HP BladeSystem c-Class enclosure

technology brief

Abstract

As server, storage and network technologies have changed, traditional rack-mounted infrastructure solutions have generated increased complexity of both infrastructure management and physical data center deployment. The HP BladeSystem c-Class enclosure is an evolution of the entire rack-mounted infrastructure. It consolidates and repackages all the supporting infrastructure elements—compute, storage, network, and power—into a single "rack-in-a-box" that accelerates the integration and optimization of the data center. This technology brief provides an overview of the HP BladeSystem c-Class enclosure, including Thermal Logic power and cooling technologies and interconnect options.

This technology brief assumes the reader is familiar with HP ProLiant server technology and has some knowledge of general BladeSystem architecture. For more information about the infrastructure components, see the HP website at www.hp.com/go/bladesystem/.

Overview of HP BladeSystem c7000 enclosure

The HP BladeSystem c7000 enclosure, announced in June 2006, is the first enclosure implemented using the BladeSystem c-Class architecture. It is optimized for enterprise data centers. It is designed to fit into standard size HP and third-party racks, and to accommodate the c-Class form-factor server/storage blades and interconnect modules. HP expects to release other enclosure sizes in the future that will meet those same requirements while being optimized for other computing environments, such as remote sites or small businesses. More information on c-Class architecture is available in the technology brief entitled "HP BladeSystem c-Class architecture" on the HP technology website at www.hp.com/servers/technology.

The BladeSystem c7000 enclosure provides all the power, cooling, and I/O infrastructure needed to support modular server, interconnect, and storage components today and throughout the next several years. The enclosure is10U high and holds up to 16 server and/or storage blades plus optional redundant network and storage interconnect modules. It includes a shared, 5 terabit per second high-speed NonStop midplane for wire-once connectivity of server blades to network and shared storage. Power is delivered through a pooled-power backplane that ensures the full capacity of the power supplies is available to all server blades.

Figure 1. HP BladeSystem c7000 enclosure – side view
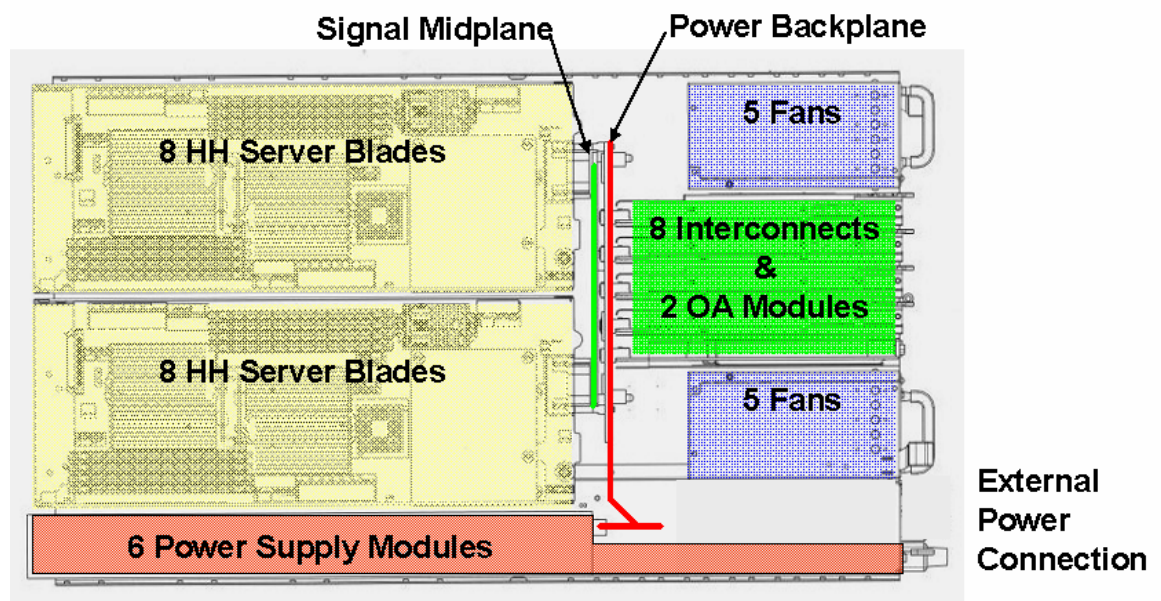
The HP BladeSystem c7000 enclosure has a fully redundant design with redundant interconnections between servers and interconnect modules. The enclosure contains a NonStop signal midplane and a power backplane that have no active components. The enclosure is available with a choice of single-phase or three-phase power subsystem, depending on the data center power infrastructure.
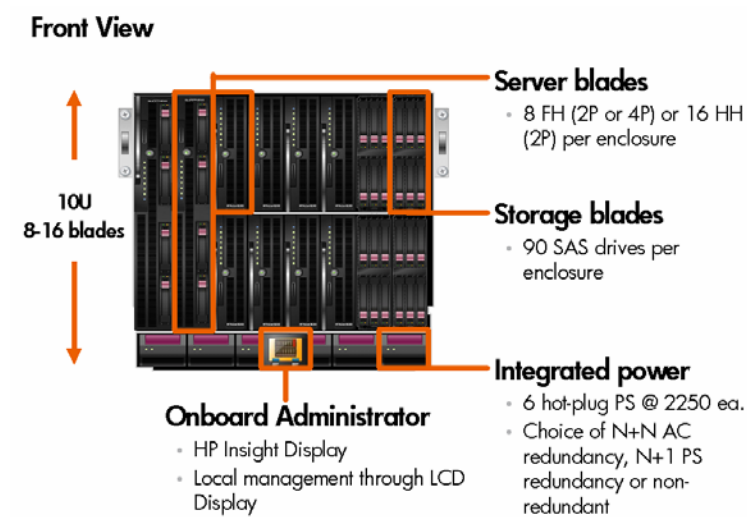
The c7000 enclosure can be populated with the following components:

• Up to 8 full-height (FH) server blades or up to16 half-height (HH) server and/or storage blades per enclosure
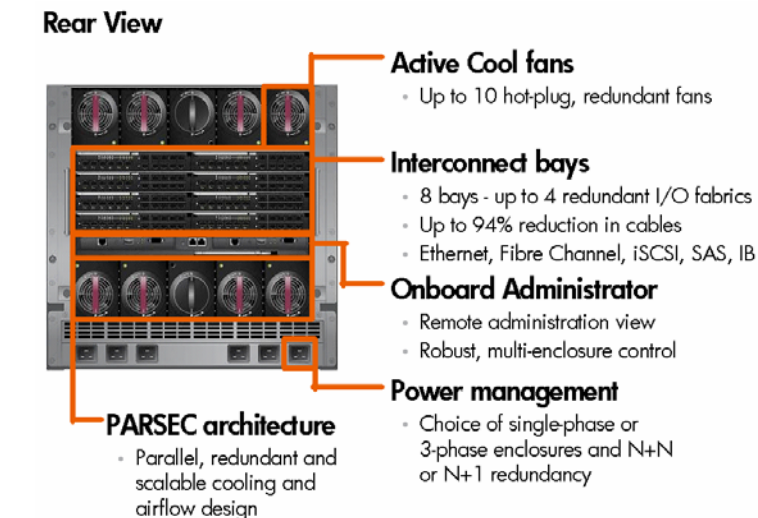
- Up to eight interconnect modules supporting a variety of network interconnect fabrics such as Ethernet, Fibre Channel (FC), InfiniBand (IB), Internet Small Computer System Interface (iSCSI), or Serial-attached SCSI (SAS) simultaneously within the enclosure
- Active Cool fan kits for a maximum of ten fans
- Up to six power supplies
- Redundant Onboard Administrator (OA) management modules (optional active-standby design)

All devices mentioned above are customer replaceable and hot-pluggable.

Figure 2. HP BladeSystem c7000 enclosure – front view

Front View

10U
8-16 blades

Server blades
- 8 FH (2P or 4P) or 16 HH (2P) per enclosure

Storage blades
- 90 SAS drives per enclosure

Integrated power
- 6 hot-plug PS @ 2250 ea.
- Choice of N+N AC redundancy, N+1 PS redundancy or non-redundant

Onboard Administrator
- HP Insight Display
- Local management through LCD Display

Figure 3. HP BladeSystem c7000 enclosure – rear view

Rear View

Active Cool fans
- Up to 10 hot-plug, redundant fans

Interconnect bays
- 8 bays - up to 4 redundant I/O fabrics
- Up to 94% reduction in cables
- Ethernet, Fibre Channel, iSCSI, SAS, IB

Onboard Administrator
- Remote administration view
- Robust, multi-enclosure control

Power management
- Choice of single-phase or 3-phase enclosures and N+N or N+1 redundancy

PARSEC architecture
- Parallel, redundant and scalable cooling and airflow design

HP Thermal Logic technologies

Thermal Logic is the term that HP uses to define the mechanical design features, built-in intelligence, and control capabilities throughout the BladeSystem c-Class that enable IT administrators to make the most of the power and thermal environments. HP Thermal Logic technology in the c-Class enclosure requires up to 30 percent less airflow and up to 50 percent less power for cooling than the equivalent number of rack servers while the c-Class components take up less rack space. It provides an instant view of power usage and temperature at the server, enclosure, or rack level. In the future Thermal Logic technology will include the ability to automatically adjust power load, workload and thermal controls to maximize performance, power and cooling capacity for each unique environment.

