c0t3d4 | 3 | Open-3 | 02 | 64 | 02 | 5B | 02 | C6 |
| | | lv_log4 | /dev/dsk/c0t3d5 | 3 | Open-3 | 02 | 65 | 02 | 5C | 02 | C7 |
| | vg_admin | lv_error | /dev/dsk/c0t3d6 | 9 | Open-9 | 00 | 70 | 00 | 8A | 00 | 8D |
| | | lv_config | /dev/dsk/c0t3d7 | 9 | Open-9 | 00 | 71 | 00 | 8B | 00 | 8E |
| Command device (alpaha108) | | | | | | 00 | 9A | | | | |
| Command device (alpaha108) | | | | | | 01 | 8C | | | | |
| Command device (alpaha109) | | | | | | 00 | 9B | | | | |
| Command device (alpaha109) | | | | | | 01 | 8D | | | | |
| Command device (alpaha155) | | | | | | | | 00 | 9A | | |
| Command device (alpaha155) | | | | | | | | 01 | 8A | | |
| Command device (alpaha156) | | | | | | | | 00 | 9B | | |
| Command device (alpaha156) | | | | | | | | 01 | 8B | | |

**Table 5 device mapping table - including BC_1**

### 1b: Map BC devices to array port

BC and CA software requires that all devices are allocated to at least one port on the array. The port does not have to be physically connected to any server or be used by any server. It can be any available port on the array.

**Note:** If necessary, the BC devices can be mapped to RCU_Target ports as long as host access to the devices is not allowed through these ports. An RCU_Target port is similar to a standard target port but with the restriction that only 128 devices can be mapped instead of the standard 255 devices. HP does not allow host access on RCU target ports because of performance impact to CA software.

HP does not recommend that a host have access to both CA and BC devices, and does not support such an installation in any cluster configuration. If it is decided to implement an additional backup server on Site 2, then this host will have access to the BC devices. Backup servers are beyond the scope of this solution.

## 1c: Design RAID Manager instance mapping

HP assumes that there is no backup server implemented in this solution and that the BC devices are allocated to an unused port on the array. This will require the application hosts to manage the BC_1 and CA_1 devices. In the example configuration, separate RAID Manager XP instances manage the CA_1 devices and the BC_1 devices. This simplifies the configuration files in some way and ensures that a configuration error in one file does not affect the availability of the other device group. The following diagram shows the required RAID Manager XP instances to mange both the CA_1 and BC_ device groups.

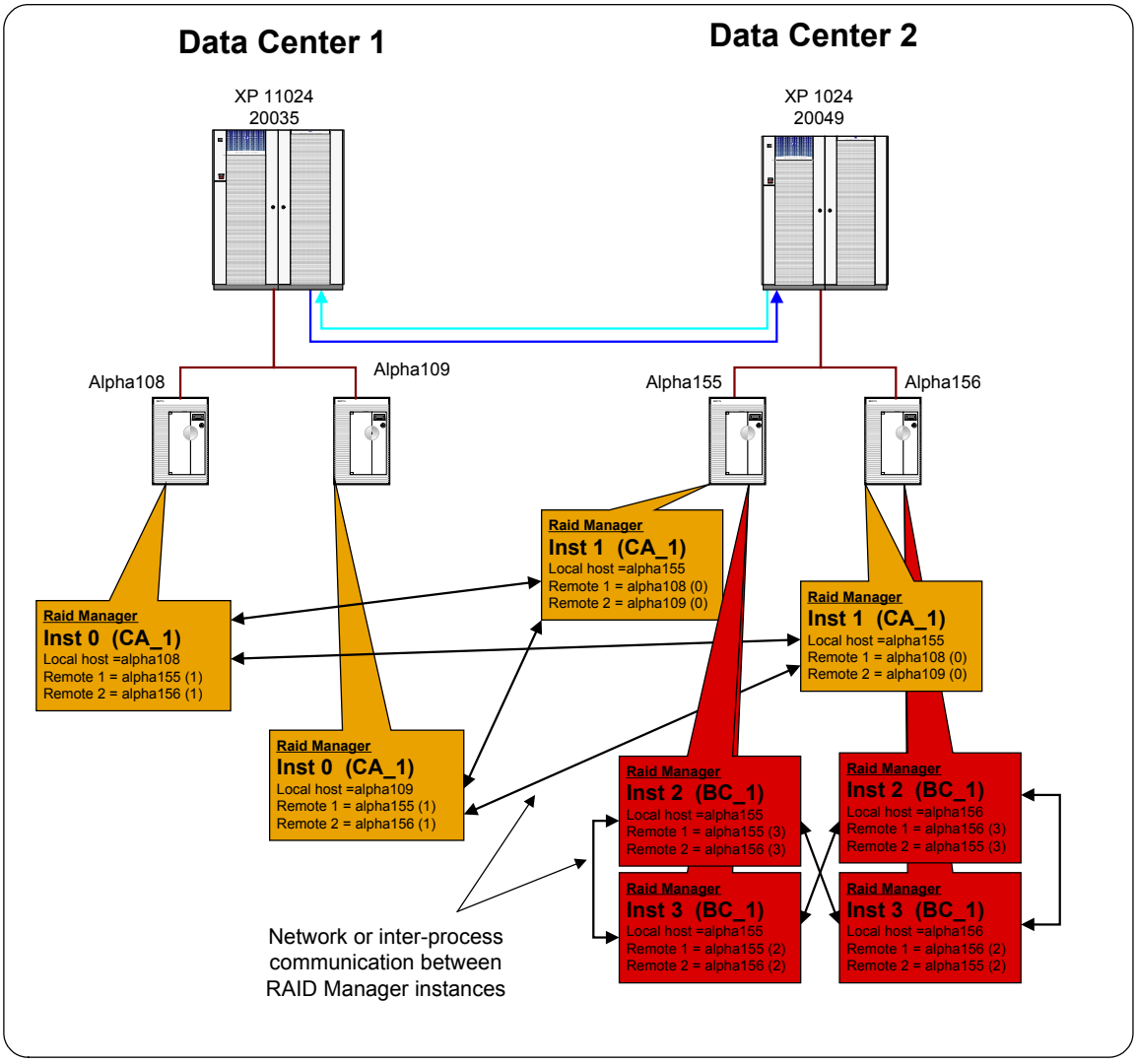



**Figure 13: RAID Manager XP instance mapping - CA_1 and BC_1**

# 2: Create the BC_1 pair

## 2a: Install RAID Manager

The hosts on Site 2 are also used for the CA_1 device pairs and therefore have RAID Manager XP already installed. If any new hosts are added to manage the BC_1 device group, ensure that the same RAID Manager XP version is installed on that host.

## 2b: Identify command devices

Using "**ioscan**" command, find the command devices allocated to the specific host (see Table 4 for the command device CU:ldev numbers). A command device is recognized by the "-CM" suffix to the device description. Find the associated device file for each command device. These command devices should be the same as the devices used for the CA_1 device groups.

## 2c: Create RAID Manager configuration files

On each host in the configuration create a temporary RAID Manager XP instance by defining only the instance variables and commands devices in the configuration file. RAID Manager XP configuration files are located in the **/etc** directory and are named **horcmxx.conf,** where the **xx** indicates the instance number. The following is an example of a temporary configuration file.

```
#/************************* For Local instance **************************/
HORCM_MON
#ip_address    service    poll(10ms)  timeout(10ms)
alpha155     horcm2     1000       3000

#/************************* For local command devices *******************/
HORCM_CMD
#dev_name       dev_name       dev_name
/dev/rdsk/c8t15d7
```

After creating the configuration file, be sure you add an entry to the **/etc/services** file to reserve a port number for each RAID Manager XP instance. The following is an example of the entries in the **/etc/services** file. Ensure that the entries are the same on all hosts in the configuration using RAID Manager XP.

```
horcm0          51000/udp   #RAIDManager instance 0
horcm1          51001/udp   #RAIDManager instance 1
horcm2          51002/udp   #RAIDManager instance 2
horcm3          51003/udp   #RAIDManager instance 3
```

Start the RAID Manager XP instance on the local host using the "**horcmstart.sh xx**" command, where **xx** is the instance number to start. After a successful startup of the RAID Manager XP instance, set the "**HORCMINST=xx**" environment variable to ensuring RAID Manager XP commands will connect to the correct RAID Manager XP instance.

Find the internal target and LUN numbers for each of the devices used by the application by using the "**raidscan -p <port_number> -fx**" command. Complete the configuration file by adding the device group name, device name, and port:Target:Lun number for each device. Define any remote instances that can manage the remote site of the device group and their services alias. The following is an example of a completed device group configuration file.

```
#ip_address    service    poll(10ms)  timeout(10ms)
alpha155     horcm2     1000      3000

#/************************* For local command devices ********************/
HORCM_CMD
#dev_name       dev_name       dev_name
/dev/rdsk/c8t15d7

#/************************** For device groups ***************************/
HORCM_DEV
#dev_group   dev_name   port#   TargetID  LU#   MU#
bc1_oracle     ora_data1  CL1-A          1     1        0
bc1_oracle     ora_data2  CL1-A          1     2        0
bc1_oracle     ora_index1      CL1-A          1     3        0
bc1_oracle     ora_index1      CL1-A          1     4        0
bc1_oracle     system1         CL1-A          1     5        0

#/*************** For remote instance definition ************************/
HORCM_INST
#dev_group   ip_address   service
bc1_oracle         alpha155    horcm3
bc_1oracle         alpha156    horcm3
```

**Note**: The configuration file for the BC_1 primary device is exactly the same as the configuration file for the CA_1 secondary devices. Copy the existing configuration file and change the services and remote instance information.

Stop and restart the instance using the "**horcmshutdown.sh xx**" and "**horcmstart.sh xx**" commands.

