



hp industry
standard servers

september 2002



white paper
16LD-0602A-WWEN

ProLiant BL e-Class C-GbE Interconnect Switch Overview

table of contents

- abstract 2
- introduction 2
- ProLiant BL e-Class network routing 2
 - same network technology, different form factor.....3
 - simplified network cabling4
 - C-GbE Interconnect Switch architecture.....5
- interconnect switch management 7
 - user account security7
 - browser-based interface8
 - console interface9
 - command line interface9
 - simple network management protocol and remote monitoring9
 - trivial file transfer protocol support10
 - switch diagnostics.....10
- network features and protocol support 12
 - enterprise class performance12
 - IEEE 802.1Q virtual local area network (VLAN)12
 - IEEE 802.1D spanning tree protocol12
 - IEEE 802.1p quality of service (QoS) packet prioritization12
 - IEEE 802.3x flow control13
 - port mirroring13
 - EtherChannel compatible trunking.....13
 - internet group management protocol (IGMP) snooping13
 - store and forward switching scheme13
 - DHCP and BOOTP protocols.....13
 - auto-negotiation of speed and duplex mode13
- network adapter teaming..... 14
- IEEE compliance 14
- ordering the C-GbE Interconnect Switch 14
- conclusion 14
- feedback 15

abstract

This white paper provides a technical overview of the ProLiant BL e-Class C-GbE Interconnect Switch option for the ProLiant BL e-Class system. The C-GbE Interconnect Switch option, which is used instead of a pair of traditional “edge” switches, includes a pair of industry standard Ethernet switches designed to dramatically reduce the number of network cables attached to the back of the ProLiant BL e-Class server blade enclosure. ProLiant BL e-Class systems are the first power-efficient, ultra-dense edge server blades engineered to address the needs of space-constrained enterprises and service providers for increased server density, rapid deployment and provisioning, and remote manageability.

The intended audience for this paper is engineers and system administrators familiar with the ProLiant BL e-Class system. For those less familiar with the ProLiant BL e-Class system, the web page at <http://www.HP.com/products/servers/proliant-bl/e-class/index.html> provides essential information.

introduction

ProLiant BL e-Class systems are the first power-efficient, ultra-dense edge server blades engineered to address the needs of space-constrained enterprises and service providers for increased server density, rapid deployment and provisioning, and remote manageability. Compaq designed the ProLiant BL e-Class system from the ground up to be an optimized solution with regard to power usage, space requirements, and performance. Thus, the ProLiant BL e-Class system provides an excellent overall solution for front-end applications in the service provider infrastructure.

Since the ProLiant BL e-Class system packages many server blades in a small space, the number of network cables within this space can quickly become overwhelming. A ProLiant BL e-Class server enclosure supports up to 20 ProLiant BL10e server blades, each with two 10/100 Mbps network ports, for a total of 40 network signals per enclosure.

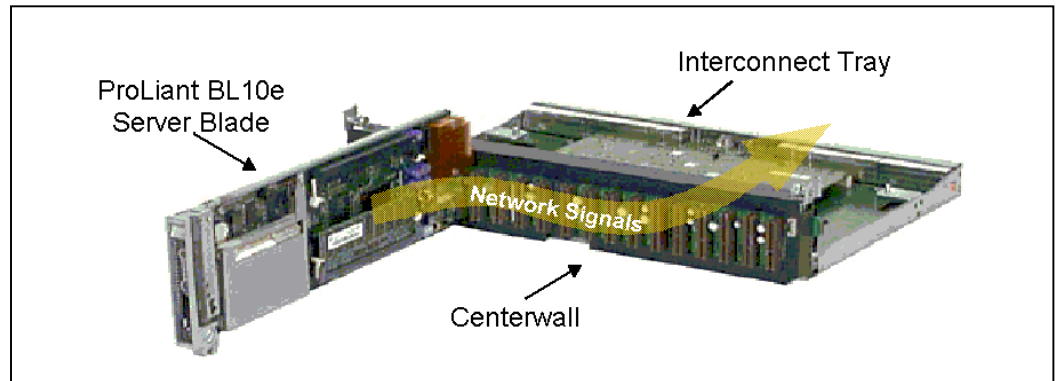
The C-GbE Interconnect Switch is one of three network connectivity options available with the ProLiant BL e-Class system. In a typical network, edge switches such as HP Procurve switches¹ are used to collect network adapter signals from the servers and concentrate them to the network. Used instead of a pair of edge switches, the C-GbE Interconnect Switch is designed to collect and significantly reduce the number of server blade network signals for direct connection to the network backbone.

ProLiant BL e-Class network routing

The ProLiant BL e-Class server enclosure supports 20 BL10e server blades, each with two NC3163 Fast Ethernet Embedded 10/100 Wake On LAN (WOL) network controllers enabled with Pre-boot eXecution Environment (PXE). Therefore, each server blade enclosure can have up to 40 network adapters. The signals from these 40 network adapters are routed from the servers via the centerwall assembly to an interconnect tray that slides into the back of the server blade enclosure (Figure 1).