HP Thermal Logic technology encompasses the following elements and capabilities:

- Active Cool fans
- Parallel Redundant Scalable Enclosure Cooling (PARSEC) design
- Instant Thermal Monitoring
- Pooled power for true N+N power redundancy
- Dynamic Power Saver mode
- Power Regulator
- Power workload balancing

Active Cool fans

Quite often, dense, full-featured, small form-factor servers use very small fans designed to provide localized cooling in the specific areas needed by the server blade. Because such fans generate fairly low airflow (in cubic feet per minute, or CFM) at medium back pressure, a single server often requires multiple fans to ensure adequate cooling. Therefore, installing many server blades together in an enclosure, with each blade containing several fans, can result in a significant amount of cost and space overhead.

A second solution for cooling is to use larger, blower-style fans that can provide cooling across an entire enclosure. Such fans are good at generating CFM, but they typically require higher power input, take up more space, make more noise, and must be designed for the maximum load in an enclosure. As a result, designers may have to sacrifice server features to allow the large, high-power fans to fit in the enclosure. Even then, ensuring adequate airflow to all the servers without leakage, over provisioning, or bypass is a challenge.

To overcome these issues for the HP BladeSystem c-Class, HP engineers designed a new type of fan based on aircraft technology that delivers both high airflow and high pressure in a small form factor that can scale to meet future cooling needs. The HP Active Cool fan technology provides the ability to optimize airflow, reduce power draw, and improve acoustic performance for any server blade configuration.

With 20 patents pending involving the Active Cool fan technology and its implementation, HP Active Cool fans are an innovative new design that can cool 16 blades using as little as 150 watts (W) of power. Active Cool fans use ducted fan technology (the fan is longer than it is wide) with a high performance motor and impeller to deliver high CFM at a high pressure. The fan includes a bell mouth inlet with a specially-designed impeller, followed by a stator section that also provides cooling fins for the motor and acoustic treatments in the rear of the fan. This design not only has cooling capacity to support blade products beyond current roadmaps, but its unique shape allows for high-pressure flow with low noise and power at even the slowest speeds.

Figure 4. Ducted fan cross-section and ducted fan blade compared to traditional server fan

Ducted fan cross-section
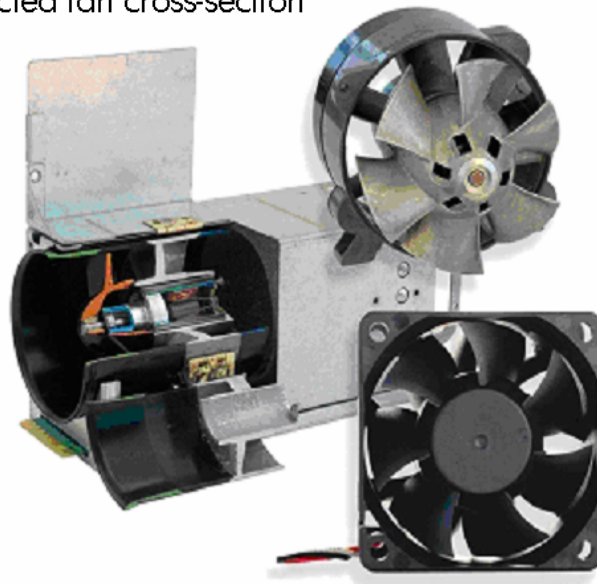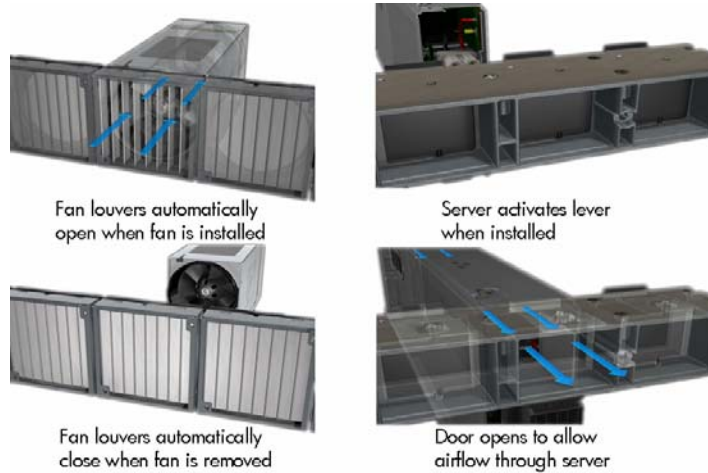
Ducted fan blade

Traditional server fan

Active Cool fans are controlled by the c-Class Onboard Administrator so that cooling capacity can be ramped up or down based on the needs of the entire system. Along with optimizing the airflow, this control algorithm allows the BladeSystem c-Class to optimize the acoustic levels and power consumption. As a result, the c-Class enclosure can accommodate full-featured servers that are 60 percent more dense than traditional rack-mount servers, while the Active Cool fans consume only 50 percent of the power typically required and use 30 percent less airflow.

HP PARSEC architecture

HP Parallel Redundant Scalable Enclosure Cooling (PARSEC) architecture is a hybrid model for cooling that combines the best of local and centralized cooling in a single system to ensure optimum airflow and cooling for all servers. Density, once a barrier to cooling, is turned into an advantage with HP Thermal Logic technologies like HP PARSEC architecture and HP Active Cool fans. With these innovations, server blades get more cooling airflow where it is needed most and use less power than traditional rack servers.
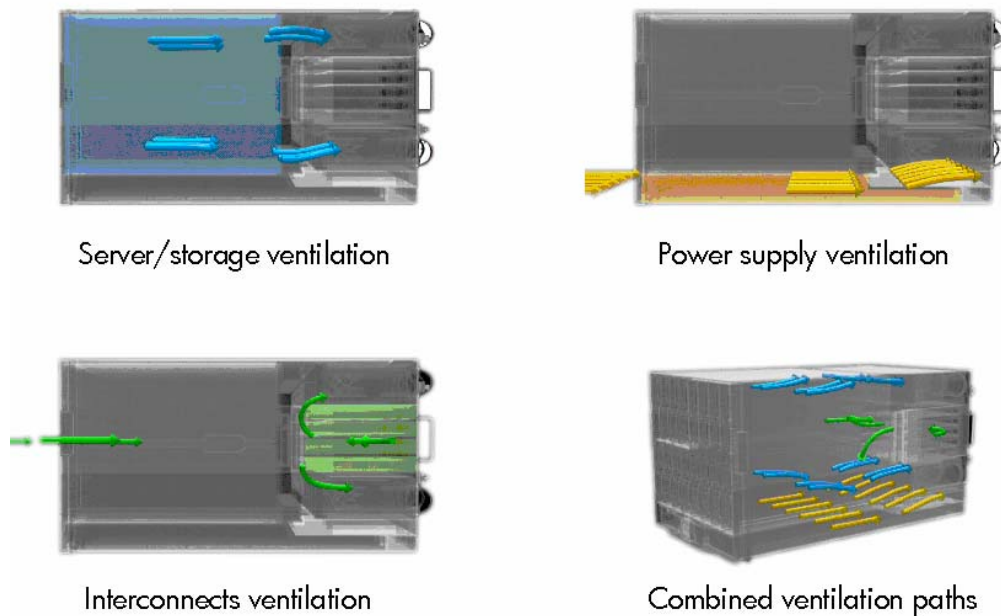
To optimize thermal design, HP developed a relatively airtight center air plenum. For example, all device bays include a shutoff door that is normally closed to prevent air leakage through that bay. When a server blade is inserted, it seals into the center air plenum docking collar and the server shut-off door opens to allow airflow across that server blade. Similarly, the fan seals into the center air plenum docking collar and each fan bay includes louvers that automatically open when a fan is installed. If a fan is uninstalled or not spinning, the pressure distribution around the fan changes. This pressure change causes the louvers to close, ensuring that cooling air is not diverted through the non-operating fan (Figure 5).

Figure 5. HP BladeSystem c7000 self sealing enclosure

Fan louvers automatically open when fan is installed

Fan louvers automatically close when fan is removed

Server activates lever when installed

Door opens to allow airflow through server

The enclosure and the components within it optimize the cooling capacity through unique mechanical designs. Airflow through the enclosure is managed to ensure that every device gets cool air and does not sit in the hot exhaust air of another device, and to ensure that air only goes where it is needed for cooling. Fresh air is pulled into the interconnect bays through a side slot in the front of the enclosure. Ducts move the air from the front to the rear of the enclosure, where it is then pulled into the interconnects and the central plenum, and then exhausted out the rear of the system (Figure 6).