## 2d: Validate the RAID Manager configuration files

RAID Manager XP only reports the status of the devices defined within the configuration file of the RAID Manager XP instance used, so you must ensure that all instance configuration files address the correct number of devices as well as the correct logical devices. The configuration file uses the target:lun number of a device to refer to a logical device in the array. It is easy to mistype a value, therefore, it is important to verify the configuration file after any changes was made to the file. Be sure you export the correct **HORCMINST=xx** environment variable as well as the **HORCC_MRCF=1** variable to access the BC device group information with the RAID Manager XP commands. Use the **pairdisplay** command to check the group definition and to ensure that all the correct logical devices are used in the configuration.

### 2e: Create the BC_1 pair

After validating the RAID Manager XP configuration file, create the device pair and start the initial copy of data from the primary to the secondary devices within the array. Connect to one of the hosts on Site 2 and set the **HORCMINST** environment variable to the instance number that will manage the primary side of the device group (this should be 2 for this example). Be sure that the **HORCC_MRCF=1** environment variable is set. Use the **paircreate** command to create the device group. To utilize the new at-time suspend options from business copy, you must specify the **-m grp** option with the **paircreate** command. This will enable the consistent suspend of the BC device group using consistency groups in the array.

### 2f: Wait for pair status

After the **paircreate** command completes the devices start to copy all data from the primary device (P-vol) to the secondary device (S-vol) this is known as the initial copy process. During initial copy all data blocks on the source device are copied to the target device, even if the block is not actively used by the application or is empty. This process is required to ensure that the devices are exact replica of each other. Using the **pairdisplay** command, wait for the pair status to reach "**PAIR**" status.

## Results

At the end of this task, you should be able to suspend the BC_1 device pairs at any time creating a stable point-in-time image of the application.

# 5 Set Up Site 3

At this point you have a fully configured, two-node metro cluster with sync CA replication between Sites 1 and 2 and a BC copy implementation on Site 2, which enables you to create a point-in-time copy of the application on Site 3. Now you need to extend the solution to the third site. To do this you must first prepare the hosts and storage array on Site 3.

## 1: Set up site 3

### 1a: Define application requirements

The application requirements at Site 3 are the same as that for Site 2. Since you will implement two copies of the application on this site you will have the same configuration as that of Site 2. HP recommends, if possible, to utilize the same configuration and simplify the solution design process. Allocate the same number and size of devices on Site 3 for the CA_2 device pair, and update the device-mapping table with this information (table 5). Here is an example of the updated device mapping table.

**Application and CA/BC mappings for Multi Site Solution**

| Application | Volume group | Logical Volume | Physical device | Size (GB) | Emulation | Site 1 Array 20035 CA_1 CU | Site 1 Array 20035 CA_1 Ldev | Site 2 Array 20049 CA_1/BC_1 CU | Site 2 Array 20049 CA_1/BC_1 Ldev | Site 2 Array 20049 BC_1/CA_2 CU | Site 2 Array 20049 BC_1/CA_2 Ldev | Site 3 Array 20040 CA_2/BC_2 CU | Site 3 Array 20040 CA_2/BC_2 Ldev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | CL1-A and CI2-A | | CL1-A and CI2-A | | CL1-B | | CL1-A and CI2-A | |
| Oracle1 | vg_archive | lv_arch1 | /dev/dsk/c0t1d0 | 14 | Open-E | 00 | 01 | 00 | 20 | 00 | A0 | 00 | 30 |
| | | | /dev/dsk/c0t1d1 | 14 | Open-E | 00 | 04 | 00 | 21 | 00 | A1 | 00 | 31 |
| | | lv_arch2 | /dev/dsk/c0t1d2 | 14 | Open-E | 00 | 08 | 00 | 22 | 00 | A2 | 00 | 32 |
| | | | /dev/dsk/c0t1d3 | 14 | Open-E | 00 | 12 | 00 | 23 | 00 | A3 | 00 | 33 |
| | vg_data1 | lv_data1 | /dev/dsk/c0t1d4 | 9 | Open-9 | 01 | 10 | 01 | 35 | 01 | B0 | 01 | 45 |
| | | | /dev/dsk/c0t1d5 | 9 | Open-9 | 01 | 12 | 01 | 36 | 01 | B1 | 01 | 46 |
| | vg_data2 | lv_data2 | /dev/dsk/c0t1d6 | 9 | Open-9 | 01 | 14 | 01 | 37 | 01 | B2 | 01 | 47 |
| | | | /dev/dsk/c0t1d7 | 9 | Open-9 | 01 | 18 | 01 | 38 | 01 | B3 | 01 | 48 |
| | vg_data3 | lv_data3 | /dev/dsk/c0t2d0 | 9 | Open-9 | 01 | 16 | 01 | 39 | 01 | B4 | 01 | 49 |
| | | | /dev/dsk/c0t2d1 | 9 | Open-9 | 01 | 18 | 01 | 3A | 01 | B5 | 01 | 4A |
| | vg_data4 | lv_data4 | /dev/dsk/c0t2d2 | 9 | Open-9 | 01 | 20 | 01 | 3B | 01 | B6 | 01 | 4B |
| | | | /dev/dsk/c0t2d3 | 9 | Open-9 | 01 | 22 | 01 | 3C | 01 | B7 | 01 | 4C |
| | vg_index | lv_index | /dev/dsk/c0t2d4 | 14 | Open-E | 00 | 30 | 00 | 43 | 00 | A9 | 00 | 53 |
| | | lv_index2 | /dev/dsk/c0t2d5 | 14 | Open-E | 00 | 31 | 00 | 44 | 00 | AA | 00 | 54 |
| | vg_logA | lv_log1 | /dev/dsk/c0t2d6 | 3 | Open-3 | 02 | 32 | 02 | 53 | 02 | C0 | 02 | 63 |
| | | lv_log2 | /dev/dsk/c0t2d7 | 3 | Open-3 | 02 | 33 | 02 | 56 | 02 | C1 | 02 | 66 |
| | | lv_log3 | /dev/dsk/c0t3d0 | 3 | Open-3 | 02 | 34 | 02 | 57 | 02 | C2 | 02 | 67 |
| | | lv_log4 | /dev/dsk/c0t3d1 | 3 | Open-3 | 02 | 35 | 02 | 58 | 02 | C3 | 02 | 68 |
| | vg_log_b | lv_log1 | /dev/dsk/c0t3d2 | 3 | Open-3 | 02 | 62 | 02 | 59 | 02 | C4 | 02 | 69 |
| | | lv_log2 | /dev/dsk/c0t3d3 | 3 | Open-3 | 02 | 63 | 02 | 5A | 02 | C5 | 02 | 6A |
| | | lv_log3 | /dev/dsk/c0t3d4 | 3 | Open-3 | 02 | 64 | 02 | 5B | 02 | C6 | 02 | 6B |
| | | lv_log4 | /dev/dsk/c0t3d5 | 3 | Open-3 | 02 | 65 | 02 | 5C | 02 | C7 | 02 | 6C |
| | vg_admin | lv_error | /dev/dsk/c0t3d6 | 9 | Open-9 | 00 | 70 | 00 | 8A | 00 | 8D | 00 | 7A |
| | | lv_config | /dev/dsk/c0t3d7 | 9 | Open-9 | 00 | 71 | 00 | 8B | 00 | 8E | 00 | 7B |
| Command device (alpaha108) | | | | | | 00 | 9A | | | | | | |
| Command device (alpaha108) | | | | | | 01 | 8C | | | | | | |
| Command device (alpaha109) | | | | | | 00 | 9B | | | | | | |
| Command device (alpaha109) | | | | | | 01 | 8D | | | | | | |
| Command device (alpaha155) | | | | | | | | 00 | 9A | | | | |
| Command device (alpaha155) | | | | | | | | 01 | 8A | | | | |
| Command device (alpaha156) | | | | | | | | 00 | 9B | | | | |
| Command device (alpaha156) | | | | | | | | 01 | 8B | | | | |
| Command device (alpaha154) | | | | | | | | | | | | 00 | 9A |
| Command device (alpaha154) | | | | | | | | | | | | 01 | 8A |

**Table 6 device mapping table - including CA_2**

### 1b: Install and configure hosts

Using the same design and installation methods as for the hosts on Sites 1 and 2, plan and install the hosts on Site 3. This host will be part of the continental cluster and will therefore be part of its own local cluster. For this reason, this node does not have to be in the same network subnet as the hosts in Sites 1 and 2.

Install the operating system with the same file system sizes as the systems in Site 1. Install all the additional applications and software on the hosts systems in the same sequence as that of Site 1, and ensure all patches are installed on the systems. If the host/hardware is not exactly the same as those in Site1 it might be necessary to install some host/hardware specific patches in addition to the patches installed on Site 1. If this host will provide quorum services to the metro cluster on Site 2, install the quorum server software and update the metro cluster with the new quorum server information.

### 1c: Design and install host-to-storage communication

The host-to-storage communication must be designed in the same way as Site 1. It is possible to have different model switches on this site as long as the switches are still supported by the SAN streams documents as well as the extender/converters used later in the installation processes. It is possible to use more or less host-to-storage connections if required, but HP recommends using the same design processes for each site to ensure similar performance characteristic on each site.

### 1d: Install and configure arrays

Using the same steps as in Site 2, design and install the array on Site 3. It will be necessary to format the array similar to the array on Site 2 to ensure that the correct number and size devices are available for the CA link. Keep in mind that this array will have both CA and BC devices for the application and therefore must have double the configuration of that in Site 1 (the configuration will be similar to that of Site 2).

Ensure that all array software licenses are installed and active on the array. Because this array will use asynchronous replication, it is necessary to have both the CA and CA extension licenses installed on this array.

### 1e: Install the application

The only installation required on Site 3 for the application is to install the basic application software on this site (if this is not shared between sites). The remaining application configuration will be completed after the CA link between Sites 2 and 3 is established and the data is copied from Site 2 to 3.

Update Table 6 ensuring that there are sufficient devices available for application on Site 3 and that all hosts on Site 3 have access to these devices.

## Results

At the end of this step you will have a fully configured host with access to the necessary devices on the array at Site 3. The application installation will be complete after the mirror process is complete.

