



hp industry
standard servers –
Linux High
Availability
solutions

May 2003



technical
white paper

Serviceguard for Linux Order and Configuration Guide

table of contents

introduction	2
what's new	2
coming soon	2
cluster configuration elements	2
system software	2
description	2
benefits	3
features	3
memory required	3
software required	3
what's included	4
HP Serviceguard for Linux toolkits	4
hardware components	4
servers	5
supported configuration with no shared mass storage	5
supported mass storage for SCSI	5
supported mass storage for Fibre Channel	6
networking	6
sample configurations	6
SCSI configuration	7
Fibre Channel configuration	10
packaged cluster example	14
power configuration	14
local regulations	15
redundant power supplies	15
site power circuits	15
uninterruptible power supply (UPS)	15
power distribution	15
power distribution testing	15
serviceguard manager	15
description	15
features	15
status definitions	16
property sheet	16
supported configurations	16
serviceguard manager hardware and software requirements	17
solutions	19
NAS 8000 with serviceguard	19
disaster-tolerant solution for Linux	19
configuration requirements	19
alternate fibre channel HBA	23
for more information	23

introduction

This document defines the server cluster configurations supported by HP for interconnection of specified HP industry standard servers and HP storage with the high-availability clustering middleware—HP Serviceguard for Linux (A.11.14). These configurations provide high server cluster availability and minimize single points of failure (SPOFs). A configuration certified and supported by HP should be adhered to. Deviations from the supported configurations can result in an inoperative cluster, an operating cluster with degraded performance, or hidden SPOFs. Therefore, deviations are not supported.

what's new

Additional support for:

- HP industry-leading ProLiant Servers are fully certified for Serviceguard for Linux environments including: ProLiant DL360 G3, DL380 G3, DL560 G1, ML350 G3, ML370 G3, and DL380 G3 Packaged Cluster
- HP StorageWorks Virtual Array 7110, EVA 3000, and EVA 5000
- Configurations for up to 16 nodes
- SuSE Linux Enterprise Server 8 powered by UnitedLinux 1.0 for Fibre Channel configurations
- HP's Linux FibreChannel HBA – FCA2214
- Application toolkits to speed deployment of Oracle and MySQL databases with Serviceguard for Linux clusters at <http://www.software.hp.com/>

coming soon

HP Serviceguard for Linux support on additional ProLiant and Itanium2 servers, HP StorageWorks RAID Array shared storage with SecurePath, and alternative mechanism to quorum service for arbitration are coming soon.

cluster configuration elements

This section identifies the hardware components, system software, network interfaces, and power distribution arrangement required to build an operational server cluster supported by HP. The various server cluster elements are listed below, and their associated configuration data is detailed in separate tables and illustrations that follow.

system software

- HP Serviceguard for Linux
- Serviceguard Manager
- Quorum Service

hardware components

- Servers for cluster
- PC or server for Quorum Service if 2-node cluster
- Storage and storage switches

sample configurations

- SCSI
- Fibre Channel
- Packaged cluster

power distribution

system software

hp serviceguard for Linux

description

HP Serviceguard for Linux (A.11.14) is specialized software for protecting critical applications from hardware and software failures. With HP Serviceguard for Linux,

multiple nodes (systems) are organized into a cluster that is capable of delivering highly available application services to LAN-attached clients.

benefits

- Maximize service uptime, ensure data integrity and minimize planned downtime - HP's robust high availability software, Serviceguard for Linux, brings the best-in-class mission critical HP-UX technologies to the Linux environment and ProLiant servers.
- Sophisticated heterogeneous cluster management and investment protection - Administer, monitor, and manage both Linux and HP-UX clusters with the Serviceguard Manager's intuitive GUI integrated with other HP management tools.
- Designed, developed, delivered and supported by HP - With HP as the single point of contact, there is no single point of failure.

features

HP Serviceguard for Linux monitors the health of various components within each node and quickly responds to failures in a way that eliminates or minimizes application downtime. Serviceguard is able to detect and respond to failures in:

- SPU
- System memory
- System and application processes
- LAN media and adapters
- Storage subsystem and adapters

For ease of management and outstanding flexibility, Serviceguard for Linux allows all of the resources needed by an application to be organized into entities called " packages." Packages consist of any resource needed to support a specific application service such as disks, network resources, and application or system processes. Packages are the entities that are managed and moved within the enterprise cluster.

Serviceguard for Linux is an important component in creating a complete and robust environment for highly available critical applications. In order to create a complete solution, the application environment should be designed to eliminate all single points of failure and to minimize the impact of various component failures.

memory required

The memory requirement for HP Serviceguard for Linux depends on the number of packages configured in the cluster. The following equation provides a rough estimate of how much lockable memory is needed:

$$\text{Total memory} = 6 \text{ MB} + 80 \text{ KB/package in cluster}$$

The total amount is needed on all nodes in the cluster, regardless of whether a given package or resource is on that node or not.

software required

One HP Serviceguard for Linux license must be ordered for each system (node).

Product number:

- If you are ordering through 1X [PN#: B9903BA]
- If you are ordering through ProLiant:
 - HP Serviceguard for Linux ProLiant Cluster (305199-B22) – one SKU for all components needed to create high availability environment
 - Software
 - Documentation
 - 2 licenses
 - Heartbeat cable
 - HP Serviceguard for Linux (307754-B22) – single license to provide ordering flexibility for clusters with an odd number of nodes or an existing cluster requiring

one additional node

Your HP salesperson or HP reseller will know which product number is appropriate.

what's included

The following deliverables are included:

- HP Serviceguard for Linux ProLiant Cluster (305199-B22)
 - 2 software licenses
 - Ethernet crossover cable (for 2 node clusters)
- HP Serviceguard for Linux (307754-B22 and B9903BA)
 - 1 software license
- 2 CDs with the following structure for all products
- 1 - Serviceguard Distributed Components CD
 - Serviceguard Manager software (A.03.00 - HP-UX, Windows and Linux)
 - ServiceGuard Manager Version A.03.00 Release Notes
 - Quorum Service software (A.02.00 - HP-UX and Linux)
 - Quorum Service Version A.02.00 Release Notes
- 2 - Serviceguard for Linux CD2 -
 - Serviceguard for Linux software (A.11.14)
 - Linux patches (Bonding – for RedHat Advanced Server 2.1 only)
 - Getting Started with HP Serviceguard Version 11.14 for Linux
 - Managing HP Serviceguard for Linux
 - HP Serviceguard for Linux Version 11.14 Release Notes

Documentation localized in English and Japanese.