Figure 6. Airflow through the HP BladeSystem c7000 enclosure—side view

Server/storage ventilation

Power supply ventilation

Interconnects ventilation

Combined ventilation paths

Redundancy is built into the design. If any fan fails, it results in only 10 to 25 percent loss in cooling capacity. As illustrated in Figure 7, c-Class enclosures are divided into four cooling zones with fans

located in each zone to provide direct cooling for blades in that zone and redundant cooling for adjacent zones. See Appendix B for more detailed fan and server population guidelines.

Figure 7. The c7000 enclosure is split into four cooling zones. Each zone can contain two full-height blades or four half-height blades. Full-height blades couple the two left or two right zones into two vertical zones. Fan 3 and fan 8 support two zones and will operate at the higher speed of the two zones.
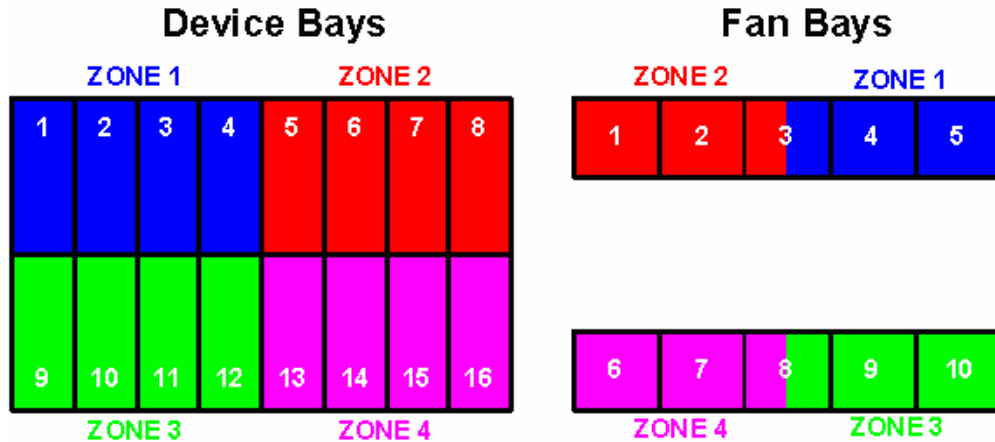
To operate c-Class blades, a minimum of four fans must be installed at the rear of the c-Class enclosure; however, up to ten fans can be installed there so that cooling capacity can scale as needs change. Using more fans allows the fans to spin slower to move the same volume of air, so each fan uses less power. Eight fans are almost always more power efficient than four fans and, as the flow rate increases, ten fans are even more efficient (Figure 8). As an added bonus, slower spinning fans create less noise.

Figure 8. Number of fans versus power draw. The circled area below indicates the point at which ten fans are more efficient than eight fans for the same airflow delivered.

Power versus CFM
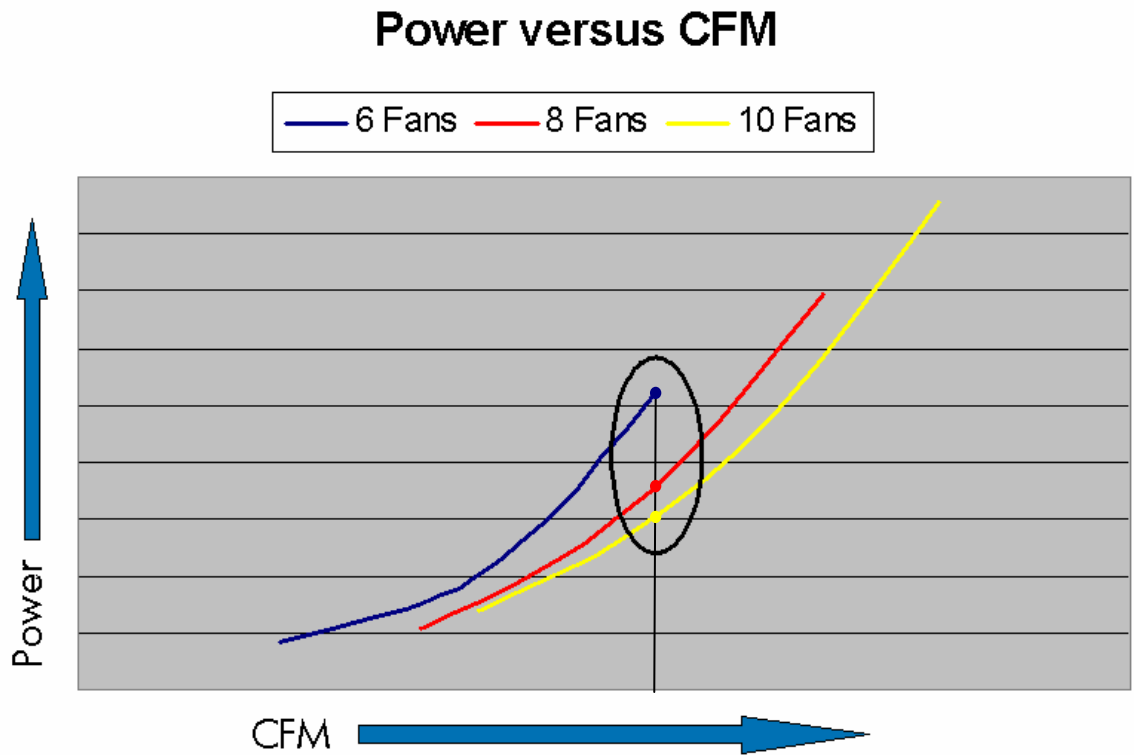
PARSEC at the server blade

The PARSEC architecture also uses precise ducting throughout the server blade to manage airflow and temperature based on the unique thermal requirements of all the critical components. The airflow is tightly ducted to make every gram of airflow count—ensuring no air bypasses the blade, so it obtains the most thermal work from the least amount of air. This concept allows much more flexibility in heatsink design choice. The heatsink design closely matches the requirements of the blade and processor architecture. For example, in the BL460 Intel® Xeon based server blade, HP was able to use a smaller, high-power processor heatsink compared to rack-mount servers. These heatsinks have vapor chamber bases, thinner fins, and tighter fin pitch compared to existing designs. This allows the largest heat transfer surface in the smallest package (Figure 9). The smaller heatsink allows more space for full-size DIMM sockets and hot plug hard drives on the server blades.

Ducting produces high pressure which allows the server blade to use less airflow, thus resulting in lower fan power required. The lower airflow requirement has the added benefit of optimizing the available data center cooling capacity that is one of the largest constraints IT facilities are up against today and in the future.

Figure 9. Processor heatsink using fully ducted design (left) vs. traditional heatsink in a 1U rack-mount server (right)

Instant Thermal Monitoring

Instant Thermal Monitoring provides a real-time view of heat, power, and cooling data. The Onboard Administrator retrieves thermal information from all server/storage blades and interconnect modules in the enclosure to ensure an optimal balance between cooling, acoustic levels and power consumption. There are more than 150 temperature sensors in a c7000 enclosure filled with half-height blades. If the enclosure's thermal load increases, the Onboard Administrator's thermal logic feature instructs the fan controllers to increase fan speeds to accommodate the additional demand. Even better, it works in reverse, using all the features of Thermal Logic to keep fan and system power at the lowest level possible. Typically, a distributed element manager queries the thermal conditions of the hardware every few minutes in a polling cycle. Onboard Administrator, on the other hand, monitors the thermal conditions on the hardware in real-time, without a delay for a polling cycle.

Power

Customers can choose from three types of HP BladeSystem c7000 enclosures to select the one that will work best with their data center power infrastructure:

- Single-phase enclosure that accepts IEC C19-C20 power cords (available worldwide for use with in-rack PDUs)
- Three-phase enclosure with a pair of US/Japan power cords with NEMA L15-30P power connectors
- Three-phase enclosure with a pair of international power cords with IEC 309, 5-Pin, 16A power connectors

The BladeSystem c-Class houses the power supplies in the same enclosure as the server/storage blades and interconnect modules. The power supply modules connect to a passive power backplane that distributes the power to all the components in a shared manner. HP sized the enclosure for 30A three-phase input lines and added power and cooling intelligence to the enclosure.

Moving the power supplies into the enclosure allowed HP to reduce the transmission distance for DC power distribution and to use an industry-standard 12V infrastructure for the BladeSystem c-Class. Using a 12V infrastructure allowed HP to eliminate several power-related components and improve power efficiency on the server blades and in the infrastructure. The control circuitry was stripped and put on the management board and fans.

NOTE

The c-Class power module in the bottom rear of the enclosure configures whether the enclosure operates on single-phase or three-phase power. This choice is available when the enclosure is ordered.

The enclosure can contain up to six 2250W self-cooled power supplies, for a maximum of 13500W output capacity per enclosure. A pair of PDUs is required for AC line redundancy. Table 1 shows other PDU options.

