HP BladeSystem c-Class architecture

Technology brief, 4th edition

Introduction

The HP BladeSystem c-Class architecture is a flexible infrastructure that makes the computing, network, and storage resources easy to install and arrange. It creates a general-purpose infrastructure that accommodates your changing business needs. This technology brief explains how the architecture supports such flexibility.

Shared cooling, power, and management resources support the flexible, modular components. This integrated aspect of the BladeSystem c-Class—especially the Onboard Administrator and Integrated Lights-Out (iLO) management tools—are key components of the BladeSystem c-Class environment. These management resources comprise a pre-boot configuration environment and support innovative technologies like Virtual Connect.

This paper includes a short description of the components within the BladeSystem c-Class and explains how they work together. But it does not discuss details about all BladeSystem products. For product details, refer to the HP BladeSystem website at www.hp.com/go/bladesystem or refer to other technology briefs at http://h18004.www1.hp.com/products/servers/technology/whitepapers/proliant-servers.html#bl.

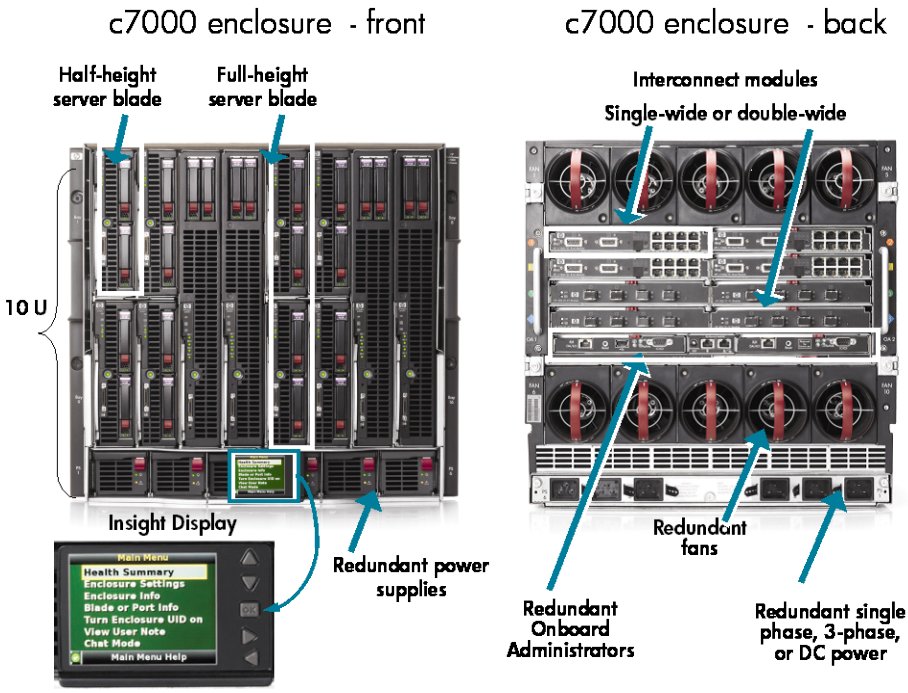
Enclosures and hardware components

The HP BladeSystem c-Class enclosure holds these components:

- ProLiant server blades
- ProLiant workstation blades
- Integrity server blades
- Storage blades
- I/O option blades
- Interconnect modules (switches, pass-thru modules, and Virtual Connect modules)
- NonStop passive signal midplane
- Passive power backplane
- Power supplies and fans
- Onboard Administrator modules

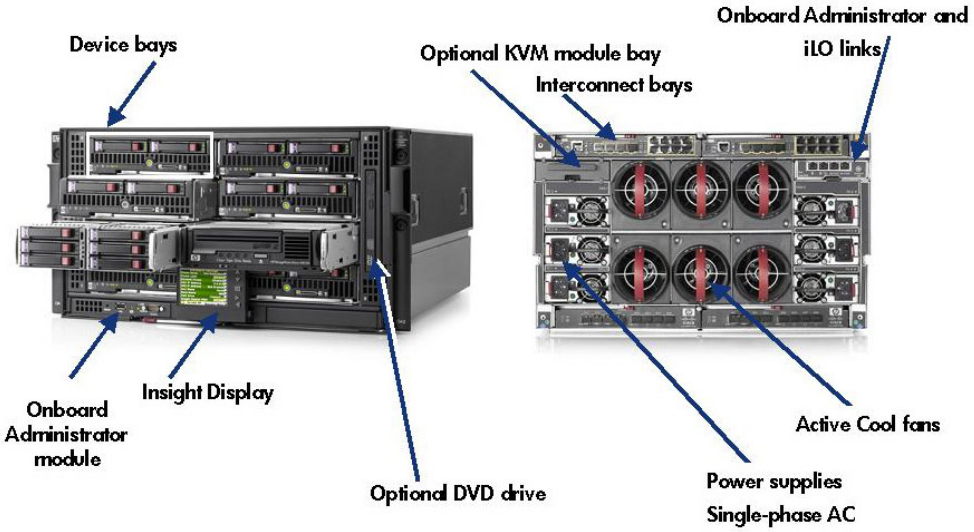
We have optimized the HP BladeSystem c7000 enclosure (Figure 1) for enterprise data centers. A single BladeSystem c7000 enclosure is 10U high. It holds up to 16 server, storage, or I/O option blades and up to 8 interconnect modules.

Figure 1: Front and back views of the BladeSystem c7000 Enclosure (single phase version)



Because of its small footprint and lower power consumption, the BladeSystem c3000 enclosure works well in smaller data centers, remote sites, or locations with power and cooling constraints. A single BladeSystem c3000 enclosure (Figure 2) is 6U high. It holds up to eight server, workstation, storage, or I/O option blades and up to four interconnect modules.

Figure 2: Front (left) and back (right) of the BladeSystem c3000 enclosure



Our BladeSystem enclosures accommodate half-height blades, full-height blades, or both. Server blades can use single-, double-, or quad-wide form factors. LAN on Motherboard (LOM) adapters and optional mezzanine cards on the server blades route network signals to the interconnect modules in the rear of the enclosure. The connections between server blades and a network fabric can be redundant.

The c7000 enclosure has eight interconnect bays that accommodate up to eight single-wide interconnect modules. They can be Virtual Connect modules, Ethernet or Fibre Channel switches, or a combination of single-wide and double-wide interconnect modules such as InfiniBand switches.

The c3000 enclosure has four interconnect bays. The bays can hold four single-wide or one double-wide and two single-wide interconnect modules.

The c-Class enclosure also holds one or two Onboard Administrator management modules. A second Onboard Administrator module acts as a redundant controller in an active-standby mode. The Insight Display panel on the front of the enclosure provides an easy way to access the Onboard Administrator locally.

The c-Class enclosures use flexible power architecture. The c7000 enclosure uses single-phase or three-phase AC or DC power inputs. The c3000 enclosure, by contrast, only uses single-phase (auto-sensing high-line or low-line) power inputs. You can connect its power supplies to low-line (100VAC to 120VAC) wall outlets. In either enclosure, you can configure power redundantly. Power supplies connect to a passive power backplane that distributes shared power to all components.

High-performance, high-efficiency Active Cool fans provide redundant cooling across the enclosure and ample cooling capacity for future needs. The fans are hot-pluggable and redundant.

