

HP High Performance Clusters LC Series Setup and Installation Guide



March 2004 (Second Edition)
Part Number 341524-002

© 2004 Hewlett-Packard Development Company, L.P.

Hewlett-Packard Company shall not be liable for technical or editorial errors or omissions contained herein. The information in this document is provided "as is" without warranty of any kind and is subject to change without notice. The warranties for HP products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.

HP High Performance Clusters LC Series Setup and Installation Guide

March 2004 (Second Edition)
Part Number 341524-002

Contents

Abstract	5
Audience Assumptions	5
Where to Go for Additional Help	5
Documentation.....	5
HP Services.....	6
Technical Support	6
Telephone Numbers	6
Limited Warranty	7
Safety Information	7
HPC LC Series Product Overview	7
HPC LC Series Cluster Components	8
Control Node.....	8
Compute Nodes.....	9
HPC Networks	9
Additional Components	10
Operating Systems and Software	10
Storage Components	10
Data CD	10
HPC LC Series Cluster Installation	11
Overview.....	11
Step 1: Physical Planning.....	11
Step 2: Receiving the HPC LC Series Cluster	11
Step 3: Positioning the Racks	12
Rack Power.....	13
Rack Cooling.....	13
Step 4: Connecting External and Inter-rack Cables	13
Power Cabling	14
Step 5: Powering On the Equipment.....	15
Step 6: Setting up the Out of Band Management Switch (LC 1000 Series)	17
Step 7: Setting up the Control Node	17
Step 8: Setting up the Compute Nodes	18
Step 9: Installing Storage System Options.....	19
For More Information	19

Abstract

This guide provides instructions and the necessary reference information required to install and set up HP High Performance Clusters LC Series (HPC LC Series) solutions. It provides information about the various solution components, as well as how to receive equipment, position racks, connect external and inter-rack cables, power on the equipment, and begin the setup process.

Some LC clusters will be shipped with storage options, as well as an operating system or application installed. This guide does not cover these cluster customizations. In these cases, this document should be used to set up the cluster hardware first. Next, the documentation supplied with the customizations, such as storage options or operating system, should be used to finish the installation or set up testing.

This document supplements the information found in the user documentation for the servers, switches and other components used in the HPC LC Series solution.

Audience Assumptions

This guide is for the person who installs, administers, and troubleshoots servers. HP assumes you are qualified in the servicing of computer equipment and trained in recognizing hazards in products with hazardous energy levels.

HP assumes that the user of this guide has advanced technical skills and has knowledge of the following topics and products:

- High performance computing concepts
- Networking skills
- Linux or Microsoft® operating system installation knowledge and experience
- High performance computing system software installation experience

Where to Go for Additional Help

Documentation

In addition to this guide, the following information sources are available:

- HP High Performance Clusters LC Series cabling guides for LC Series 1000, 2000, and 3000
- User documentation for the cluster components

Documentation for most HP High Performance Clusters LC Series components (servers, switches, options) is included with the cluster shipment.

The Myrinet switch documentation, however, is not included with the shipment. Instead, it can be downloaded from the following website:

<http://www.myricom.com>

HP Services

HP offers a variety of installation care packs and professional services. There are highly skilled professionals to help you install the cluster hardware, storage options, operating system, or applications. HP can also provide customized consulting and integration services to meet specific customer needs.

To learn more about what services are available for your HPC LC Series solution, please visit the following:

- HP Services at <http://www.hp.com/hps/>
- HP Care Pack Services at <http://www.hp.com/hps/carepack>

Technical Support

To ensure the best possible support for this HPC LC Series product, contact HP Services using the telephone numbers listed below.

HP technical support will take calls on the cluster solution and will perform fault isolation. If the problem is determined to be in the Cyclades product, HP will pass the call along with our analysis to Cyclades for correction of the problem.