Table 1. PDU options for c7000 enclosure

| Model	Plug Type	Capacity	# of Enclosures Supported
24 Amp PDU (NA/JPN)	L6-30P	4992 VA @ 208V	1
32 Amp PDU (Intl)	IEC-309 3-Pin 32A	7360 VA @ 230V	1 or 2 depending on load
40 Amp PDU (NA/JPN)	Hubbell CS8265C (L6-50P equiv.)	8320 VA @ 208V	1 or 2 depending on load
40 Amp PDU (Intl)	IEC-309 3-Pin 63A	9200 VA @ 230V	1 or 2 depending on load
17.3 kVA S348 Monitored, single input, 3-phase, 60A (NA/JPN)	IEC-309 4-Pin 63A	17.2 kVA @ 208V	Up to 3 @ 5700 VA
22 kVA S332 Monitored, single input, 3-phase, 32A (Intl)	IEC-309 5-Pin 32A	22 kVA @ 230V	Up to 4 @ 5500 VA

Pooled power

All the power in the enclosure is provided as a single pool that any blade can access. This provides maximum flexibility when configuring the power in the system so that customers can choose what level of redundancy with which to operate. Because this power design has no zones, it facilitates both N+N and N+1 power modes, which future-proofs the enclosure for higher power requirements, if needed. Therefore, looking forward at least 5 years, HP believes there is sufficient power capacity to handle future power-hungry devices.

If needed, customers can constrain the maximum BTUs per enclosure and rack to enable the enclosure to fit in an existing rack power envelope.

The power supplies (Figure 10) have three redundancy modes: AC redundant, power supply redundant, and no redundancy mode.

Figure 10. HP BladeSystem c7000 power supplies
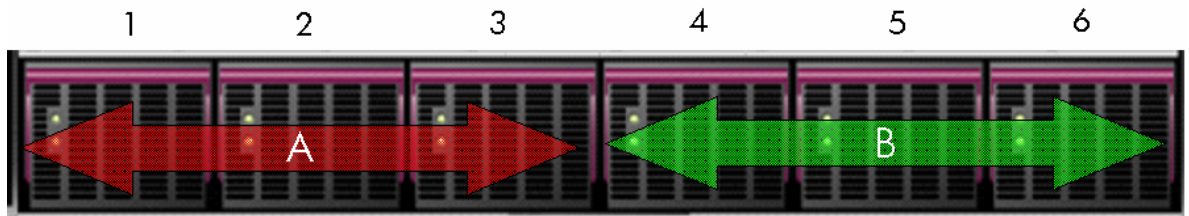
If N+N AC redundancy is configured, then total power available is the amount from the A or B side containing fewer power supplies (3 + 3 configuration = 6750W). In this configuration, N power supplies are used to provide power and N are used to provide redundancy, where N can equal 1, 2 or 3. Any number of power supplies from 1 to N can fail without causing the enclosure to fail. When correctly wired with redundant AC line feeds, this will also ensure that an AC line feed failure will not cause the enclosure to power off.

If N+1 power supply redundancy is configured, then the total power available is defined as total power available less one power supply (5+1 configuration = 11250W). Up to six power supplies can be installed, with one power supply always reserved to provide redundancy. In the event of a single power supply failure, the redundant power supply will take over the load of the failed power supply. An AC line feed failure, or failure of more than one power supply, will cause the system to power off.

If no power redundancy is configured, the total power available is defined as the power available from all power supplies installed (6 power supplies installed = 13500W). Any power supply or AC line failure will cause the system to power off.

The Onboard Administrator manages power allocation rules of various components and can limit overall power capacity for the enclosure. More information on power management is available in the technology brief entitled "Managing the HP BladeSystem c-Class" available at this URL: www.hp.com/servers/technology.
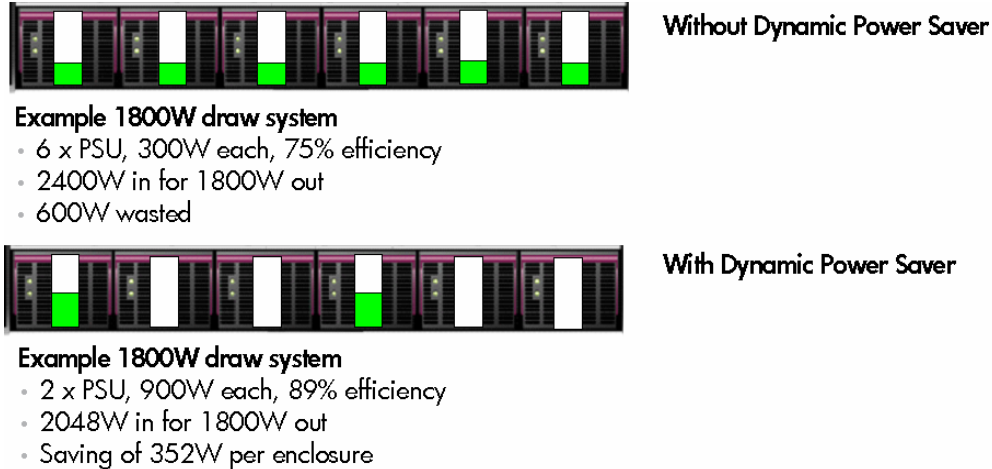
Dynamic Power Saver mode

Dynamic Power Saver mode provides power load shifting for maximum efficiency and reliability. Dynamic Power Saver technology, first introduced with the BladeSystem p-Class1U power enclosure, maximizes power supply efficiency to provide real customer power savings and, therefore, money savings. Power supply efficiency is simply a measure of AC watts input in versus DC watts out, so at 50 percent efficiency, 2000W in would equal 1000W out. The difference is wasted energy, which is costly and generates unnecessary heat.

The Dynamic Power Saver mode is active by default since it saves power in the majority of situations. When enabled, Dynamic Power Saver mode saves power by running the required power supplies at a higher rate of utilization and putting unneeded power supplies in a standby mode. Dynamic Power Saver uses the fact that most power supplies will operate inefficiently when lightly loaded and more efficiently when heavily loaded. A typical power supply running at 20 percent load could have an efficiency as low as 60 percent, but at 50 percent load it could be up to 90 percent efficient, providing a significant saving in power consumption.

In the first example in Figure 11, without Dynamic Power Saver, power demand is low and spread inefficiently across six power supplies. In the second example, with Dynamic Power Saver active, the power load is shifted to two power supplies for more efficient operation. The remaining power

supplies are placed in a standby condition when the power demand from the server enclosure is low. When power demand increases, the standby power supplies instantaneously deliver the required extra power. This enables the power supplies to operate at optimum efficiency, with no affect on redundancy.

Figure 11. Example of power consumption with and without Dynamic Power Saver

Without Dynamic Power Saver

Example 1800W draw system
- 6 x PSU, 300W each, 75% efficiency
- 2400W in for 1800W out
- 600W wasted

With Dynamic Power Saver

Example 1800W draw system
- 2 x PSU, 900W each, 89% efficiency
- 2048W in for 1800W out
- Saving of 352W per enclosure

Dynamic Power Saver is enabled by the Onboard Administrator module. When this feature is enabled, the total enclosure power consumption is monitored in real-time and automatically adjusted with changes in demand.

NOTE
In redundant environments a minimum of two power supplies are always active and the maximum load that can be reached on any PSU is 50 percent. Once the 50 percent load is reached, another two power supplies are activated to ensure that redundancy is maintained at all times.

Power Regulator and power workload balancing

HP's ProLiant Power Regulator provides Integrated Lights-Out (iLO)-controlled speed stepping for Intel x86 processors. The Power Regulator feature improves server energy efficiency by giving CPUs full power for applications when they need it and reducing power when they do not. This power management feature enables ProLiant servers with policy-based power management to control CPU power state. Power Regulator can be configured for continuous, static low power mode or for Dynamic Power Savings mode in which power is automatically adjusted to match application CPU utilization demand.

Power workload balancing maximizes performance per watt used. Power workload balancing uses the HP Power Regulator technology to manage power at an enclosure level to ensure that power usage stays within defined power caps. Using power caps, system administrators can constrain the maximum BTUs per enclosure and rack to enable the enclosure to fit in an existing rack power envelope. A simple power cap will allow devices to power on until power usage reaches the specified power cap and then will prevent any more devices from powering on. It will be possible to manage power to future server blades so that the blade itself will not exceed a given power value, no matter what application load is applied. In this future environment, any power-managed device could

power on and the administrator would raise and lower caps on the individual server blades to ensure that the enclosure stays under the enclosure-level power cap while maintaining optimum performance for the user applications. Power workload balancing will be implemented in a future revision of the OA firmware but is not available at the date of this publication.

Additional information on the HP Power Regulator is provided in the paper titled "Power Regulator for ProLiant," available at www.hp.com/servers/power-regulator.