NOTE: Updated documentation available from <http://docs.hp.com/hpux/ha>

HP Serviceguard for Linux toolkits

Available from <http://www.software.hp.com> at no additional charge are MySQL, Apache, NFS, Samba, and SendMail Toolkits for Linux, which enable quick implementation of leading applications into a Serviceguard for Linux cluster. The toolkits provide a set of shell scripts that allow you to configure Serviceguard packages that start, stop, and monitor these highly available applications and easily customize for your environment. The toolkit for Oracle (T2376A) is available for purchase from the same site. You cannot use Serviceguard Toolkits without the associated applications. The applications must be installed on your Linux system before you install, configure, and test your package. (When installing the Red Hat Linux distribution, select the "NFS Server" option for use with the NFS Toolkit. The Red Hat installation program will install both NFS kernel and utility.)

hardware components

The following hardware components are supported:

servers

- Supported HP industry standard servers within the cluster
 - DL380 G2
 - DL380 G2 Packaged Cluster
 - DL360 G3
 - DL380 G3
 - DL380 G3 Packaged Cluster
 - DL580 G2
 - DL560 G1
 - ML350 G3
 - ML370 G3
- Supported Quorum Service platforms
 - Any HP server or PC that supports Red Hat 7.1 and above, RedHat AS 2.1, or SuSE SLES8.

The Quorum Service (QS) acts as tie breaker when a failure in a cluster leaves exactly 50% of the nodes running and must be installed and run on a system outside 2-node clusters. The QS is a separate application provided on the "Serviceguard Distributed Components" CD. The node on which the QS is running must have network connectivity to the clusters for which it is providing services. Ensure that the connection with each node is independent of the cluster heartbeat connection so that both are not likely to fail at the same time. The QS is supported for 2 to 8 node configurations, required for a 2-node configuration and recommended for a 3- to 4-node configuration. The QS in a Linux environment can manage up to 50 heterogeneous Linux and HP-UX clusters and 100 nodes. The QS and clustered nodes may be on different subnets.

For more information on the QS, please see "Arbitration For Data Integrity in MC/ServiceGuard Clusters" whitepaper at <http://www.docs.hp.com/hpux/ha/index.html>.

supported
configuration with no
shared mass storage

Certain applications do not require shared storage. Serviceguard supports configurations without shared storage as well as those with shared storage.

supported mass
storage for SCSI

SCSI HBAs

- Any integrated Smart Array Controller on supported servers
 - Requires firmware version 2.38 and CCISS driver version 2.4.34
- Smart Array 5i Plus Controller (264030-B21)
 - Requires firmware version 2.38 and CCISS driver version 2.4.34
- Smart Array 532 Controller (225338-B21)
 - Requires firmware version 2.38 and CCISS driver version 2.4.34

Smart Array Cluster Storage (201724-B21)

- Requires firmware version 1.70

Supported drives – any drive supported in the selected storage subsystem see each product's QuickSpecs

<http://www5.compaq.com/products/quickspecs/Division/Division.html> (World Wide version)

http://www5.compaq.com/products/quickspecs/North_America/North_America.html (North America/US version)

supported mass storage for Fibre Channel

- HBA: FCA2214 (281541-B21): Requires driver 6.0.4 available on the HP Support page (Follow these links – HP.com > Downloads > Compaq Downloads > Disk Arrays > Adapters & Switches > FCA2214 > additional drivers and updates)
- Fibre Channel storage subsystems: HP StorageWorks va7100, va7400 (HP18 firmware recommended), HP StorageWorks va7110 and va7410 (A100 firmware recommended), HP StorageWorks va7110 (A100 minimum firmware required), HP StorageWorks EVA 3000 and EVA 5000, and HP StorageWorks xp48, xp128, xp512, xp1024 disk arrays
- Disk drives: Any drive supported in the selected storage subsystem.
- Fibre Channel switch: Any switch supported with the selected storage subsystems.

NOTE: Please contact your HP Storage representative for the latest firmware revision and support information for the EVA and XP systems and for switch connectivity support on Linux.

networking

HP Serviceguard for Linux supports redundant networks, even though these are not shown in the diagrams. The redundant client network connections allow local failover from one NIC to another in case of a network path failure. These connections must be on the same subnet. Several LAN interfaces (NICs) can be grouped together in a process known in Linux as channel bonding. Channel bonding can be configured in High Availability mode and Load Balancing mode. In the bonded group (HA mode), one interface is used to transmit and receive data, while the others are available as backups. If one interface fails, another interface in the bonded group takes over. In load balancing mode, which allows all interfaces to transmit data in parallel, in an active/active arrangement. In this case, high availability is provided by the fact that the bond still continues to function (with less throughput) if one of the component LANs should fail. It is strongly recommended to use channel bonding in each critical IP subnet to achieve highly available network services. Failover from one NIC to another prevents a package, or the entire system, from failing over to another system. This minimizes impact to the application and to the users.

Supported NICs are:

- Integrated NICs on the various supported servers
- Compaq NC3134 Fast Ethernet NIC 64 PCI Dual Port 10/100 (138603-B21)
- Compaq NC6136 Gigabit Server Adapter, 64-bit/66-MHz, PCI, 1000 SX (203539-B21)
- Compaq NC7770 PCI-X Gigabit Server Adapter (244948-B21)

sample configurations

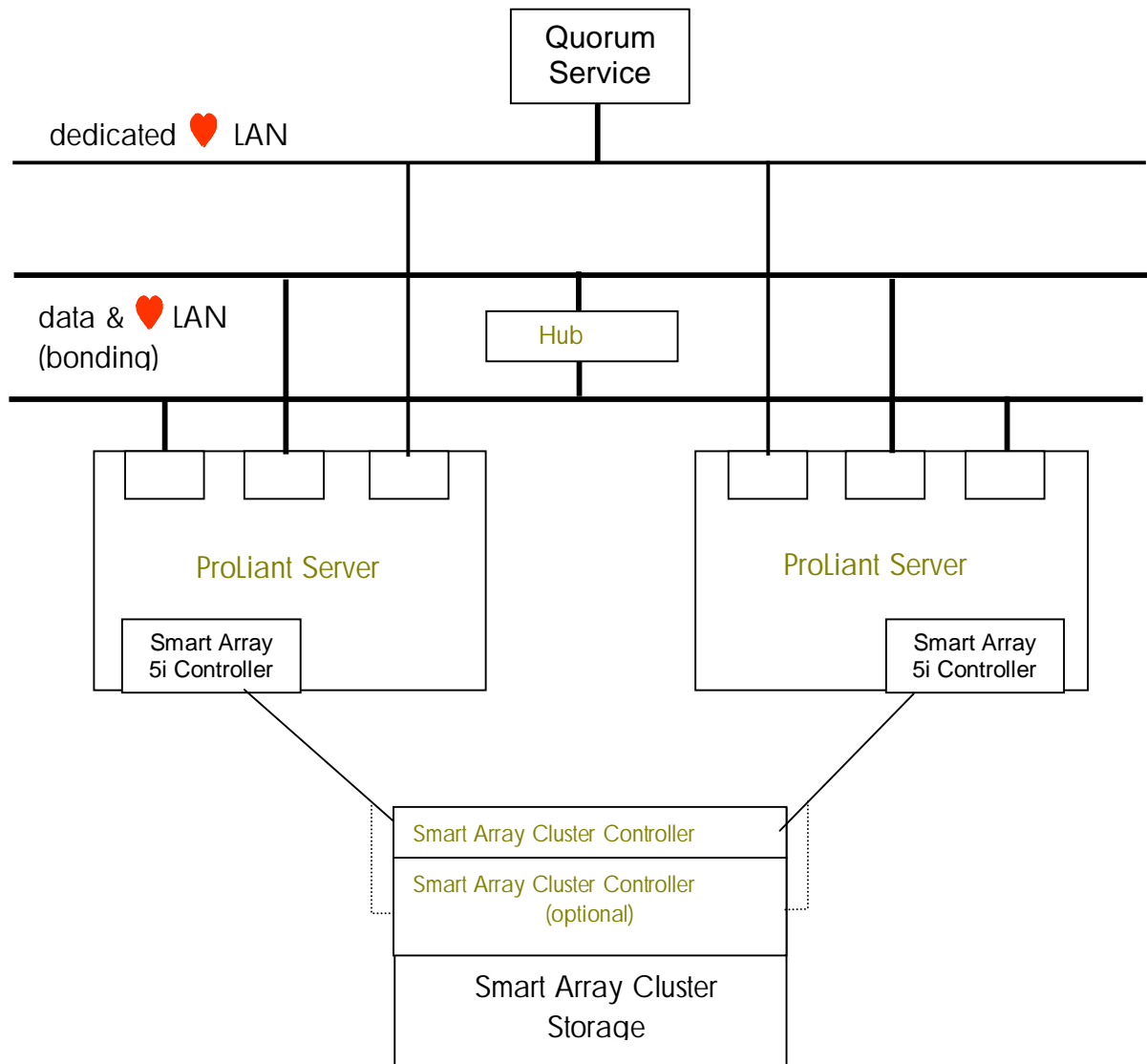
The following sample configurations can be used as a guideline to help configure a cluster. There are specific configuration examples and a description of the process used to create them.