# 6 Implement the CA_2 Link Between Sites 2 and 3

For the long distance link, HP recommends a dedicated private virtual circuit with guaranteed service level agreements on latency, packet loss, and bandwidth. Minimum usable throughput on the link must be 0.5 MB/s or higher. HP recommends at least 1 MB/s. This will determine the throughput on the long-distance link and will effect the time to copy data to the remote site. Adjust the minimum values to the size and performance of the application.

FC-to-IP converters are available in 1x1 (one fiber input to one network output) or 2x2 (two fiber inputs to two network outputs) units. When using 2x2 units, a higher level of redundancy is achieved as well as some load balancing. However, HP recommends using at least two converter units on each side for redundancy. The network link can either be shared by the two units (using network switches) or each pair can have its own dedicated network.

## 1: Design and install physical/logical links between sites

### 1a: Identify "Initiators" and "RCU Targets"

The physical communication between the arrays requires a physical path from an "Initiator" port on one array to an "RCU Target" port on the other array. "Initiator" ports are sender or talker ports and initiate communication while the "RCU Target" port is a receiver or listener port. The communication path from an "Initiator" to an "RCU Target" provides one-way communication from one array to another. To communicate in both directions and to fail over applications automatically between arrays, it is necessary to create an additional "Initiator" to "RCU Target" path in the opposite direction. The ability of a port to act as an "Initiator" or "RCU Target" is determined by the processor controlling the ports. On some of the older arrays, a processor managed two adjacent ports and therefore both the ports managed by the processor must be set to the same function.

Carefully plan the "Initiator" to "RCU Target" configurations. Pay attention to high availability and any performance requirements. The same switches can be shared with host and CA communication as long as the host and CA ports are zoned in separate zones. When using switches, the switch combines the "Initiator" and "RCU Target" communication into one ISL that must be extended between sites. Ensure that the ISL has sufficient bandwidth available for the CA traffic and does not overload the ISL with multiple "Initiators." Although the ISL cannot be zoned only for CA traffic, it is not advisable and not supported to have host access over the ISL between sites. Allowing a host on Site 2 to access storage on Site 3 can become problematic during site failures, and allowing hosts to perform large transactions over the ISL used for CA communications can negatively impact the CA communication.

It is possible to share the "Initiator" and "RCU Target" ports on Site 2 with both the short-distance CA *1 link and the lon- distance CA*2 link. In doing this, the switch fabric will span all three sites. This may complicate the switch management. HP recommends that when spanning the three sites with one fabric, the host access to the fabric is minimized or the fabric is dedicated to CA traffic.

### 1b: Define intersite links

Because the long-distance requirements between Sites 2 and 3, it is not possible to implement direct fiber optic connections between the sites. The latency and signal strength of fiber optic transmission does not allow for these long distances. For this reason, you must convert the fiber optic protocol to a network protocol and propagate the I/O using standard network technologies. A dedicated virtual circuit allows for specific performance characteristics on this network. At the remote site, the network protocol is converted to optical again and completes the link between the sites.

After the physical link between the sites is installed, it is possible to connect the converter systems to the network link and ensure that the communication between converter systems is operational. This test can be performed with the "**ping**" command on the converters. The command will also indicate the average response time on the network. The next step is to connect the switches to the converters with each switch on its own port. It is important to document the port used on the converter since this will determine the connection for the remote switch.

Before connecting the switches to the converters it is necessary to merge the two (or three) switch configurations and zoning information since the connections will create one central fabric between the switches. It is recommended to perform the connection with the application down since the merging of the zones may disrupt host-to-storage communications.

The following diagram shows the interconnections between Site 1, 2, and 3 using DWDM extender and FC-to-IP converters. For redundancy the two WAN links must be routed by different routs. The configuration creates two independent fabric configurations between the switches. Each fabric provides an "Initiator" to "RCU Target" communication channel in both directions.

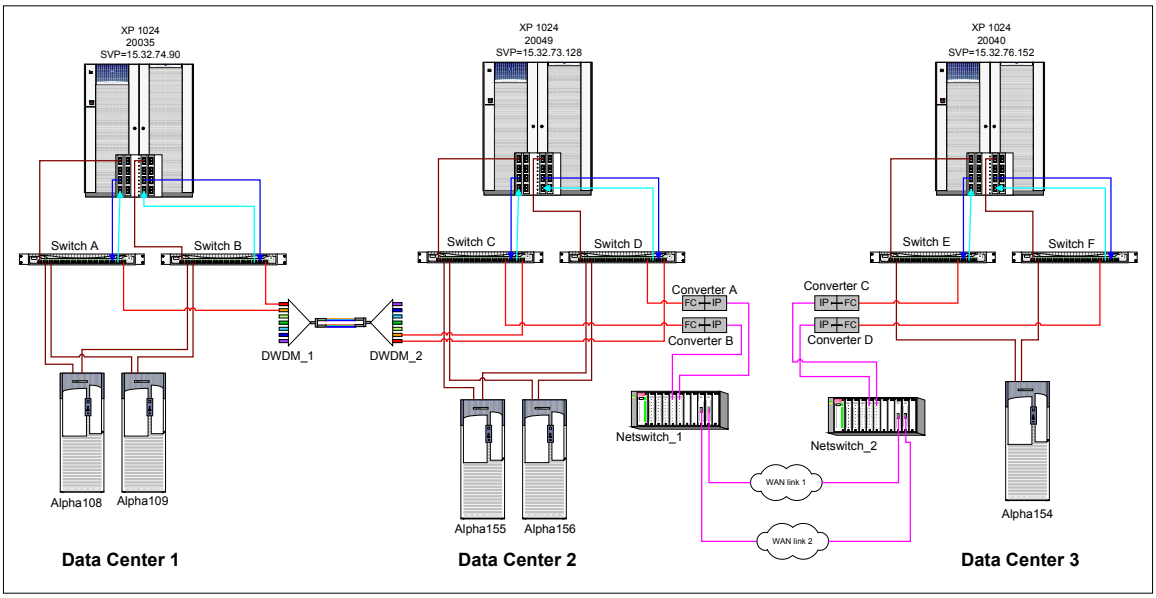



**Figure 14: Physical connections between Site 1,2 and 3**

## 1c: Design switch zoning

When connecting the switches on Sites 1 and 2 with those on Site 3, a single fabric will be established between the sites with shared configuration and zoning information. To create the shared fabric, you must ensure that the configuration setting on all switches are similar and that each switch has its own unique domain ID number. This can be performed by telnet into the switch and changing the configuration settings.

**Note:** Changing the domain ID number on the switch may invalidate any zone configuration on the switch.

The best way to ensure that the configurations are the same is to download the configuration from each switch and compare all parameters. Change all the parameters to a common set, and then upload the new configuration. To complete this operation, the switches must be disabled. It is possible to continue operations through one set of switches while upgrading the other set by using alternative paths set on the host. However,

to reduce any risk of interrupting production systems, HP recommends that you schedule downtime for this operation.

**Note:** To change configuration settings, the switch must be offline, which suspends all hosts-to-storage operations.

To merge the zone information, it is necessary to delete all zone configurations on one of the switches (preferably the one on Site 3), then complete the interconnection of the switches and redefine the zones deleted earlier. The new configuration will now be shared between both switches. You must define switch port zoning to ensure that the host systems do not see any devices at the other site. The use of either hard port zoning or worldwide name zoning is allowed. ISL links cannot be zoned only for CA access because ISL ports are open to be used by any port that is in a zone covering multiple switches. The following diagram shows the completed zone configuration for the three sites. Note that there are two independent fabrics available between the sites.

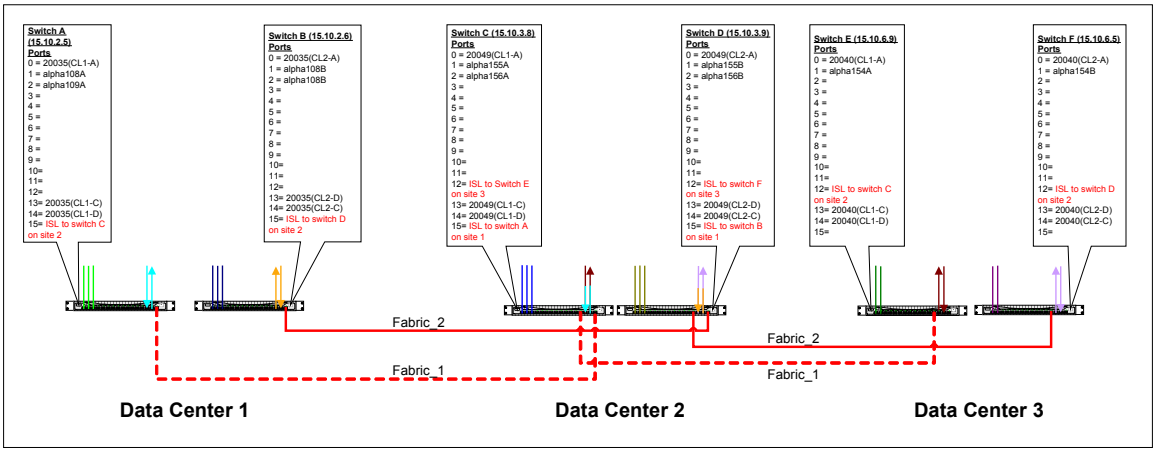



**Figure 15: Sample zone mapping between Site 1,2 and 3**

**Fabric 1** (Switch_A (1) + Switch_C (3) + Switch_E (5))

| Ref | Zone Name | Switch | Port | Function |
|---|---|---|---|---|
| | DC1_HP1 | Switch_A | 0<br>1<br>2 | HP hosts in DC 1 Primary link |
| | DC2_HP1 | Switch_C | 0<br>1<br>2 | HP hosts in DC 2 Primary link |
| | DC3_HP1 | Switch_E | 0<br>1 | HP hosts in DC 3 Primary link |
| | CA1_DC1_DC2 | Switch_A<br>Switch_C | 13<br>14<br>13<br>14 | CA link between DC1 and DC2 |
| | CA3_DC2_DC3 | Switch_C<br>Switch_E | 13<br>14<br>13<br>14 | CA link between DC2 and DC3 |

