HP XC6000 Cluster

Reference guide

HP XC6000 Cluster overview

At a glance

The Linux®-based HP XC6000 Cluster is a parallel supercomputer that scales from 17 to 256 nodes, and the architecture is designed to scale to higher node counts. It is based on 2-way HP Integrity rx2600 servers acting as dedicated service nodes and performing management and administrative functions within the cluster, and 2-way HP Integrity rx2600 servers acting as application nodes. The HP Integrity rx2600 servers are based on Intel® Itanium® 2 processors. These clusters are available in standard 17-, 34-, 68-, 128-, and 256-node configurations and utilize the high-bandwidth, low-latency Quadrics Elan 3 interconnect and the HP ProCurve 2650 Ethernet switches. These 10/100 Ethernet switches are used for an administrative/boot network. In addition to the standard configurations, HP XC6000 Clusters can also scale in increments of four application nodes—starting at a minimum of 16 application nodes—to meet customer compute needs.

Each Linux-based HP XC6000 Cluster is composed of these elements:

- HP Integrity rx2600 servers—dual-processor systems that act as service nodes and dedicated compute/application nodes. One service node is required for every 16 compute/application nodes, and a service node is designated as the administrative/boot node for the cluster. Service nodes also take on other administrative roles and act as I/O nodes for file access.
- HP ProCurve Switch 2650—a 48-port 10/100 Ethernet interconnect—subnet and admin network
- High-speed, low-latency Quadrics Elan 3 communications interconnect; the Quadrics Elan 4 communications interconnect will be phased in when available in early 2004
- Local embedded service node storage plus optional SAN storage, using HP StorageWorks Modular SAN Array 1000 (MSA1000) or HP StorageWorks Enterprise Virtual Array 3000 (EVA 3000)
- HP Rack 10000 Series 42U rack enclosures
- One rackmount keyboard/monitor is required on XC6000 Cluster systems. On systems with two or more service nodes, one 4-port KVM switch is required to connect to up to four service nodes per configuration.

Figure 1. An HP XC6000 Cluster based on the Intel Itanium architecture with 128 nodes (256 processors), a 128-port Quadrics Elan 3 high-speed interconnect, and MSA1000 SAN storage array is shown below. The configuration consists of seven 17-node compute building (CBB) block racks, one 9 node partial CBB rack, and a utility building block rack with an Elan 3 switch chassis, TFT display, KVM switch, admin switch, and MSA1000 SAN storage array.

HP XC6000 system architecture

The XC6000 Cluster has a parallel system architecture that supports the management and execution of multiple parallel and serial applications on the XC cluster systems. The XC system architecture presents the user and the administrator with key single-system traits. Users see the XC system as a single system for login access, resource access, and job execution. Administrators control the system from a single service node, using it to perform tasks such as system management, performance monitoring, and hardware diagnostics, thus avoiding the complexity and difficulty of standard cluster administration. The XC system operates with a single root file system (the exception is the administration node, which, for reliability reasons, is managed independently). Because the system files are not scattered across multiple locations, there is only one set of configuration data to manage. Problems normally associated with managing version skew within a cluster cannot occur.

The XC6000 Cluster is a full-service system. Unlike many other clustered systems, the system services are organized to enhance the performance of tactical applications that require tightly coupled, synchronous cooperation between processes running on multiple nodes. This enhancement is accomplished by offloading the system services to specialized nodes, called service nodes, which provide all of the operating, administrative, and file services that are normally expected from a full-service system.

In the XC6000 architecture, applications are executed on one or more specialized application/compute nodes. Operations that are not of direct benefit to the application are migrated to other system components external to the application node. Within the application node, single processors are dedicated to the execution of an application's individual process. Virtually all services provided on the application node are dedicated to the application, which reduces or eliminates the context switching and resource contention normally found on general-purpose systems.

HP XC6000 Cluster systems are designed to support the XC cluster architecture as well as the XC cluster software. In addition, the XC cluster systems are easy to configure, build, install, and support. To facilitate configuration and scaling, they are made up of modular building blocks: compute building blocks (CBBs), interconnect building blocks (IBBs), utility building blocks (UBBs), and storage building blocks (SBBs).

Compute building blocks

One XC6000 Cluster system compute building block (CBB) rack contains up to 16 HP Integrity rx2600 server application nodes, one Integrity rx2600 server service node, and one HP ProCurve 2650 subnet 10/100 switch. Multiple CBB racks are connected together to expand the size of the cluster to up to 255 nodes.

The XC6000 Cluster system compute building blocks (CBBs) consist of the following:

- An HP Rack 10000 Series 42U rack with server slide kits
- Three power distribution units (PDUs) and cable management assemblies
- Up to 16 dual-processor Integrity rx2600 server application nodes (minimum of four application nodes, except in the first CBB rack)
- One dual-processor Integrity rx2600 server service node (one service node is needed for every 16 application nodes)
- One HP ProCurve Switch 2650 network switch and rackmounting kit
- Network cabling to connect the nodes to the HP ProCurve switch
- Integrity rx2600 server application nodes must all have processors with the same speed and the same memory capacity

The application nodes and service nodes are wired internally to the CBB rack to reduce external cabinet cabling. Each of the application and service nodes is also connected to the high-speed Quadrics Elan 3 interconnect switch with a PCI Elan 3 adapter card. Each application node MP (maintenance processor) and administration LAN port is connected to the 10/100 48-port HP ProCurve Switch 2650 switch in the cabinet to form a sub-network. Each service node in the rack will have an administration LAN connection to the HP ProCurve subnet. There is one HP ProCurve Switch 2650 in an XC6000 CBB rack. The switch in one CBB is linked with the Gigabit link in its adjacent rack to form a sub-network of up to 32 application nodes. Two CBB racks are used to make up a 32 application node/2 service node sub-network. Each application node requires a single disk for swapping, and the disks in all the application nodes must be the same.