Before calling, be prepared to provide the following information:

- Product number of cluster or failing device
- Serial number of cluster or failing device
- Node name or host name
- Operating system
- Question or problem statement
- Contact name and phone number
- Access number, obligation ID number, or system handle provided with your contract (for systems covered under an HP support agreement)

Telephone Numbers

- Americas region:
 - United States and Canada: 800-345-3746
 - Argentina, Paraguay, Uruguay, Chile, Peru, Bolivia: 54-11-4779-4779, 54-11-4779-4787
 - Brazil: 55-11-4689-2620 (Warranty or standard support), 55-11-4689-2014 (Premium or Gold support)

Additionally, each LC Cluster Series offers the choice of three cluster interconnect types:

- Myrinet
- Gigabit Ethernet
- 10/100 Fast Ethernet

The flexible solutions are defined and ordered with the help of the Design and Configuration guide. The guide provides more than 50 reference designs as cluster starting points that can be further customized to your needs. The guide assists with the server and interconnect selection and then helps determine the other required cluster components. The guide takes the guesswork out of ordering a cluster because it lists the needed components right down to the quantities and lengths of the cables needed for each solution.

Each LC Series also offers a packaged 32 node cluster. This reduces the order process to ordering just two part numbers: one for the cluster and one for the interconnect (Gigabit or Myrinet).

Each cluster ships with a Configuration Resource Kit containing a documentation CD, a data CD, hardcopies of selected documentation, and additional cable labels to assist with future cluster expansion.

The reference designs can easily be customized to support any node count cluster up to 128 nodes. This setup guide covers the hardware installation of these reference design configurations.

The reference designs can be further customized with, for example, the addition of storage options or an operating system. This setup guide does not cover these items. Refer to the documentation that comes with these options for setup information.

HPC LC Series Cluster Components

Each HPC LC Series cluster contains one control node, a collection of compute nodes, interconnects, rack(s), and rack infrastructure. It may also contain optional operating systems, software, and storage components.

Control Node

One ProLiant DL380 or DL145 server functions as the control node in each reference design. The control node is used as the interface to the user community via the public LAN for job dispatch, control, monitoring, and job completion within the cluster. The control node serves as the only access point to the compute nodes in the cluster.

The control node models and options available for use in the cluster are listed in the Design and Configuration Guide.

Compute Nodes

Depending on which configuration was ordered, the cluster contains ProLiant DL140, ProLiant DL145, or ProLiant DL360 server compute nodes. The compute nodes perform the basic work unit of the cluster. More compute nodes are added to the cluster to increase performance.

The current LC Series cluster reference designs scale to support up to 128 nodes per cluster. The server models and options available for use in the cluster are listed in the Design and Configuration Guide. In Myrinet solutions, each compute node will have a PCI Myrinet adapter installed.

HPC Networks

Each HPC LC Series cluster includes multiple networks:

- Out of Band Management network

The Out of Band Management network provides cluster management capability not available via the In Band Management network.

Cyclades terminal servers are used for Out of Band Management on ProLiant DL140 server based systems. HP ProCurve 2650 switches are used for Out of Band Management on ProLiant DL145 server based systems using IPMI. ProCurve 2650 switches are used for Out of Band Management on DL360 server based systems using iLO.

- Cluster interconnect network

The cluster interconnect network is the main data network that connects all of the compute nodes for cluster interprocessor communication (IPC) and message passing interface (MPI) functions. This network can be either a Myrinet, Gigabit Ethernet, or 10/100 Fast Ethernet network.

Various types of ProCurve or Myrinet switches are used for the cluster interconnect network, depending on the overall type and size of the HPC solution.

- Management network (In Band Management)

The management network (In Band Management network) is used for overall cluster management using a standard Ethernet connection. This network is also used to connect storage systems to the cluster in the reference designs.

Gigabit Ethernet ProCurve 2848 or 10/100 Fast Ethernet ProCurve 2650 switches are used for the management network, depending on the cluster configuration.

- Public network

The control node connects the cluster to the public network (WAN interface).

Additional Components

Each HPC LC Series reference design also comes equipped with an HP TFT 5600 RKM (integrated keyboard, monitor, and mouse), Power Distribution Units, and extra network cables for external network connectivity. All components are integrated and pre-cabled into HP 10000 series 42U racks.

All of the internal rack network cables are labeled with a descriptive cable label to facilitate the identification process of each cable connection. The HPC LC Series cabling guides for LC 1000, LC 2000, and LC 3000 illustrate the point-to-point connections of each network cable and describe the cable label nomenclature in detail.