**Fabric 2** (Switch_B (3) + Switch_D (4) + Switch_F (6) )

| Ref | Zone Name | Switch | Port | Function |
|---|---|---|---|---|
| | DC1_HP2 | Switch_B | 0<br>1<br>2 | HP hosts in DC 1 Secondary link |
| | DC2_HP2 | Switch_D | 0<br>1<br>2 | HP hosts in DC 2 Secondary link |
| | DC3_HP2 | Switch_F | 0<br>1 | HP hosts in DC 3 Secondary link |
| | CA2_DC1_DC2 | Switch_B<br>Switch_D | 13<br>14<br>13<br>14 | CA link between DC1 and DC2 |
| | CA4_DC2_DC3 | Switch_D<br>Switch_F | 13<br>14<br>13<br>14 | CA link between DC2 and DC3 |

**Figure 16: Switch zone information**

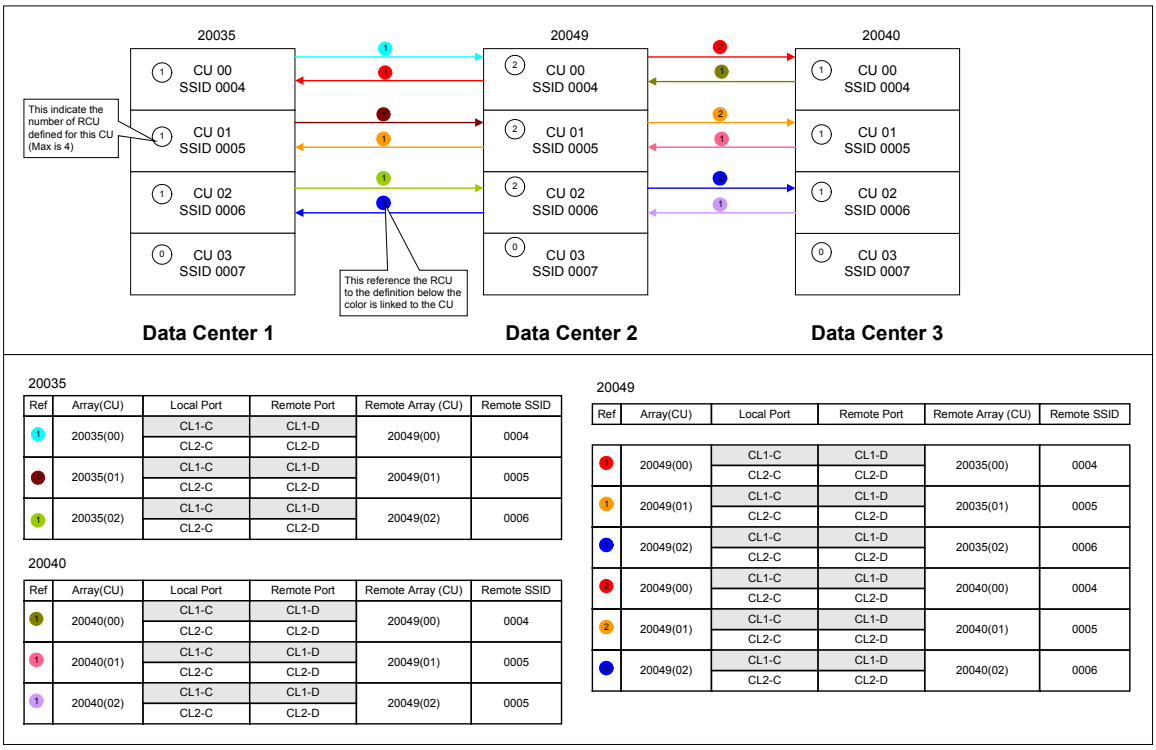
## 1d: Design CU/RCU mapping

On the XP array, you format physical devices (array group) with certain emulation types and logical sizes, and then allocate these logical devices to a control unit (one CU can address 255 logical devices). Then assign a logical device from the CU to a host port with a specific target and LUN number. Use devices from several CUs rather than only one unit since most of the devices in one CU are normally located on the same physical drive. When using CA to replicate data between arrays, it is necessary to create a logical path

between arrays called a remote control unit (RCU). An RCU must be configured between a source CU (initiator) and a target CU to create a CA device pair using devices controlled by the specific CUs. For full failover capabilities, configure a reverse RCU from the target array to the source array. Any CU can only support four RCU definitions. An RCU definition can define one or more physical paths between the arrays as communication paths for the device pairs. All I/O will be shared between these paths, and failure of one path does not affect the pair status, only the mirror performance.

**Note:** To ensure optimal performance, HP recommends that alternative paths between the arrays have the same performance characteristics. The use of different technologies to create the links between arrays might negatively affect the CA performance.

To replicate the entire application, you must configure RCU links for every CU the application is using. Using the application to device mapping (Table 6), determine all the CUs in each array that require a logical path between them. In the mapping of the devices groups (Table 6), you took great care to ensure that you pair devices from the same CU group throughout the configuration. This may not always be possible, but HP recommends staying within the same CU for each device pair and using the BC devices to switch to another CU within a site. In the following sample, an RCU link exists between CU 0 on array 20035 and CU 0 on array 20049. A reverse RCU link exists between CU 0 on 20049 and CU 0 on array 20035. In the same way an RCU exists between CU 0 on array 20049 and CU 0 on array 20040 with a reverse RCU from CU 0 on 20040 to CU 0 on 20049. Similar RCU links exist for CU 01 and 02. The array at Site 2 now has two RCU defined for each CU. The maximum for a CU is four RCUs.




20035

| Ref | Array(CU) | Local Port | Remote Port | Remote Array (CU) | Remote SSID |
|---|---|---|---|---|---|
| 1 | 20035(00) | CL1-C | CL1-D | 20049(00) | 0004 |
| | | CL2-C | CL2-D | | |
| 1 | 20035(01) | CL1-C | CL1-D | 20049(01) | 0005 |
| | | CL2-C | CL2-D | | |
| 1 | 20035(02) | CL1-C | CL1-D | 20049(02) | 0006 |
| | | CL2-C | CL2-D | | |

20040

| Ref | Array(CU) | Local Port | Remote Port | Remote Array (CU) | Remote SSID |
|---|---|---|---|---|---|
| 1 | 20040(00) | CL1-C | CL1-D | 20049(00) | 0004 |
| | | CL2-C | CL2-D | | |
| 1 | 20040(01) | CL1-C | CL1-D | 20049(01) | 0005 |
| | | CL2-C | CL2-D | | |
| 1 | 20040(02) | CL1-C | CL1-D | 20049(02) | 0005 |
| | | CL2-C | CL2-D | | |

20049

| Ref | Array(CU) | Local Port | Remote Port | Remote Array (CU) | Remote SSID |
|---|---|---|---|---|---|
| 1 | 20049(00) | CL1-C | CL1-D | 20035(00) | 0004 |
| | | CL2-C | CL2-D | | |
| 1 | 20049(01) | CL1-C | CL1-D | 20035(01) | 0005 |
| | | CL2-C | CL2-D | | |
| 1 | 20049(02) | CL1-C | CL1-D | 20035(02) | 0006 |
| | | CL2-C | CL2-D | | |
| 2 | 20049(00) | CL1-C | CL1-D | 20040(00) | 0004 |
| | | CL2-C | CL2-D | | |
| 2 | 20049(01) | CL1-C | CL1-D | 20040(01) | 0005 |
| | | CL2-C | CL2-D | | |
| 2 | 20049(02) | CL1-C | CL1-D | 20040(02) | 0006 |
| | | CL2-C | CL2-D | | |

**Figure 17: Sample RCU configuration for Site 1,2 and 3**

**Note:** Each CU have a SSID number defined on the array. This number is required when creating the RCU.

### 1e: Configure RCUs

**With the physical paths between the arrays configured and installed and the logical paths designed, it is time to create the RCUs between the arrays. This is done using the Command View array console software or the internal laptop of the array (SVP). Ensure that all ports are configured as Initiators and RCU Targets as defined, using the port configuration menus. Create the RCU specifying the remote array serial numbers and CU numbers as well as the SSID of the remote CU. Specify all paths to the remote array as well as the minimum paths that are required to keep the link operational. HP recommends that the minimum path is set to 1 to ensure that the CA link remains operational during link failures until there is no more link available between the arrays. Check the RCU status to ensure that all paths show operational status. Repeat the process for each CU on the source array as well as for all CUs on the target array, creating logical paths in both directions.**

## Results

At the end of this task you have a physical and logical path configured between the array at Site 2 and the array at Site 3. At this stage you are ready to start the replication process between these arrays.

# 2: Create the CA_2 pair

At this stage, it is assumed that the physical connections between Sites 2 and 3 exist and that all RCUs have been configured. Creating the pair relationship between devices can be performed from either the array's service processor (SVP) or from the host using RAID Manager XP. Since the RAID Manager XP configuration is required for the cluster configuration and for management of the solution, HP recommends using the RAID Manager XP to configure the device pairs.

You can only create the CA_2 pair when the BC_1 pair is in "suspend" status. Use the RAID Manager XP **`pairsplit`** command to suspend the BC_1 pair before starting this process.

## 2a: Install RAID Manager

If the RAID Manager XP software was not installed during the host installation then it must be installed at this point. Follow the instruction in the "readme" file distributed with the software for installation procedures. Ensure that the latest version of RAID Manager XP is installed and that the version is at least version 01.11.00 or later. The RAID Manager XP software is required to be installed on all hosts in the configuration on Sites 1, 2, and 3.

## 2b: Design RAID Manager instance mapping

When creating device pairs with RAID Manager XP, you must configure at least two instances of RAID Manager XP, one instance to manage the primary devices (P-vol) and one instance to manage the secondary devices (S-vol). In the case of CA pairs the two instances will be located on two different hosts: one at Site 2 and one at Site 3. RAID Manager XP can automatically attempt communications with another instance if the first instance fails to respond. In the example, there are two hosts on Site 2 and one on Site 3, each able to run an RAID Manager XP instance. The two hosts on Site 2 each have an instance that can manage the primary devices of the pair and attempt to communicate with the instances on Site 3.