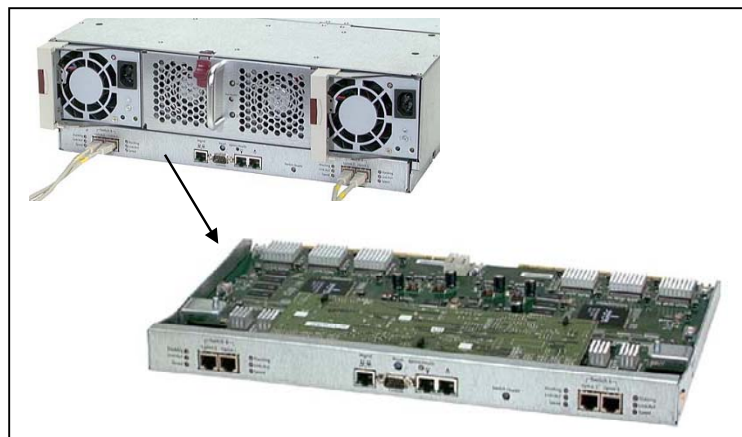
¹ Information about HP Procurve switches is available at <http://www.hp.com/rnd/index.htm>.

figure 1: ProLiant BL e-Class network routing



The C-GbE Interconnect Switch is a ProLiant BL e-Class interconnect tray option that uses a pair of redundant, non-blocking, managed layer 2 Ethernet switches (Figure 2). The interconnect switch provides up to a 40-to-1 reduction in network cables while offering network redundancy and a full set of enterprise-level features. This significantly reduces the time required to deploy, manage, and service ProLiant BL e-Class systems.

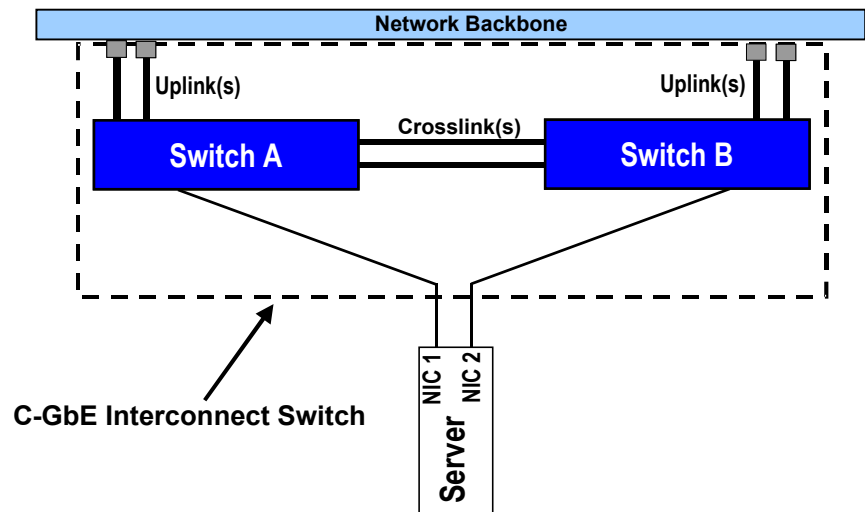
Figure 2: ProLiant BL e-Class server blade enclosure (back view) with the C-GbE Interconnect Switch option



**same network
technology, different
form factor**

In a typical server network configuration designed with redundancy, two or more network interface controller (NIC) ports are used per server. These NIC ports are routed to two separate edge switches. One or more connections are commonly made between the edge switches for added availability. The edge switches are used to collect NIC signals (usually 10/100) from the servers for aggregation to the network backbone, typically via gigabit uplinks. The BL e-Class C-GbE Interconnect Switch and blade architecture accomplish the same thing using the same technology but in a different form factor (Figure 3).

figure 3: typical redundant server network configuration

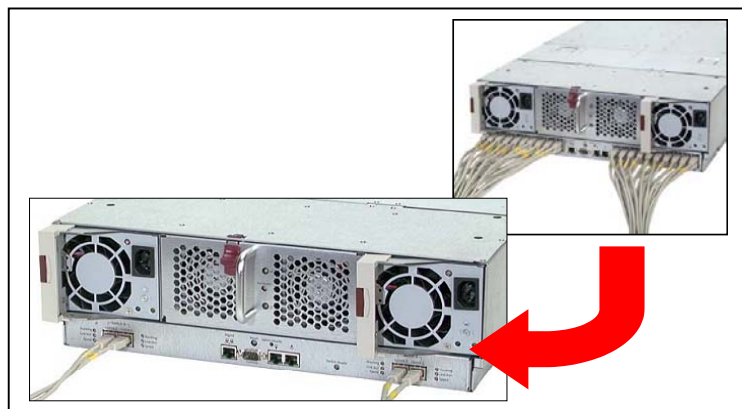


With the C-GbE Interconnect Switch option, the edge switches and connections have been moved inside the BL e-Class server blade enclosure. The same network technology is used and the network configuration and administration remain unchanged. That is, the interconnect switch is configured and administered in the same manner as any other industry-standard Ethernet switch. The interconnect switch is an industry-standard managed layer 2 switch compatible with other industry-standard switches.

simplified network cabling

As mentioned earlier, each 3U server blade enclosure supports up to 40 network adapters. Therefore, a single fully configured ProLiant BL e-Class server blade enclosure would require 40 network cables, and a 42U rack loaded with fully configured BL e-Class server blade enclosures would require a total of 560 network cables. A C-GbE Interconnect Switch collects all 10/100T network signals from the server blades within an enclosure and concentrates them to one to four 10/100/1000T Gigabit Ethernet uplinks for connection to the network backbone (Figure 4). Concentrating cables 40 to 1 results in the maximum of a 98 percent reduction in cables.