The configurations described do not include any networking equipment. As shown in the following diagrams, a hub (or switch) is necessary to connect the two client LANs. The customer must then connect the client LANs to the remainder of their network infrastructure.

LOCALIZATION: These examples use North American part numbers. Please refer to the QuickSpecs for the various products or talk to your sales representative to determine the part numbers to order in your country. Also, any power calculations to determine PDU connections in this example use 120V. Please determine your connections based on local voltage and safety laws.

SCSI configuration

This is HP's recommended cluster configuration for SCSI with a redundant network for data.



sample SCSI configuration

The following table is an example of a 2-node SCSI configuration with a DL380 G2 and a DL580 without any UPS systems. There are enough PDUs to support the redundant power as described earlier. The customer can add 2 UPS systems that each support about 18A at 120V if desired for this particular configuration.

While a configuration requires a minimum of 2 LAN connections per server (redundant LAN connections carrying both heartbeat and public LAN) this example shows 3 LAN connections each, adding a separate heartbeat connection. The redundant connections carry the public LAN and act as a backup for heartbeat as well. In the various examples, integrated NICs are used when possible. No other LAN components are shown in any of the configuration examples since this is expected to go into an existing network infrastructure.

The table for this example has four sections, one for each server, a storage section and a section for the rack components. Additional memory has been added for servers in these examples. This can be changed to suit the needs of the specific environment.

The first server is a dual-processor DL380 G3 with 2.8GHz Xeon processors. Internal storage is configured with a SmartArray 5i controller to provide RAID1 capability for the two internal SCSI drives. The same SmartArray 5i also provides connection to the SmartArray Cluster Storage. The customer could decide to change the number of internal drives. The optional redundant power supply and fans are also included; the customer has the option of eliminating these if desired.

The second server is a 4-way DL580 G2 with 2GHz Xeon processors. The internal SmartArray 5i is used for internal RAID storage while a SmartArray 532 is used to connect to the SmartArray Cluster Storage.

Both servers are configured with extra memory. Customers can configure memory to meet their needs.

A basic DL320 is used to run the Quorum Service. Any server or PC that supports RedHat Professional 7.3, RedHat Advanced Server 2.1 or SuSE SLES8 can be used for this function.

The shared storage is the SmartArray Cluster Storage with the redundant controller.

A flat panel monitor/keyboard is used in this configuration with a 4-port console switch. Any of these components can be changed at the customer's discretion.

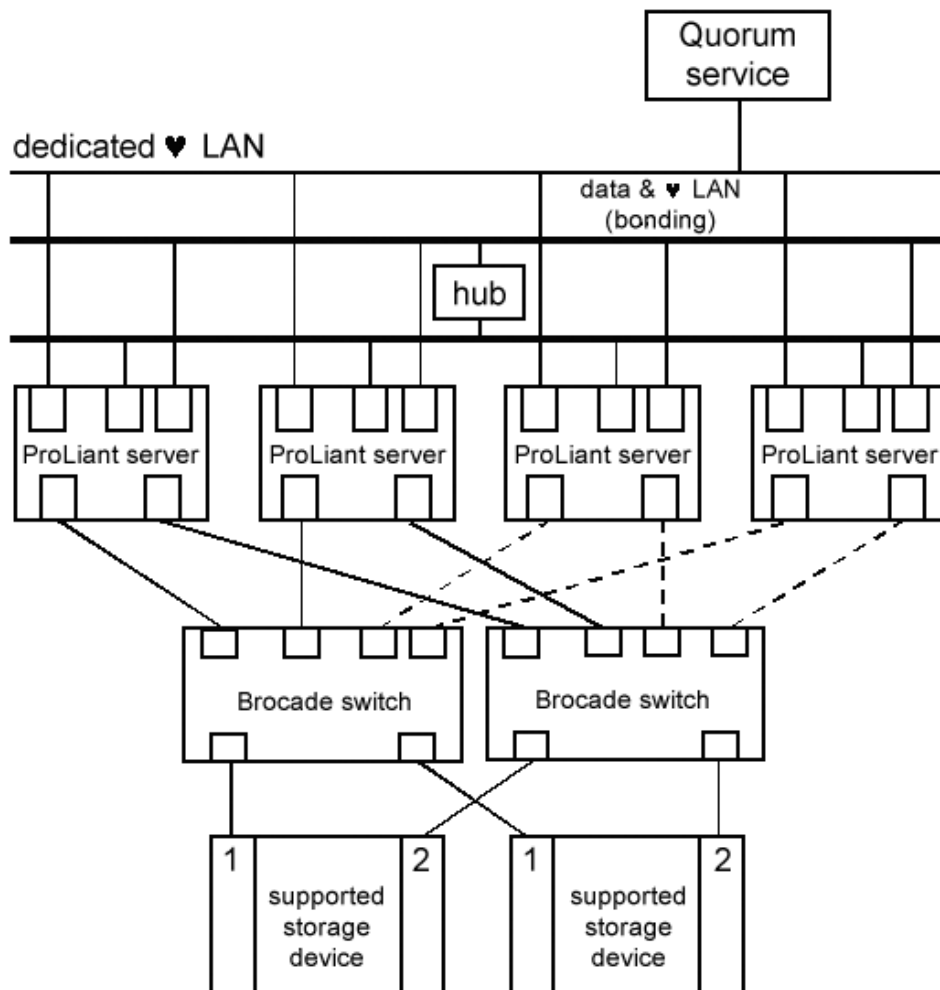
table 1. sample SCSI configuration

description	product number	quantity
Select ONE of the following 3 rows for the Serviceguard licenses		
HP Serviceguard for Linux on PL1X	B9903BA	2
HP Serviceguard for Linux ProLiant Cluster on ProLiant	305199-B22	1
HP Serviceguard for Linux on ProLiant	307754-B22	2
Select ONE of the following 3 rows for the Serviceguard support		
Serviceguard Software Updates	325544-002	2
Serviceguard application 9 x 5 support & software updates	325542-002	2
Serviceguard application 24 x 7 support & software updates	325543-002	2

