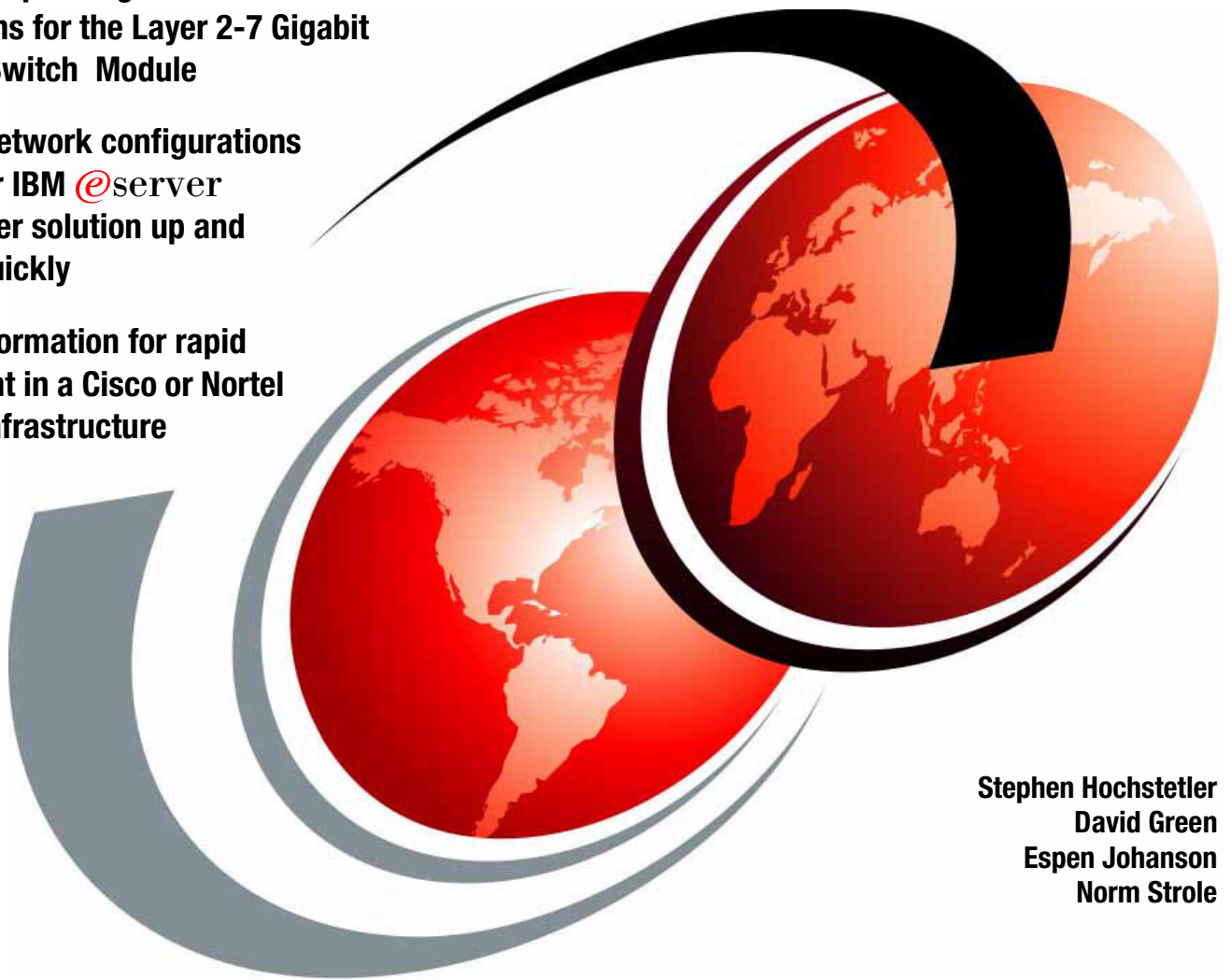


IBM @server BladeCenter Layer 2-7 Network Switching

Step-by-step configuration instructions for the Layer 2-7 Gigabit Ethernet Switch Module

Working network configurations to get your IBM @server BladeCenter solution up and running quickly

Helpful information for rapid deployment in a Cisco or Nortel network infrastructure



Stephen Hochstetler
David Green
Espen Johanson
Norm Strole



International Technical Support Organization

**IBM @server BladeCenter
Layer 2-7 Network Switching**

January 2004

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (January 2004)

This edition applies to IBM @serverBladeCenter (product number 8677-1XX).