The service nodes in the compute rack are also connected to an administration switch that is in the utility rack. The MP and LAN ports of the service nodes are connected to an HP ProCurve Switch 2650 administration switch. Each application node is required to have one disk and can have up to three disks, which must be the same size and speed, and each application node must have one disk for swapping. Service nodes are required to have one disk and can have up to three disks, which must be the same size and speed, and each application node must have one disk for swapping as well as for file storage. Each service node will also have a Gigabit LAN PCI interface for external user communication.

One service node in the XC cluster system is designated the master administrative node and is responsible for booting the system and other administrative functions. Select service nodes are identified to provide I/O and file system access. These nodes will have additional disks or an optional Fibre Channel (FC) host bus adapter to connect to an external storage subsystem (MSA1000 or EVA 3000). Application nodes are added four at a time in a CBB rack, with up to 16 application nodes per CBB rack, always with one service node.

Interconnect building blocks

The XC6000 Cluster system interconnect building blocks (IBBs) consist of an HP 10000 Series 42U rack, two power distribution units, and rack accessory kits with up to two Quadrics Elan 3 node- or link-level switches (chassis, with switch node-level and link-level cards) to create an interconnect switch for more than 128 nodes. IBB racks are not required for systems with 128 nodes or less. The application and service nodes in the CBB racks are connected to the Quadrics Elan 3 switch cards in the IBB rack. The IBB racks are connected to form a large switch fabric that allows more than 128 nodes to be connected together. Quadrics Elan 3 16-port switch node cards are added to the node-level chassis to provide the necessary switch ports to connect the application and service nodes. Quadrics Elan 3 16-port uplink cards are added to the link-level chassis to provide the necessary switch fabric in a federated switch configuration. A clock box is also required for a federated switch configuration.

Utility building block

The XC6000 Cluster system utility rack, or utility building block (UBB), is used in configurations with 128 nodes or less to house the 128-port Quadrics chassis, the system TFT console, the administrative HP ProCurve 2650 network switch, and the optional MSA1000 SAN. One UBB is required for each XC6000 system. (For 256-node cluster systems, node 256, a service node, is mounted in the UBB rack.) The UBB consists of these components:

- An HP 10000 Series 42U rack
- A Quadrics Elan 3 128-port chassis (in smaller XC6000 Cluster systems—those with 128 nodes or less—only one Quadrics Elan 3 128-port switch chassis is required, and this chassis is mounted in a utility rack)
- Two power distribution units

- An HP ProCurve 2650 system administration LAN switch (for configurations with two or more service nodes)
- An optional MSA1000 storage array subsystem (maximum of two shelves)
- A rackmount keyboard/TFT monitor with a four-port KVM switch (KVM switch required when two or more service nodes are used) to connect to up to four service nodes
- For 256-node cluster systems, node 256, a service node, is mounted in the UBB rack

Storage building blocks

HP XC6000 Cluster system storage building blocks (SBB) are optional EVA 3000 storage subsystems that come in their own HP 10000 Series 42U racks. The EVA storage subsystem is connected to the I/O service nodes in the cluster with a PCI host bus adapter (HBA). Using two I/O service nodes with a single HBA is the recommended way to connect the EVA 3000 storage subsystem. Use the EVA 3000 storage configuration rules for ordering the HBA for the service node. EVA storage subsystems are the design target of the Lustre file systems. More configuration information will be available in the future to configure Lustre file systems.

System options

Placement of the CBB, IBB, UBB, and optional SBB is dependent on the layout of the customer site, and inter-cabinet cabling must be taken into account. The factory will build to a default layout if site-specific information is not available. It is recommended that site layout information be provided at the time of system purchase. System options for the XC6000 Cluster consist of application- and service-node memory content, application- and service-node disk content, a TFT console with a KVM switch, SAN storage, and EVA 3000 and MSA1000 storage. These options are supported by the XC6000 Cluster system software. Options not specified in this reference guide will need to be reviewed on a case-by-case basis for integration into the system. Equipment that is not approved will be shipped as non-integrated material.

Node count

HP XC6000 Cluster systems have a minimum of 16 application nodes and one service node. Both are configured with a minimum of 1 GB of memory per CPU. Application nodes can be added in a CBB rack, four at a time, up to a total of 32, with one service node for every 16 application nodes. An XC6000 Cluster system can scale to up to 256 total nodes (240 application nodes and 16 service nodes). Larger XC6000 Cluster systems (beyond 256 nodes) can be designed and configured; contact HP for more information.

Integration, services, and ordering information

HP XC6000 Cluster systems are fully integrated into racks at the factory and tested to assure proper system operation before being shipped to the customer site. XC6000 Cluster software is loaded and tested. Inter-cabinet cabling is labeled to facilitate installation at the customer site. XC6000 Cluster systems also come with hardware and software documentation as well as the necessary HP XC System Software licenses, compiler tools, and software media kit. Special considerations in the XC hardware design have been made so that there is proper cooling of the major components in the cabinets. Special cable and cable management assemblies are included for proper inter- and intra-cabinet cabling as well as to promote adequate air flow and maintenance of the equipment. Special mounting brackets for the application nodes, network switches, interconnect chassis, and cooling baffles are also included to allow the most efficient use of the rack space. These considerations increase system reliability, allow the system to be factory integrated and shipped to the customer site without damage, and facilitate the installation of the system at the customer site.