description	product number	quantity
HP ProLiant DL380R03 P2800-512KB, 512 MB	301111-001	1
Intel Xeon 2.80 GHz-512KB Processor Option Kit	257915-B21	1
1024MB of Advanced ECC PC2100 DDR SDRAM DIMM Memory Kit (2 x 512 MB)	300679-B21	1
NC7770 PCI-X Gigabit Server Adapter	244948-B21	2
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	2
Hot Plug AC Redundant Power Supply Module (NEMA cord)	313054-001	1
DL380 G3 Redundant Fan Option Kit	293048-B21	1
HP ProLiant DL580R02 X2.0 GHz-2MB, 2048 MB	202176-001	1
Intel Xeon X2000-2MB Processor Option Kit	307276-B21	2
NC3134 Fast Ethernet NIC 64 PCI Dual Port 10/100	138603-B21	1
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	2
Smart Array 532 Controller	225338-B21	1
Quorum Service PC/server – not connected to shared storage		
ProLiant DL320R01 P1133-512 KB, 128MB, ATA 40 GB with CD/Floppy	251280-002	1
Smart Array Cluster Storage	201724-B21	1
Smart Array Cluster Storage Controller	218252-B21	1
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	12
Rack Components		
Rack 10642 Pallet	245161-B21	1
Sidewall Panel Kit for Rack 10642	246099-B21	1
600mm Stabilizer Kit for the Rack 10000	246107-B21	1
PDU 24Amp Low, NA/JPN/LA, 4 extension bars, each with 8(5-15) outlets	252663-D71	2
TFT5600 RKM	221546-001	1
Server Console Switch 4 Port	400336-001	1
CPU to Console Switch Cable 12'	110936-B21	3

Fibre Channel configuration

This is HP's recommended cluster configuration for Fibre Channel with a redundant network for data.



sample Fibre Channel configuration

The following sample Fibre Channel configuration uses four servers in the cluster, with four different server configurations to highlight the possibilities rather than as a configuration that a customer would generally implement. Up to eight nodes are supported in a Fibre Channel configuration. A PC or server to run the Quorum Service is not required for this configuration so is not included.

The key items that need to be in the servers are the correct number and type of LAN connections and the Fibre Channel host bus adapters (HBAs). This example shows only the HP Virtual Array 7100 (va7100) for shared storage and shows only one. Multiple va7100s can be attached to the Fibre Channel switches. The switch shown is the HP Brocade 2400, which has eight ports and is available from HP. The va7100 that is listed includes the base enclosure with dual power supplies; two controllers, each with 512 MB mirrored cache; and two optical GBICs. Other versions of the va7100 are available. Details are on the website at

http://www.hp.com/products1/storage/products/disk_arrays/index.html

The rack components were also selected using the online Web configuration tool as a guideline. HP IA-32 Server Rack Assistant was also used because of its ability to detail the power calculation. These tools are available at

<http://www.compaq.com/RackBuilder>

This rack was designed to use 240V power. There are two PDUs in this example. The localization options change based on country. The power cords will likely change based on country and the details of power connections in your location. No details are shown for LAN connection to switches. White papers are available on the Procurve section of the HP website (<http://www.hp.com/rnd/index.htm>), describing how to set up your LAN environment for the best availability.

table 2. sample fibre channel configuration

description	product number	quantity
Select ONE of the following 3 rows for the Serviceguard licenses		
HP Serviceguard for Linux on PL1X	B9903BA	4
HP Serviceguard for Linux ProLiant Cluster on ProLiant	305199-B22	2
HP Serviceguard for Linux on ProLiant	307754-B22	4
Select ONE of the following 3 rows for the Serviceguard support		
Serviceguard Software Updates	325544-002	2
Serviceguard application 9 x 5 support and software updates	325542-002	2
Serviceguard application 24 x 7 support and software updates	325543-002	2
HP ProLiant DL380R03 P2800-512KB, 512 MB	301111-001	1
Intel Xeon 2.80 GHz-512KB Processor Option Kit	257915-B21	1
1024MB of Advanced ECC PC2100 DDR SDRAM DIMM Memory Kit (2 x 512 MB)	300679-B21	1
NC7770 PCI-X Gigabit Server Adapter	244948-B21	1
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	2
Hot Plug AC Redundant Power Supply Module (NEMA cord)	313054-001	1
DL380 G3 Redundant Fan Option Kit	293048-B21	1
HP StorageWorks fca2214 fibre channel hba for Linux	281541-B21	2
2m LC-LC Multi-Mode Fibre Channel Cable	221692-B21	2
ProLiant DL380R02 P1400-512KB, 256 MB	253864-001	1
Pentium III 1400-512KB processor Option Kit	201099-B21	1
1024-MB PC133-MHz Registered ECC SDRAM Memory Option Kit	201694-B21	1
NC7770 PCI-X Gigabit Server Adapter	244948-B21	1

description	product number	quantity
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	2
Compaq Hot Plug Redundant Power Supply Module	225011-001	1
Hot Plug Redundant Fan Option Kit	225012-B21	1
HP StorageWorks fca2214 fibre channel hba for Linux	281541-B21	2
2m LC-LC Multi-Mode Fibre Channel Cable	221692-B21	2
HP ProLiant DL580R02 X2.0 GHz-2MB, 2048 MB	202176-001	1
Intel Xeon X2000-2MB Processor Option Kit	307276-B21	2
NC7770 PCI-X Gigabit Server Adapter	244948-B21	2
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	2
HP StorageWorks fca2214 fibre channel hba for Linux	281541-B21	2
2m LC-LC Multi-Mode Fibre Channel Cable	C7529A	2
HP ProLiant DL580R02 X2.0 GHz-2MB, 2048 MB	202176-001	1
Intel Xeon X2000-2MB Processor Option Kit	307276-B21	2
NC7770 PCI-X Gigabit Server Adapter	244948-B21	1
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	2
HP StorageWorks fca2214 fibre channel hba for Linux	281541-B21	2
2m LC-LC Multi-Mode Fibre Channel Cable	221692-B21	2
Storage Subsystem		
Virtual Array 7100 bundle	A6262AE	1
HP SP Install/Start-Up va7100 Dual Ctrl	H7550A or H7550E	1
Enterprise Class 36GB 15K-rpm FC HDD	A6193A	12
Disk slot filler panel	A6198A	3
RITTAL 9K CAB RAIL KIT	A5672A	1
HP Storaeworks SAN Switch 2/8-EL	258707-B21	2
2 Gb/s Small Form Factor Pluggable SW Transceiver Kit	221470-B21	12
2m LC/SC Multi-Mode Fibre Channel Cable	221691-B21	4