This document created or updated on January 29, 2004.

© Copyright International Business Machines Corporation 2004. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	vii
Trademarks	viii
Preface	ix
The team that wrote this Redpaper	ix
Become a published author	x
Comments welcome	x
Chapter 1. Introduction to the IBM @server BladeCenter Layer 2-7 Gb Ethernet Switch Module	1
1.1 Introduction to GbESM: Capabilities, features, functions	2
1.2 Switch management and operating concepts	6
1.2.1 Switch management and control	6
1.2.2 Switch operating concepts	7
1.3 Ports and performance features	8
1.4 Switch and network management	8
1.5 Network cables	9
1.6 Supported network standards	9
Chapter 2. GbESM management and initial configuration	11
2.1 GbESM management through the management module	12
2.1.1 Establishing a physical connection to the management module	13
2.1.2 Using the management module Web interface to initially configure GbESM	13
2.2 GbESM management using the GbESM Web interface	19
2.2.1 Basic GbESM configuration examples	20
2.2.2 Configuring the GbESM via the Browser-Based Interface	21
2.2.3 Configuring the switch via the CLI	22
Chapter 3. Layer 2-7 GbE Switch Module for IBM @server BladeCenter functionality 31	
The OSI reference model	32
3.1 Layer 1: Physical layer	33
3.2 Layer 2: Data link layer	33
3.3 Layer 3: Network layer	43
3.4 Layer 4: Transport layer	54
3.5 Layer 5: Session layer	61
3.6 Layer 6: Presentation layer	61
3.7 Layer 7: Application layer	61
Chapter 4. Layer 2 configuration guidelines	71
4.1 Configuring multi-port trunks	72
4.1.1 Port trunking example	73
4.2 Configuring IEEE 802.1Q VLANs	75
4.2.1 VLANs and port VLAN ID numbers	76
4.2.2 VLAN topologies and design considerations	80
4.3 Configuring the Spanning Tree Protocol	83
Chapter 5. Deploying the Layer 2-7 GbE Switch Module for IBM @server BladeCenter in a Cisco environment	85
5.1 Introduction	86

5.2	Architecture summary	86
5.2.1	Datacenter networks introduction	86
5.2.2	Common Cisco components	87
5.3	Guidelines, rules, and comments	89
5.3.1	Cleaning and resetting systems	90
5.3.2	Rules for attaching the eServer BladeCenter to a Cisco infrastructure	92
5.4	Preliminary information on configuration examples	96
5.5	Configuration examples	104
5.5.1	Single GbESM, single link to a single Cisco switch	104
5.5.2	Single GbESM, single link to two Cisco switches	109
5.5.3	Single GbESM, four port static aggregation to a single Cisco switch	116
5.5.4	Single GbESM, dual port static aggregation to two Cisco switches	124
5.5.5	Dual GbESMs, each with a single link to the same Cisco switch	131
5.5.6	Dual GbESMs with a single link to two different Cisco switches	132
5.5.7	Dual GbESMs each with one link to separate Cisco switches	133
5.5.8	Dual GbESMs with 4 port static aggregation to different Cisco switches	141
5.5.9	Dual GbESMs with two static aggregation to two Cisco switches	141
5.6	Troubleshooting GbESM connections to Cisco devices	156
5.6.1	Troubleshooting specifics	156
Chapter 6. Deploying the Layer 2-7 GbE Switch Module for IBM @server BladeCenter in a Nortel environment		159
6.1	Nortel Networks feature descriptions	161
6.2	Limitations of configuration examples	162
6.3	Preliminary configuration for examples	162
6.3.1	General recommendations	162
6.3.2	IP addresses used in example configurations	163
6.3.3	Base configuration options common to all examples	164
6.3.4	Basic configuration procedures	165
6.3.5	Base configuration tasks for the GbESM	166
6.3.6	Base configuration tasks for BayStack 380-24T	169
6.3.7	Base configuration tasks for Passport 8600	173
6.4	Configuration examples	175
6.4.1	Single GbESM with link aggregation and routing to single BayStack 380-24T	175
6.4.2	Validation of GbESM configuration	179
6.4.3	Validation of BayStack 380-24T configuration	179
6.4.4	Single GbESM with link aggregation to dual BayStack 380-24Ts	181
6.4.5	Validation of GbESM configuration	184
6.4.6	Validation of BayStack 380-24T configuration	185
6.4.7	Dual GbESMs with two port aggregation to Dual BayStack 380-24Ts	186
6.4.8	Validation of GbESM configuration	190
6.4.9	Validation of BayStack 380-24T configuration	190
6.4.10	Dual GbESMs with four port aggregation to dual BayStack 380-24Ts	192
6.4.11	Validation of GbESM configuration	196
6.4.12	Validation of BayStack 380-24T configuration	196
6.4.13	Dual GbESMs with four port aggregation, each to a single Passport 8600	198
6.4.14	Validation of GbESM configuration	201
6.4.15	Validation of Passport 8600 configuration	201
6.4.16	Dual GbESMs with four port SMLT to dual passport 8600s	203
6.4.17	Validation of GbESM configuration	208
6.4.18	Validation of Passport 8600 configuration	208
6.5	Troubleshooting GbESM connections to Nortel Networks devices	210