**HP BladeSystem Power Sizer**

The HP BladeSystem Power Sizer is a tool that assists facilities teams and IT staff in sizing their power and cooling infrastructure to meet the needs of an HP BladeSystem solution. The BladeSystem Power Sizer is based on actual component-level power measurements of a system stressed to maximum capability. The sizer allows a customer to select the type and number of components within each server blade and enclosure and to see the effect of changes on the power consumption and heat loading.

Values obtained from the BladeSystem Power Sizer tool are based on worst case loads and are intended for facility planning purposes only. Actual power consumption will vary with application type, application utilization and ambient temperature. The BladeSystem Power Sizer is available at the following URL: http://www.hp.com/go/bladesystem/powercalculator.

# I/O Infrastructure and interconnect options

A key component of the c7000 enclosure is the I/O infrastructure—essentially, a NonStop signal midplane that provides the internal wiring between the server or storage blades and the interconnect modules. The NonStop signal midplane is an entirely passive board that takes advantage of serializer/deserializer (SerDes) technology to support multiple I/O protocols. The term *passive* means there are no active electrical components on the board. On one side of the board are the sixteen connectors for the server/storage blades. Internal traces in the printed circuit board link them to the eight connectors on the other side of the board for the interconnect modules (Figure 12).

Figure 12. Illustration showing both the server blade connectors and the interconnect module connectors
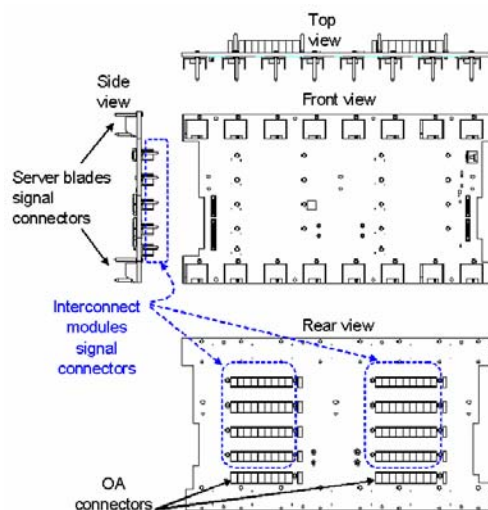
By taking advantage of the similar four-wire differential transmit and receive mechanism, the NonStop signal midplane can support either network-semantic protocols (for example, Ethernet, Fibre Channel, and InfiniBand) or memory-semantic protocols (PCI Express), using the same signal traces. Figure 13 illustrates how the physical lanes can be logically "overlaid" onto sets of four traces. Interfaces such as Gigabit Ethernet (1000-base-KX) or Fibre Channel need only a 1x lane, or a single set of four traces. Higher bandwidth interfaces, such as InfiniBand DDR, will need to use up to four lanes.

Figure 13. Traces on the NonStop signal midplane can transmit many different types of signals, depending on what I/O fabrics are used. The right-hand side of the diagram represents how the signals can be "overlaid" onto the same traces.
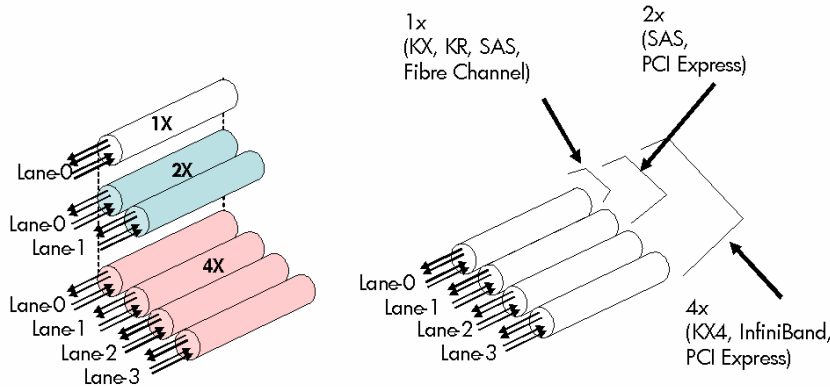
The NonStop signal midplane of the enclosure has eight 200-pin connectors to support eight individual switches or four double bay switches or a combination of the two. It provides the flexibility of 1x, 2x, or 4x connections from the server blade mezzanine cards which provide connectivity to the interconnect bays. The rear of the enclosure includes eight interconnect bays that can accommodate eight single or four redundant interconnect modules. All interconnect modules plug directly into these interconnect bays. Each HP BladeSystem c-Class enclosure requires two interconnect switches or two pass-thru modules, side-by-side, for a fully redundant configuration.

Switches and pass-thru modules

The first implementation of the BladeSystem c-Class includes two different pass-thru modules (HP 1-Gb Ethernet Pass-Thru Module and HP 4-Gb Fibre Channel Pass-Thru Module), two Ethernet switches (HP GbE2c Ethernet Blade Switch and Cisco Catalyst 3020 Blade Switch), and one Fibre Channel switch (Brocade 4-Gb SAN Switch).

Future implementations will include high-bandwidth fabrics that are gaining popularity in the market. The HP website (www.hp.com/go/bladesystem/interconnects) contains the most up-to-date information about the c-Class interconnect modules.

Switches offer customers a traditional approach to administering the network. The primary value in blade switches is cable consolidation via high-speed uplinks and the shared blade power and cooling infrastructure.

Ethernet and Fibre Channel pass-thru modules are available when direct one-to-one connections between servers and LAN or SAN are required. HP Ethernet and Fibre Channel pass-thru modules provide 16-port transparent, 1:1 port connectivity between the server and an external switch.

The 1-Gb Ethernet Pass-Thru Module for HP BladeSystem c-Class is designed for those desiring an unmanaged direct connection between each server blade within the enclosure and an external network device such as a switch, router or hub. The module is unmanaged. Basic status is monitored by the Onboard Administrator. The ports do not auto-negotiate speed. The speed on each port is fixed at 1-Gb.

The 4-Gb Fibre Channel Pass-Thru Module for HP BladeSystem c-Class is designed for those customers desiring a direct connection between each server blade within the enclosure and external Fibre Channel SAN switches, directors, or SAN controllers. The modular design easily installs into the I/O bays of the c-Class enclosure. Installing a pair of Fibre Channel pass-thru modules provides redundant paths for connection to the SAN. The Fibre Channel pass-tThru module provides a direct 1/2/4 Gb auto-negotiated Fibre Channel link from each server blade to the SAN for the HP BladeSystem c-Class and an iLO link for management.

Refer to the *HP BladeSystem Onboard Administrator User Guide*[1] for specific pass-thru port connections for each blade. Connections differ by blade type.

NOTE

HP discourages the use of the Ethernet Pass-Thru Module. Because the blades are connected via SerDes to the I/O bays, and SerDes Ethernet does not have an auto-negotiation protocol, a switch is required to connect to 10/100 networks outside of the enclosure. The NICs themselves are capable of different modes of operation, but the outbound wiring to which they are connected is not auto-negotiation friendly. Note that this is an Ethernet limitation. The Fibre Channel ports do auto-negotiate, even in the pass-thru configuration.

Server blades

Server blades for the BladeSystem c-Class are built according to standard form-factors referred to as half-height and full-height. The enclosure can hold either full-height or half-height blades or a combination of the two. It is preconfigured with device bay shelves to house the half-height server and/or storage blades. The shelf must be removed to accommodate the full-height server blade.

For I/O connectivity, every server ships with at least two Ethernet connections built into the server. To maintain flexibility, the server blades use optional mezzanine cards to provide additional I/O fabric connections such as Gigabit Ethernet, InfiniBand, and Fibre Channel.

Storage blades

Storage blades provide another option besides internal disk drives or SAN connectivity. The c-Class enclosure supports two types of blade storage solutions: direct-attach storage blades and shared storage blades. For mechanical compatibility, the storage blades use the same half-height form factor as the server blades.

A direct-attach storage blade holds up to six SAS or SATA drives and must be paired with an adjacent server blade in the same zone. This is because the physical connection between the direct-attach storage blade and its adjacent server blade is a specific PCIe lane across the NonStop midplane that connects the adjacent bays together. Half-height server blade and direct-attach storage blade must be in bays 1/2, 3/4, 5/6, 7/8, 9/10, 11/12, 13/14, or 15/16, in either order. Refer to Appendix B for bay number details. The pair cannot be in bays that are not adjacent within a zone

---

[1] http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00705292/c00705292.pdf

(between two metal walls). The direct-attach storage blade is equipped with a Smart Array controller to enable hardware-based RAID configurations. A mezzanine (mezz) card is not required to connect a half-height server blade to the direct-attach storage blade.

For a full-height server blade, a mezz card is required to connect to the direct-attach storage blade. It must be in the mezz 3 connector. This is to allow full use of the interconnect bays with mezz 1 or mezz 2 cards and to be consistent with the half-height blades, in case customers want to mix half-height and full-height server blades in the same enclosure. Because the direct-attach storage blade must be in the bottom bay when used with a full-height blade, a blank may be attached above the storage blade to block the empty upper bay, or a half-height server blade may be inserted into the upper bay. For this usage, the storage blade should be installed before the half-height server blade is installed, and the half-height server blade should be removed before the storage blade is removed.