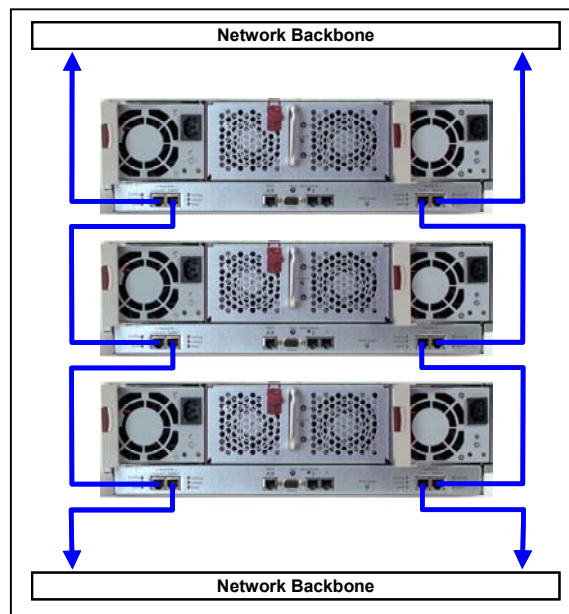
figure 4. ProLiant BL e-Class C-GbE Interconnect Switch cable reduction



This feature makes it possible to use just one uplink per server enclosure and a total of only 14 connections for all 560 network adapters in a 42U rack. Although using just one uplink per enclosure provides the maximum network cable reduction, system administrators may choose to use from two to four uplink ports per enclosure to increase bandwidth, to separate network and management data onto physically isolated ports, or to add redundant connections to the network backbone.

Since the C-GbE Interconnect Switch uses industry-standard, managed layer 2 switch technology, interconnect switches from different server enclosures and other vendor switches all within the same rack or location may be linked together. This reduces network cabling even further and provides the desired level of oversubscription. For example, Figure 5 shows a configuration with three fully populated BL e-Class server blade enclosures using the interconnect switch option. The C-GbE Interconnect Switches are linked, or daisy chained together, in redundant configuration via the four gigabit uplinks that are, in turn, linked to the network backbone. Each enclosure contains 40 network adapters and has a maximum bandwidth of 4 Gbps, for a total bandwidth of 12 Gbps for entire system of three enclosures. However, since uplinks are daisy chained together, the maximum system throughput is limited to 4.0 Gbps in this example. This configuration creates a 3x oversubscription (12 Gbps versus 4 Gbps). Nevertheless, it reduces the overall network cables at the system or rack level from 120 to 4, a 97 percent cable reduction, while still maintaining two sets of redundant connections to the network backbone.

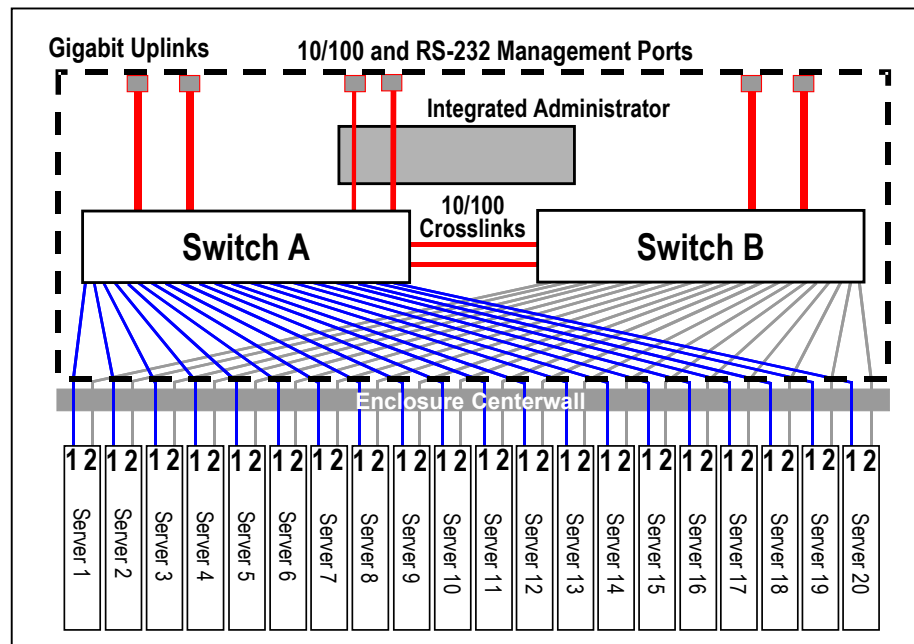
figure 5: ProLiant BL e-Class C-GbE Interconnect Switch linking with oversubscription