Related publications	211
IBM Redbooks	211
Other publications	211
Online resources	211
How to get IBM Redbooks	212
Index	213

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law. INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:


This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

@server™
eServer™
@server™
ibm.com®

xSeries®
BladeCenter™
IBM®
Redbooks™

Redbooks (logo) ™
Tivoli®

The following terms are trademarks of other companies:

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Nortel Networks, the Nortel Networks Globemark, Unified Networks, and How the world shares ideas are trademarks of Nortel Networks Corporation.

Cisco, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries. All other trademarks mentioned in this document are the property of their respective owners.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Other company, product, and service names may be trademarks or service marks of others.

Preface

This IBM® Redpaper will help you install, tailor, and configure the new IBM @server BladeCenter™ and the Layer 2-7 GbE Switch Module (GbESM) in various network environments.

In this Redpaper, we discuss the features and functions of the GbESM and how it is managed. We also introduce some networking terms and concepts to the users and administrators of the IBM @server BladeCenter product. This material is meant to give the non-networking professional an overall view of the switching and bridging environment, not to substitute for in-depth training on networking fundamentals.

We then provide step-by-step instructions for establishing your initial network configurations to ensure your Layer 2-7 GbE Switch Module is configured correctly and is operable so you can begin immediate communications across the network. We base this discussion on actual working configurations that we built and tested in our labs.

Also featured in this Redpaper are several configuration examples of deploying the Layer 2-7 GbE Switch Module in Cisco and Nortel environments.

The team that wrote this Redpaper

This Redpaper was produced by a team of specialists from around the world working at the International Technical Support Organization, Raleigh Center.

Stephen Hochstetler is a Consulting I/T Specialist at the International Technical Support Organization, Austin Center. He writes extensively and teaches IBM classes worldwide on all areas of systems management and Linux clusters. Before joining the ITSO four years ago, Stephen worked in the Tivoli® Services Organization as a Network Management Specialist. He is a certified IBM IT Specialist and an ITIL Service Manager.

David Green is a Staff Engineer at IBM in Research Triangle Park, North Carolina. He works in BladeCenter ecosystem development. He worked on the development of the Layer 2-7 GbE Switch Module and the Optical Passthrough Module for BladeCenter. He has a Bachelor of Science degree in Information Systems from UNC-Greensboro. His areas of expertise include IBM eServer™ BladeCenter, Fibre Channel, SANS, and networking.