HP XC6000 Cluster systems are installed by HP Services. They come with both installation services and HP Consulting and Integration (C&I) system startup services as part of the solution. These services must be included in the initial system order. A one-year onsite service, parts, and labor warranty comes with each cluster. The service provided by the warranty provides coverage Monday–Friday from 9:00 a.m. to 5:00 p.m. Enhanced services are available to meet customers' needs. HP Services Consulting and Integration support is required to help customers with system startup. Optional training and application/solution development is available.

The XC6000 Clusters can be ordered through the High-Performance Technical Computing (HPTC) Competency Centers within each region. These centers have comprehensive menus and tools for developing customer-specific configurations based on the XC cluster architecture. XC6000 Clusters can be ordered through the region HPTCD Competency Centers and will be delivered through the local Integration Center. The XC6000 Cluster solution will not be available through the HP Watson configuration tool.

This reference guide provides descriptions of some typical configurations based on total node count: 17, 34, 68, 128, and 256 nodes. As noted earlier, other configurations between 17 and 256 nodes are supported as standard offerings and are configured by determining the number of necessary compute building blocks, and adding application nodes in increments of four.

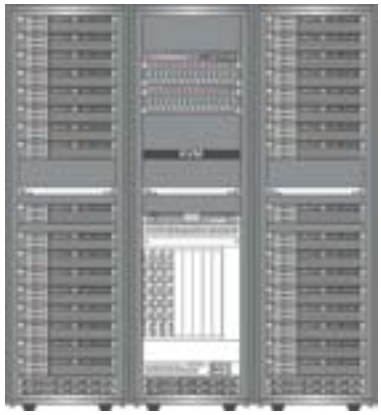
HP XC6000 Clusters: options

17-node cluster

Figure 2. The 17-node Linux cluster based on the Intel Itanium processor consists of one CBB and one UBB building block. The Cluster solution integrates one HP Integrity rx2600 64-bit service node server, 16 HP Integrity rx2600 64-bit application node servers, one HP ProCurve Switch 2650 48-port Ethernet switch, one 32-port Quadrics Elan 3 high-speed interconnect, two HP Rack 10642 (42U) cabinets, one TFT rackmount keyboard/monitor, and an optional MSA1000 storage subsystem. An EVA 3000 storage subsystem is available as an optional storage building block.

34-node cluster

Figure 3. The 34-node Linux cluster based on the Intel Itanium processor consists of two compute building block (CBB) racks and one utility building block (UBB) rack. The CBB racks each include one HP Integrity rx2600 64-bit service node server, 16 HP Integrity rx2600 64-bit application node servers, and one HP ProCurve Switch 2650 48-port Ethernet subnet switch in an HP 10642 42U rack. The UBB rack consists of one 48-port Quadrics Elan 3 high-speed interconnect, a TFT rackmount keyboard/monitor, one 4-port KVM switch, one HP ProCurve Switch 2650 admin switch, and an optional MSA1000 storage subsystem. An optional EVA 3000 storage subsystem is also available.



68-node cluster

Figure 4. The 68-node Linux cluster based on the Intel Itanium processor consists of four compute building block (CBB) racks and a utility building block (UBB) rack. The CBB racks each include one HP Integrity rx2600 64-bit service node server, 16 HP Integrity rx2600 64-bit application node servers, and one HP ProCurve Switch 2650 48-port Ethernet subnet switch in an HP 10642 42U rack. The UBB rack consists of one 80-port Quadrics Elan 3 high-speed interconnect, one TFT rackmount keyboard/monitor, one 4-port KVM switch, one HP ProCurve Switch 2650 admin switch, and an optional MSA1000 storage subsystem. An optional EVA 3000 storage subsystem is also available.

128-node cluster

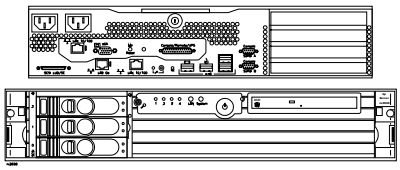
Figure 5. The 128-node Linux cluster based on the Intel Itanium processor consists of seven full compute building block (CBB) racks, one partial CBB rack, and a utility building block (UBB) rack. The full CBB racks each include one HP Integrity rx2600 64-bit service node server, 16 HP Integrity rx2600 64-bit application node servers, and one HP ProCurve Switch 2650 48-port Ethernet subnet switch in an HP 10642 42U rack. The partial CBB rack has only eight application nodes and one service node. The UBB rack consists of one 128-port Quadrics Elan 3 high-speed interconnect, one TFT rackmount keyboard/monitor, one 4-port KVM switch, one HP ProCurve Switch 2650 admin switch, and an optional MSA1000 storage subsystem. An optional EVA 3000 storage subsystem is also available.