General-purpose flexible design

The BladeSystem c-Class enclosure supports many blades and interconnect device options:

- ProLiant server blades using AMD or Intel x86 processors
- ProLiant workstation blades
- Integrity server blades
- StorageWorks storage blades
- Tape blades
- PCI-X or PCI Express (PCIe) expansion blades

BladeSystem interconnect modules support a variety of networking standards:

- Ethernet
- Fibre Channel
- Fibre Channel over Ethernet (FCoE)
- InfiniBand
- iSCSI
- Serial Attached SCSI (SAS)

The architecture of the c-Class enclosure provides a basis for broad solutions, including the HP CloudSystem Matrix (www.hp.com/go/matrix). Matrix supports shared IT infrastructure services by integrating pools of computing, storage, and networking capabilities with management tools.

The c-Class enclosure interoperates and connects with other HP infrastructure pieces, including external storage components such as DAS (direct-attached storage), NAS (network attached storage), and SAN (storage area network) solutions (www.hp.com/go/blades/storage).

The design supports flexibility:

- Blade form factors that can scale vertically or horizontally—half-height or full-height blades and single-, double-, or quad-wide blades
- Interconnect module form factors that can scale—single-wide or double-wide modules
- Uplinks to connect up to seven enclosures
- Signal midplane that allows flexible use of I/O signals, supporting multiple fabrics using the same traces

Scalable blade form factors

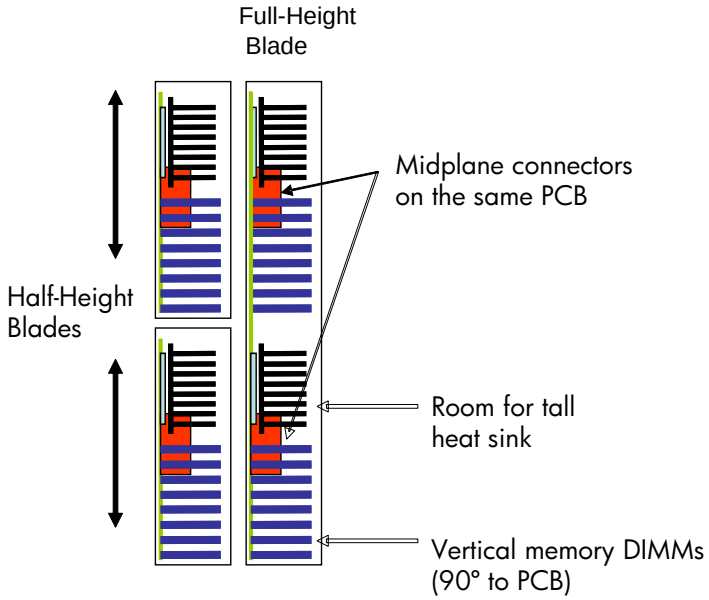
The half-height and full-height blade form factors, which scale blades vertically in the c7000 enclosure and horizontally in the c3000 enclosure, provide several benefits. They include reduced cost, increased reliability, and improved ease-of-use. Placing full-height form factors, half-height form factors, or both, in the same enclosure lets you exploit its space more efficiently. (Figure 3) For example, you can fill the enclosure with high-performance full-height server blades, or you can use a mixture of the two form factors.

The size of the full-height blades allows enough room to include two signal connectors on the same printed circuit board (PCB) plane for reliable and simple connectivity to the NonStop signal midplane. A removable, tool-less divider holds the half-height blades in the enclosure. See "Technologies in the HP BladeSystem c7000 Enclosure" or "HP BladeSystem c3000 Enclosure" at http://h18004.www1.hp.com/products/servers/technology/whitepapers/proliant-servers.html#bll for configurations.

The thickness of a device bay t provides several advantages over narrower or wider alternatives:

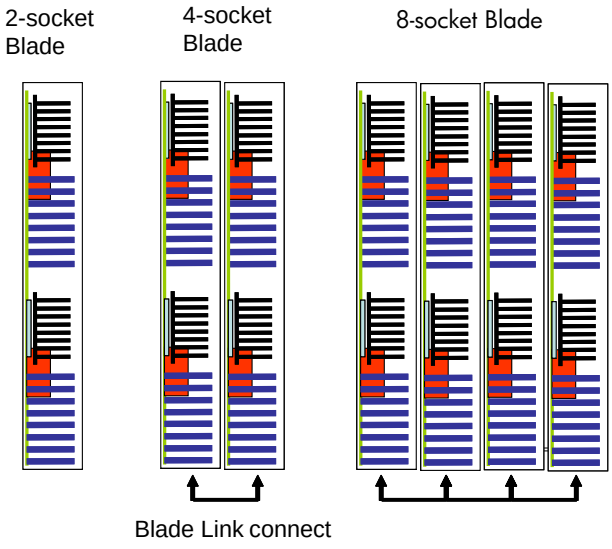
- It holds industry-standard components
- It holds enough blades to amortize the cost of the shared enclosure infrastructure (power supplies and fans)
- It uses cost-effective, standard-height DIMMs in the server blades
- It uses vertical, rather than angled, DIMM connectors to give better signal integrity and more room for heat sinks. The vertical DIMM connectors also allow more DIMM slots per processor and provide better airflow across the DIMMs.

Figure 3: BladeSystem c-Class form factors, scaling vertically with half-height and full-height blades



You can use two full-height bays in combination for a full-height double-wide blade form factor. For example, HP Integrity server blades can combine multiple blades to create 2-, 4-, or 8-socket systems, as shown in Figure 4. Each base server blade has a Blade Link connector. The Blade Link connector joins selected QPI (Quick Path Interconnect) ports among processors, the required clock signals, and side band signals for the system to operate as a scale-up multiprocessor system. Blade Link is available only with Integrity blades.

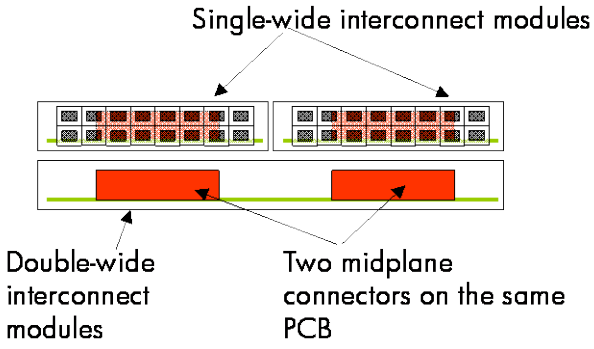
Figure 4: Horizontal scaling with Integrity blades and Blade Link connections



Scalable interconnect form factors

The interconnect bays also scale in a single-wide and double-wide form factor for efficient use of space and improved performance. A single interconnect bay can accommodate two single-wide interconnect modules in a scale-out configuration or a larger, higher-bandwidth double-wide interconnect module (Figure 5).

Figure 5: Single-wide and double-wide interconnect form factors



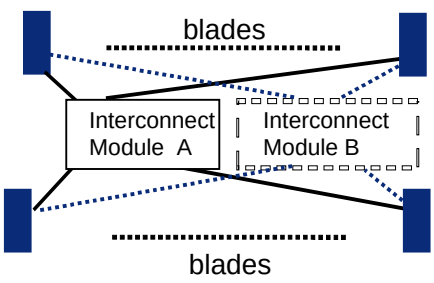
The scalable interconnect form factor provides similar advantages as the scalable device bays:

- It supports the maximum number of interconnect modules.
- It allows enough space in a double-wide module to include two signal connectors on the same PCB plane. This affords reliable and simple connectivity to the NonStop signal midplane.