When mapping RAID Manager XP instances, try to keep them in sequence; for example, 0 = Application, 1 = CA_1, 2 = BC_1, and so on. RAID Manager XP runs on any host attached to the array. RAID Manager XP does not need to see all the devices, just the command device for the array.

The following diagram shows the RAID Manager XP instances required to create and manage the CA_1, BC_1, and CA_2 device pairs among Sites 1, 2, and 3.

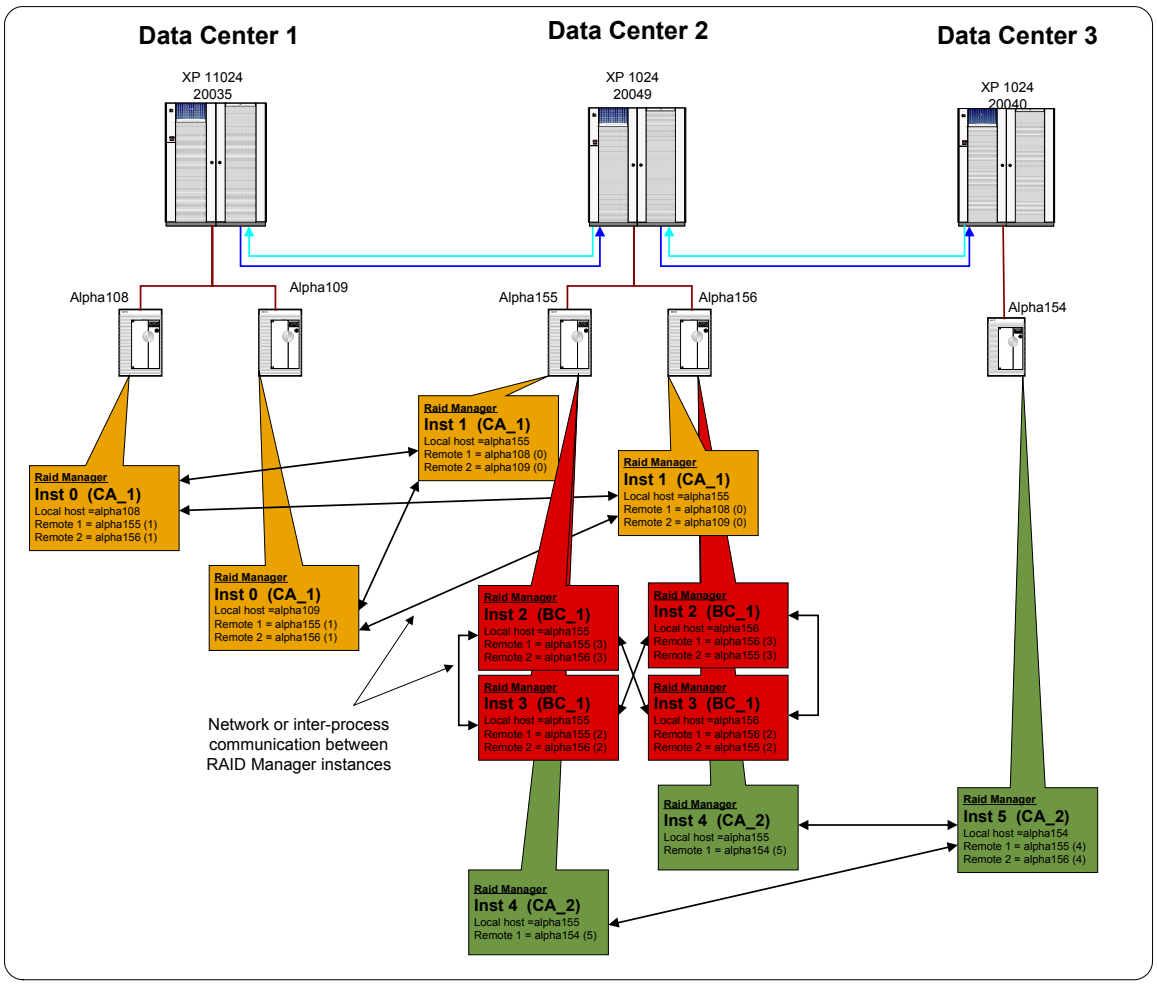



**Figure 18: RAID Manager instance for CA_1, BC_1, and CA_2 pairs**

## 2c: Identify command devices

Using **"ioscan"** command find the command devices allocated to the specific host (see Table 6 for the command device CU:ldev numbers). A command device is recognized by the "-CM" suffix to the device description. Find the associated device file for each command device.

## 2d: Create RAID Manager configuration files

On each host in the configuration, create a temporary RAID Manager XP instance by defining only the instance variables and commands devices in the configuration file. RAID Manager XP configuration files are located in the **/etc** directory and are named **horcmxx.conf,** where **xx** indicates the instance number. The following is an example of a temporary configuration file.

*#/************************ For Local instance **************************/*

*HORCM_MON*

*#ip_address service poll(10ms) timeout(10ms)*

*alpha154 horcm5 1000 3000*

```
#/************************* For local command devices ********************/
HORCM_CMD
#dev_name      dev_name      dev_name
/dev/rdsk/c3t15d7
```

After creating the configuration file, be sure to add an entry to the **/etc/services** file to reserve a port number for each RAID Manager XP instance. The following is an example of the entries in the **/etc/services** file. Be sure the entries are the same on all hosts in the configuration using RAID Manager XP.

```
horcm0          51000/udp   #RAIDManager instance 0
horcm1          51001/udp   #RAIDManager instance 1
horcm2          51002/udp   #RAIDManager instance 2
horcm3          51003/udp   #RAIDManager instance 3
horcm4        51004/udp   #RAIDManager instance 4
horcm5          51005/udp   #RAIDManager instance 5
```

Start the RAID Manager XP instance on the local host using the "**horcmstart.sh xx**" command, where **xx** is the instance number to start. After a successful startup of the RAID Manager XP instance set, the "**HORCMINST=xx**" environment variable ensuring RAID Manager XP commands will connect to the correct RAID Manager XP instance.

Find the internal target and LUN numbers for each of the devices used by the application by using the "**raidscan -p <port_number> -fx**" command. Complete the configuration file by adding the device group name, device name, and port:Target:Lun number for each device. Define any remote instances that can manage the remote site of the device group and their services alias. The following is an example of a completed device group configuration file.

```
#ip_address   service   poll(10ms)  timeout(10ms)
alpha154    horcm5    1000      3000

#/************************* For local command devices ********************/
HORCM_CMD
#dev_name      dev_name      dev_name
/dev/rdsk/c3t15d7

#/************************** For device groups ***************************/
HORCM_DEV
#dev_group   dev_name   port#  TargetID  LU#   MU#
ca2_oracle     ora_data1  CL1-A          4     1           0
ca2_oracle     ora_data2  CL1-A          4     2           0
ca2_oracle     ora_index1     CL1-A          4     3           0
ca2_oracle     ora_index1     CL1-A          4     4           0
ca_2oracle     system1        CL1-A          4     5           0
```

```
#/*************** For remote instance definition ************************/
HORCM_INST
#dev_group    ip_address   service
ca2_oracle          alpha155    horcm4
ca2_oracle          alpha156    horcm4
```

Stop and restart the instance using the "**horcmshutdown.sh xx**" and "**horcmstart.sh xx**" commands.

## 2e: Validate the Raid Manager configuration files

RAID Manager XP only reports the status of the devices defined within the configuration file of the RAID Manager XP instance used, so you must ensure that all instance configuration files address the correct number of devices as well as the correct logical devices. The configuration file uses the target:lun number of a device to refer to a logical device in the array. It is easy to mistype a value. therefore, it is important to verify the configuration file after any changes are made to the file. Use the **pairdisplay** command to verify the group definition and to ensure that all the correct logical devices are used in the configuration.

## 2f: Create the CA_2 pair

After validating the RAID Manager XP configuration file, you can create the device pair and start the initial copy of data from Site 2 to Site 3. Connect to the hosts on Site 2. Set the **HORCMINST** environment variable to the local instance number (this should be 4 for this example). Use the **paircreate** command create the device group. When creating the device group, it is necessary to specify the fence level of the device group. For the CA_2 link, the fence level should be **Async**.

## 2g: Wait for pair status

After the paircreate command completes, the devices will start to copy all data from the local site (Site 2) to the remote site (Site 3). This is known as the initial copy process. Depending on the performance of the links between the sites and the total disk space used by the application, this process can take between a few hours to a number of days to complete. During initial copy all data blocks on the source device are copied to the target device, even if the block is not actively used by the application or is empty. This process is required to ensure that the devices are exact replica of each other. Using the **pairdisplay** command, wait for the pair status to reach "**PAIR**" status.

# Results

At the end of this task, you have a replica of the point-in-time copy of the production data available on Site 3.

# 7 Create the Cluster for Site 3

## 1: Install Cluster software

If the cluster software was not installed during the host installation, then you must install the software at this point. If software was installed earlier, check the installed software on all hosts.

### 1a: Install ServiceGuard

The first step is to install the base ServiceGuard software on each host in the configuration that will be part of the cluster on Site 3. Follow the installation instructions in the product documentation. While the data is being synchronized between Sites 2 and 3, you can perform the cluster installation on the hosts. When the device pair reaches "pair" status, the cluster configuration can be completed. Be sure you install all required patches for the base cluster software on all hosts as well as any hardware specific patches for cluster usage.

### 1b: Install MetroCluster XP-CA toolkit

Install the MeroCluster and XP-CA toolkit on all hosts in the configuration. Be sure you install all the latest MetroCluster patches on each host. The XP-CA toolkit will be used to manage the local CA devices during application startup.

## 2: Configure the cluster

The following are the main steps in the process of configuring a cluster with XP-CA toolkit. For detail instructions, refer to the appropriate installation manuals.