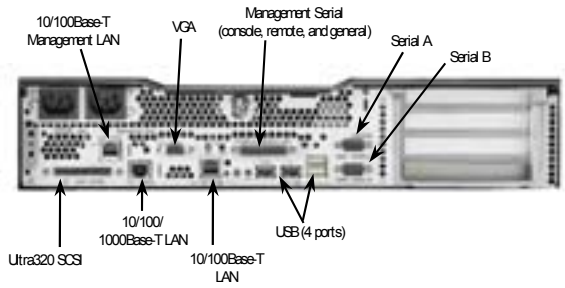
256-node cluster

Figure 6. The 256-node Linux cluster based on the Intel Itanium processor consists of 15 full compute building block (CBB) racks, three IBB interconnect building block racks with 256 switch ports, and a utility building block (UBB) rack. The full CBB racks each include one HP Integrity rx2600 64-bit service node server, 16 HP Integrity rx2600 64-bit application node servers, and one HP ProCurve Switch 2650 48-port Ethernet subnet switch in an HP 10642 42U rack. The IBB racks each include two Elan 3 chassis. The UBB rack consists of one TFT rackmount keyboard/monitor, one 4-port KVM switch, an HP ProCurve Switch 2650 admin switch, a service node, and an optional MSA1000 storage subsystem. An optional EVA 3000 storage subsystem can also be included. Alternative cabinet layouts are possible to fit the site floor plan.

HP Integrity rx2600 server—service/administrative and application nodes

Figure 7. HP Integrity rx2600 server





The service, application, and I/O nodes of the XC6000 Cluster are dual-processor HP Integrity rx2600 servers based on the Intel Itanium architecture. They consist of Integrity rx2600 server packages with these components:

- One HP Integrity rx2600 64-bit 2U server chassis with two Itanium 2 processors. Four different processors are offered: 1.0 GHz with 1.5 MB cache, 1.3 GHz with 3 MB cache, 1.4 GHz with 1.5 MB cache, and 1.5 GHz with 6 MB cache.
  - Processor speed of application and service nodes must be the same
- 1, 2, 4, or 8 GB DDR SDRAM memory (minimum memory per node is 2 GB)
  - Application nodes must have the same memory content
- MP video option (required for sub-network booting)
- CD/DVD optical drive required in the service node, not required in the application node
- N+1 power supply required in the service node, not required in the application node
- Four PCI option slots
- One embedded 10/100BT Ethernet LAN port
- One embedded 10/100/1000BT Ethernet LAN port
- Two universal serial bus (USB) ports
- One hard drive required in each service node and each application node
  - Application nodes must all have the same size and speed disk configuration
- Gigabit Ethernet adapter required in the service node
- FC host bus adapter required in service nodes acting as I/O nodes connected to SAN storage

HP Integrity rx2600 server with single 1.5 GHz/6 MB processor; a second processor; as well as the addition of memory, embedded disk storage, and PCI adapters are required
HP Integrity rx2600 server with single 1.3 GHz/3 MB processor; a second processor; as well as the addition of memory, embedded disk storage, and PCI adapters are required
HP Integrity rx2600 server package with single 1.4 GHz/1.5 MB processors; second processor; as well as the addition of memory, embedded disk storage, and PCI adapters are required
HP Integrity rx2600 server package with single 1.0 GHz*/1.5 MB processor; a second processor; as well as the addition of memory, embedded disk storage, and PCI adapters are required
MP (EIC) card (required for application and service nodes) is included in the rx2600 packages
1.5 GHz/6 MB Intel Itanium 2 processor
1.3 GHz/3 MB Intel Itanium 2 processor
1.4 GHz/1.5 MB Intel Itanium 2 processor
1.0 GHz/1.5 MB Intel Itanium 2 processor
DVD/CD-ROM drive (service node only)
Memory (minimum 2 GB memory per application node and per service node, supporting memory interleaving—2 banks of equal-size memory)
Memory—1 GB chip spare PC2100 DDR SDRAM memory quad (4 x 256 MB)
Memory—2 GB chip spare PC2100 DDR SDRAM memory quad (4 x 512 MB)
Memory—4 GB chip spare PC2100 DDR SDRAM memory quad (4 x 1024 MB)
Memory—8 GB chip spare PC2100 DDR SDRAM memory quad (4 x 2048 MB)
Up to 3 hard drives (note: one hard drive mandatory for swapping; if multiple drives are selected, they must be the same size and speed)
Hard drive—36 GB Ultra320 SCSI 10,000 rpm (1") hot-swap hard disk drive
Hard drive—73 GB Ultra320 SCSI 10,000 rpm (1") hot-swap hard disk drive
Hard drive—142 GB Ultra320 SCSI 10,000 rpm (1") hot-swap hard disk drive
Quadrics QSNet-1 PCI adapter (required in service and application nodes)
Gigabit Ethernet PCI LAN adapter (required in service node)
SAN HBA FC 2214 PCI adapter (required in service node when rx2600 is connected to SAN storage)

*Note: The XC6000 Cluster with 1.0 GHz rx2600 servers is available by special request, please contact HP XC6000 Product Management for more information.

Quadrics Elan 3 high-speed interconnect

Each Quadrics Elan 3 high-speed interconnect consists of these components:

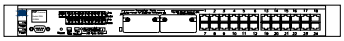
- Quadrics Elan 3 128-port node-level chassis
- Quadrics Elan 3 128-port uplink-level chassis—used in federated switch configurations (IBB only)
- One to eight Quadrics Elan 3 16-port line cards to connect to the application and service nodes per node-level chassis
- One to eight Quadrics Elan 3 16-port uplink line cards per link-level chassis—used in federated switch configurations
- Quadrics Elan 3 clock box—used in federated switch configurations
- Quadrics Elan 3 link cables (varying lengths)
- Quadrics Elan 3 clock cables (varying lengths)
- One Quadrics Elan 3 high-speed interconnect switch is mounted in the utility rack to support up to a 128-node XC6000 Cluster
- Multiple Quadrics Elan 3 high-speed interconnect switches are mounted in the interconnect racks (IBB) (2 per rack) to support XC6000 Clusters with more than 128 nodes in a federated switch configuration