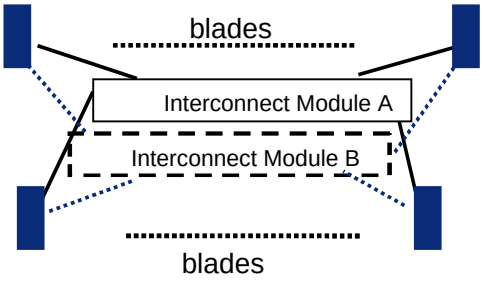
Star topology

The device bays and interconnect bays connect in a fan-out, or star, topology centered around the interconnect modules. The exact topology depends on your configuration and the enclosure. For example, if you place two single-wide interconnect modules side-by-side (Figure 6a), the architecture is a dual-star topology. Each blade has redundant connections to the two interconnect modules. If you use a double-wide interconnect module, it is a single star topology, providing more bandwidth to each server blade. Figure 6b shows the redundant configuration using double-wide interconnect modules.

Figure 6: The blade-to-interconnect topology differs depending on your configuration.



6a



6b

NonStop signal midplane provides flexibility

The BladeSystem c-Class uses a high-speed, NonStop signal midplane that supports multiple high-speed fabrics. It is unique because it can use the same technology for the physical traces to transmit Ethernet, Fibre Channel, InfiniBand, or SAS signals. As a result, you can fill the interconnect bays with a variety of interconnect modules, depending on your application needs.

Physical layer similarities among I/O fabrics

The basis of serialized I/O protocols such as Ethernet, Fibre Channel, SAS, and InfiniBand is a physical layer that uses multiples of four traces with a SerDes (serializer/deserializer) interface. The backplane Ethernet standards of 1000 Base KX, 10G-Base-KR, and the 8Gb Fibre Channel standard use a similar four-trace SerDes interface. Consolidating and sharing the traces between different protocols creates an efficient midplane design.

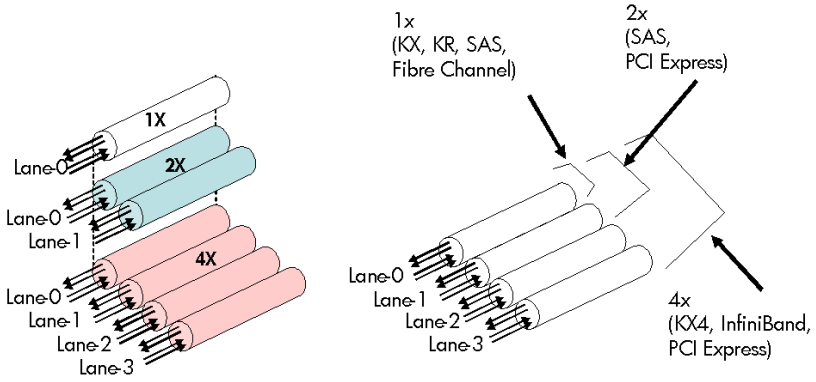
Figure 7 illustrates the logical overlay of the physical lanes onto sets of four traces. Interfaces such as GbE or Fibre Channel need only a 1x lane (a single set of four traces). Higher bandwidth interfaces, such as InfiniBand, need up to four lanes. Your choice of network fabrics dictates whether the interconnect module form factor needs to be single-wide (for a 1x/2x connection) or double-wide (for a 4x connection).

By logically overlaying the traces, we avoid:

- Replicating traces on the NonStop signal midplane to support each type of fabric
- Large numbers of signal pins for the interconnect module connectors.

This simplifies the interconnect module connectors and uses midplane real estate efficiently.

Figure 7: Lanes (sets of four traces) in the signal midplane support 1-, 2- , or 4-lane standard protocols.



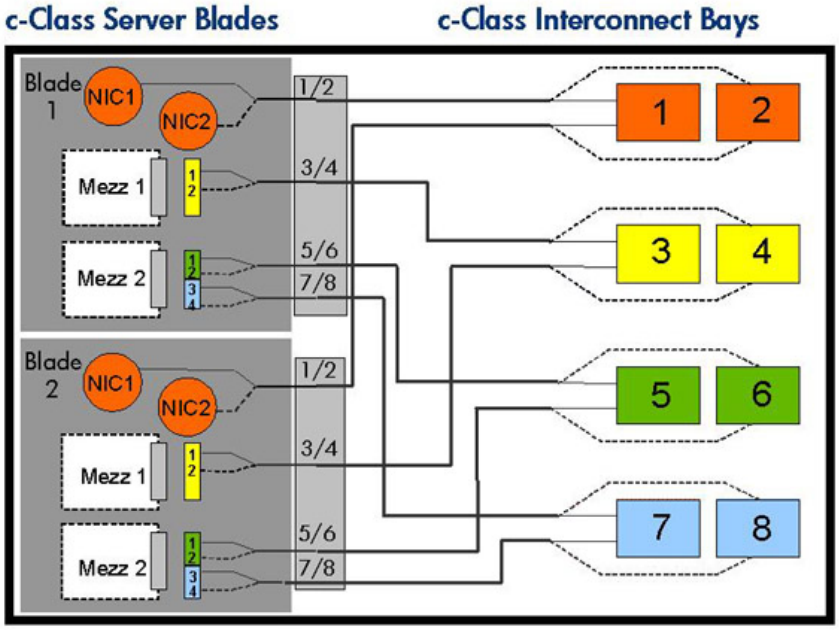
Connectivity between blades and interconnect modules

The c-Class server blades use mezzanine cards to connect to various network fabrics. The mezzanine cards on the server blades connect to the interconnect modules through independent traces on the NonStop signal midplane.

Connections differ depending on the enclosure.

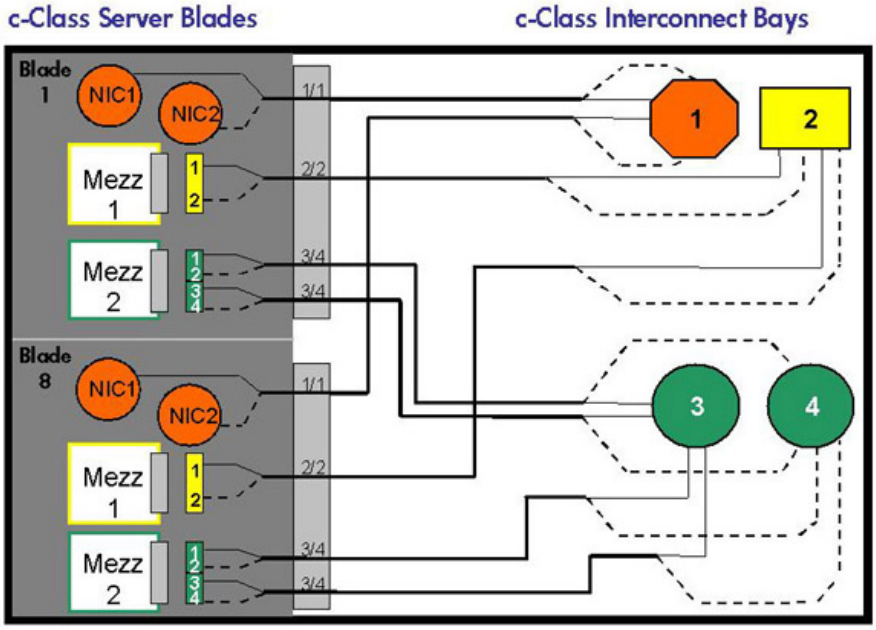
We designed the c7000 enclosure to include fully redundant connections between the server blades and interconnect modules. Figure 8 shows how half-height server blades connect to the interconnect bays in the c7000 enclosure.