Operating Systems and Software

Each HPC LC Series cluster can be ordered with Linux or Microsoft operating systems installed. Additionally, application software can be installed on the cluster. This setup guide does not cover setup or installation of operating systems and software, however.

Storage Components

Each HPC LC Series cluster can be customized with storage options. These options will require setup and installation steps that are not covered by this guide. Use the documentation that comes with these options for setup information.

Data CD

The cluster ships with a data CD which contains useful information on your cluster components that can save time in answering questions about the hardware in the cluster. The CD includes information on each cluster server such as: rack serial number where installed, server serial number, iLO DNS name, iLO MAC address, NIC IDs and NIC MAC addresses.

HPC LC Series Cluster Installation

Overview

Installation of your HPC LC Series cluster includes these general steps, described in more detail on the following pages:

1. Physical planning
2. Receiving the HPC LC Series cluster
3. Positioning the racks
4. Connecting external and inter-rack cables
5. Powering on the equipment
6. Setting up the Out of Band Management switch (LC 1000 Series only)
7. Setting up the control node
8. Setting up the compute nodes
9. Installing storage system options (if applicable)

Step 1: Physical Planning

Physical planning for your HPC LC Series deployment is one of the first things that must be considered before beginning the installation. You must ensure that you have enough physical space, adequate power and ventilation. You should also consider providing a backup power source such as an Uninterruptible Power Supply (UPS). A properly designed computer room has adequate ventilation and cooling for racks with servers and storage devices and has the appropriate high-line power feeds installed. For more information on datacenter design and planning, please refer to Technology Brief TC030203TB at the link below. This technology brief describes trends affecting datacenter design, explains how to determine power and cooling needs, and describes methods for cost-effective cooling.

Technology Brief TC030203TB can be downloaded from:

http://wwss1pro.compaq.com/support/reference_library/viewdocument.asp?countrycode=1000&prodid=137&source=tc030203tb.xml&dt=21&docid=15719

Step 2: Receiving the HPC LC Series Cluster

The HPC LC Series cluster components are shipped fully integrated in 42U racks. Depending on the size of the cluster ordered, the shipment can consist of from one to four 42U racks. More racks could be contained in the shipment if additional storage options are ordered as well. Every configuration is shipped fully integrated with easy to read cable labels to facilitate the cabling process.

Rack Power

Each fully configured rack in an LC Series cluster reference design comes with three 24-Amp (Americas region) or 32-Amp (European regions) High Voltage Power Distribution Units (PDUs). The data center will need to be configured to support this amount of power and power cabling. Power units supplied for other regions or to unique customer specifications may come with alternative PDUs. This should be reviewed prior to receiving the cluster to be sure that the data center is properly equipped.

Rack Cooling

The racks in each HPC LC Series solution draw cool air in through the front and exhaust warm air out of the rear. To ensure continued safe and reliable operation of the equipment, place the system in a well-ventilated, climate-controlled environment. A minimum of 50 inches (127 cm) is needed in front of the rack for adequate cooling and servicing. A minimum of 30 inches (76.2 cm) is required behind the rack for adequate cooling and servicing, but 50 inches (127 cm) are recommended. The HPC LC Series solutions should be placed in data centers with an adequate air-conditioning system to handle continuous operation of this solution. The maximum allowable ambient operating temperature for the LC Series Clusters is 35°C (95°F).

Please review the documentation for each of the components within your HPC LC Series solution to learn more about the recommended ambient temperatures. Component placement in the rack is very important to ensure proper cooling. Larger cluster interconnect switches, for example, must be located at the bottom of the rack to allow for additional cool air. It is also very important to keep components installed in the servers as recommended, such as hard disk drives, CD-ROM drives, or their blanks, to ensure proper airflow through the server.

Step 4: Connecting External and Inter-rack Cables

All of the network cables within each rack are labeled for easy identification. The HPC LC Series cabling guides explain and illustrate the cabling connections for each of the HPC LC Series solutions in detail.

All cables whose two endpoints reside in a single rack will already be connected when delivered to the customer's site.