### 2a: Create the cluster

The first step in configuring the cluster is to create a basic cluster configuration file identifying all nodes and hardware components in the cluster. Use the **`cmquerycl`** command to create a template configuration file with all the appropriate node information added to the file. Edit the file created by the **`cmquerycl`** command, and modify the cluster name and package configuration. Check the network interface definitions, and ensure that the correct network is used for the heartbeat. After all modification is complete, check the configuration file using the **`cmcheckconf`** command. After a successful check, the configuration can be applied using the **`cmapplycnf`** command. All the cluster commands can be executed from any node in the cluster; however, selecting one node and the configuration node is recommended.

### 2b: Start the cluster

When the **`cmapplyconf`** successfully delivers the cluster configuration to all nodes in the cluster, the cluster can be started using the **`cmruncl`** command. Use the **`cmviewcl`** command to check the cluster status, and ensure all nodes and network interfaces are available. Edit the **`/etc/rc.config.d/cmcluster`** file and enable automatic startup of the cluster services on each of the nodes. This ensures that the node automatically rejoins the cluster after a failure or reboot.

## 3: Configure cluster package

In the cluster configuration the application and all application resources are grouped together in a package. The package can then be moved from one node to the other either manually or automatically. Package configuration exists out of two sections: a package configuration file and a package script file.

### 3a: Create package configuration file

Use the **cmcreatepkg** command to create the package configuration file. The file is normally located in the **/etc/cmcluster/<package_name>/<package_name>.ascii** file. This file defines the package name, nodes that can run the application, failover policies, resources, and startup values. After editing the default package configuration file, use the **cmapplyconf** command to add the package to the cluster configuration. Any changes to this file must be applied to the cluster using the **cmapplyconf** command. The package configuration file will be similar to the one in the metro cluster environment, but automated startup should be disabled for this package.

### 3b: Create package script file

Use the **cmcreatepkg** command to create the default package script file for the package. Edit this file and add all the volume groups and logical volumes to the file that the application has to use. Customize the resources and application startup/shutdown commands. This is the script that is executed as soon as the application starts on the node to ensure that all resources are available for the application. The script by itself does not check the CA mirror status, and requires the CA mirror configuration file to enable mirror checking. If the package would be started without the CA toolkit configuration file, the status of the mirror pairs will not be checked and the application may fail to start. For more detail on the package startup script, refer to the cluster documentation. This file is normally located in the package directory **/etc/cmcluster/<package_name>** and is called **<package_name>.sh.** This file must be manually copied to all nodes in the cluster. Any changes must be done to all nodes in the cluster. The package script will be similar to the one in the metro cluster with possible different values for the IP interfaces in the package.

### 3c: Create CA-XP toolkit configuration file

The CA-XP toolkit configuration file indicates to the package script that some hardware replication is in place and forces the script to first check the mirror status before starting the application. The configuration file is located in the package directory and is called **<packge_name>_caxp.env.** The configuration file has a number of parameters that determine application startup condition.. Refer to the toolkit documentation for a detailed description of each parameter. The file also contains the RAID Manager XP instance number and device group name that manage the application data.

## 4: Prepare application and package for cluster

In this step, ensure that all application resources are available on all the hosts.

### 4a: Distribute volume group information

During the metro cluster installation all the hosts have been configured with the volume group information. The volume group detail is now part of the point-in-time image on the BC_1 devices and has been copied to Site 3 using the CA_2 device group. When the CA_2 device group reaches pair status, the volume group information is available on Site 3 and must be added to the OS on Site 3.To enable the application to run on the nodes in the cluster on Site 3, import the volume group information on all the local nodes. Most volume groups are self contained, meaning that the volume group ID is written in the reserve areas of the devices, although custom naming conventions are not maintained on the device. Use the **vgexport -s -p -m <mapfile.map> <vg_name>** command on the primary host to create a map file for each volume group. The map file contains the volume group id (provided by the –s option) and a map of custom names to logical segments on the devices. The –p option will prevent the volume group from being permanently exported from the system. Copy the map file to all hosts in data center 3. To be able to read volume group ID from the local devices on each host, suspend the CA link with read/write access to all devices. Use the RAID Manager XP **pairsplit** command with a **-rw** option to enable all the CA_2 devices for read/write access and to suspend mirroring of the devices. Create the device files for the volume groups using the **mknod** command, and then use the **vgimport** command to import the volume group information on each node.

**Note:** Do not activate the volume groups on more than one system at any time. Doing this will cause file system corruption.

After the volume group definitions are imported to all nodes, ensure that the volume group is disabled on all nodes and reactivate the CA device mirroring using the **`pairresync`** command.

### 4b: Enable volume group for cluster awareness

The local volume is a replication of the productions system, and the production system already has the cluster aware bit set. The cluster aware bit wrote a cluster ID to the volume group to identify the cluster that controls the device group. The local cluster is different from the metro cluster, therefore it would not be possible to activate the local device unless the cluster ID is changed on the device. To change the cluster ID, use the **`vgchange -c n`** command to remove the cluster ID and then **`vgchange -c y`** command to write the new ID. The package is now ready to be started on this node.

**Note:** Since the CA_2 devices are only suspended, a resume of the CA_2 device pair will overwrite the cluster ID with that from the metro cluster. During the continental cluster application startup the continental cluster software will change the cluster ID.

## 5: Test the cluster

Testing the application in this environment is not easy since starting the application in production mode will result in the takeover on the CA_2 link, and recovering from this takeover can take some time. It is possible to start the application on the local devices without test for CA functions by renaming the **`caxp.env`** file and then starting the application. Do not change any data during this test since all changes will be lost after the CA_2 device pair is resumed.

## Results

At the end of this step you have successfully installed and configured a local cluster on Site 3.

# 8 Create the Continental Cluster

## 1: Install continental cluster software

If the continental cluster software was not installed during the host installation, you must install the software at this point. If software was installed earlier, verify the installed software on all hosts in data center 1, 2, and 3.

## 2: Update the package configuration

To ensure that the package will not start automatically when the cluster is formed, disable automated startup of the package on all cluster nodes. This can be done by editing the package configuration file and setting the AUTO_RUN parameter to NO. Use the **`cmapplyconf`** command to apply this change to the cluster. Perform this in both the metro cluster configuration and the local cluster on Site 3.

## 3: Create monitor package

Using the template scripts provided with continental cluster software, create a monitor package in the cluster on Site 3. This package will monitor the metro cluster on Sites 1 and 2. If necessary a monitor package can be added on the metro cluster to monitor the cluster on Site 3. The template scripts are located in the **`/opt/cmconcl/scripts`** directory. See continental cluster software manuals for more detail.

## 4: Configure the cluster

Create the continental cluster config file using the **`cmqueryconcl -C`** command. Edit the configuration file and update all the clusters that must be monitored as well as the hosts in the cluster. Example output of this section of this file:

*CONTINENTAL_CLUSTER_NAME multi_site_cc*

*CLUSTER_NAME multi_site_1*
*CLUSTER_DOMAIN rose.hp.com*
*NODE_NAME alpha108*
*NODE_NAME alpha109*
*NODE_NAME alpha155*
*NODE_NAME alpha156*
*MONITOR_PACKAGE_NAME ccmonpkg_1*
*MONITOR_INTERVAL 60 SECONDS*

*CLUSTER_NAME multi_site_2*
*CLUSTER_DOMAIN rose.hp.com*
*NODE_NAME alpha154*
*MONITOR_PACKAGE_NAME ccmonpkg_2*

*MONITOR_INTERVAL 60 SECONDS*

Define the packages to be monitored and the package names in each cluster. Add this information to the continental cluster configuration file.

*RECOVERY_GROUP_NAME fs1*
*PRIMARY_PACKAGE multi_site_1/fs1*
*RECOVERY_PACKAGE multi_site_2/fs1_dc3*

Finally, add the monitor intervals and messages to the configuration file.

*CLUSTER_EVENT multi_site_1/UNREACHABLE*
*MONITORING_CLUSTER multi_site_2*
*CLUSTER_ALERT 0 MINUTES*
*NOTIFICATION CONSOLE "multi_site_1 is unreachable 0 minutes notification"*

Verify the configuration file using the **`cmcheckconcl -v -C <config_file>`** command, then apply the config file using the **`cmapplyconcl`** command.

# 5: Test the cluster

Use the **`cmviewconcl`** command to check the configuration.

*# cmviewconcl*

*CONTINENTAL CLUSTER multi_site_cc*

*RECOVERY CLUSTER multi_site_2*

*PRIMARY CLUSTER STATUS EVENT LEVEL POLLING INTERVAL*
*multi_site_1 up normal 1 min*

*PACKAGE RECOVERY GROUP fs1*

*PACKAGE ROLE STATUS*
*multi_site_1/fs1 primary up*
*multi_site_2/fs1_dc3 recovery down*

# Results

At the end of this task, you have a multi-site configuration with both metro and continental cluster software ensuring that the application can only run at one site and one host at any time.

# 9 Implement the BC_2 BusinessCopy at Site 3

At this stage, you have a multi-site implementation with cluster solution in place. During the resync processes for data from Site 2 to 3, the S-vol of the CA_2 pair is invalid leaving Site 3 vulnerable to any disasters. HP recommends implementing a BC copy on Site 3 to maintain a save copy of data on Site 3 at all times. The following step describes the process of creating the business copy at Site 3.

## 1: Plan BC_2 configuration

### 1a: Define application requirements

In the CA_2 pair, you paired each one of the devices used by the application on the second site with a device on the third site. You now have to pair each one of these third site devices with another device on this site and create a business copy pair with these devices. The first step in this process is to identify all the devices that will be used in the BC device pair and add them to the table created with the CA_2 pair (Ttable 6). Identify devices of the same emulation and size and record there CU:Ldev numbers in the table. After this step the updated table will look like the following.