description	product number	quantity
Rack Components		
Rack 10642 Pallet	245161-B21	1
Sidewall Panel Kit for Rack 10642	246099-B21	1
600mm Stabilizer Kit for the Rack 10000	246107-B21	1
PDU 16Amp High, WW (12 recpt, detached cord)	252663-D71	2
TFT5600 RKM	221546-001	1
Server Console Switch 4 Port	400336-001	1
CPU to Console Switch Cable 12'	110936-B21	4

packaged cluster
example

This example shows how a packaged cluster can be used to form the basis of a SCSI cluster. This is nearly the same as the SCSI example above; with the added convenience of ordering the cluster components through one simple part number. The comments and explanations in the previous SCSI section are the same for this configuration.

NOTE: Rack and rack power components are not shown in this example.

The extra LAN connection described in the other SCSI example is not shown here. No other LAN components are shown in any of the configuration examples since this is expected to go into an existing network infrastructure.

table 3. sample packaged cluster configuration

description	product number	quantity
Serviceguard license		
HP Serviceguard for Linux ProLiant Cluster on ProLiant	305199-B22	1
DL380 Generation 3 (G3) packaged Cluster	305370-001	1
Select ONE of the following 3 rows for the Serviceguard support		
Serviceguard Software Updates	325544-002	2
Serviceguard application 9 x 5 support and software updates	325542-002	2
Serviceguard application 24 x 7 support and software updates	325543-002	2
Options for servers in packaged cluster (split equally to each server)		
Intel Xeon 2.8 GHz-512KB	257915-B21	2
1024MB of Advanced ECC PC2100 DDR SDRAM		
DIMM Memory Kit (2 x 512 MB)	300679-B21	2
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	4
Hot Plug AC Redundant Power Supply Module (NEMA cord)	313054-001	2
DL380 G3 Redundant Fan Option Kit	293048-B21	2
Options for the Smart Array Cluster Storage in packaged cluster		
Smart Array Cluster Storage Controller	218252-B21	1
HP 36.4GB 15K U320 Universal Hard Drive, 1"	286776-B22	12
Quorum Service PC or server – not connected to shared storage (any Linux certified PC or server)		
ProLiant DL320R01 P1133-512 KB, 128MB, ATA 40 GB		
with CD/Floppy	251280-002	1

power configuration

To ensure high availability, it is recommended that the power source be configured in such a way that there should not be power loss to 50% or more of the nodes in the cluster, including the computer running the Quorum Service.

Implementing the recommended power configuration will improve the availability of the cluster. In many cases, failure of a server or storage component from loss of AC power can be prevented. This prevents package failover and restart or storage re-mirroring. This section can also be used as a guideline in case the recommendations cannot be followed.

- local regulations
 - Local codes and regulations that relate to power must be followed, even if they differ from these recommendations.

- redundant power supplies
 - Most servers and storage equipment have optional or included redundant power supplies. These should be used and configured as described later in this guide.

- site power circuits
 - It is recommended that the cluster use two or three power circuits. This will prevent loss of one circuit from impacting more than 50% of the nodes of the cluster.

- uninterruptible power supply (UPS)
 - It is recommended that a UPS be used for each power circuit or that a site UPS be used. If a site UPS is used, then analysis should be done to ensure that loss of one circuit does not cause the failure of more than 50% of the nodes.

- power distribution
 - Multiple power distribution units (PDUs) should be used. They should be connected to at least two different power circuits. When a component (server, switch, etc.) has two power cords, then each should be connected to PDUs that are powered from different circuits.
 - If a component has more than two power cords, then they can be connected to different PDUs that are all powered from separate power circuits. If this is not possible, then all of the power cords of the component should be connected to the same PDU. In this case, follow the guideline of ensuring that there is not power loss to 50% or more of the servers in the cluster.
 - Power calculations for a PDU should be done with the assumption that another PDU or power circuit in the cluster could be down. For example, if a server that has two power supplies has one power cord connected to a PDU, then assume that all of the power for the server needs to be supplied by that PDU. This accounts for PDU, power supply, and power circuit failures.

- power distribution testing

It is important to test a cluster after it is configured. Part of that test should include resilience to power circuit, PDU, or UPS failure. Each of these components should be “failed” in turn to ensure that the cluster would not fail.

- serviceguard manager

- description

Serviceguard Manager (A.03.00) is a graphical user interface that provides a tool to display and manage Serviceguard clusters maintaining high availability. Using Serviceguard Manager, operators see color-coded, graphically intuitive icons to get the big-picture view of multiple clusters so that they can proactively manage the clusters, nodes, and applications.

- features
 - Administration functions enable users to perform tasks including:
 - Run, halt, or move a package, including moving a package via drag and drop
 - Run or halt a node or a cluster
 - Change node switching and package switching

Easy-to-understand confirmation dialogues are provided to inform users about the consequences of administrative actions. Users are shown the specific Serviceguard

command used and progress messages that track the ongoing administrative operation.

- Monitoring functions enable users to act on changing status:
 - Color-coded, graphically-intuitive icons visually present topology, online status, and configuration information for multiple clusters, their member nodes, and packages
 - SNMP event browser displays all SNMP traps generated by Serviceguard nodes being monitored
 - Alerts icon on the toolbar can show you the most critical problem in all the cluster objects on the map
 - Auto-Refresh provides automatic updates, including the property sheets, and can be enabled or disabled; refresh intervals can also be tuned to user requirements
- Saving status and configuration “snapshots” enables user to perform support or analysis functions.
- Dynamic scoping of clusters allows users to specify which clusters they want to monitor and administer.
- Manage clusters from different subnets. One instance of Serviceguard Manager can display up to 10 connected sessions.
- Comprehensive online help is available.
- Ease of use minimizes training to remotely display clusters from multiple management stations running HP-UX 11.x, Red Hat Linux 7.1 or later, Microsoft® Windows NT® 4.0 (Service Pack 5 or later), Windows 2000 Professional Edition or Windows XP Professional Edition.
- Flexibility of product allows it to be standalone or integrated with OpenView and Servicecontrol Manager.


status definitions


Serviceguard Manager uses color-coded icon borders and badges to show status information about the clusters, nodes, and packages.


The color-coded icon borders can be green, yellow, or red to indicate that the managed object is up, in a warning state, or down.

A status badge appears to the right of the managed object and shows critical status information. The icon border of the managed object with badges would be yellow, indicating a warning state. Types of status badges include:

 or  Indicates that the cluster or package is starting or halting.