If your HPC solution is comprised of multiple racks, then there will be some inter-rack cabling. That is, some of the cables from one rack will need to be connected to components in another rack. Following the cabling guide and using the cable labels, connect these cables. One end of these cables will be connected to a component in a rack. The other end will be coiled and secured in the rack for shipping purposes. You will need to unpack the free cable end and then connect it to its destination in the other rack. The cable lengths provided with the solution are planned to make the connections as follows:

-
- Switch-to-switch connections are routed through the sides of each rack.
 - Server- to-interconnect switch connections will be routed down from the server to the floor, over to the rack with the switch, and then up from the floor to the switch. If a Myrinet cluster was ordered, care must be taken when routing the fiber Myrinet cables to prevent cable damage. The minimum bend radius, or the smallest internal radius possible on a corner or bend, is 1.5 inches (3.81 cm).

Each HPC reference design includes a 20-foot (6.196 m) cable to connect the control node's Gigabit Ethernet NIC to the LAN. It also includes a 20-foot (6.196 m) cable to make the iLO port connection. These cables are not labeled and are not shipped in a rack, but are included in the cluster accessories packaging. Refer to the cabling wiring diagrams in the cabling guide for more information.

IMPORTANT: Refer to the HPC LC Series cabling guides for LC 1000, LC 2000, and LC 3000 for details on the power and network cabling connections, cable label nomenclature, wiring diagrams, and rack layouts for each HPC LC Series solution.

The cluster reference designs assume you will be connecting the control node of your HPC LC Series solution to an external DHCP server for setup. The DHCP server must be provided by the customer and be made available to the system before proceeding to use iLO on the control node. It is also assumed that you are planning to use the control node as a DHCP server for the other cluster components. Refer to the server's iLO documentation shipped with the cluster for more information.

Power Cabling

Each 24A or 32A PDU shipped in the reference design LC Series solutions is comprised of one PDU Control Unit supporting up to four eight-receptacle Extension Bars (power strips).

You will notice in your solution that some Control Units are not installed with all four of the Extension Bars, and that not all eight receptacles of each Extension Bar are used. This is to avoid current overload of the PDU or Extension Bar. General guidelines for these components are to limit each Extension Bar to five or fewer components and each PDU to 14 or fewer components.

Each Extension Bar has its own breaker switch. Furthermore, the PDU Control Unit has one breaker switch for each Extension Bar. The factory installs the Control Units in a 0U orientation and sets the breaker switches on the Control Unit to On but sets the breaker switches on the individual Extension Bars to Off.

Before plugging in the main power cable of each PDU Control Unit to the power source, check each Extension Bar to make sure that the breaker switch is set to Off. Once you have verified the breakers are Off then connect to the power source.

Step 5: Powering On the Equipment

After the power cables are connected, you are ready to power on the equipment.

IMPORTANT: Before powering on the servers, review Table 1: Factory System Settings to learn about the factory system settings for specific cluster components that are pre-configured before the HPC cluster is shipped to the customer.

1. Turn on each Extension Bar breaker switch.
2. Check to see that the servers come up to a standby state.
3. Check the switches to see that they have powered up. Some switches have a separate on/off switch while others power up immediately when power is applied.
4. Verify that the TFT5600 keyboard/monitor/mouse unit has powered up.
5. If any of the above components do not power up or come up to a standby condition, check the power switches to make sure they are On and check the power cords to ensure that they did not come loose during shipping.

Table 1: Factory System Settings

ProLiant DL380 Server	
BIOS settings for all LC Series configurations	<ul style="list-style-type: none">• The operating system setting is set to Linux unless a Microsoft® Windows® operating system is ordered with the solution• Hyper-threading is disabled
BIOS Settings for LC 1000 Series Clusters (DL140 based) to set up the Serial Console/EMS Support	<ul style="list-style-type: none">• EMS Console = Local• BIOS Serial Console Port = COM1:• BIOS Serial Console Baud Rate = 19200, 8, n, 1• Terminal Emulation Mode = VT100
ProLiant DL140 Server	
BIOS Settings	<ul style="list-style-type: none">• Hyper-Threading is disabled• NIC2 is set to the default PXE NIC
Remote Access	<ul style="list-style-type: none">• Remote Access via the serial port is set up (Advanced => Remote Access Configuration)• Remote Access = Enabled• Serial Port Mode = 19200 8,n,1• Flow Control = None• Redirection after BIOS POST = Boot Loader• Terminal Type = VT100
Boot Settings	<ul style="list-style-type: none">• Boot Settings Are Configured (Boot => Boot Settings Configuration)• Quick Boot = Disabled• Quiet Boot = Disabled