Management network and console interconnect

The HP ProCurve Switch 2650 functions as the administration network and sub-network switch for the XC6000 Cluster. These 48-port Ethernet switches are reliable network components. The HP ProCurve Switch 2650 includes these features:

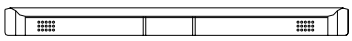
- 48 10/100 ports
- 2 Gigabit link ports
- 10/100 auto-sensing per port automatically detects and sets the speed for any 10Base-T or 100Base-TX device.
- MDI/MDI-X cascade port makes it easy to add other hubs to the network.
- A comprehensive LED display with per-port indicators provides an at-a-glance view of status, activity, and speed.
- Automatic polarity correction and auto-partitioning on all ports: helps find and fix common network problems.
- Built-in bridge automatically connects 10 MB/s and 100 MB/s devices without requiring additional products, modules, or stacking to operate correctly.
- An HP ProCurve 2650 switch is connected to a second HP ProCurve switch (sub-network switch) to form a sub-network consisting of 32 application nodes and two service nodes. Large clusters will have multiple sub-networks for each set at 34 nodes.
- One HP ProCurve 2650 switch (admin switch) is used in configurations with two or more service nodes to connect the MP and administration port of the service nodes in the cluster (located in the UBB rack).
- An HP ProCurve Rackmount kit is required to mount the HP ProCurve 2650 switch into the 10642 rack to enable integration of the switch for shipment.

Figure 8. HP ProCurve 2650 10/100 switch



Monitors and keyboards

Figure 9. HP rackmount TFT5600 keyboard/monitor and display

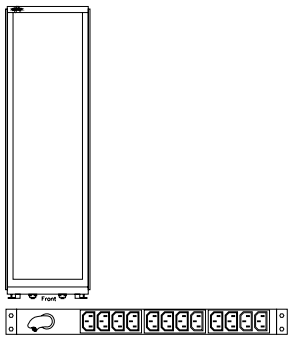


A rackmount keyboard/monitor (RKM) is required with the XC6000. The RKM with TFT display is attached to the HP Integrity rx2600 service node that is designated as the cluster system administration node. The TFT display is mounted in the utility rack (UBB) and provides local user access to the cluster. A four-port KVM switch is required to connect the RKM to multiple service nodes to allow for failover connectivity to other service nodes in cluster systems with two or more service nodes. The TFT display is connected directly to the single service node in the system when there are less than two service nodes in the cluster.

Cabinet and power distribution unit

HP Rack 10642 (42U)—rack cabinet with front and rear doors mounted on a shock pallet
Blanking panels (graphite)
Side panel 42U HP Rack 10000 Series (all)
Power distribution unit (PDU)—HV 24 amp North America/Japan & HV 32 amp international (3 per CBB and 2 per UBB and IBB)
Power cord, terminated, 3-conductor, SPT-2, IEC320-C13
HP ProCurve Rackmount kit
Quadrics Cable Management Accessory kits
Quad Rack Accessory Kit

Figure 10. Power distribution unit—PDU, 24 amp-High, North America, Japan; PDU, 32 amp-High, international. Three PDUs are required for each compute/service node cabinet; two PDUs are required for each utility rack and interconnect rack.



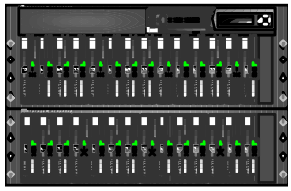
SAN storage

HP StorageWorks Modular SAN Array 1000

The HP StorageWorks Modular SAN Array 1000 (MSA1000) is offered as optional external SAN storage for the XC6000 Cluster. Select one host bus adapter for each designated service I/O node. The required system I/O bandwidth will determine the number of required I/O service nodes connected to the SAN storage subsystem. The MSA1000 storage subsystem is mounted in the UBB cabinet with the Quadrics Elan 3 chassis in XC clusters with 128 nodes or less. In XC clusters with more than 128 nodes, the MSA1000 storage subsystem is mounted in the utility cabinet. Up to 4 TB of storage is supported per MSA1000 subsystem. The MSA1000 support up to two shelves.

If larger capacity or higher bandwidth is required, two MSA1000 storage subsystems can be configured. A ratio of one MSA for one I/O node is recommended for balanced performance in RAID 5 configuration.

Figure 11. HP StorageWorks Modular SAN Array 1000



The HP StorageWorks Modular SAN Array 1000 includes these components:

- One MSA1000 controller with 256 MB cache
- One MSA1000 Fibre Channel I/O module with 2 Gb SFP short wave transceiver
- Redundant hot-pluggable power supply/blower assemblies
- Universal rackmounting kit
- MSA1000 support CD and documentation
- Serial cable
- Two power cables
- Two 3-foot very high density cable interconnect (VHDCI) to VHDCI SCSI cables

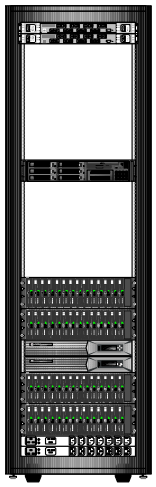
Options:

- MSA1000—one controller, 14 disk slots
- Dual bus disk shelf with 14 slots—SW4454
- MSA1000 SAN switch 2/8 (integrated; includes four 2 GB SFPs, 7 ports available, 4 transceiver equipped
- HP StorageWorks Enclosure Model 4454R (one maximum per MSA1000)
- 2 Gb SFP short-wave transceiver kit
- 36.4 GB hot-pluggable Ultra320 universal hard drive, 10,000 rpm (1")
- 72.8 GB hot-pluggable Ultra320 universal hard drive, 10,000 rpm (1")
- 146.8 GB hot-pluggable Ultra320 universal hard drive, 10,000 rpm (1")
- 36.4 GB hot-pluggable Ultra320 universal hard drive, 15,000 rpm (1")
- 72.8 GB hot-pluggable Ultra320 universal hard drive, 15,000 rpm (1")
- 2 m LC-LC multimode Fibre Channel cable
- 5 m LC-LC multimode Fibre Channel cable
- 15 m LC-LC multimode Fibre Channel cable
- 30m LC-LC Multi-Mode Fibre Channel cable
- 50m LC-LC Multi-Mode Fibre Channel cable

HP StorageWorks Enterprise Virtual Array 3000

HP StorageWorks Enterprise Virtual Array 3000 (EVA 3000) is offered as optional external SAN storage for the XC6000 Cluster. Select one host bus adapter for each designated service I/O node. The required system I/O bandwidth will determine the number of required I/O service nodes connected to the SAN storage subsystem. The EVA 3000 storage subsystem comes mounted in its own HP Rack 10000 Series cabinetry. Up to 8 TB of storage is supported per EVA 3000.

Figure 12. HP StorageWorks Enterprise Virtual Array 3000



The EVA 3000 storage array has these components:

- One 3U controller assembly with two HSV100 controllers that have redundant power supplies
- Two M5114 3U dual-redundant FC loop 14-bay disk enclosures
- 42U graphite storage cabinet with appropriate mounting rails and power
- Virtual controller software for HSV100 dual controllers
- Eight disk drives—select from several available models

Requirements**:**

- HP factory integration (required with each EVA3000 storage subsystem)
- HP field installation—XC6000 installation includes the CBB, IBB, and UBB racks; installation of the EVA3000 storage subsystem is not included

Options:

- 42U EVA cab 60 Hz (based on Series 10000 rack system)
- 42U EVA cab 50 Hz (based on Series 10000 rack system)
- 2C2D EVA 3000-C/8 x 36 GB/15 HDD with Foundation Service solution
- 2C2D EVA 3000-C/8 x 72 GB/10 HDD with Foundation Service solution
- 2C2D EVA 3000-C/8 x 72 GB/15 HDD with Foundation Service solution
- 2C2D EVA 3000-C/8 x 146 GB/10 HDD with Foundation Service solution
- M5314 FC drive enclosure
- 36 GB/10K dual–port 2 Gb FC-AL 1" drive
- 36 GB/15K dual–port 2 Gb FC-AL 1" drive
- 72 GB/10K dual–port 2 Gb FC-AL 1" drive
- 146 GB/10K dual–port 2 Gb FC-AL 1" drive
- 72 GB/15K dual–port 2 Gb FC-AL 1" drive
- SAN Switch 2/8-EL
- SAN Switch 2/16
- 2 Gb SFP transceiver kit

- 2m LC-LC multimode Fibre Channel cable
- 5m LC-LC multimode
- 15m LC-LC multimode Fibre Channel cable
- 30m LC-LC multimode Fibre Channel cable
- 50m LC-LC multimode Fibre Channel cable
- HP StorageWorks Virtual Controller Software v3.0 media kit for dual HSV100 controllers
- Storage Management Appliance III
- FCA2214 2 GB FC host adapter (**note:** the XC6000 v1.0 Cluster software supports the FCA2214 FC host adapter on the Integrity rx2600 server as a special offer; the Integrity rx2600 server does not normally support this adapter outside of XC6000 Cluster solutions)

Hardware documentation

HP XC6000 Cluster systems come with two hardware documentation kits. Each hardware documentation kit includes both a CD (contains a soft copy of the *User System and Installation Guide* and hardware release notes) and a hard copy of the *User System and Installation Guide*. Two sets of HP Integrity rx2600 server documentation are also included. Additional copies of the XC6000 Cluster hardware documentation kits can be ordered separately.

HP XC System Software and documentation

HP XC System Software v1.0 is a fully integrated and supported cluster management software solution developed by HP. It consists of a Linux kernel, cluster management software, Platform LSF scheduling software, license management software, message-passing interface, MLIB tools, RPMs, and system management modules. The HP XC System Software kit includes binaries, sources, and documentation on CD; hard-copy documentation; and a software use license. Two software licenses are required for each cluster system.

- HP XC System Software v1.0 base license kit—comes with a base unlimited software license, one software media kit, administrator documentation kit, and user documentation kit
- HP XC System Software v1.0 CPU license—one license per CPU is required (must be ordered in addition to base license kit); comes in 1, 2, 16, 64, 128, 256, and 512 CPU license packages

Additional software media kits, administrator documentation kits, and user documentation kits may be ordered separately.

Development tools for XC6000

- **Fortran and C/C++ compilers** from Intel are provided at added cost from Intel or a reseller to use with the HP XC System Software. The toolkits include the compiler license, documentation, and release notes.
- **TotalView** parallel debugger from Etnus, Inc. is provided at added cost from the vendor or a reseller.
- **Vampir/Vampirtrace** MPI performance analysis tool from Pallas is provided at added cost from the vendor or a reseller.