Figure 8: Redundant connection of half-height server blades to the interconnect bays in the c7000 enclosure



With the c3000, you can use a single Ethernet switch or redundant Ethernet switches in interconnect bays 1 and 2. Figure 9 shows how half-height server blades connect to the interconnect bays in the c3000 enclosure.

Figure 9: Connection of half-height server blades to the interconnect bays in the BladeSystem c3000 enclosure



For more information about connections, review the enclosure quick start guides at http://h18004.www1.hp.com/products/blades/components/c-class-tech-installing.html.

To provide these connections, c-Class architecture must supply a mechanism to match the mezzanine cards on the server blades to the interconnect modules. The electronic keying mechanism in the Onboard Administrator helps your system administrators correct potential fabric mismatch conditions when they configure each enclosure. Before any server blade or interconnect module can power up, the Onboard Administrator, assisted by an iLO processor on each server blade, queries the mezzanine cards and interconnect modules to see if they are compatible. If the Onboard Administrator detects a configuration problem, it gives a warning and provides information to correct the problem.

Blade-to-blade connectivity

The NonStop signal midplane allows you to use more modular components than previous generations of blade systems. We can develop components in the blade form factor and connect them across the NonStop signal midplane—front-to-back or side-to-side. The architecture supports front-to-back modularity by connecting mezzanine cards in the server blades at the front of the enclosure to the matching interconnect modules in the rear of the enclosure.

There are four dedicated lanes between each side-by-side pair (odd/even) of device bays. This lets you connect optional storage or I/O devices to a server blade through the midplane. The Onboard Administrator disables these side-to-side links when they are unusable, such as when two server blades reside in adjacent device bays.

Some examples of storage and I/O device options using the dedicated links to a server blade include:

- StorageWorks SB40c Storage Blade that consist of a RAID controller and additional drives
- StorageWorks Ultrium Tape Blades that hold LTO-2, LTO-3, or LTO-4 Ultrium tape cartridges
- BladeSystem PCI Expansion Blade that holds two off-the-shelf PCI-X or PCIe cards

High bandwidth and performance

BladeSystem c-Class enclosure capabilities allow you to change your configuration as the demand for power and bandwidth grows:

- Blade form factors enable server-class components (discussed in the "General-purpose flexible design" section)
- High-bandwidth NonStop signal midplane
- Separate power backplane

NonStop signal midplane scalability

The NonStop signal midplane supports signal rates of up to 10 gigabits per second (Gb/s) per lane (each lane consists of four SerDes transmit/receive traces). Each half-height server bay has 16 lanes delivering the cross-sectional bandwidth to conduct up to 160 Gb/s per direction.

In a BladeSystem c7000 enclosure, the aggregate bandwidth between device bays and interconnect bays is up to 5 terabits per second (Tb/s) across the NonStop signal midplane. Calculate the aggregate backplane bandwidth as follows: 160 Gb/s x 16 device bays x 2 directions = 5.12Tb/s. It is bandwidth between the device bays and interconnect bays. It does not include additional traffic capacity between interconnect modules or cross-bay connections. Current half-height blades use 12 of the 16 lanes to the interconnect bays.

One of the areas our engineering teams focused on was high-speed signal integrity. Getting this level of bandwidth between bays required special attention to high-speed signal integrity:

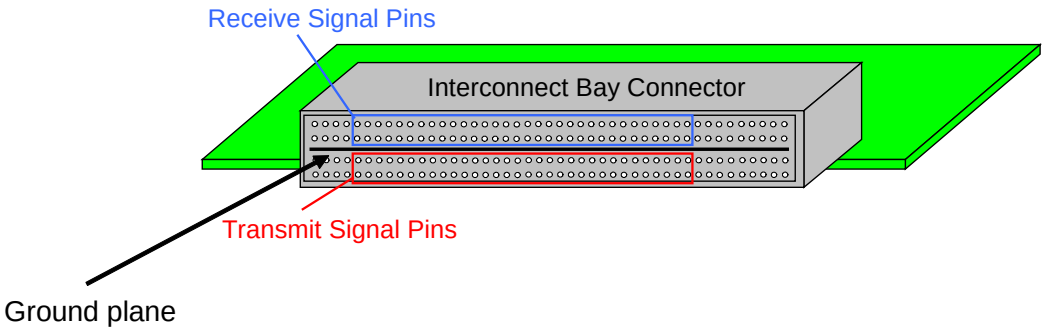
- Using general best practices for signal integrity to minimize end-to-end signal losses across the signal midplane
- Moving the power into an entirely separate backplane to independently optimize the NonStop signal midplane
- Providing a method to set optimal signal waveform shapes in the transmitters, depending on the topology of the end-to-end signal channel

Best practices

To ensure high-speed connectivity among all blades and interconnect modules, we leveraged our many years of experience in designing HP Superdome computers. Specifically, our engineers paid special attention to

- Controlling the differential signal impedance along each end-to-end signal trace across the PCBs and through the connector stages
- Using a ground plane to isolate, receive, and transmit signal pins (see Figure 10)
- Keeping signal traces short to minimize losses
- Routing signals in groups to minimize signal skew

Figure 10: Separation of the transmit and receive signal pins in the interconnect bay connector

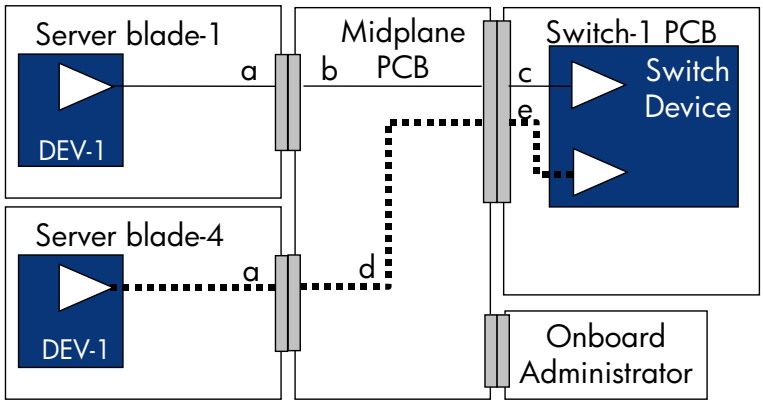


For more information about what we did to ensure signal integrity, see "Electrical signal integrity considerations for HP BladeSystem" available at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01712559/c01712559.pdf.