Shared storage blades are not available as of this writing. They will be implemented as an option in the future.

Mezzanine cards

The mezzanine cards can fit onto any ProLiant or Integrity c-Class server blade. HP designed the server blades to use two types of mezzanine cards, referred to as Type-I and Type-II, to connect to the various interconnect fabrics. Type-I and Type-II mezzanine cards differ only in the amount of power required and their physical size. See Table 2 for interoperability of mezz cards with mezz connectors in half-height and full-height server blades.

Table 2. Interoperability of Type-I and Type-II cards with mezz connectors in half-height and full-height server blades

Server blade type	Mezzanine 1 connector	Mezzanine 2 connector	Mezzanine 3 connector (Full-height only)
Half-height server blade options	Type-I mezz card	Type-I mezz card	
	Type-I mezz card	Type-II mezz card	
Full-height server blade options	Type-I mezz card	Type-I mezz card	Type-I mezz card
	Type-I mezz card	Type-I mezz card	Type-II mezz card
	Type-I mezz card	Type-II mezz card	Type-I mezz card
	Type-I mezz card	Type-II mezz card	Type-II mezz card

Type-II mezzanine cards have slightly more power available to them and are slightly longer. Both Type-I and Type-II mezzanine cards use the same 200-pin connector part, but more pin definitions are used for some implementations of Type-II mezzanine cards. At this writing, all mezzanine cards have PCIe x4 interface, except the InfiniBand DDR single-port HCA, which has an x8 interface. At this writing, all mezz 1 connectors are Type-I connectors and all mezz 2 and mezz 3 connectors are Type-II connectors with a PCIe x8 interface. All mezzanine cards to date are Type-I cards and, therefore, work in all mezzanine connectors. Future Type-II mezzanine cards will only work in Type-II connectors.

NOTE
The InfiniBand 4x DDR single-port mezzanine card will NOT operate at its full data rate of 20Gbps in the initial mezz 1 connectors with PCIe x4. This is because the host interface on the mezz 1 connectors is PCIe x4 at 10Gbps bandwidth. However, since the host interface for the mezz 2 or mezz 3

connectors is PCIe x8 at 20Gbps bandwidth, this InfiniBand mezzanine card will operate at its full data rate of 20Gbps in mezz 2 or mezz 3 connectors or in future PCIe x8 mezz 1 connectors.

Virtual Connect

With the c-Class architecture HP introduced a new type of interconnect technology: Virtual Connect. As it is implemented in the c-Class architecture, Virtual Connect technology provides virtualized server I/O connections to the Ethernet (LAN) or Fibre Channel (SAN) networks. Virtual Connect technology virtualizes the server-edge so that networks can communicate with pools of HP BladeSystem servers as opposed to the conventional one-to-one relationship.

Virtual Connect consists of hardware (the Virtual Connect module) and a management agent that runs on the Virtual Connect module. Like Ethernet and Fibre Channel switches, the Virtual Connect modules slide into the switch bays of the c7000 enclosure. The Ethernet module is necessary in order to install Fibre Channel because the Virtual Connect Manager software runs on a processor on the Ethernet module. The Ethernet module has sixteen 1-GbE downlinks to servers (connected across the NonStop signal midplane), eight 1-GbE uplinks to the network (RJ45 copper Ethernet connectors), two 10-GbE connectors (for copper CX4 cables), one 10-GbE internal inter-switch link (across the NonStop signal midplane) for a failover connection between Virtual Connect modules. The Fibre Channel module has sixteen 4-Gb Fibre Channel downlinks to servers and four 1/2/4-Gb auto-sensing Fibre Channel uplinks to the network.

Full details about Virtual Connect technology are available in the technology brief entitled "HP Virtual Connect technology implementation for the HP BladeSystem c-Class" on the HP technology website at www.hp.com/servers/technology.

Fabric connectivity and port mapping

Because the connections between the blade bays and the interconnect module bays are hard-wired through the NonStop signal midplane, the server mezzanine cards must be matched to the appropriate type of interconnect module. For example, a Fibre Channel mezzanine card must be placed in the mezzanine connector that connects to an interconnect bay holding a Fibre Channel switch. To simplify the installation of the various mezzanine cards and interconnect modules, the Onboard Administrator uses an "electronic keying" process to detect any mismatch between the mezzanine cards and the interconnect modules.

Interconnect bays 1 and 2 are reserved for Ethernet switches or pass-thru modules supporting server LAN on Motherboard (LOM) NIC connections to ports on the Ethernet switch or pass-thru module. Supported bays for additional Ethernet switch modules include unpopulated bays 3/4, 5/6, or 7/8. Redundant switches must be configured adjacent to one another in bays 3/4, 5/6, or 7/8.

The HP BladeSystem c7000 enclosure allows for easily connecting the ports of embedded devices to the interconnect bays. For port mapping purposes, it does not matter in which bay a server blade is installed. The mezzanine connectors always connect to the same interconnect bays.

Several port types are referenced in Tables 3 and 4:

- Examples of 1x ports are 1-Gb Ethernet (1-GbE) pass-thru modules and FC interconnect modules.
- An example of a 2x port is a SAS interconnect module.
- Examples of 4x ports are 10-GbE pass-thru modules and InfiniBand interconnect modules.
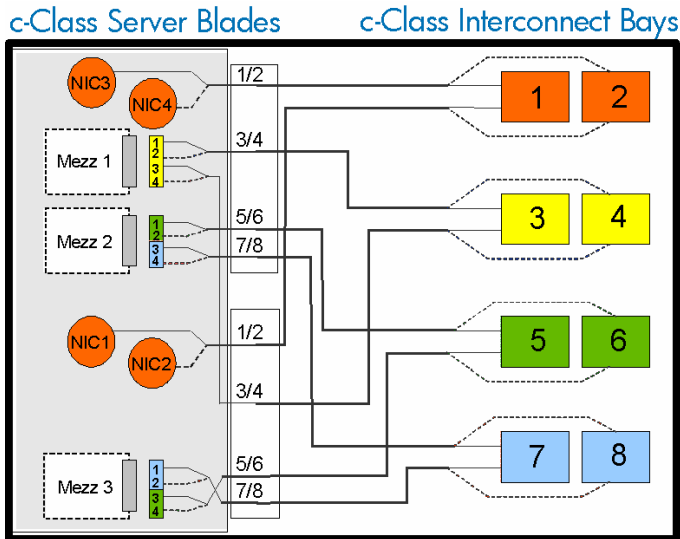
Figure 14. HP BladeSystem c7000 port mapping half-height blades to interconnect bays
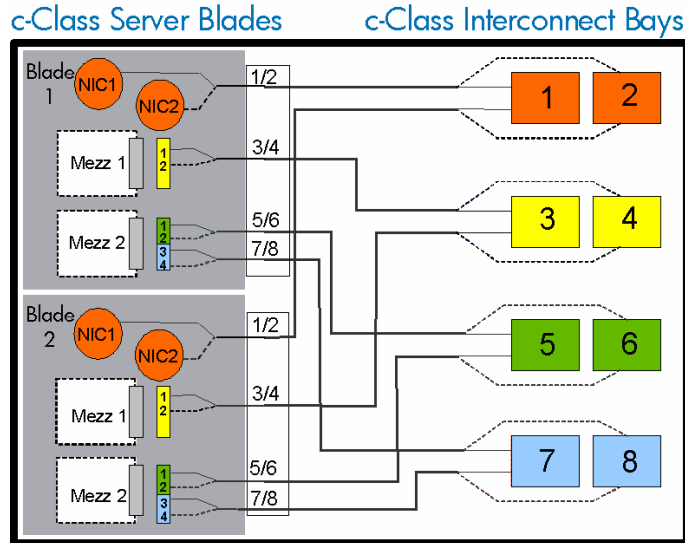
Table 3. Port mapping for half-height server blades

Server blade port	Port number	Connects to interconnect bay	Comment
NIC	NIC 1	1	For 2 single-wide Ethernet interconnect modules
	NIC 2	2	
Mezz 1 connector, Option 1	1x/2x port 1	3	For 2 single-wide interconnect modules
	1x/2x port 2	4	
Mezz 1 connector, Option 2	4x port 1	3 + 4 (together)	For 1 double-wide interconnect module
Mezz 2 connector, Option 1	1x/2x port 1	5	For 4 single-wide interconnect modules
	1x/2x port 2	6	
	1x/2x port 3	7	
	1x/2x port 4	8	
Mezz 2 connector, Option 2	4x port 1	5 + 6 (together)	For 2 double-wide interconnect modules
	4x port 2	7 + 8 (together)	