C-GbE Interconnect Switch architecture

The interconnect switch features a fully redundant architecture that includes switch redundancy (dual switch modules) and redundant paths to the network ports on the server blades. This architecture provides a highly available, fully meshed design to the network backbone (Figure 6).

figure 6. ProLiant BL e-Class C-GbE Interconnect Switch architecture



Each switch module (labeled Switch A and Switch B) directly connects to one of the two NC3163 NICs on each ProLiant BL10e server blade (labeled as NIC 1 and 2) via the centerwall of the server blade enclosure. Each switch module includes two 10/100/1000T Gigabit Ethernet uplink ports. Besides providing a completely non-blocking network configuration, these uplinks can physically isolate data and provide redundant connections to the backbone. IEEE 802.1D spanning tree protocol is supported to eliminate potential problems caused by redundant networking paths.

The two switch modules are further connected through redundant 100 Mbps crosslinks that permit traffic between the switch modules. These crosslinks make it possible to see any server NIC from any uplink port. As a result, any single uplink port may be used to connect to all 40 server network connections for the 40-to-1 network cable reduction. On a heavily utilized system, however, using a single uplink port for all 40 NICs can cause a traffic bottleneck. For example, using only uplink 1 on switch module A forces all traffic to and from NIC 2 on all server blades to travel over the crosslinks (a 200-Mb path). This path to the server blade NICs is intended primarily as a failover route and generally is not used as a primary path. Instead, for optimal performance, at least one uplink port on each switch could be used. In general, utilizing just one uplink is more applicable for systems with only a few server blades or limited network traffic.

The entire C-GbE Interconnect Switch option is powered and cooled by the 1+1 redundant hot-plug power supplies and the 2+2 redundant hot-plug fans that ship standard with the ProLiant BL e-Class server blade enclosure.

The redundant architecture of the C-GbE Interconnect Switch allows system administrators to configure the network for continued access to each server blade in case any of the following components should fail:

- interconnect switch module
- switch within the network backbone
- server blade network controller

- server blade NIC-to-interconnect switch port or connection
- uplink port and uplink port connection and cable
- interconnect switch cross-link port or connection
- power and fan failure

For example, if switch module A were to fail for any reason, network access to all servers would still be possible on NIC 2 via the uplink ports on switch module B. Similarly, if access to the uplinks on switch module A were inaccessible, network traffic could be sent through switch module B directly to NIC 2 and also to NIC 1 via the switch module crosslinks. In this case, access to all 40 network controllers is possible, but NIC 1 traffic is via the two 10/100 Mbps crosslinks. As explained earlier, these crosslinks are intended primarily as a failover route. Such a failover scenario should be viewed as a temporary situation if the crosslink bandwidth could cause a bottleneck.

interconnect switch management

Since the interconnect switch option is an industry-standard managed layer 2 switch, it is configured and administered in the same manner as any other industry-standard Ethernet switch. The interconnect switch includes a default configuration to assist with installation.

Both a browser-based and a menu-driven console interface are provided to configure, manage, and monitor the interconnect switch on a per port basis. An industry-standard scriptable command line interface (CLI) is planned as a future, no-cost software upgrade. SNMP (simple network management protocol) and RMON (remote monitoring) manageability are further supported. Even an SNMP-based scripting utility is available to execute customized configuration templates on multiple switch modules. Any combination of the switch-to-server blade network ports, gigabit uplink ports, and cross connection ports can be disabled, enabled, configured, and monitored as required. Out-of-band and in-band access are supported locally and remotely from anywhere on the network. Switch administration is possible via any uplink port and the 10/100T and serial management ports provided (Figure 7 and Table 1).

user account security

To secure the management interfaces, the switch administrator can configure multiple password-protected accounts with various levels of access. Access can be restricted to a limited number of management stations by configuring a list of IP addresses of those stations that can access the interconnect switch. Telnet sessions and the console management interface can be also configured with a time-out period.