Espen Johanson is a IT Architect in Network & Connectivity Services, IBM Norway. His areas of expertise include WAN, LAN, Routing (RIP, OSPF, EIGRP, BGP, ISIS, and so forth), Bridging, DLSW, QoS, Content Delivery Networking, Optical Networking and Network Design. He is a Cisco Certified Internetworking Expert #9705 as well as a Cisco Certified Design Professional. Espen works mainly on high availability infrastructure solutions including network assessment, design, and implementation. Stephen Hochstetler

Norm Strole is a Senior Technical Staff Member in Research Triangle Park, North Carolina. He has 23 years of experience in the networking field and has been with IBM for 30 years. He holds a Ph.D. degree in Electrical Engineering from Duke University. Norm's areas of expertise include local area network systems and LAN switching. He has written extensively on token-ring networking and has published papers in the *IBM Journal of Research and Development* and the *IBM Systems Journal*.

Thanks to the following people for their contributions to this project:

Ishan Sehgal
IBM, RTP, NC

Ulises Fabre
Shailesh Naik
Nortel Networks, Santa Clara CA

Scott Lorditch
Nortel Networks

Mark Goodgion
IBM Web Hosting Services, Raleigh NC

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, Business Partners and/or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this Redpaper or other Redbooks™ in one of the following ways:

- ▶ Use the online **Contact us** review redbook form found at:

ibm.com/redbooks

- ▶ Send your comments in an Internet note to:

redbook@us.ibm.com

- ▶ Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. JN9B Building 003 Internal Zip 2834
11400 Burnet Road
Austin, Texas 78758-3493



Introduction to the IBM @server BladeCenter Layer 2-7 Gb Ethernet Switch Module

This chapter provides an introduction to the IBM @server BladeCenter Layer 2-7 GbE Switch Module (GbESM). It also gives a brief overview of important network concepts and the different methods for managing the switch.

1.1 Introduction to GbESM: Capabilities, features, functions

The IBM @server BladeCenter Layer 2-7 GbE Switch Module is a four port gigabit switch module that can be installed in the IBM @server BladeCenter Type 8677. The @server BladeCenter supports up to four GbESMs installed in the chassis at any one time. The @server BladeCenter ships with support for two GbESMs. An Ethernet daughter card option is available for the blades to support two additional GbESMs, for a total of four.

The GbESMs are hot-swappable subsystems that provide Ethernet switching capabilities within a BladeCenter chassis. The primary purpose of the switch module is to provide Ethernet connectivity among the processor blades, management modules, and the external network infrastructure.

The BladeCenter chassis supports a minimum of one switch installed in one of the slots in the rear of the chassis. The blades each ship with two 1-gigabit full-duplex links that correspond to slots 1 and 2 in the rear of the chassis. The IBM @server BladeCenter will function with a single switch installed. However, two independent switch modules installed in module slots 1 and 2 of the chassis are required for redundancy. With this configuration, each of the server blades in the BladeCenter chassis is then able to use either of its network interfaces.

The GbESM has 18 ports that can be configured by the user. Ports 1 through 14 on the switch module are internal gigabit ports that correspond to server blades 1 through 14, respectively. Ports 17 through 20 are external 10/100/1000 Mbps ethernet ports for connection to the external network infrastructure. These ports are identified as Ext1, Ext2, Ext3, and Ext4 in the switch configuration menus and are labeled 1 through 4 (from top to bottom) on the switch module. There are also two ports for dedicated use by the management modules. These ports, 15 and 16, are on a dedicated management VLAN (4095) and have a dedicated management interface (Interface 128) configured. These ports appear in the Web or telnet management interfaces to the GbESM. However, the ports, VLAN, and interface cannot be configured by a user. This prevents changes to the ports that could cause access to the management modules to fail. If this happened, the switch could no longer communicate with the management modules and the management modules would no longer be able to manage the switch. When the switch is first installed in the IBM @server BladeCenter, by default, the switch can *only* be managed through one of these interfaces. The external interfaces are disabled for security reasons. If the ports were enabled, a user who knows the default TCP/IP address and user name for the switch could access the switch and render it inoperable and unmanageable before the switch was configured.

Essential information such as the machine type and serial number are located on the identification label on the side of the GbESM. You will need this information when you register the Layer 2-7 GbE Switch Module with IBM. See Figure 1-1 on page 3 for the location of the identification label.