continued

Table 1: Factory System Settings *continued*

Linux OS Settings	
Completed if Linux OS is shipped on the cluster	<ul style="list-style-type: none">• The Bootloader and OS Configuration for the control node and each compute node are set up for Out of Band management via a serial console.• Modified <code>/etc/inittab</code> to spawn <code>agetty</code> for <code>/dev/ttyS0</code> in runlevels 2, 3, 4, and 5• Modified <code>/etc/securetty</code> by adding the line <code>ttyS0</code>• Configured the boot loader to use serial line• Configured the Linux kernel to use serial console• Disabled the X server: such as modified <code>/etc/inittab</code> to select runlevel 3 as default• Set up the SysRq functionality
ProLiant DL360 Server	
BIOS settings for all LC Series configurations	<ul style="list-style-type: none">• The operating system setting is set to Linux unless a Microsoft® Windows® operating system is ordered with the solution.• Hyper-threading is disabled
iLO settings	<ul style="list-style-type: none">• The iLO DNS name is set to match the iLO cable label for each node.• The iLO username and password are set to “Administrator” on each node.• The iLO server label tags for each node are removed <p>NOTE: By default, iLO is set to obtain an IP address from a DHCP server. The DHCP server can be the cluster control node or some other DHCP server provided by the customer. The DHCP server must be made available to the systems before proceeding to use iLO for initial setup. Refer to the server’s iLO documentation that is shipped with the cluster for more information.</p>
ProLiant DL145 Servers	
BIOS settings	<ul style="list-style-type: none">• Hyper-threading is disabled• NIC2 is set to the default PXE NIC

continued

Table 1: Factory System Settings *continued*

ProCurve Switches	
Trunking	Trunking between switches is set up on systems that have multiple racks to improve performance. The ports used for this purpose are the highest numbered switch ports. The exact trunks are specified in the cabling tables and wire diagrams.
VLANs	<p>VLANs are set up on systems that use that same ProCurve switch for the management and cluster interconnect networks. These are the 16-node reference GigE and Fast Ethernet systems that expand to 22 nodes.</p> <ul style="list-style-type: none">• The GigE system uses a ProCurve 2848. Ports 1-22 are used for the interconnect network and ports 23-45 are used for the management network. Port 48 is used for the OOB switch connection.• The Fast Ethernet system uses a ProCurve 2650. Ports 1-22 are used for the interconnect network and ports 25-47 are used for the management network. Port 48 is used for the OOB switch connection. <p>Refer to the ProCurve documentation for additional information on trunks and VLANs.</p>

Step 6: Setting up the Out of Band Management Switch (LC 1000 Series)

LC 1000 Series clusters use Cyclades terminal servers that must be set up and configured at the customer's site. HP recommends that you follow Cyclades' Quick Start process which is detailed in the *AlterPath Console Server User Guide*. This guide is included with the cluster shipment. It is also available from the Cyclades website at www.cyclades.com.

Step 7: Setting up the Control Node

The control node is connected to the TFT5600 monitor in the solution so you can directly control and monitor this server.

1. First, power up the control node and confirm that it passes the power on self test (POST).
2. Set up the control node's iLO configuration if desired. The control node will need to be connected to a DHCP server supplied by the customer to complete the iLO setup.
3. If the cluster was shipped with an operating system installed, the server will boot to the OS. Once it has been confirmed that the server boots properly you can continue to set up the control node to your operating system and applications specifications. Refer to the operating system and software vendor's installation instructions for additional information.
4. Visually check each NIC Link light on the server and the switches to which they connect to verify that each cable has a link established.
5. If the server does not boot properly, you can try the following as part of the troubleshooting process:
 - a. Look for error messages during POST.