figure 7.: C-GbE Interconnect Switch external panel

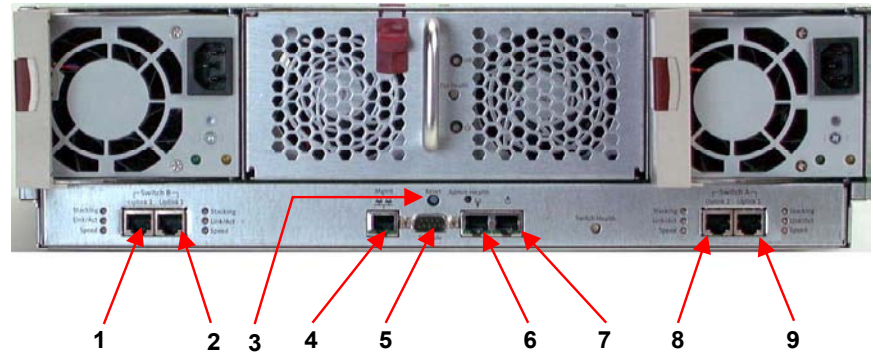


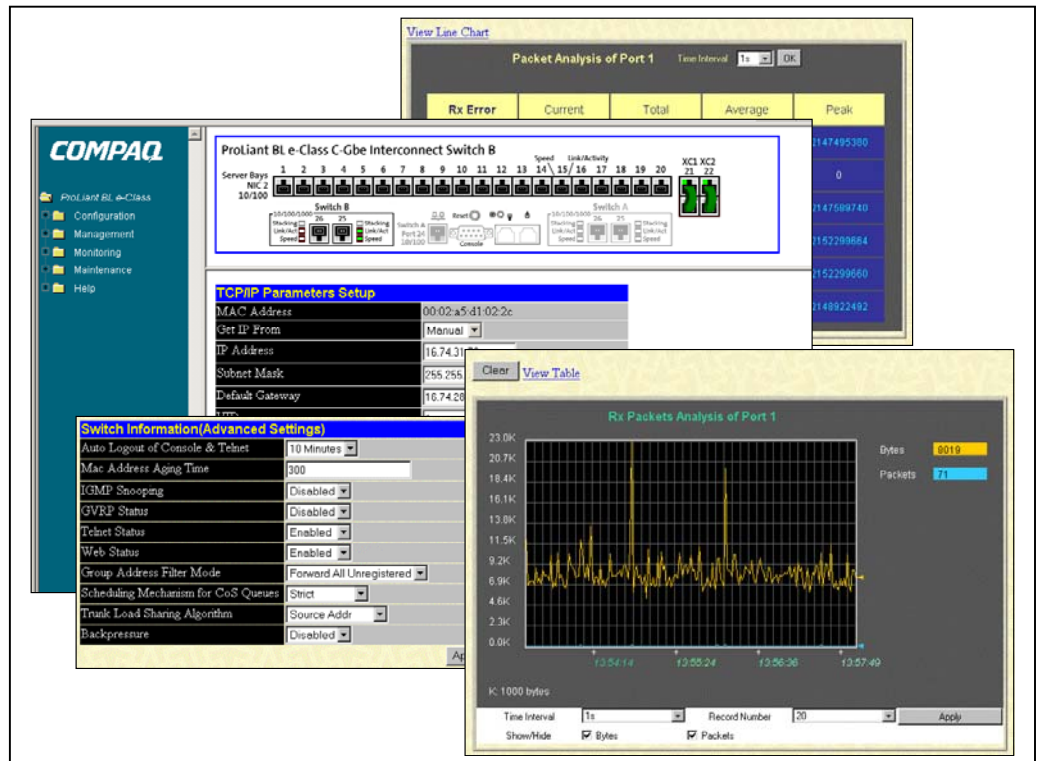
table 1. key to figure 7 callouts

item no.	item description
1	Switch B 10/100/1000T Gigabit Ethernet Uplink Port 26, RJ-45 Connector
2	Switch B 10/100/1000T Gigabit Ethernet Uplink Port 25, RJ-45 Connector
3	Combined interconnect switch and Integrated Administrator Reset Button
4	Combined interconnect switch and Integrated Administrator 10/100T Ethernet Management Port 24, RJ-45 Connector
5	Combined interconnect switch and Integrated Administrator Console Port, DB-9 Serial Connector
6	Reserved for future use
7	Reserved for future use
8	Switch A 10/100/1000T Gigabit Ethernet Uplink Port 26, RJ-45 Connector
9	Switch A 10/100/1000T Gigabit Ethernet Uplink Port 25, RJ-45 Connector

browser-based interface

The included browser-based interface is accessible using a web browser over a TCP/IP network (Figure 8). Access is possible via any Ethernet switch port, such as the gigabit uplinks. Access is also possible through the 10/100T Ethernet port provided on the Integrated Administrator.

figure 8. ProLiant BL e-Class C-GbE Interconnect Switch browser-based interface



console interface

The included console management interface provides the same configuration options as the browser-based interface, plus the added benefit of an out-of-band connection (Figure 9). There are three methods of accessing this interface:

- After starting a console session to the Integrated Administrator, either locally or remotely, a user can connect to the interconnect switch with the connect switch [A/B] command.
- A user can start an interconnect switch console session via the included RS-232 console port, without first going through the Integrated Administrator.
- A console Telnet session is supported over the network via the 10/100T Integrated Administrator port or any switch Ethernet port.