Figure 15. HP BladeSystem c7000 port mapping for full-height blades to interconnect bays

c-Class Server Blades   c-Class Interconnect Bays

Table 4. Port mapping full-height server blades

Server blade port	Port number	Connects to interconnect bay	Comment
NIC	NICs 1 and 3	1	For 2 single-wide Ethernet interconnect modules
	NICs 2 and 4	2	
Mezz 1 connector, Option 1	1x/2x port 1, 3	3	For 2 single-wide interconnect modules for 4-port connectivity
	1x/2x port 2, 4	4	
Mezz 1 connector, Option 2	4x port 1	3 + 4	For 1 double-wide interconnect module for 1-port connectivity. Use mezz connector 2 or 3 for 2-port 4x connectivity.
Mezz 2 connector, Option 1	1x/2x port 1	5	For 4 single-wide interconnect modules
	1x/2x port 2	6	
	1x/2x port 3	7	
	1x/2x port 4	8	
Mezz 2 connector, Option 2	4x port 1	5 + 6	For 2 double-wide interconnect modules
	4x port 2	7 + 8	
Mezz 3 connector, Option 1	1x/2x port 1	7	For 4 single-wide interconnect modules
	1x/2x port 2	8	
	1x/2x port 3	5	
	1x/2x port 4	6	
Mezz 3 connector, Option 2	4x port 1	7 + 8	For 2 double-wide interconnect modules
	4x port 2	5 + 6	

Port mapping differs slightly between full-height and half-height server blades due to the support for additional mezzanine cards on the full-height version. HP has simplified the process of mapping mezzanine ports to switch ports by providing intelligent management tools via the Onboard Administrator and HP Insight Manager software.

Onboard Administrator

Onboard Administrator is a management controller module that resides within the BladeSystem c-Class enclosure. The Onboard Administrator works with the iLO 2 management processors on each server blade to form the core of the management architecture for BladeSystem c-Class. HP customers have the option of installing a second OA board in the BladeSystem c7000 enclosure to act as a completely redundant controller in an active-standby mode. If the standby OA board determines that the active Onboard Administrator is not functioning properly, it shuts down the active Onboard Administrator and assumes the active role itself. This redundancy can be particularly useful in remote site situations.

There are three ways for IT technicians and administrators to access the Onboard Administrator:

1. The Insight Display is a rack-mounted interactive display device described in more detail in a section below.
2. The web GUI uses event-driven, push technology. No screen refresh is necessary to view failures and/or events. If an event occurs, it is pushed to the web GUI and updated immediately. The GUI provides seamless integration with HP OpenView, Insight Control Data Center Edition and Insight Control Linux® Edition.
3. Finally, administrators who commonly use scripting commands can access the Onboard Administrator through a command-line interface (CLI).

Onboard Administrator collects system parameters related to thermal and power status, system configuration, and managed network configuration. It manages these variables cohesively and intelligently so that IT personnel can configure the BladeSystem c-Class and manage it in a fraction of the time that other solutions require.

One of the major advantages of the BladeSystem c-Class is its flexibility for allowing customers to configure the system in virtually any way they desire. However, this flexibility brings with it some complexity to ensure that the BladeSystem c-Class is configured correctly. The configuration logic powers up the interconnect modules first. Server blades are not powered up until the Onboard Administrator has verified that the configuration is correct. If there is a configuration issue, the Insight Display indicates what the issue is and possible remedies. To assist in the configuration and setup process for the IT administrator, the Onboard Administrator verifies four attributes for each server blade and interconnect module as they are added to the enclosure:

- Electronic keying—The OA automatically queries all mezzanine cards and interconnect modules as they are deployed to check that the I/O fabric types match. If they do not, the OA issues a warning with suggested corrective action.
- Power—The OA ensures that there is sufficient power available to power up a blade or interconnect module.
- Cooling—The OA makes sure there is sufficient cooling capacity for the blade or interconnect module.
- Location—The OA makes sure that blades, Active Cool fans, and power supplies are in the correct locations to receive proper cooling and support the chosen power configuration. For example, if the administrator is installing only two server blades, they must be in server bays 1, 2, 9, or 10. Similarly, fans must go into fan bays 4, 5, 9, and 10.

If a configuration problem exists, the Onboard Administrator gives diagnostic help so that the IT technician can easily fix the problem by looking at the Insight Display.

Onboard Administrator also includes logic to help manage multiple enclosures in a rack. This helps the IT administrator understand information such as power conditions and blade configurations in multiple enclosures. The Onboard Administrator allows single point access for up to four enclosures. Thus, the administrator can use a single sign-on to log into a single Onboard Administrator and use the web GUI to graphically view and manage all the c-Class components within the linked enclosures. For example, an IT administrator could automatically propagate management commands—such as putting an upper limit on power levels for all server blades—throughout the linked enclosures.

Onboard Administrator retrieves thermal information from the components in the enclosure. If the enclosure's thermal load increases, the Onboard Administrator's thermal logic feature instructs the fan controllers to increase fan speeds to accommodate the additional demand. The speed of the individual fans can be adjusted to reduce noise and power consumption, and to compensate for airflow differences within the enclosure. The performance of each subsystem is proactively monitored, and any failures or warnings can be reported to the system log and to broader infrastructure management tools like HP Systems Insight Manager (when SNMP is enabled). The Onboard Administrator manages subsystem failure by taking appropriate action to maintain the enclosure's ability to operate, including adjusting fan speed or reducing power consumption.

The Onboard Administrator uses sophisticated power measurement sensors to accurately monitor exactly how much power is being consumed and how much power is available. Because Onboard Administrator uses real-time measured power data instead of maximum power envelopes, customers can deploy as many servers and interconnect modules as possible for the available power.

Configuring the HP BladeSystem c-Class to fit into a network is now a much simpler and more automated process. The Onboard Administrator significantly enhances the network infrastructure management by offering two methods to initially configure the IP addresses of the server blade iLO 2 ports and the interconnect module management ports via the management network: DHCP or Enclosure Bay Static IP Addressing. This configuration capability is managed through a single point, the Onboard Administrator, as opposed to managing each iLO or interconnect module individually.

More information about the Onboard Administrator is available in the technology brief entitled "Managing the HP BladeSystem c-Class" at this URL: www.hp.com/servers/technology.

Insight Display

The Insight Display, Figure 16, is an ever-ready, rack-mounted information exchange device with access to all of the Onboard Administrator's setup, management, and troubleshooting features. It provides a quick and easy-to-use facility that enables the rack technician to initially configure the enclosure. It also provides information about the health and operation of the enclosure. The Insight Display is effective from a mechanical standpoint because it is big enough to allow the technician to see ample amounts of information, and it can slide back and forth to allow access to the power supplies.

Figure 16. Insight Display

When the enclosure is powered up for the first time, the Insight Display launches an installation wizard to guide the user through the configuration process. To identify the enclosure, the enclosure unit identification (UID) LED and the background of the Insight Display are illuminated blue when the enclosure is initially powered on. The Installation Wizard automatically turns on the enclosure UID at the beginning of the installation and turns it off when the installation is complete. After configuring the enclosure, the Insight Display verifies that there are no installation or configuration errors.

More information about the Insight Display is available in the technology brief entitled "Managing the HP BladeSystem c-Class" at this URL: www.hp.com/servers/technology.

Recommendations

HP recommends the single-phase enclosure models for most situations. Most typical c-Class customer configurations that HP builds at the time of this writing (for example, 16 Blades with 2 CPUs each, 4 to 8 GB of RAM, Ethernet, Fibre Channel) do not exceed 6 kVA for a worst case load. Under normal operating conditions, with all components at 100 percent load, approximately 4 to 4.5 kVA is used. Most customer systems will probably consume less than 4 kVA, as all components will not be at 100 percent load simultaneously. Actual measurements of real customer applications show numbers between 3 and 3.5 kVA. If the three-phase enclosure were used to build a rack with four enclosures, 8 x 30A three-phase power whips would be required and would supply roughly 70 kVA to the rack, of which no more than 18 to 24 kVA would be used.

The three-phase c-Class enclosure is generally useful in only a few situations:

• When there is only one enclosure and will never be more than one enclosure
• When the configuration is an extreme one (for example, 16 blades with 2 high-power CPUs, 32 GB of RAM, all mezz connectors and switch bays filled, running high-performance technical computing loads)
• When the enclosure is a drop-in replacement for a p-Class solution where the three-phase power already exists
• When the customer plans to maintain the enclosure for a long period of time and to update the internal components with newer ones that could take advantage of full three-phase power capacity

23

A more efficient use of power can be accomplished using three-phase power for delivery to the rack and single-phase power in the rack. This can be achieved by using the S332 or S348 PDU to run 2 x 60A three-phase power to the rack, connected to three or four single-phase enclosures in the rack. If power capping technology is used in this scenario, the enclosure power consumption will never exceed 4.3 kVA and thus will not exceed the capacity of the PDU, even if an AC line were to fail.