Channel topology and equalization settings

Even when using best practices, insertion and reflection losses can degrade high-speed signals transmitted across multiple connectors and long PCB traces. Insertion losses, such as conductor and dielectric material losses, increase at higher frequencies. Impedance discontinuities, primarily at connectors cause reflection losses. To compensate for these losses, shape the transmitter's signal waveform by selecting signal equalization settings. But a transmitter's equalization settings depend on the end-to-end channel topology and the type of component sending the signal. Both topology and the transmitting component can vary in the BladeSystem c-Class because of the flexible architecture and the use of mezzanine cards and NICs or other embedded I/O devices. As shown in Figure 11, the topology for device 1 on server blade 1 (a b c) is different from the topology for device 1 on server blade 4 (a-d-e). So, a link configuration mechanism in the Onboard Administrator (assisted by iLO on each server blade) identifies the channel topology for each device and configures the proper equalization settings for that device.

Figure 11: Different instances require different equalization settings



Signal midplane provides reliability

To provide high reliability, we designed the NonStop signal midplane as a completely passive board. It has no active components along the high-speed signal paths. The PCB consists primarily of traces and connectors. While there are a few components on the PCB, they are limited to passive devices that are unlikely to fail. The only active device is an EEPROM (Electrically Erasable Programmable Read-Only Memory) utility, which the Onboard Administrator uses to get information such as the midplane serial number. If this device fails, it does not affect the NonStop signal midplane. The solutions include the HP NonStop S-series, core networking switches from HP Networking, Cisco, and Juniper Networks, and core SAN switches from Brocade and Cisco.

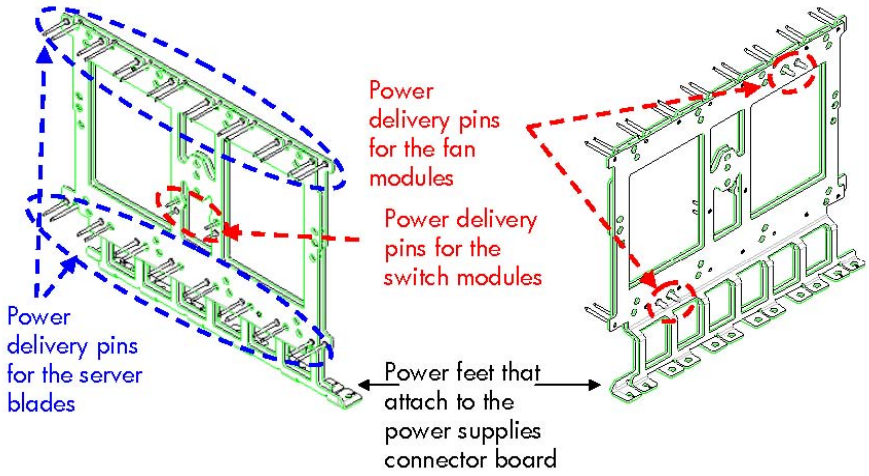
Separate power backplane

The PCB power backplane is separate from the NonStop signal midplane. This design improves the signal midplane by reducing its PCB thickness, reducing electrical noise from the power components that would affect high-speed signals, and improving the thermal characteristics. These design choices result in reduced cost, improved performance, and improved reliability.

Power backplane scalability and reliability

To minimize losses while distributing power, the power backplane has solid copper plates and integrated power delivery pins (Figure 12). Solid copper plates reduce voltage drops, produce high current density, and provide high reliability.

Figure 12: c-Class power backplane and the power delivery pins



Power and cooling architecture with HP Thermal Logic

With the BladeSystem c-Class architecture, HP engineers consolidated power and cooling resources to manage them efficiently within an enclosure. Every level of the architecture contains Thermal Logic technologies: processors, server blades, Active Cool fans, and the c-Class enclosure itself. Thermal Logic, which refers to the mechanical features and control capabilities throughout the design, gives

you the ability to see exactly how and where you're using power, and it allows you to optimize the enclosure for your power and cooling environment.

Through the Onboard Administrator controller, you can access real-time power and temperature data to monitor your environment. Onboard Administrator allocates power to the device bays based on the specific configuration of each blade. As you insert blades into the enclosure, the Onboard Administrator discovers each one and allocates power, based on measured maximum power. For ProLiant servers, the BIOS performs a power test during POST and adjusts the power allocation.

Onboard Administrator also lets you adjust operating conditions to meet your data center requirements. You can maximize performance based on your power and cooling budgets and reduce the need for expensive power and cooling upgrades.

Server blades and processors

HP Power Regulator and Insight Control power management software let you measure the power consumption of each blade and control the active processor power on ProLiant and Integrity servers. This technology takes advantage of processor performance states to scale power to meet performance requirements. For more information about HP Power Regulator technology, see these websites: www.hp.com/servers/power-regulator for ProLiant servers and www.hp.com/go/integritythermallogic for Integrity servers.

Visit this website for additional information about Insight Control power management: http://h18013.www1.hp.com/products/servers/management/ipm/index.html.

Precise ducting throughout the server blade manages airflow and temperature based on the unique thermal requirements of all critical components. This ensures that no air bypasses the server blade, which gives you the most thermal work from the least amount of air. This concept allows flexibility in heat sink design. Our engineers have designed heat sinks that closely match the requirements of the server blade and processor architecture.

Most important, c-Class server blades incorporate intelligent management processors (iLO for ProLiant server blades, Integrity iLO for Integrity server blades) that provide detailed thermal information for every blade. The iLO devices forward management information to the Onboard Administrator so you can access it through the Onboard Administrator interface.

Enclosure

At the enclosure level, Thermal Logic provides a number of advantages:

- Dynamic Power Saver mode to operate power supplies at high efficiencies
- Active Cool Fans to minimize power consumption
- Mechanical design features to optimize airflow
- Enclosure-level Dynamic Power Capping to set power limits

Dynamic Power Saver Mode

Most power supplies operate inefficiently when lightly loaded and more efficiently when heavily loaded. Dynamic Power Savings mode saves power by running the required power supplies at a higher load and putting unneeded power supplies in a standby mode. When power demand increases, the standby power supplies deliver the required extra power. As a result, the enclosure operates at optimum efficiency with no impact on redundancy.

Active Cool fans

Some small form-factor servers such as blade or 1U servers use very small fans to provide localized cooling in specific areas. Because these fans generate low airflow (measured in cubic feet per minute, or CFM) at medium backpressure, a single server often needs multiple fans to cool it.

Blower-style fans provide cooling across an entire enclosure. These fans are good at generating high airflow, but have the following drawbacks:

- They usually need more power.
- They produce more noise.
- They do not adjust for load. They run to cool for the highest load in an enclosure, constantly generating high airflow.

An alternative to the very small fan and blower-style fan is the HP Active Coolfan. The Active Cool fan delivers high airflow and high pressure in a small form factor. It uses ducted fan technology with a high-performance motor and impeller to deliver high CFM at high pressure. Use the Active Cool 100 fan in a c3000 enclosure and the Active Cool 200a fan in a c7000 enclosure.

The Onboard Administrator controls the thermal logic algorithms of the Active Cool fans. Onboard Administrator ramps up or lowers the cooling capacity based on the needs of the entire system. Along with optimizing airflow, the fan's control algorithm optimizes acoustic levels and power consumption. Even as new fans become operational, the Onboard Administrator thermal logic algorithms remain unchanged. Microprocessors inside the Active Cool fans translate the thermal logic algorithms into the actual rotor speeds required for the fan.

Active Cool fans deliver better performance than other fans in the server industry because of their mechanical design and the control algorithm. By aggregating the cooling capabilities of a few high-performance fans, we reduced the overhead of having many, localized fans for each server blade. This simplifies cooling and lowers the cost of the entire architecture.