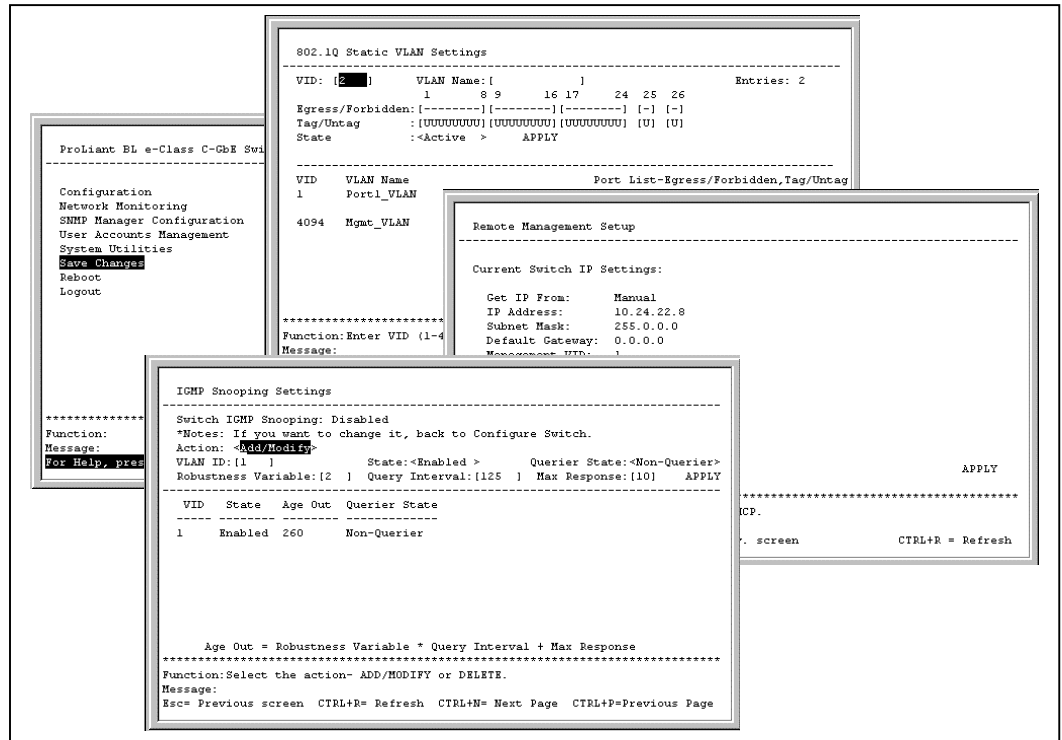
command line interface

An industry-standard scriptable command line interface (CLI) is planned for a future, no-cost firmware upgrade.

simple network management protocol and remote monitoring

Each switch module can be configured and monitored remotely from an SNMP/RMON-based network management station such as Insight Manager 7 and HP Openview. An interconnect switch SNMP scripting utility and recommended example scripts are included. The switch modules support industry-standard SNMP management information bases (MIBs), Compaq switch MIBs, and RMON groups 1 (statistics), 2 (history), 3 (alarm), and 9 (event) for fault detection, configuration, and monitoring of switch functionality.

figure 9. ProLiant BL e-Class C-GbE Interconnect Switch console interface



Information about Insight Manager 7 is available on the HP website at <http://www.compaq.com/products/servers/management/cim7-description.html>.

Information about HP Openview is available on the HP website at <http://www.hp.com/products1/softwareproducts/software/openview/index.html>.

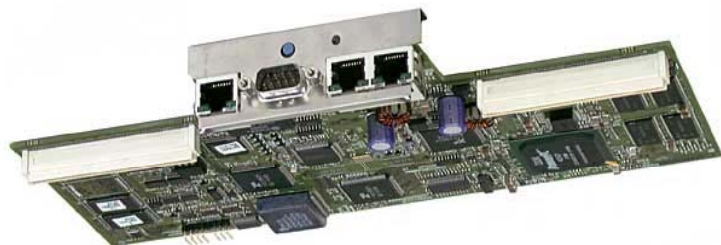
trivial file transfer protocol support

The trivial file transfer protocol (TFTP) feature of the interconnect switch permits a TFTP server to upload a copy of the interconnect switch configuration file that can be saved and downloaded to the original or different interconnect switch. This provides a means to rapidly deploy multiple systems with similar configurations and provides backup and restore capabilities. In addition, a history log can be uploaded from the interconnect switch to the TFTP server.

switch diagnostics

The C-GbE Interconnect Switch option includes a removable Integrated Administrator module (Figure 10). Information about the Integrated Administrator module is available at <http://www.compaq.com/products/servers/proliant-bl/e-class/integrated-admin.html>.

figure 10. ProLiant BL e-Class Integrated Administrator



The ProLiant BL e-Class Integrated Administrator provides a single management console for the efficient management of a server blade enclosure and its accompanying server blades. This includes automatic health monitoring of the interconnect switch with SNMP trap generation and system LED status updates. External LEDs are provided for enclosure and interconnect switch status and for link and speed on each gigabit uplink RJ-45 connector (Figure 11 and Table 2). An emergency enclosure shut-down feature is included in case of critical system temperature caused by the interconnect switch or other enclosure component.