NOTE

NA/JPN three-phase enclosures use a different type of three-phase power than International enclosures.

HP also recommends the following for configuring BladeSystem c-Class enclosures:

- Use Virtual Connect or managed switches to reduce cabling and management overhead.
- Use the HP BladeSystem Power Sizer.
- Use at least eight fans; use ten fans for optimum power and cooling.
- For all server blades other than BL685c, the InfiniBand 4x DDR single-port mezzanine card will work in mezz 1, but it will work better in mezz 2 or mezz 3. For BL685c, the InfiniBand 4x DDR single-port mezzanine card will work equally well in mezz 1, mezz 2 or mezz 3 connectors.

Summary

The HP BladeSystem c7000 enclosure is the foundation of a new modular computing architecture that consolidates and simplifies infrastructure, reduces operational cost and delivers IT services more effectively. Designed from the start to deliver a bold agenda for the future of blade computing, it integrates HP innovations to tackle the toughest problems facing today's IT infrastructures—cost, time, energy and change. Thermal Logic technologies provide the mechanical design features, built-in intelligence, and control capabilities throughout the BladeSystem c-Class that enable IT administrators to make the most of the power and thermal environments. The shared, high-speed NonStop midplane and pooled-power backplane in the enclosure accommodate new bandwidths and new technologies. The Onboard Administrator supplies an intelligent infrastructure to provide essential power and cooling information and help automate the management of the infrastructure. The BladeSystem c7000 enclosure provides all the power, cooling, and I/O infrastructure to support c-Class modular servers, interconnects, and storage components, today and throughout the next several years.

Appendix A. Acronyms in text

The following acronyms are used in the text of this document.

Table A-1. Acronyms

Acronym	Acronym expansion
AC	Alternating current
CFM	Cubic feet per minute
CLI	Command line interface
DC	Direct current
DDR	Double data rate
FC	Fibre channel
HBA	Host bus adapter
HCA	Host channel adapter
IB	InfiniBand
iLO	Integrated Lights Out
iSCSI	Internet Small Computer System Interface
kVA	Kilovolt-ampere
LAN	Local area network
LCD	Liquid crystal display
LOM	LAN on Motherboard
NIC	Network interface card
OA	Onboard Administrator
PARSEC	Parallel, redundant, scalable enclosure cooling
PCIe	PCI Express
PDU	Power distribution unit
PSU	Power supply unit
SAN	Storage area network
SAS	Serial attached SCSI
TOE	TCP/IP Offload Engine
U	Unit of measurement for rack-mount equipment (U is 1.75 inches or 4.44cm)
UID	Unit identification
VA	Volt-ampere
W	Watt

Appendix B. Fan and server population guidelines

Figure B-1. HP BladeSystem c7000 enclosure – Fan population guidelines. For correct operation, fans and server blades must be installed in the correct bays. The Onboard Administrator will ensure that fans and blades are correctly placed before allowing systems to power on.
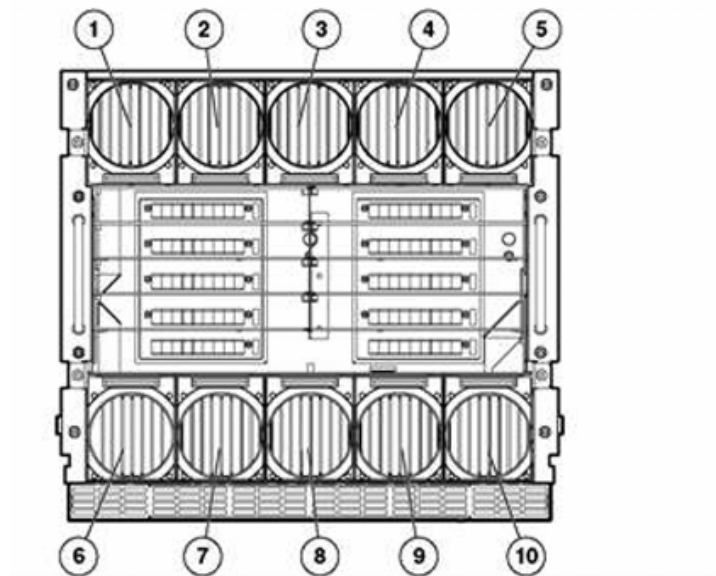

Number of fans	Fan bays populated (Zone)	Device bays supported
4	4, 5 (Zone 1) and 9, 10 (Zone 3)	Two devices located in bays 1, 2, 9 or 10
6	3, 4, 5 (Zone 1) and 8, 9, 10 (Zone 3)	1-4, 9-12
8	1, 2, 4, 5, 6, 7, 9, 10 (all four zones)	1-4, 9-12, 5-8, 13-16
10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10 (all four zones)	1-4, 9-12, 5-8, 13-16

Figure B-2. HP BladeSystem c7000 enclosure – Half-height server blade bay numbering. Half--height servers should be populated from top and bottom from left to right from the front of the enclosure. So the first two half-height servers would be placed in bays 1 and 9, the second two half-height servers would be placed in bays 2 and 10 and so on until the enclosure is full. Full-height servers should be populated from left to right.
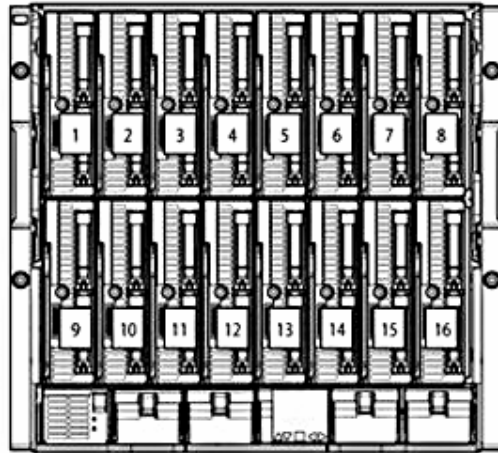
Figure B-3. Mixed configuration – Full-height and half-height population rules. The c7000 enclosure is divided into four partitions, or quadrants, by the vertical support metalwork. Within each partition a removable divider is used to support half height devices. To install a full-height blade in any partition,this divider must be removed. As a consequence, a partition can only contain either full-height server blades or half-height server blades. The half-height SB40c storage blade must be installed within the same partition as the partner server blade. If the SB40c is partnered with a full-height server blade, it must be installed in the lower bay adjacent to the full-height server and the server requires the HP SB40c PCI-express pass-thru mezz card in the mezz 3 connector. There is a half-height bracket that must be installed on top of the SB40c used with a full-height server blade in order to plug a half-height server or blade blank above the SB40c.
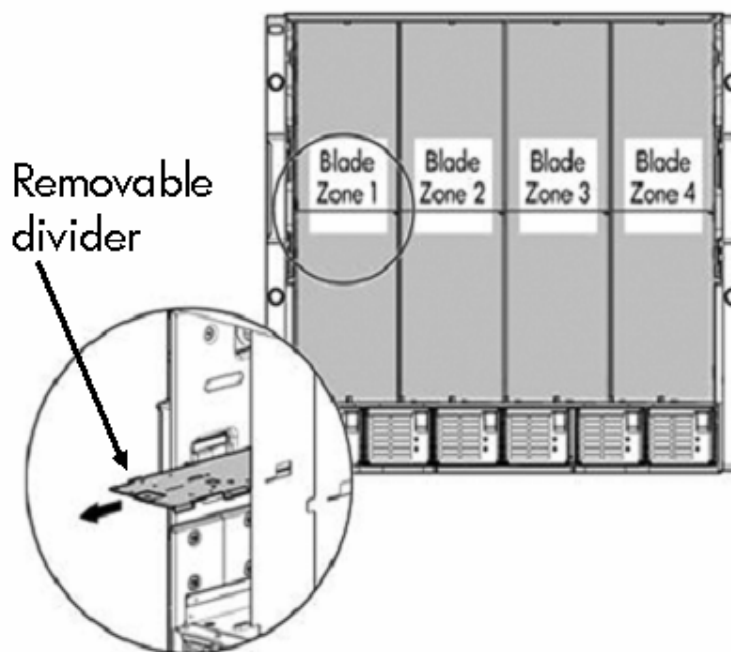
For more information

For additional information, refer to the resources listed below.

Resource description	Web address
General HP BladeSystem information	http://www.hp.com/go/bladesystem/
HP BladeSystem c-Class documentation	http://h71028.www7.hp.com/enterprise/cache/316735-0-0-0-121.html
HP BladeSystem c-Class Enclosure Setup and Installation Guide	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00698286/c00698286.pdf
HP BladeSystem c-Class interconnects	www.hp.com/go/bladesystem/interconnects
Technology briefs about HP BladeSystem	http://h18004.www1.hp.com/products/servers/technology/whitepapers/proliant-servers.html
HP BladeSystem Power Sizer	http://www.hp.com/go/bladesystem/powercalculator

Call to action

Send comments about this paper to TechCom@HP.com.

hp invent