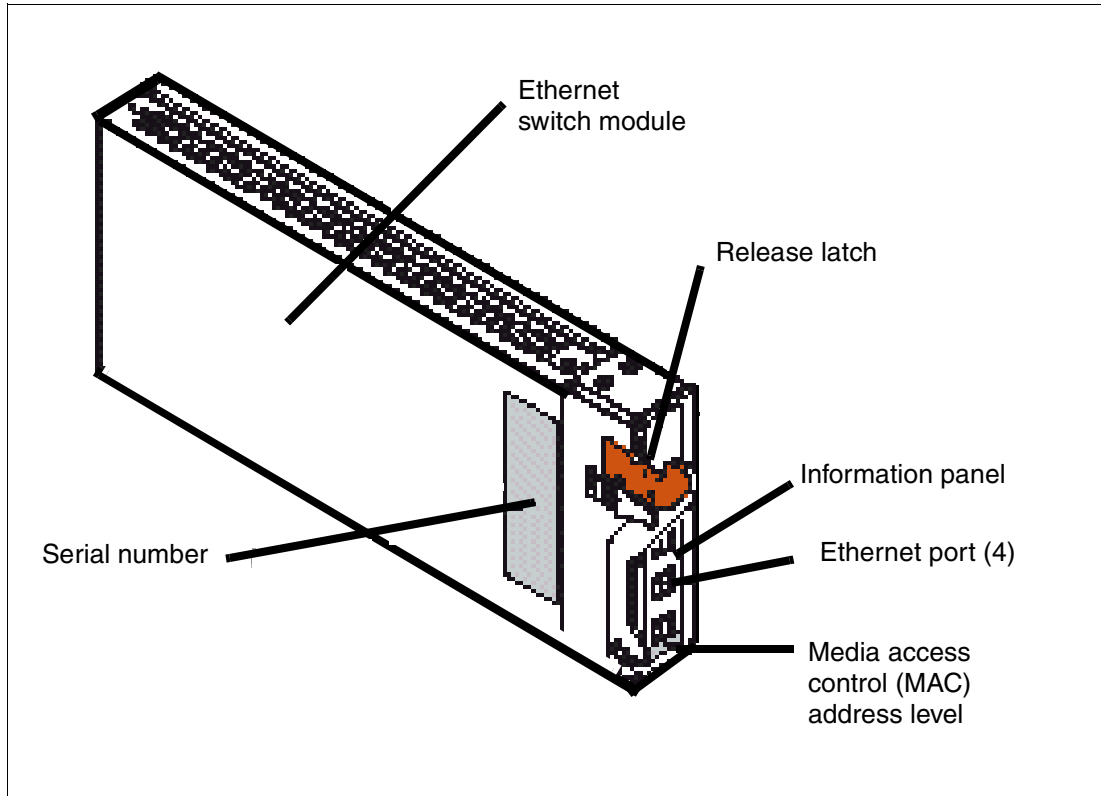


Figure 1-1 IBM BladeCenter Layer 2-7 GbE Switch Module

Note: The MAC address is also located on a separate label on the information panel under the external Ethernet port connectors.

The GbESM can be managed via telnet or a Web interface. Both the Web and telnet interfaces can be started by accessing the switch directly or starting a session from the management module's Web interface.

Note: The external ports on the switch must be enabled, and enabled for management as well, to manage the switch directly without going through the management module. Because the management interface is on VLAN 4095 and cannot be moved, a separate interface from the management interface must be created with an IP address valid on the production network to manage the switch directly.

The management interface on each switch by default is assigned a TCP/IP address that corresponds to the module slot the switch is installed in. Table 1-1 on page 4 lists the module slots and corresponding IP addresses. Figure 1-2 on page 4 shows the locations for the slots in the BladeCenter chassis. The default addresses can be changed to addresses on your management network. However, the management module IP addresses and the GbESM management interface addresses must all be on the same subnet. If the management network is a different network than the production network, a router is required to access the management network. As stated in the previous note, if you wish to manage the switch directly from the production network, an interface must be created with an IP address valid on the production network.