figure 11: Interconnect switch external panel LEDs

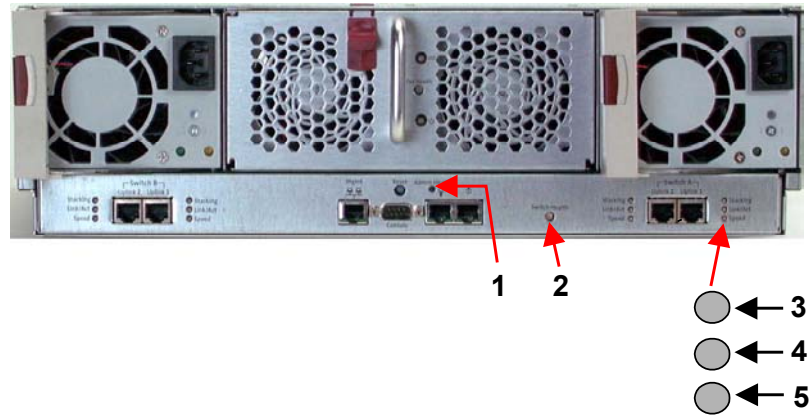


table 2. key to figure 14 LEDs

LED	description	color	status
1	Integrated Administrator Health	green amber red off	Enclosure on, Integrated Administrator health good Integrated Administrator health degraded Integrated Administrator health critical Enclosure off, Integrated Administrator health good
2	Interconnect Switch Health	green amber red off	Enclosure on, interconnect switch health good Interconnect switch health degraded Interconnect switch health critical Enclosure off or booting, interconnect switch health good
3	Stacking Status*	green yellow off	Base unit Slave unit No stacking
4	Link Activity	green flashing green yellow off	Network link Network activity Port disabled No network link
5	Link Speed	green yellow off	100 Mbps 1000 Mbps 10 Mbps or no network link

*Stacking functionality support is being considered for a future release.

Additional diagnostic features include a power-on self test (POST) built into the interconnect switch boot-up process and support for port mirroring. Port mirroring is a diagnostic and debugging tool that provides a method to monitor network traffic on a switch port by sending a copy of the data to a different port.

More information about the ProLiant BL e-Class Integrated Administrator is available online at <http://www.compaq.com/products/servers/proliant-bl/e-class/integrated-admin.html>. For more detailed information about the administration capabilities of the C-GbE Interconnect Switch, see the *ProLiant BL e-Class C-GbE Interconnect Switch User Guide*, available online at http://www.pro.compaq.com/support/reference_library/viewdocument.asp?countrycode=1000&prodid=5315|ProLiant+BL+e-Class+C-GbE+Interconnect+Switch&source=263682-001.xml&dt=264.

network features and protocol support

enterprise class performance

The interconnect switch option uses industry-standard managed layer 2 switches compatible with other industry-standard switches from popular network companies. It offers many enterprise-level network features and supports several industry standards.

The interconnect switch includes the following:

- nonblocking, full wire speed on all ports
- 16.8 Gbps switching fabric (8.4 Gbps per switch module) with 12 million packets per second aggregate maximum throughput (64-byte packets)
- 16,382 MAC addresses with MAC address based filtering and MAC address packet priority configuring
- 64 MB main memory, 2 MB flash memory, and 32 MB packet buffer memory

IEEE 802.1Q virtual local area network (VLAN)

Each switch supports up to 64 port-based 802.1Q tagged VLANs with GVRP dynamic VLAN registration.² VLANs simplify administration and planning, enhance performance by conserving bandwidth, and improve security by limiting traffic to specific domains.

IEEE 802.1D spanning tree protocol

The support for the spanning tree protocol allows the blocking of links that form loops between switches in a network. When multiple links between switches are detected, a primary link is established. Duplicated links are blocked from use and become standby links. If the primary link fails, the standby link is activated. One instance of STP is supported per each switch, also known as mono-spanning tree. STP switch parameters can be configured independently at the switch and port level. Each switch can automatically find the STP root bridge on the network; if not, it will act as the root bridge for the STP domain. Also, the switch supports STP fast forward or bypass mode allowing a port to skip the standard STP mode and enable itself directly in forward mode.

IEEE 802.1p quality of service (QoS) packet prioritization

Class of Service (CoS) for packet prioritization allows switch administrators to set priority levels on each switch for forwarding packets. Each switch supports four classes of traffic (buffers or queues) for implementing priority. The interconnect switch allows administrators to map eight priority levels to four classes. Traffic from a specific server port can be given priority over packets from other devices according to this range of priority levels. For example, with multiple packets in a buffer, the packet with the highest priority would be forwarded first, regardless of when it was received.

² GVRP stands for Generic VLAN Registration Protocol.

IEEE 802.3x flow control

The support for flow control enables each switch to signal network components to slow down or stop transmission of data during periods of peak network traffic. This provides a means to overcome switch overload by smoothing out traffic flow and eliminating risk of lost data.

port mirroring

The interconnect switch allows the user to mirror a switch port to another port for network monitoring and troubleshooting purposes. This technology offers a way for network packet analyzers to view the traffic moving through the switches by providing a copy of the traffic that is currently being passed through any other port.