**Application and CA/BC mappings for Multi Site Solution**

| Application | Volume group | Logical Volume | Physical device | Size (GB) | Emulation | Site 1 Array 20035 CA_1 CL1-A and CI2-A CU | Ldev | Site 2 Array 20049 CA_1/BC_1 CL1-A and CI2-A CU | Ldev | BC_1/CA_2 CL1-B CU | Ldev | Site 3 Array 20040 CA_2/BC_2 CL1-A and CI2-A CU | Ldev | BC_2 CL1-B CU | Ldev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Oracle1 | vg_archive | lv_arch1 | /dev/dsk/c0t1d0 | 14 | Open-E | 00 | 01 | 00 | 20 | 00 | A0 | 00 | 30 | 00 | B0 |
| | | | /dev/dsk/c0t1d1 | 14 | Open-E | 00 | 04 | 00 | 21 | 00 | A1 | 00 | 31 | 00 | B1 |
| | | lv_arch2 | /dev/dsk/c0t1d2 | 14 | Open-E | 00 | 08 | 00 | 22 | 00 | A2 | 00 | 32 | 00 | B2 |
| | | | /dev/dsk/c0t1d3 | 14 | Open-E | 00 | 12 | 00 | 23 | 00 | A3 | 00 | 33 | 00 | B3 |
| | vg_data1 | lv_data1 | /dev/dsk/c0t1d4 | 9 | Open-9 | 01 | 10 | 01 | 35 | 01 | B0 | 01 | 45 | 01 | C0 |
| | | | /dev/dsk/c0t1d5 | 9 | Open-9 | 01 | 12 | 01 | 36 | 01 | B1 | 01 | 46 | 01 | C1 |
| | vg_data2 | lv_data2 | /dev/dsk/c0t1d6 | 9 | Open-9 | 01 | 14 | 01 | 37 | 01 | B2 | 01 | 47 | 01 | C2 |
| | | | /dev/dsk/c0t1d7 | 9 | Open-9 | 01 | 18 | 01 | 38 | 01 | B3 | 01 | 48 | 01 | C3 |
| | vg_data3 | lv_data3 | /dev/dsk/c0t2d0 | 9 | Open-9 | 01 | 16 | 01 | 39 | 01 | B4 | 01 | 49 | 01 | C4 |
| | | | /dev/dsk/c0t2d1 | 9 | Open-9 | 01 | 18 | 01 | 3A | 01 | B5 | 01 | 4A | 01 | C5 |
| | vg_data4 | lv_data4 | /dev/dsk/c0t2d2 | 9 | Open-9 | 01 | 20 | 01 | 3B | 01 | B6 | 01 | 4B | 01 | C6 |
| | | | /dev/dsk/c0t2d3 | 9 | Open-9 | 01 | 22 | 01 | 3C | 01 | B7 | 01 | 4C | 01 | C7 |
| | vg_index | lv_index | /dev/dsk/c0t2d4 | 14 | Open-E | 00 | 30 | 00 | 43 | 00 | A9 | 00 | 53 | 00 | B9 |
| | | lv_index2 | /dev/dsk/c0t2d5 | 14 | Open-E | 00 | 31 | 00 | 44 | 00 | AA | 00 | 54 | 00 | BA |
| | vg_logA | lv_log1 | /dev/dsk/c0t2d6 | 3 | Open-3 | 02 | 32 | 02 | 53 | 02 | C0 | 02 | 63 | 02 | D0 |
| | | lv_log2 | /dev/dsk/c0t2d7 | 3 | Open-3 | 02 | 33 | 02 | 56 | 02 | C1 | 02 | 66 | 02 | D1 |
| | | lv_log3 | /dev/dsk/c0t3d0 | 3 | Open-3 | 02 | 34 | 02 | 57 | 02 | C2 | 02 | 67 | 02 | D2 |
| | | lv_log4 | /dev/dsk/c0t3d1 | 3 | Open-3 | 02 | 35 | 02 | 58 | 02 | C3 | 02 | 68 | 02 | D3 |
| | vg_log_b | lv_log1 | /dev/dsk/c0t3d2 | 3 | Open-3 | 02 | 62 | 02 | 59 | 02 | C4 | 02 | 69 | 02 | D4 |
| | | lv_log2 | /dev/dsk/c0t3d3 | 3 | Open-3 | 02 | 63 | 02 | 5A | 02 | C5 | 02 | 6A | 02 | D5 |
| | | lv_log3 | /dev/dsk/c0t3d4 | 3 | Open-3 | 02 | 64 | 02 | 5B | 02 | C6 | 02 | 6B | 02 | D6 |
| | | lv_log4 | /dev/dsk/c0t3d5 | 3 | Open-3 | 02 | 65 | 02 | 5C | 02 | C7 | 02 | 6C | 02 | D7 |
| | vg_admin | lv_error | /dev/dsk/c0t3d6 | 9 | Open-9 | 00 | 70 | 00 | 8A | 00 | 8D | 00 | 7A | 00 | 7D |
| | | lv_config | /dev/dsk/c0t3d7 | 9 | Open-9 | 00 | 71 | 00 | 8B | 00 | 8E | 00 | 7B | 00 | 7E |
| Command device (alpaha108) | | | | | | 00 | 9A | | | | | | | | |
| Command device (alpaha108) | | | | | | 01 | 8C | | | | | | | | |
| Command device (alpaha109) | | | | | | 00 | 9B | | | | | | | | |
| Command device (alpaha109) | | | | | | 01 | 8D | | | | | | | | |
| Command device (alpaha155) | | | | | | | | 00 | 9A | | | | | | |
| Command device (alpaha155) | | | | | | | | 01 | 8A | | | | | | |
| Command device (alpaha156) | | | | | | | | 00 | 9B | | | | | | |
| Command device (alpaha156) | | | | | | | | 01 | 8B | | | | | | |
| Command device (alpaha154) | | | | | | | | | | | | 00 | 9A | | |
| Command device (alpaha154) | | | | | | | | | | | | 01 | 8A | | |

**Table 7 device mapping table - including BC_1**

### 1b: Map BC devices to array port

BC and CA software requires that all devices are allocated to at least one port on the array. The port does not have to be physically connected to any server or be used by any server. It can be any available port on the array.

**Note:** If necessary, the BC devices can be mapped to RCU_target ports as long as host access to the devices is not allowed through these ports. An RCU_Target port is similar to a standard target port but with the restriction that only 128 devices can be mapped instead of the standard 255 devices. HP does not allow host access on RCU target ports because of performance impact to CA software.

HP does not recommend that a host have access to both CA and BC devices, and does not support such an installation in any cluster configuration. If it is decided to implement an additional backup server on Site 2, then this host will have access to the BC devices. Backup servers are beyond the scope of this solution.

## 1c: Design RAID Manager instance mapping

We assume that a backup server is not implemented in this solution and that the BC devices are allocated to an unused port on the array. This will require the application hosts to manage the CA_2 and BC_2 devices. In the example configuration, separate RAID Manager XP instances manage the CA_1, BC_1, CA_2, and BC_2 devices. This simplifies the configuration files and ensures that a configuration error in one file does not affect the availability of the other device group. The following diagram shows the required RAID Manager XP instances to mange the CA_1, BC_1, CA_2, and BC_2 device groups.

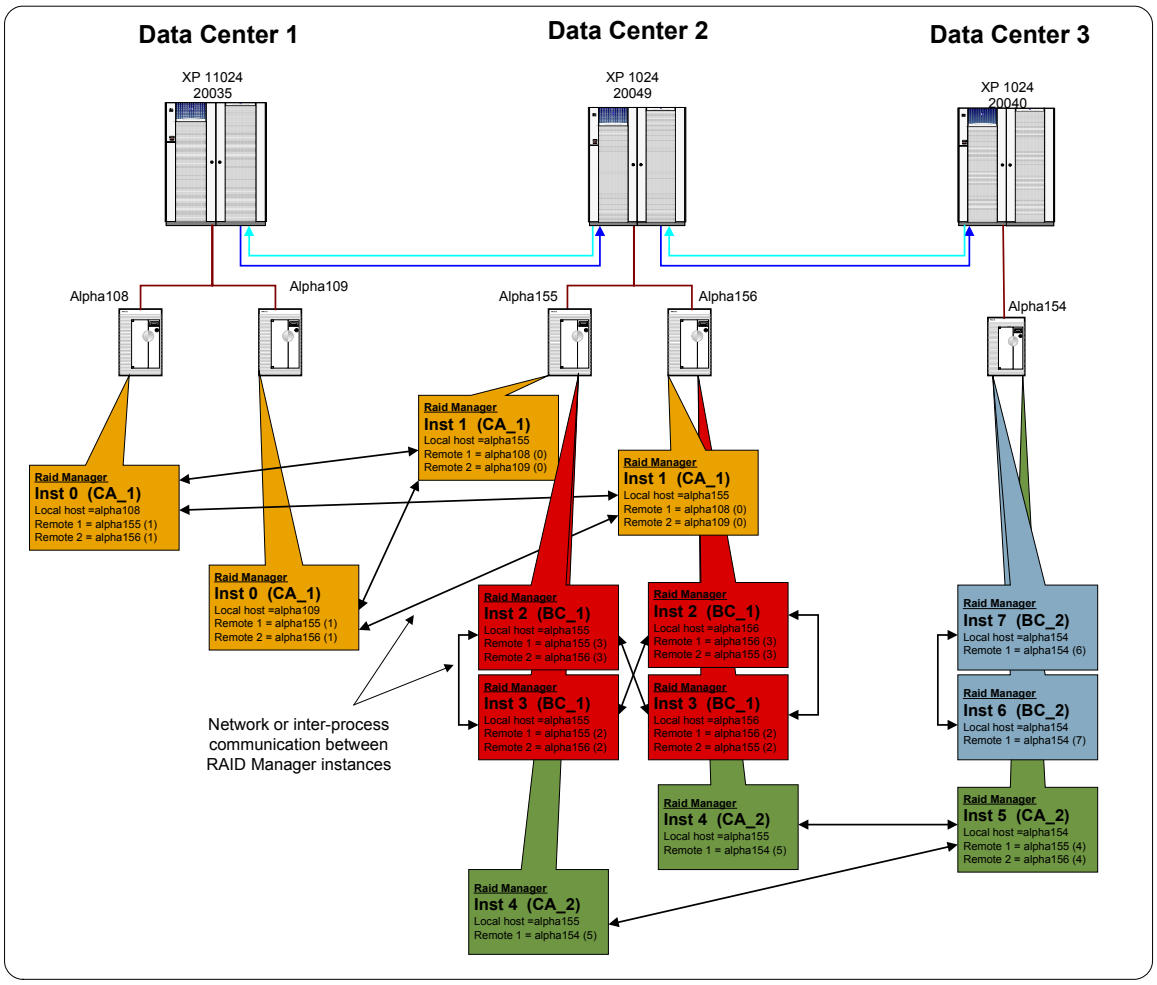



**Figure 19: RAID Manager XP instance mapping - CA_1, BC_1, CA_2, and BC_2**

# 2: Create the BC_1 pair

## 2a: Install RAID Manager

The hosts on Site 3 are also used for the CA_2 device pairs and therefore have RAID Manager XP already installed. If any new hosts was added to manage the BC_2 device group, be sure the same RAID Manager XP version is installed on that host.