Table 1-1 Default IP address listing

Bay Number	TCP/IP Address
1	192.168.70.127
2	192.168.70.128
3	192.168.70.129
4	192.168.70.130

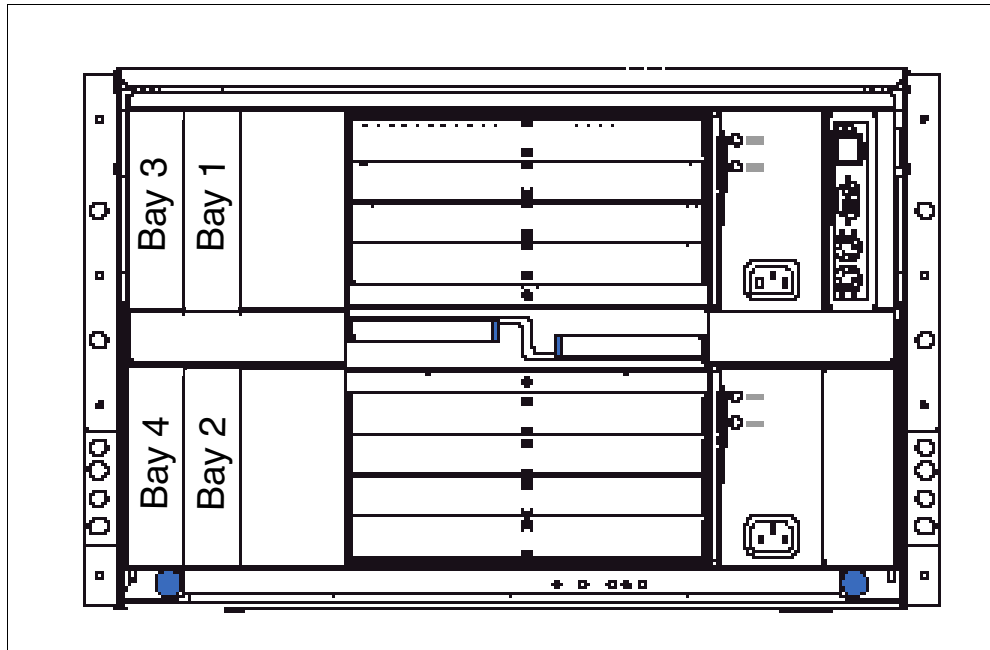


Figure 1-2 Bay locations rear of the IBM eServer BladeCenter

Support for a GbESM installed in bays 3 and 4 requires an Ethernet daughter card option to be installed in any blade that uses those switches. The GbESM also contains an information panel with status LEDs on the rear of the switch. The status LEDs include an OK light, error LED and link and activity lights for each of the external ports on the switch. Figure 1-3 on page 5 shows the locations of the LEDs on the information panel.

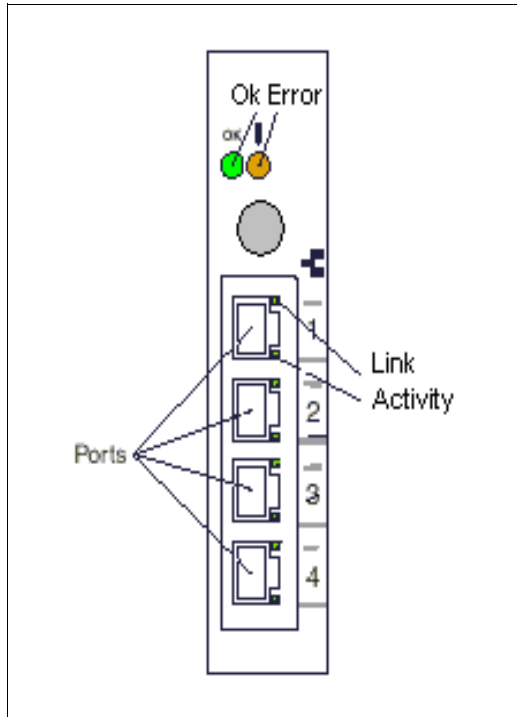


Figure 1-3 Information panel

OK (power-on): This green LED is located above the four external 10/100/1000 Mbps ports on the information panel. When this LED is on, it indicates that the switch module has passed the power-on self-test (POST) and is operational.

! (Ethernet switch error): This amber LED is located next to the OK (power-on) LED on the information panel. This LED indicates that the switch module has a fault. If the switch module fails to POST or power on, this fault LED will be lit.