 Indicates that one or more packages are down in a cluster. You can drill down on the cluster to determine which packages are down with red borders.

 Indicates that one or more packages do not have a node for failover. You can access property sheets to determine if this single point of failure is due to a node being down, disabled package auto-run, or disabled node switching.

 Indicates a package was never configured to have a failover node.

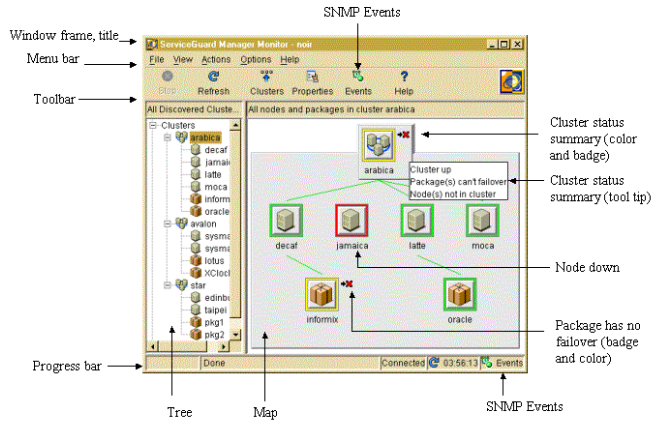
property sheet

A key feature of Serviceguard Manager is the ability to view detailed information from a Property Sheet accessible through the Properties window to learn more about a particular cluster, package, or node such as status, configuration, and number of nodes and packages.

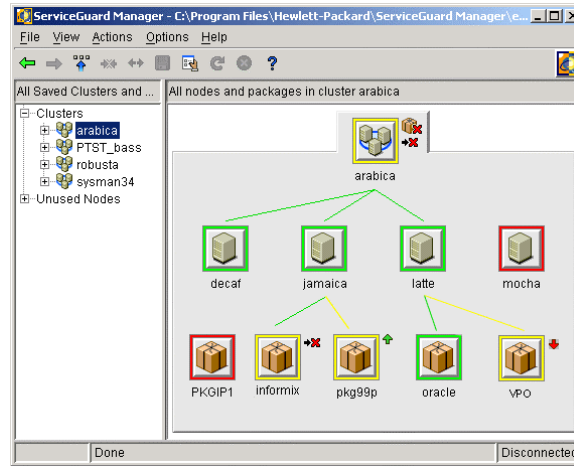
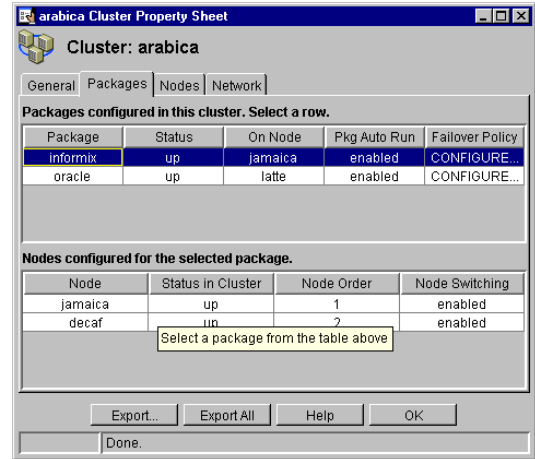
supported configurations

- Serviceguard version A.11.13 or later
- For Serviceguard Manager to support cluster administration, at least one Serviceguard Cluster Object Manager must exist in the cluster (installed with Serviceguard A.11.12 or later for HP-UX, or Serviceguard A.11.14 for Linux)

figure 1. serviceguard manager



cluster property sheet



serviceguard
manager hardware
and software
requirements

hp-ux management
station

- HP-UX 11.0 or later
- 256 MB of available memory
- 65 MB of available hard disk space under /opt
- 1 MB of additional hard disk space under /usr
- 1 MB of available hard disk space under /etc/opt/OV if HP OpenView has been installed
- 15 MB of additional hard disk space under /var for log files

PC management station

- Windows NT 4.0 (Service Pack 5 or later) or Windows 2000 Professional Edition
- Intel® Pentium® processor, 200 MHz or higher
- 128 MB of memory
- 150 MB of virtual memory
- 25 MB of available hard disk space
- SVGA or higher-resolution video adapter (800 × 600 and 256 colors minimum)
- 1 MB of available hard disk space if OpenView is installed
- 15 MB for log files when Serviceguard Manager is running
- Up to 30 MB of additional disk space is required during the installation process

Linux management station

- Red Hat Linux 7.1 (or later) and kernel version 2.4.2 (or later) or SuSE SLES8 / UL 1.0 and kernel version 2.4.19 (or later)
- Intel Pentium processor, 200 MHz or higher
- 128 MB of memory
- 100 MB of available hard disk space under /usr/local
- 15 MB of additional hard disk space available for log files created under /usr/local/sgmgr/log

patches

- To utilize the administration functions with version A.11.13 of Serviceguard installed on the server, patches PHSS_24678 and PHSS_22175 must be applied on the server
- For each Serviceguard node, install SNMP master agent patch PHSS_26724 for HP-UX 10.x or PHSS_26725 for HP-UX 11.x
- Serviceguard Manager installs its own JRE (Java Runtime Environment). Check the following websites for operating system patches required for JRE version 1.3.1.
 - On HP-UX, check patches at: <http://www.us-support.external.hp.com> or <http://www.europe-support.external.hp.com>
 - On Linux, check patches at: <http://www.redhat.com>
 - On Windows PC, check Service Pack at: <http://microsoft.com/ms.htm>

NOTE: For the latest version of Serviceguard Manager, a free non-orderable product, download from www.software.hp.com. The product is also shipped with each Serviceguard order on the Serviceguard Distributed Components CD.

solutions

NAS 8000 with serviceguard

HP expands its existing network-attached storage (NAS) offerings and continues to deliver industry-leading storage solutions with the release of its new NAS 8000 series of solutions designed to provide scalable, easy-to-manage, heterogeneous NAS solutions for its VA and XP arrays, as well as NAS integration into its SAN solutions for flexible file-sharing. These solutions offer a NAS-optimized OS for file serving and cluster technology—from Serviceguard for Linux—for mission-critical needs.

For more information regarding this solution, go to:

<http://www.hp.com/products1/storage/nas/8000/index.html>

disaster-tolerant solution for Linux

HP's metropolitan cluster for Linux is an integrated solution, utilizing Serviceguard for Linux and the Cluster Extension XP software (CLX) for the HP XP disk arrays.

CLX is host-based software that integrates Continuous Access (CA) XP remote mirroring with Serviceguard failover/failback operations.