Mechanical design to optimize airflow

By design, each c-Class enclosure lets fresh, cool air flow over all the server blades (in the front of the enclosure) and all the interconnect modules (in the back of the enclosure). HP optimizes the cooling capacity across the enclosure by optimizing airflow and minimizing leakage using a central plenum, self-sealing louvers surrounding the fans, and automatic shut-off doors surrounding the device bays.

A dedicated side slot in the front of the enclosure pulls fresh air into the interconnect bays. Ducts move the air from the enclosure's front to the rear, where the air is then pulled into the interconnect modules and the central plenum, and then exhausted from the rear of the system.

Each power supply module has its own fans, optimized for the airflow characteristics of the specific power supplies. The exhaust air does not interfere with the airflow path of the server blades or interconnect modules, because the power supplies are in a separate part of the enclosure.

Because the enclosures have separate cooling zones, the Active Cool fans cool their own zone and provide redundant cooling for the rest of the enclosure. One or more fans can fail and still leave enough fans to cool the enclosure adequately. The number of fans that can fail depends upon the number of blades, the number of fans, and the location of the blades. The Onboard Administrator reports thermal subsystem status and redundancy level, and it updates the system log and alert HP SIM when the thermal subsystem status changes.

Enclosure Dynamic Power Capping

With Enclosure Dynamic Power Capping, you can set a power cap for an entire BladeSystem enclosure, including all the individual server blades. When you use this capability, the Onboard

Administrator monitors and adjusts the power allocation to the server blades, making sure not to exceed the enclosure power cap.

Enclosure Dynamic Power Capping also allows the reallocation of power among the server blades over time. Using the blade power budget as its limit, the Onboard Administrator software uses a sophisticated algorithm to increase the power caps of busy individual servers consuming more power while it decreases the caps for server blades using less power. Thus, the Onboard Administrator makes the best use of power among the server blades in the enclosure while maintaining overall power consumption below the enclosure power cap.

For more information, see the technology brief "HP Power Capping and HP Dynamic Power Capping for ProLiant servers" at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01549455/c01549455.pdf.

Configuration and management technologies

An intelligent infrastructure lies at the heart of the BladeSystem c-Class. It combines embedded management capabilities in the hardware and integrated management software to streamline operations and increase productivity. The BladeSystem intelligent infrastructure makes essential power and cooling data available. It helps you automate the management of the infrastructure. It lets you control network connections across the enclosure.

These components create the intelligent infrastructure:

- iLO management processors integrated onto server blades
- Onboard Administrator
- Interconnect module management such as the HP Virtual Connect Manager (discussed in the "Virtual Connect" section)

In addition, the embedded management information feeds into these higher-level management tools:

- Virtual Connect Enterprise Manager (discussed in the "Virtual Connect" section)
- HP Insight Control
- HP Insight Dynamics

For more information, see the technology brief "Management architecture of HP BladeSystem c-Class systems" at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00814176/c00814176.pdf

Interactions among the BladeSystem's firmware components, BIOS, iLO, Onboard Administrator, Virtual Connect Manager, Smart Array, and NICs can be complex.. To help with this, we have designed a one-step update process for BladeSystem Firmware Release Sets. It provides a collection of ProLiant BladeSystem firmware. We test it as a solution stack and release it regularly on a single DVD ISO image. The HP Smart Update Manager is the software engine that drives the firmware update process. You can download the Smart Update DVD by going to the website at www.hp.com/go/foundation. Another good resource is the "BladeSystem ProLiant Firmware Management Best Practices" paper at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c02049593/c02049593.pdf.

Integrated Lights-Out

Each ProLiant and Integrity server blade designed for the BladeSystem c-Class includes an iLO or Integrity iLO management processor. Regardless of a server blade's operating condition, the iLO processor lets you manage the blade remotely and securely. The iLO processor provides access to a remote console, virtual power button, and system management information such as hardware health, event logs, and configuration. The iLO processor also monitors thermal and operational conditions within each server blade and forwards this information to the Onboard Administrator. For specific

product information, go to the iLO website www.hp.com/go/ilo and the Integrity iLO website: www.hp.com/go/integrityilo.

Onboard Administrator

The Onboard Administrator acts as the brains of the c-Class enclosure. It enables you to manage server blades or switches within the enclosure. The Onboard Administrator communicates with an iLO processor on each server blade to form the core of the BladeSystem management architecture. The Onboard Administrator performs configuration steps for the enclosure, collects thermal and power data, and enables run-time management and configuration of the enclosure components. It informs technicians of problems within the enclosure through email, SNMP, or the Insight Display.

The Onboard Administrator monitors:

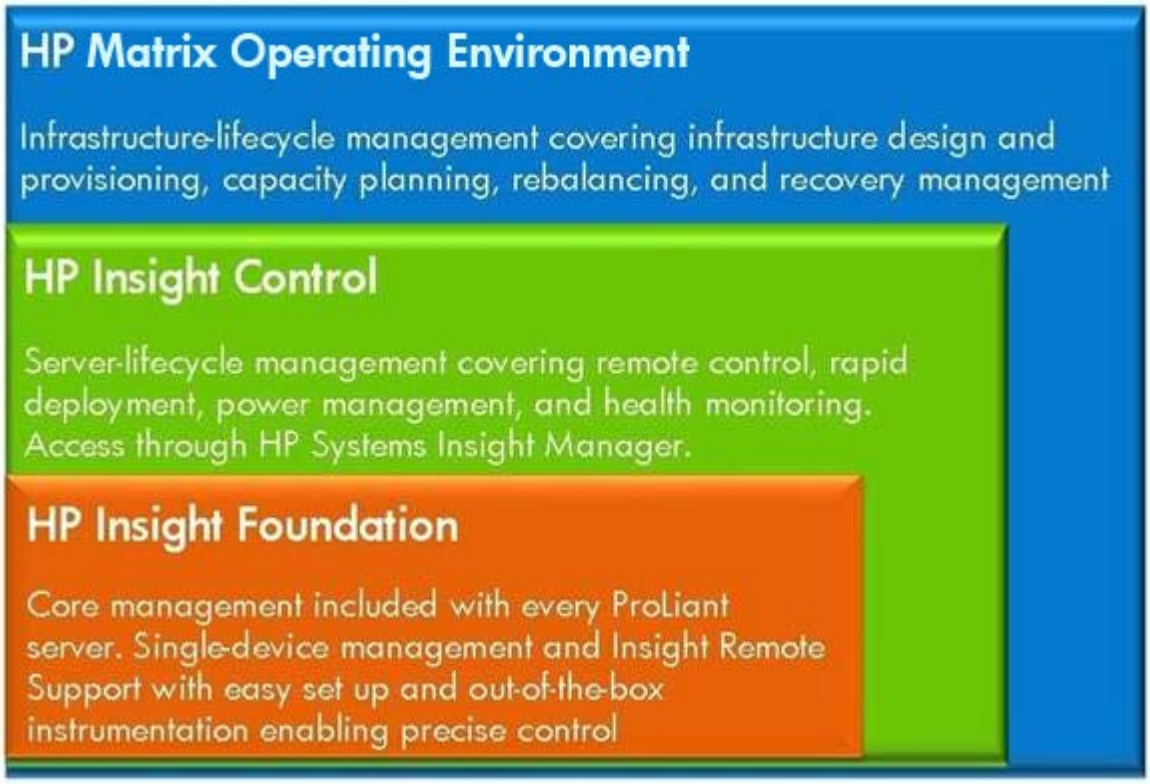
- Thermal conditions: If the thermal load increases, the Onboard Administrator's thermal logic feature increases fan speeds to accommodate the additional demand.
- Power allocation guidelines and power capacity limits for various components: The Onboard Administrator determines power consumption and availability. Because the Onboard Administrator uses measured power data, you can maximize the use of the available power.
- Enclosure Dynamic Power Caps: If you have configured Enclosure Dynamic Power Capping, the Onboard Administrator monitors and adjusts the power allocated to each blade to make sure your servers don't exceed the enclosure power cap.
- Hardware configurations: The Onboard Administrator assists in the configuration and setup process. As you add server blades and interconnect modules to the enclosure, the Onboard Administrator verifies electronic keying of interconnects and mezzanine cards as well as the location of components. The electronic keying mechanism ensures that the interconnect modules and mezzanine cards are compatible. It determines the signal topology and sets appropriate equalization levels on the transmitters to ensure best signal reception by the receivers.
- Network control capabilities: The Onboard Administrator provides tools to identify and assign IP addresses automatically for the c-Class components on existing management networks (for components supporting Dynamic Host Configuration Protocol).
- Enclosure linking: When appropriately configured, the Onboard Administrator allows single point access, or single sign-on. You can log in to a single Onboard Administrator and view and manage c-Class components in up to seven linked enclosures.