## 2b: Identify command devices

Using **"ioscan"** command find the command devices allocated to the specific host (see Table 7 for the command device CU:ldev numbers). A command device is recognized by the "-CM" suffix to the device description. Find the associated device file for each command device. These command devices should be the same as the devices used for the CA_2 device groups.

## 2c: Create RAID Manager configuration files

On each host in the configuration, create a temporary RAID Manager XP instance by defining only the instance variables and commands devices in the configuration file. RAID Manager XP configuration files are located in the **/etc** directory and are named **horcmxx.conf,** where **xx** indicates the instance number. The temporary configuration file will look like this.

```
#/************************* For Local instance **************************/
HORCM_MON
#ip_address    service    poll(10ms)  timeout(10ms)
alpha154     horcm6     1000       3000

#/************************* For local command devices *******************/
HORCM_CMD
#dev_name       dev_name       dev_name
/dev/rdsk/c9t15d7
```

After creating the configuration file, be sure to add an entry to the **/etc/services** file to reserve a port number for each RAID Manager XP instance. The following is an example of the entries in the **/etc/services** file. Ensure that the entries are the same on all hosts in the configuration using RAID Manager XP.

```
horcm0          51000/udp   #RAIDManager instance 0
horcm1          51001/udp   #RAIDManager instance 1
horcm2          51002/udp   #RAIDManager instance 2
horcm3          51003/udp   #RAIDManager instance 3
horcm4          51004/udp   #RAIDManager instance 4
horcm5          51005/udp   #RAIDManager instance 5
horcm6          51006/udp   #RAIDManager instance 6
horcm7          51007/udp   #RAIDManager instance 7
```

Start the RAID Manager XP instance on the local host using the "**horcmstart.sh xx**" command, where **xx** is the instance number to start. After a successful startup of the RAID Manager XP instance set the "**HORCMINST=xx**" environment variable to ensure the ensuing RAID Manager XP commands connect to the correct RAID Manager XP instance.

Find the internal target and LUN numbers for each of the devices used by the application by using the "**raidscan -p <port_number> -fx**" command. Complete the configuration file by adding the device group name, device name, and port:Target:Lun number for each device. Define any remote instances that can manage the remote site of the device group and their services alias. The following is an example of a completed device group configuration file.

```
#ip_address    service    poll(10ms)  timeout(10ms)
alpha154     horcm6     1000       3000


#/************************* For local command devices ********************/
HORCM_CMD
#dev_name       dev_name       dev_name
/dev/rdsk/c9t15d7


#/************************** For device groups ***************************/
HORCM_DEV
#dev_group    dev_name    port#   TargetID  LU#   MU#
bc2_oracle      ora_data1  CL1-A          1     1        0
bc2_oracle      ora_data2  CL1-A          1     2        0
bc2_oracle      ora_index1      CL1-A           1     3        0
bc2_oracle      ora_index1      CL1-A           1     4        0
bc2_oracle      system1         CL1-A           1     5        0


#/*************** For remote instance definition ************************/
HORCM_INST
#dev_group    ip_address   service
bc2_oracle         alpha154    horcm7
```

**Note**: The configuration file for the BC_2 primary device is exactly the same as the configuration file for the CA_2 secondary devices. Copy the existing configuration file and change the services and remote instance information.

Stop and restart the instance using the "**horcmshutdown.sh xx**" and "**horcmstart.sh xx**" commands.

## 2d: Validate the RAID Manager configuration files

RAID Manager XP only reports the status of the devices defined within the configuration file of the RAID Manager XP instance used, so you must ensure that all instance configuration files address the correct number of devices as well as the correct logical devices. The configuration file uses the target:lun number of a device to refer to a logical device in the array. It is easy to mistype a value, therefore, it is important to verify the configuration file after any changes was made to the file. Be sure you export the correct

**HORCMINST=xx** environment variable as well as the **HORCC_MRCF=1** variable to accesses the BC device group information with the RAID Manager XP commands. Use the **pairdisplay** command to check the group definition and to ensure that all the correct logical devices are used in the configuration.

### 2e: Create the BC_2 pair

After validating the RAID Manager XP configuration file, create the device pair and start the initial copy of data from the primary to the secondary devices within the array. Connect to one of the hosts on Site 3, and set the **HORCMINST** environment variable to the instance number that will manage the primary side of the device group (this should be 2 for this example). Ensure that the **HORCC_MRCF=1** environment variable is set. Use the **paircreate** command to create the device group.

### 2f: Wait for pair status

After the **paircreate** command completes the devices start to copy all data from the primary device (P-vol) to the secondary device (S-vol). This is known as the initial copy process. During initial copy all data blocks on the source device are copied to the target device, even if the block is not actively used by the application or is empty. This process is required to ensure that the devices are exact replica of each other. Using the **pairdisplay** command, wait for the pair status to reach "**PAIR**" status.

## Results

At the end of this task, you will be able to suspend the BC_2 device pairs at any time creating a stable point-in-time image of the application and enabling a save copy of the data on Site 3.

# 10 Install and Configure the Multi-Site DT Management Tools

At this point, you have a fully installed multi-site solution with replication enabled between all three sites. The management of the configuration is very difficult and cycling the data from Site 2 to 3 is still a manual process. To improve the management of the configuration HP has developed some tools to assist in the multi-site administration as well as performing the cycling of the data.

For details on installing and using the MSDT Tools, see the *HP Multi-Site DT Management Tool User Guide*.

## 1: Install the tools

Install the Tools on every host involved in the solution and any additional host that might be used to manage the solution. The remaining tasks can be performed from any host where the Tools are installed. Download the tools from the storage tools Web site http://storagetools.lvld.hp.com/xp/

The toolset is distributed in UNIX "TAR" format. Copy the tar file to a temporary directory and extract the contents using the **`tar xvf <filename>`** command. Install the toolset using the **`install_hpux_msdt`** install script extracted from the tar image. The script will prompt for all host IP addresses that will form part of the multi-site hosts. This list will create the access list to the host agent software. The script will also prompt for the outbound IP address on the local host. If the host has more than one IP address, indicate which one to use for communication to other hosts.

Install the toolset on all hosts before starting the configuration file steps.

## 2: Create the MSDT Tools configuration file

Create the configuration file for the MSDT toolset using the sample files provided in the **`/etc/opt/hpmsdt/conf/sample`** directory. Configure the application section specifying all the hosts that can run the application. Create configuration details for each of the device groups specifying all Meta information for each device group. Add any custom parameters to the configuration file. For more detail on creating the configuration file, refer to the *HP Multi-Site DT Management Tool User Guide*.

## 3: Validate the configuration file

Use the **`msdt_verify`** command to validate the configuration file.

## 4: Distribute the configuration file

Use the **`dist_conf`** command to distribute the configuration file to all hosts defined in the configuration file. Ensure that all the hosts in the configuration have received the configuration file correctly.

## 5: Create custom scripts

Create **`pre_exec`**, **`post_exec`**, and **`find_app`** scripts to interface with the customer's applications. Sample scripts are available in the **`/etc/opt/hpmsdt/exec/sample`** directory. Theses scripts are used during the cycle process to prepare the application for the creation of the point-in-time image. Distribute the scripts to all hosts using FTP or remote copy commands.

## 6: Test/customize the disp_conf script

Run the **Disp_conf** script for the application and check the output. The **disp_conf** script displays the configuration information for all device groups defined in the tool configuration file. Ensure that the **disp_conf** script can obtain status information from all the hosts in the configuration and that it displays the status of the configuration correctly. In some cases it may be required to customize the script for the specific customer need.

## 7: Test/customize the cycle_pair script

The **cycle_pair** script is designed to cycle the data from Site 2 to 3. The script checks the current status of all device groups, and if the status is equal to the current status, the script starts the cycling processes. The first step in the process is to create the point-in-time copy of the data, then the script moves the data to Site 3. Customization of this script may be required for special customer needs. The script always starts and stops with the same device pair status. This allows for the next cycle pair run to start with the correct device status.

## 8: Schedule cycle_pair script

When the **cycle_pair** script is functioning and can repeatedly be re-run without errors, it is time to schedule the script to run on a regular interval. The cycle time was defined in the detail design phase of the project. Any scheduler can be used to schedule the cycle pair process. Be sure that the scheduler will notify the correct people when there is a failure in the cycle pair process. Any failure events must be handled as soon as possible and the cycle process must be repeated. This will influence the currency of the data on Site 3.

## Results

At the end of this step you implemented the tools to manage the multi-site solution and created the scripts to manage data cycling.

# 11 Verify the Solution

At this stage, the solution installation is completed. It is important to verify that all configuration files are updated and correct on all hosts. To ensure proper functioning of the cluster solutions, it is necessary to perform some cluster failover testing as wells as application testing.

The test and level of test will depend on each customer. HP recommends repeating some of these tests on a regular basis to ensure proper functioning of the hardware and software in the solution.

See the *MSDT Implementation Blueprint: Administration Guide* for guidance on some of the day-to-day tasks as well as example test that can be performed in the solution.