Integrating CLX with Serviceguard for Linux allows packages to automatically fail-over:

- Among local nodes that are attached to the same XP series array.
- Between one system that is attached locally to its XP frame and another remote node that is attached locally to another XP frame.

configuration requirements

The supported disaster-tolerant configuration has the following requirements:

cluster topology and geography

- The current maximum number of nodes supported is eight. A Quorum Service node is required to provide quorum service to the cluster. It provides arbitration services for the cluster when a cluster partition is discovered. A node running the Quorum Service cannot be a member of any cluster to which it is providing cluster quorum services.
- The recommended configuration implements three data centers, with the third data center housing the Quorum Service.
- The minimum supported configuration is two data centers, with one of the data centers housing the Quorum Service. In this configuration, if the data center housing the Quorum Service is down, the nodes in the second data center will not be able to form the cluster, as there is no quorum. At the minimum, the Quorum Service should be in a separate room with its own power circuit.
- The maximum distance between the data centers is currently limited either by the maximum distance supported for the networking type or by the CA link being used, whichever is shorter—but no more than 100 kilometers.
- DWDM (Dense Wave Division Multiplexing) device can be used for the network and data replication links to increase the distance up to 100 kilometers between data centers.

cluster networking links

- The supported network interfaces used for cluster heartbeat are 10Base-T and 100Base-T.
- There must be less than 200 milliseconds of latency in the cluster heartbeat network between the data centers.
- No routing is allowed for the cluster heartbeat network between the data centers.
- There must be at least two alternately routed cluster heartbeat links between the data centers to prevent the "backhoe problem." The "backhoe problem" can occur when all cables are routed through a single trench and a tractor on a construction job severs all cables, disabling all communications between the data centers.

data replication
continuous access (CA)
links

- One of the heartbeat links has to be a dedicated link. The other heartbeat link can be shared with the application network link.
- There must be at least two alternately routed CA links between the two primary data centers.
- The supported distance for CA is varied depending on the link types. For more details on CA connectivity options, see the "Continuous Access XP Extension and Performance" white paper at http://xpslpgrms.corp.hp.com/whitepapers/xp512_whitepapers/CAET112701b.doc

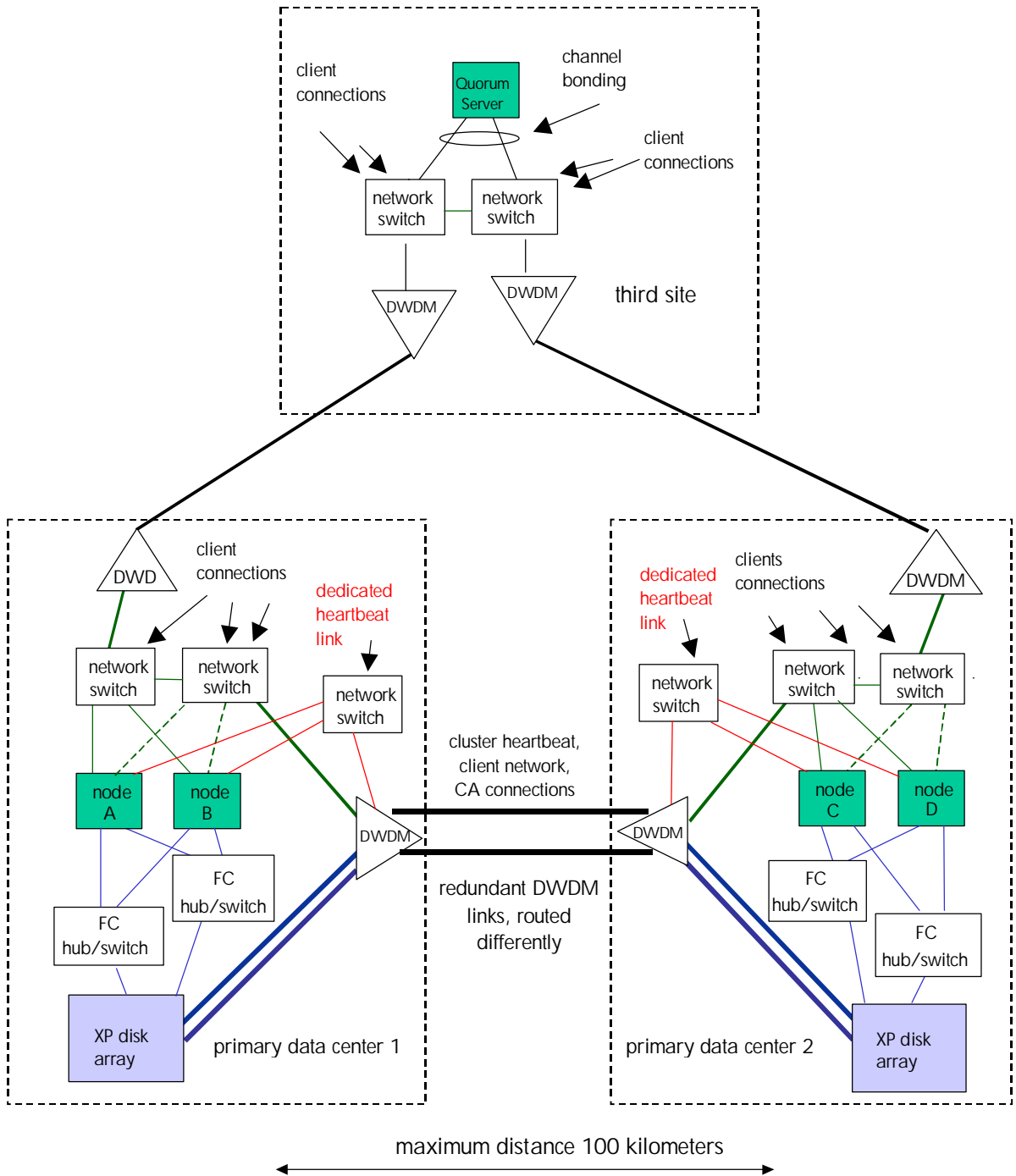
DWDM links for both
networking (cluster
heartbeat and
application network)
and CA

- The maximum distance supported between the data centers for a DWDM configuration is 100 kilometers.
- Both the networking (cluster heartbeat and application network) and CA links can go through the same DWDM box. A separate DWDM box is not required.
- The fiber-optic links between the DWDM boxes must be "dark fiber" links, non-switched circuits. They must be alternately routed between the two primary data centers.
- For the highest availability, it is recommended to have two separate DWDM boxes (in each data center) used for the links between each data center. However, since most DWDM boxes are typically designed to be fault tolerant, it is acceptable to use only one DWDM box (in each data center) for the links between each data center. If a single DWDM box is used, a minimum of one active and one redundant standby fiber link feature of the DWDM box must be configured. When using ESCON for CA, note that the ESCON timeout is shorter than the DWDM link failover time. Therefore, a minimum of two active fiber links on the DWDM box must be configured.

Quorum service

- Since only one IP address can be configured for a Quorum Service, it is suggested to configure the LAN used for the Quorum Service IP address with two or more physical LAN cards using Channel Bonding to improve the availability of the Quorum Service if a LAN card failure occurs. A node running the Quorum Service cannot be a member of any cluster to which it is providing cluster quorum services.
- The Quorum Service can provide quorum services for a maximum of 50 clusters and 100 nodes.