An Insight Display screen on the front of each c-Class enclosure provides quick access to the Onboard Administrator. For example, when you power up the enclosure, the Insight Display launches an installation wizard to guide you through the configuration process. After initial configuration, the Insight Display provides feedback and advice if there are any installation or configuration errors. You can access management menus, display User Notes, and chat with a remote administrator using the Insight Display.

HP Insight Control and HP Matrix Operating Environment

Since many BladeSystem users have more than one enclosure, it makes sense to have a way to manage an environment centrally with tens or hundreds of BladeSystem enclosures. The optional HP Insight Control software lets you manage and control HP server and storage resources through its central management console, HP Systems Insight Manager (Figure 13).

Figure 13: Relationship between Insight Control and Matrix Operating Environment



Insight Control provides an integrated management solution for physical and virtual servers. Insight Control can integrate with Microsoft System Center and VMware vCenter Server environments if your IT environment has already standardized on those management platforms. For Linux-centric environments, we offer Insight Control for Linux.

Matrix Operating Environment is optional, advanced infrastructure-management software that builds on Insight Control and Systems Insight Manager. It lets you instantly adjust to varying business demands. Insight Dynamics can help consolidate workloads, perform energy-aware capacity planning, and provision infrastructure

For more information, see www.hp.com/go/insightcontrol and www.hp.com/go/insightdynamics.

HP Virtual Connect and Virtual Connect Flex-10 technology

The BladeSystem c-Class architecture reduces the complexities of connection management in a blade environment through Virtual Connect technology. Virtual Connect virtualizes the server connections to LANs and SANs so you can configure connections only once. When installing a server blade, you can assign the necessary LAN and SAN connections for that server. There's no longer a need to wait until LAN and SAN administrators are available to reconfigure connections for a server.

With Virtual Connect Flex-10 technology, you can allocate a 10 GbE connection into up to four connections and adjust the bandwidth of each connection. With Virtual Connect FlexFabric modules, one of the four connections can be allocated to a storage connection (FlexHBA) to support Fibre Channel over Ethernet or iSCSI traffic.

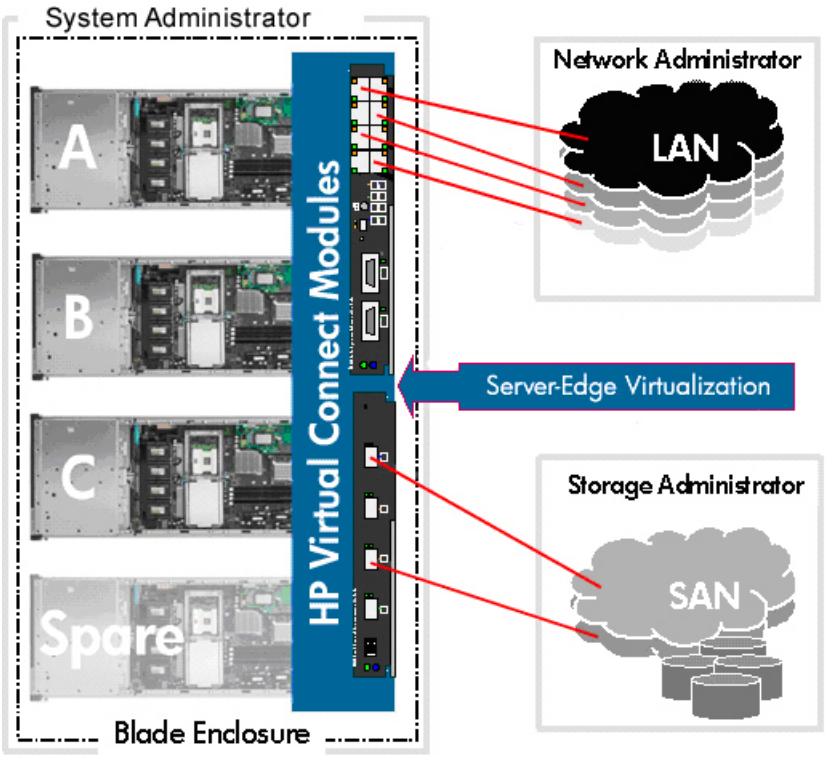
Virtual Connect

Virtual Connect includes a set of interconnect modules and embedded software that implements server-edge virtualization. It puts an abstraction layer between the servers and the external networks (Figure 14).

Virtual Connect uses connection profiles in combination with dynamic pools of unique MAC and WWN addresses to establish server connections to LANs and SANs. The server connection profiles contain MAC, WWN, and boot-from-SAN definitions assigned to BladeSystem enclosure bays and not to individual servers. The physical server in each bay uses the MAC and WWN assignments in the bay profile instead of its default NIC or HBA addresses. Even if you replace a server, the MAC and WWN assignments for the device bay remain constant, and the change is invisible to the network.