EtherChannel compatible trunking

The interconnect switch IEEE 802.3ad static port trunking feature allows several ports to be grouped together and act as a single logical link called a trunk. This feature provides a bandwidth that is a multiple of a single link's bandwidth. It also improves reliability, since a configurable way of load balancing is automatically applied to the ports in the trunked group. A link failure within the group causes the network traffic to be directed to the remaining links in the group. The trunking implementation supports Cisco EtherChannel link aggregation (Fast EtherChannel, Gigabit EtherChannel).

internet group management protocol (IGMP) snooping

For efficient use of bandwidth, IGMP snooping manages multicast traffic in the interconnect switch by allowing directed switching of the IP multicast traffic. The interconnect switch can use IGMP snooping to configure switch module ports dynamically, so that IP multicast traffic is forwarded only to those ports associated with IP multicast hosts.

IGMP snooping allows the switch module to recognize IGMP queries and reports sent between network stations or devices and an IGMP host that belong to a specific multicast group. When enabled for IGMP snooping, the switch module can open or close a port to a specific device based on IGMP messages passing through the module. This feature further limits unnecessary broadcasts.

store and forward switching scheme

The interconnect switch provides a store and forward switching scheme that allows each packet to be buffered (stored) before it is forwarded to its destination. This scheme improves reliability in networks with high traffic. Packets that cannot be forwarded are saved immediately, rather than dropped, and packets behind them are less likely to be dropped in periods of heavy usage.

DHCP and BOOTP protocols

An interconnect switch can be configured to obtain an IP Address from a DHCP or BOOTP server during the boot process.³ The IP settings can be manually configured through the console interface. The IP settings are also configurable from the other interfaces, such as the web; but since the connection is based on an IP address for these interfaces, users would have to reconnect with the newly assigned IP address.

auto-negotiation of speed and duplex mode

The interconnect switch automatically identifies and matches its port speed and duplex mode of an attached device. This feature simplifies deployment in mixed 10, 100, and 1000 Mbps networks.

³ DHCP stands for Dynamic Host Configuration Protocol. BOOTP stands for Bootstrap Protocol.

network adapter teaming

Switch considerations for teaming network adapters are the same with the C-GbE Interconnect Switch option as with any other network switch. The ProLiant BL10e server includes two NC3163 network controllers that support all three types of teaming:

- Network Fault Tolerance (NFT)
- Transmit Load Balancing (TLB)
- Switch-assisted Load Balancing (SLB), formerly known as Fast EtherChannel/Gigabit EtherChannel (FEC/GEC) teaming

The two NC3163 NIC ports on each server are split across the two switches for redundancy (that is, the NICs are routed to two separate switches); see Figure 6. Both NFT and TLB support teamed NICs that are routed to separate switches, so both these types of teaming can be implemented using the C-GbE Interconnect Switch. However, SLB teaming requires that the switch support trunking (link aggregation) and that all the NICs in the team be connected to the same switch. The ProLiant BL e-Class Interconnect Switch does support trunking; but since the two NICs are routed to different switches, SLB cannot be implemented when using the C-GbE Interconnect Switch.

For additional information about teaming ProLiant network adapters, see the teaming overview and white papers located at

<http://www.compaq.com/products/servers/networking/teaming.html>.

IEEE compliance

The ProLiant BL e-Class C-GbE Interconnect Switch complies with the following IEEE specifications:

- 802.3 Ethernet
- 802.3u 100Base-X Ethernet
- 802.3ab 1000Base-T Gigabit Ethernet
- 802.3ad Trunking (Static)
- 802.3x Ethernet Flow Control
- 802.1D Spanning Tree Protocol
- 802.1p Quality of Service (QoS) Packet Prioritization
- 802.3ac and 802.1Q Port-based Tagged VLANs and VLAN Extensions
- 802.1Q Trunking Encapsulation

ordering the C-GbE Interconnect Switch

Each ProLiant BL e-Class server enclosure requires one interconnect tray — either a patch panel or an interconnect switch. The ProLiant BL e-Class C-GbE Interconnect Switch tray, part number 243283-B21, includes all components necessary for one server enclosure.

The kit includes:

- ProLiant e-Class C-GbE Interconnect Switch tray with two layer 2 switches
- four 10/100/1000T Gigabit Ethernet uplinks
- removable Integrated Administrator
- 10/100T and serial management ports
- quick install card
- system utilities and user documentation (downloadable from the web)

conclusion

The ProLiant BL e-Class system is the first in a line of modular, blade servers for the data center. As enterprises move to increasingly modular architectures, requiring quick change and rapid deployment, the ProLiant BL e-Class system provides industry-leading performance in a power-efficient, space saving design.

The ProLiant e-Class C-GbE Interconnect Switch is available as one of three interconnect tray options for the ProLiant BL e-Class system. The interconnect switch is a pair of fully redundant, managed layer 2 Ethernet switches that consolidate up to forty 10/100 Mbps network adapters to one to four 10/100/100T Gigabit Ethernet uplinks. The switch is

ideal for users who desire greatly decreased network cabling for a significant reduction in the time required to deploy, manage, and service ProLiant BL e-Class systems.

feedback

Please direct comments regarding this communication to the ISS Technology Communications Group at this Internet address: TechCom@HP.com

The information in this document is subject to change without notice.

09/2002

P/N 16LD-0602A-WWEN