HP factory rack integration

The factory rack integration provided by HP for the XC6000 Cluster includes the following:

- Factory integration must be ordered for each XC6000 system at the time of the intial system order (check with your local Competency Center for factory integration center charges)
- Staging and integration of the HP Integrity rx2600 servers, storage devices, and peripheral devices

- Review of customer order for required licenses, cables/cable lengths, software revisions, and hardware
- Thorough test of the system, running extensive diagnostics and system exercisers for an extended period of time to reasonably facilitate a problem-free installation and ongoing reliability
- Configuration of Linux cluster includes these services:
  - Consolidation and routing of all material across the customer purchase orders for integration; technical edit and verification of the configuration against the order
  - Verification of Hardware configuration
  - Cabling of entire cluster, labeling of all inter-cabinet cabling
  - Configuration of cluster interconnect—Quadrics Elan 3
  - Configuration of each node's MP port and network port
  - Configuration of service and administration nodes
  - Loading of Linux software on the system and layered products
  - Configuration of disk partitions
  - Setup of IP addresses for systems and all network equipment
  - Loading of any custom configuration files
  - Verification of cluster configuration
- System labeling, including color-coding or point-to-point labels of all cables, device labels, and system node name labels
- Configuration of external storage, SAN storage, or local embedded file system storage

Mandatory HP field installation

All HP XC6000 customers are required to have onsite customer installation of the system provided by HP technicians. This includes the unpacking of the equipment, inspecting the equipment for damage, positioning and joining the cabinets together as required, cabling up the system (power and data cables), and verifying proper operation of the equipment and running diagnostics. There is a charge for the XC6000 Cluster cabinets, as well as a separate charge for EVA Storage cabinet installation (using the normal EVA installation charges). The field installation activities are focused on setting up the equipment and not on the software installation. HP Consulting and Integration (C&I) activities cover the setup and startup of the XC6000 Cluster management software. The installation for each cluster is ordered at the time of the initial system purchase and must be included on the order.

Mandatory HP Consulting and Integration Services

All HP XC6000 customers are required to have onsite customer integration management and systems software knowledge transfer sessions provided by HP technicians. This must be ordered at the time of the initial system purchase. Following are the required startup services and optional services available for each size of XC6000 Cluster system. Optional XC Cluster user training is also available.

- Required startup services
  - One-day onsite systems software knowledge transfer for less than 18 nodes
  - Two-day onsite systems software knowledge transfer for 18–34 nodes
  - Three-day onsite systems software knowledge transfer for 35–69 nodes
  - Five-day onsite systems software knowledge transfer for 70–136 nodes
  - Ten-day onsite systems software knowledge transfer for 137–272 nodes

- Optional services—recommended for less than 34 nodes
  - Cluster Integration Management
  - Cluster Systems QuickStart
  - Cluster Applications QuickStart
- Optional services—recommended for 34–135 nodes
  - Cluster Integration Management
  - Cluster Systems QuickStart
  - Cluster Applications QuickStart
- Optional services—recommended for 136–272 nodes
  - Cluster Integration Management
  - Cluster Systems QuickStart
  - Cluster Applications QuickStart
- Optional training
  - Cluster systems administration course
  - Cluster applications migration course

Contact HP C&I Services for more details and information on configurations for additional C&I Service offerings.

HP customer support

HP customer support provides onsite hardware break/fix support and remote remedial software call-center support. Customer support offsite software services include level 1 and 2 support:

- Level 1 is defined as everyday user/system administration issues.
- Level 2 is defined as issues relating to installation and configuration problems, along with other problems not solvable by following the vendor-supplied documentation.

The HP software support team will work in parallel with the appropriate vendor and development groups to address level 3 and 4 support elevations:

- Level 3 elevations typically require that patches and modifications be generated by the vendor to resolve deficiencies in the product.
- Level 4 elevations deal with enhancements in the functionality of the product that will typically be included in future releases.

Software contracts can be tailored to meet any customer needs, including remedial break/fix, migration and upgrade planning, and a full suite of proactive deliverables. These services are available through HP Care Packs—to cover all needed levels of services for a limited number of instances—or through software contracts, from the Bronze up to the Platinum or Custom levels of service.

Technical specifications

HP Integrity rx2600 server—head node, management server, and compute node

Dimensions (h x w x d)		3.4 x 19.0 x 26.8 in./ 8.6 x 48.5 x 68.1 cm
Weight	Maximum	56 lb./25 kg
	No drive	38.6 lb./17.5 kg
Input requirements (per power supply)	Range line voltage	100 to 240 VAC
	Nominal line voltage	100–127 V/200—240 V
	Rated input current	7.2A @ 120 V; 3.6A @ 220 V shared across inputs (power supply)
	Rated input frequency	50 to 60 Hz
	Rated input power	Typcial power input: 600 W Maximum power input: 714 W
BTU rating	Maximum heat dissipation Typical heat dissipation	4375 BTU/hr. 1950 BTU/hr.
Power supply output power (per power supply)	Theoretical maximum power dissipation Typical maximum power dissipation	1350 W 600 W
Temperature range	Operating	41°F to 95°F/5°C to 35°C
	Shipping	–40°F to 158°F/–40°C to 70°C
Relative humidity (non-condensing)	Operating	15% to 80%
	Non-operating	8.5% to 85%
Acoustic noise	Idle (fixed disk drives spinning)	
	L_{WAd} (bels)	6.4
	Maximum configuration	
	L_{WAd} (bels)	7.2

Quadrics Elan 3 high-speed interconnect chassis

Maximum BTUs per hour (with line cards in all slots)	N/A
Power (max.)	1000 W
Max input current (100 to 127 V)	6 amps
Max input current (200 to 240 V)	3.25 amps
Height (1U = 1.75 inches = 44.45 mm)	22.7 in./57.7 cm/13U
Width	17.3 in./44 cm
Depth	36 in./91.5 cm
Weight without line cards	N/A
Weight with line cards in all slots	154 lb./70 kg
Temperature (operating)	41° F to 95° F/5° C to 35° C
Temperature (storage)	N/A
Relative humidity (operating)	10% to 90% non-condensing
Power input voltage	220 to 250 VAC 50/60Hz