For more information, see the technology brief "HP Virtual Connect technology for the HP BladeSystem c-Class" at
http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00814156/c00814156.pdf

Figure 14: HP Virtual Connect technology provides server-edge virtualization



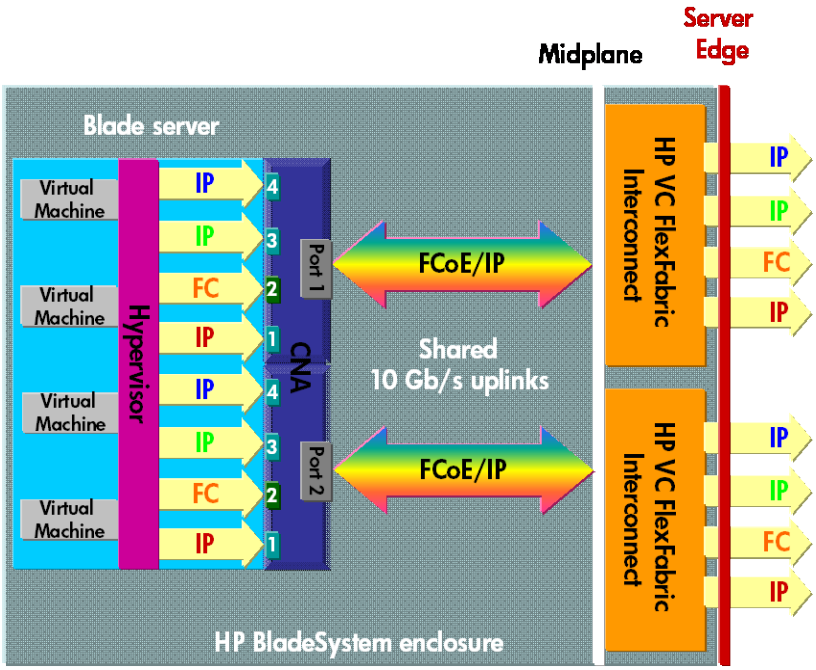
Virtual Connect Flex-10

Virtual Connect Flex-10 technology is a hardware-based technology. It lets you separate the bandwidth of a single 10GbE port into four physical network connections (FlexNIC) or three physical NIC devices and one physical HBA device. Through the Virtual Connect management interface, you can allocate the bandwidth in 100 Mb increments up to 10 Gb/s. You can also change the allocated bandwidth of an active FlexNIC or FlexHBA without rebooting the server. Each blade server can have up to four times as many network connections without adding NIC or HBA cards or switches.

Virtual Connect FlexFabric

HP Virtual Connect FlexFabric interconnect modules and FlexFabric adapters extend Flex 10 technology to include data and storage traffic within each 10 Gb server connection. This allows you to allocate LAN and SAN fabrics of a single 10 GbE data stream into four separate connections—one connection for each server port can be allocated to storage (FlexHBA) for either FCoE or iSCSI. You can adjust the bandwidth and routing information of each connection (Figure 15). A FlexFabric adapter functions as a standard NIC, a Flex-10 NIC, or a converged network adapter (CNA). To aggregate the LAN and SAN data, the FlexFabric adapter encapsulates Fibre Channel frames as FCoE or uses iSCSI along with the Ethernet LAN data. You can configure FlexFabric adapters to support either FCoE or iSCSI, but not concurrent streams of both. When the data stream enters the Virtual Connect FlexFabric interconnect module, the converged LAN and SAN traffic separate. Data streams leaving the BladeSystem enclosure use traditional Ethernet and storage (FC or SCSI) protocols. As a result, Flex-10 technology with FlexFabric adapters and VC FlexFabric modules provides a significant reduction in cabling, switches, and required ports at the server edge.

Figure 15: Using FlexFabric adapters with Fibre Channel (FCoE) and IP traffic



FlexFabric ports and dual interconnects configured for redundancy

Managing Virtual Connect

The Virtual Connect Manager, which controls single Virtual Connect domains, is a simple web console built into the firmware of Virtual Connect Ethernet modules. You can also access the Virtual Connect Manager through the Onboard Administrator. This option can be used for environments with up to four enclosures.

HP Virtual Connect Enterprise Manager is a software application that simplifies the management of large BladeSystem environments using Virtual Connect to control LAN and SAN connectivity. It lets you manage multiple Virtual Connect Domains using a single console application.

Availability technologies

The BladeSystem c-Class enclosure uses redundant configurations to eliminate single points of failure, and its architecture reduces the risk of component failures and the time required for changes. The c-Class enclosure employs multiple signal paths and redundant hot-pluggable components to provide maximum uptime for components in the enclosure.

Redundant configurations

The B c-Class enclosure minimizes the chances of a failure by providing redundant power supplies, cooling fans, and interconnect modules. For example, customers have the option of using power supplies in an N+N redundant configuration or an N+1 configuration. You can place the interconnect modules side-by-side for redundancy. In addition, you can use redundant Onboard Administrator modules in an active-standby configuration.

The c-Class architecture provides redundant paths using multiple facility power feeds into the enclosures, blade-to-interconnect bay connectivity, and blade-to-enclosure manager connectivity. Because all c-Class components are hot pluggable, you can quickly reestablish a redundant configuration in the event of a failure.

Reliable components

We took every opportunity in the c-Class architecture to design for reliability, especially for critical components considered single points of failure. Our customers sometimes ask if the NonStop signal midplane for the BladeSystem c-Class enclosure could be a single point of failure because it is not replicated. We designed the midplane to mitigate that risk (as described in the "General-purpose flexible design"_section).

In the unlikely event that an Onboard Administrator module fails, server blades and interconnect modules continue to operate. The module can be removed and replaced without affecting operations of the server blades and interconnect modules.

A component's operating temperature can play a significant role in reliability. The mechanical design of the BladeSystem c-Class enclosures minimizes the operating temperature of components in all critical areas. The Active Cool fan design and the Onboard Administrator's thermal monitoring of the entire system add to the lifespan of components by making sure that the fans cool the entire enclosure. Because of its unique fan blade, housing, motor windings, bearings, and drive circuit, the Active Cool fan provides higher reliability than typical server fans.

Reducing configuration time

Several important technologies in the BladeSystem c-Class reduce the amount of time needed to replace, upgrade, and configure systems: Onboard Administrator and the related Insight Display, Virtual Connect technology, and hot-plug devices.

With the intelligence of the Onboard Administrator and the easy-to-use Insight Display panel, you can configure and troubleshoot systems in minutes, rather than hours or days. Adopting Virtual Connect technology removes administrative burdens from LAN and SAN administrators because they are not required to change their network setup every time a server blade configuration changes. It is quick and easy to migrate the network service from a failed server blade to a functional server blade.

Finally, the fans, power supplies, interconnect modules, Onboard Administrator modules, server blades, and storage blades are hot pluggable, making it easy to repair or upgrade your systems.

Conclusion

Increasingly, enterprise customers require greater levels of efficiency and responsiveness in their IT environments. The HP BladeSystem c-Class architecture provides a flexible, simplified infrastructure to support their changing business needs. Every level of the architecture contains Thermal Logic technologies—processors, server blades, Active Cool fans, and the c-Class enclosure itself—letting you optimize the enclosure for your power and cooling environment. Shared cooling, power, and management resources, especially the Onboard Administrator and iLO management tools, support the modular components.

For more information

Resource description	Web address
HP BladeSystem Technical Resources	http://h71028.www7.hp.com/enterprise/cache/316735-0-0-0-121.html
HP BladeSystem Power Sizer	www.hp.com/go/bladesystem/powercalculator
HP BladeSystem website	www.hp.com/go/bladesystem/
HP Power Regulator for ProLiant	http://h18000.www1.hp.com/products/servers/management/ilo/power-regulator.html
ProLiant Servers Technology Communications	http://h18013.www1.hp.com/products/servers/technology/whitepapers/proliant-servers.html#bl
"Innovation Envelope: Hot Chips in Blades," Kevin Leigh, presented at the Hot Chips 21 conference, August 23-25, 2009, Stanford University	http://www.hotchips.org/archives/hc21/

Send comments about this paper to TechCom@HP.com


Follow us on Twitter: http://twitter.com/ISSGeekatHP

TC1109880, December 2011