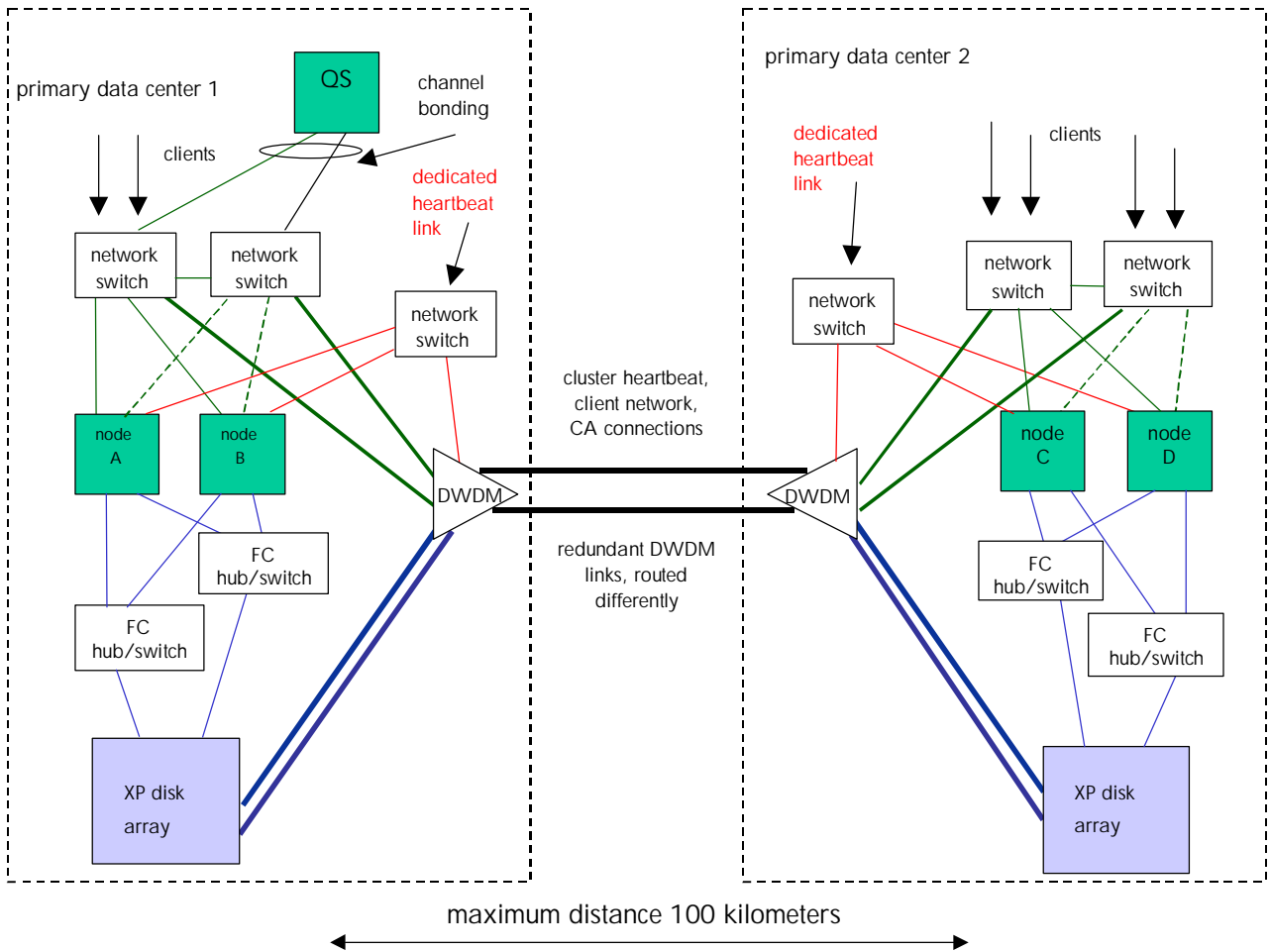
figure 2. recommended configuration with three data centers



In the example above, the DWDM boxes connected between the two Primary Data Centers are configured with multiple dark fiber links. The DWDM links are used for both networking and CA data replication. If the DWDM box is fault tolerant, only one is required in each data center. If it's not fault tolerant, two are required in each data center. The Quorum Service site distance can be up to 100 kilometers away from either Primary Data Center since DWDM links are also used. It is also supported to connect to

the Arbitrator data center without DWDM links, as long as the distance is within the allowed specifications for the network type being used.

figure 3. minimum supported configuration with two data centers



In the example above, the Quorum Service is in data center 1. Note that if data center 1 is down, the cluster nodes in data center 2 will not be able to form a new cluster. Therefore, this configuration is not recommended.

Hardware requirements

system platform for cluster nodes

Industry standard servers

- All servers supported in Serviceguard for Linux configurations

system platform for Quorum Service

- Any HP PC or server certified for Red Hat Profession 7.1 and above or Red Hat Advanced Server 2.1

hp disk array requirements

- HP Disk Array xp48 and xp512 with microcode version 01.11.22.00/00 or later
- HP Disk Array xp128 and xp1024 with microcode version 21.01.02.00/00 or later

software requirements

The solution required the following software versions:

- Red Hat Advanced Server 2.1, which uses a Linux 2.4.9-e3 kernel
- HP Serviceguard for Linux Version A.11.14
- Quorum Server Version A.02.00
- RAID Manager XP Version 01.09.02 or later
- HP Cluster Extension XP for Serviceguard for Linux Version 1.03.00
- HP Continuous Access XP
- HP Linux Disaster Recovery Data Replication Enabler
 - download from http://www.software.hp.com/HA_products_list.html

alternate fibre
channel HBA

NOTE Emulex Fibre Channel PCI host bus adapter; models LP8000-F1, LP8000-N1, LP9000-N1, LP9002L-F2, LP9002L-X2, LP952 are supported but are not available from HP. The customer must order these from their reseller. Special care needs to be taken when ordering cables as there are two types of Fibre Channel connectors supported in this set of HBAs.

If an Emulex FC HBA is used then use driver 4.20i or later.

The Emulex driver is delivered in arbitrated-loop mode. If switches are set in fabric mode, then the driver parameters need to be changed to match. This is documented at the Emulex Web site: <http://www.emulex.com/>. To find the detailed instructions select "Drivers, software, and manuals" under "quick links". Then select by operating system: Linux, and on that page select Fibre Channel drivers. Select "configure driver" for the appropriate driver version (the version supported by the selected storage subsystem). This document details how to change the various parameters.

for more
information

For more information regarding this solution, go to:

http://www.hp.com/products1/storage/products/disk_arrays/xpstorage/sw/cluster/index.html

To learn more about HP's other high availability products and solutions, please visit:

<http://www.hp.com/products1/unix/highavailability/>

and for ProLiant and Serviceguard for Linux:

<http://www.hp.com/servers/proliant/highavailability/linux/serviceguard/>