HP ProCurve Switch 2650 48-port 10/100 switch

Dimensions (h x w x d)	17.4 x 12.8 x 1.7 in./44.0 x 32.5 x 4.4 cm
Weight (single power system)	9.8 lb./ 4.44 kg
Temperature (operating)	32° F to 131° F/0° C to 55° C
Temperature (storage)	–40° F to 158° F/–40° C to 70° C
Humidity	Operating—15 to 95% @ 104° F/40° C, non-condensing Non-operating—15 to 90% @ 149° F/65° C, non-condensing
Input voltage	100 to 127 VAC/ 200 to 240 VAC 50/60 Hz
Maximum current	1.5 A
Power	36 W
Heat dissipation	123 BTU/hr.

HP Rack 10642 (42U)—base cabinet (empty)

	Total cabinet	**Shipping**
Height	78.7 in./199.90 cm	86.22 in./219 cm
Depth	39.69 in./100.82 cm	48 in./121.92 cm
Width	24 in./60.96 cm	32 in./81.28 cm
Weight	253 lb/114.8 kg	325 lb/147.4 kg
PDU (qty. varies with cabinet)	3 PDU for CBB rack and 2 PDU for UBB and IBB rack	
Color	Doors: graphite metallic; frame: carbon	

HP StorageWorks Modular SAN Array 1000

Refer to MSA1000 specifications on the HP StorageWorks Web site.

The MSA1000 storage array is mounted in the XC6000 utility rack when configured with an XC6000 Cluster.

HP StorageWorks Enterprise Virtual Array 3000

Refer to EVA 3000 specifications on the HP StorageWorks Web site.

The EVA 3000 comes configured and mounted in its own standalone 10000 series cabinets when configured with an XC6000 Cluster.

HP XC6000 compute building block rack

The compute building block has one 10642 rack, 16 HP Integrity rx2600 application nodes, one rx2600 service node, one HP ProCurve 2650 switch, three PDUs, and cabling (maximum configuration).

Height	**Total cabinet** 78.7 in./199.90 cm	**Shipping** 86.22 in./219 cm
Depth	39.69 in./100.82 cm	48 in./121.92 cm
Width	24 in./60.96 cm	32 in./81.28 cm
Weight (est.)	1200 lb./545 kg	1272 lb./578 kg
Power connection—3 PDUs	NA/Japan—220 to 240, 24 amp, 50/60 Hz (each PDU)	Int—220 to 240, 32 amp, 50/60 Hz (each PDU)
Power consumption (est.)	10,236 W (prelim.)	
Heat dissipation (est.)	32,610 BTU/hr. (prelim.)	
Temperature range (est.)	50° to 95° F/10° to 35° C operating (prelim.)	–22° to 122° F/–30° to 50° C non-operating (prelim.)
Relative humidity (est.) (non-condensing)	15% to 80% operating (prelim.)	5% to 95% non-operating (prelim.)

HP XC6000 utility building block rack

The utility building block has one 10642 rack, one Quadrics chassis, one TFT display, one HP ProCurve 2650 admin switch, one StorageWorks MSA1000 Storage Array, two PDUs, and cabling (maximum configuration).

Height	**Total cabinet** 78.7 in./199.90 cm	**Shipping** 86.22 in./219 cm
Depth	39.69 in./100.82 cm	48 in./121.92 cm
Width	24 in./60.96 cm	32 in./81.28 cm
Weight (est)	600 lb./272 kg	672 lb./305 kg
Power connection—2 PDUs	NA/Japan—220 to 240, 24 amp, 50/60 Hz (each PDU)	Int—220 to 240, 32 amp, 50/60 Hz (each PDU)
Power consumption (est.)	2750 W (prelim.)	
Heat dissipation (est.)	8500 BTU/hr. (prelim.)	
Temperature range (est.)	50° to 95° F/10° to 35° C operating (prelim.)	–22° to 122° F/–30° to 50° C non-operating (prelim.)
Relative humidity (est.) (non-condensing)	10% to 90% operating (prelim.)	5% to 95% non-operating (prelim.)

HP XC6000 interconnect building block rack

The interconnect building block has one 10642 rack, two Myricom chassis, and two PDUs (maximum configuration).

Height	**Total cabinet** 78.7 in./199.90 cm	**Shipping** 86.22 in./219 cm
Depth	39.69 in./100.82 cm	48 in./121.92 cm
Width	24 in./60.96 cm	32 in./81.28 cm
Weight (est)	700 lb./318 kg	772 lb./350 kg
Power connection—2 PDU	NA/Japan—220 to 240, 24 amp, 50/60 Hz (each PDU)	Int—220 to 240, 32 amp, 50/60 Hz (each PDU)
Power consumption (est.)	2500 W (prelim.)	
Heat dissipation (est.)	7200 BTU/hr. (prelim.)	
Temperature range (est.)	50° to 95° F/10° to 35° C operating (prelim.)	–22° to 122° F/–30° to 50° C non-operating (prelim.)
Relative humidity (est.) (non-condensing)	10% to 90% operating (prelim.)	5% to 95% non-operating (prelim.)

Appendix—XC6000 product menu

Refer to the HP XC Web site at **www.hp.com/techservers/clusters/xc_clusters.html** for the latest XC6000 product menu, and consult the HPTC Competency Center in your region for ordering information.

5982-1780EN, 03/25/2004