Ethernet link: This green link status LED is located at the top of each external 10/100/1000 Mbps port. When this LED is lit on a port, it indicates that there is a connection (or link) to a device on that port.

Ethernet activity: This green activity LED is located at the bottom of each external 10/100/1000 Mbps port. When this LED blinks on a port, it indicates that data is being received or transmitted (that is, activity is occurring) on the port.

Depending on the application, the external Ethernet interfaces can be configured to meet a variety of requirements for bandwidth or function. The IBM @server BladeCenter Layer 2-7 GbE Switch Module has been pre-configured with default parameter settings that can be used with most typical installations. However, all Layer 2-7 GbE Switch Modules will need a few basic parameter settings initially, such as a TCP/IP address for management, security access and control parameters, and basic setup of the external ports for link aggregation.

This high performance switch is ideally suited for networking environments that require superior microprocessor performance, efficient memory management, flexibility, and reliable data storage. Performance, reliability, and expansion capabilities were key considerations in the design of the GbESM. These design features make it possible for you to customize the system hardware to meet your needs today, while providing flexible expansion capabilities for

the future. If you have access to the World Wide Web, you can obtain up-to-date information about the GbESM and other IBM server products on the following Web site:

<http://www.ibm.com/eserver/xseries/>

User guides, drivers, and firmware updates can all be found at this site.

1.2 Switch management and operating concepts

This section provides a brief overview of several of the features of the GbESM and how the switch is managed. This section also covers some of the concepts of networking. The features and concepts, as well as the management of the switch, are covered in greater detail later in this Redpaper.

1.2.1 Switch management and control

The switch supports two management user interfaces. A Web server is included on the switch. It is invoked through the management module (MM) Web interface control utility. The other primary interface is a telnet interface that can also be invoked through the management module control utility. As previously stated, by default the external ports on the switch are disabled. Initial access to the switch must be through the management module to enable the external ports and configure an interface so the switch can be accessed from the network. Once the network settings are configured, it is no longer necessary to go through the management module control interface. The switch can be accessed either through the management module control utility or directly from the network.

It is not necessary to use the MM browser interface to access the GbESM; any standard telnet client or browser can be used. In some cases telnet may be deliberately disabled to force the use of SSH. In order to use the telnet option in the MM browser interface, the station where the browser is installed must have a Java Runtime Environment (JRE) installed.

After a TCP/IP address has been assigned to the GbESM, you can perform many different management and control tasks. These tasks fall in the following categories:

- ▶ Configuration of switch parameters
 - Switch TCP/IP address
 - Default gateway
 - General switch information: switch location, contact, system name
- ▶ Remote management setup
- ▶ Network monitoring
 - SNMP and traps
 - View port statistics
 - Monitor data traffic
- ▶ Switch maintenance

More information on these tasks and specific instructions on configuring the switch are given later in this Redpaper.

1.2.2 Switch operating concepts

This section is a brief introduction to several of the concepts necessary to understanding how the switch functions. A more in-depth explanation of each of the concepts listed follows later in this Redpaper.

Packet forwarding

The switch stores mapping information from destination MAC addresses to the destination port in a forwarding table. This information is then used to forward packets to specific ports. This capability reduces network congestion and frees the switch from having to forward traffic to all ports.

Spanning Tree Protocol

Spanning Tree Protocol provides a mechanism to block links between switches that form loops within the network. This could be caused by an incorrectly cabled network or a desire by the network designer to have redundant links between switches. If multiple links between switches are detected, Spanning Tree configures one link as the primary and blocks the secondary. In the event of the primary link failing, the secondary link is activated automatically.

VLANs

A virtual local area network (VLAN) is a logical network topology configured on the physical network layout. A VLAN can be used to combine any collection of blade servers in a chassis into a logical network segment. VLANs typically correspond to TCP/IP subnets on the network. VLANs also divide a network into broadcast domains. Any broadcast traffic generated on a VLAN will be kept on that VLAN. This keeps broadcast traffic off the network and, depending on the type of broadcast traffic, within the BladeCenter chassis. VLAN concepts, which are explained in greater detail later in this paper, include IEEE 802.1Q VLAN, packet forwarding as it pertains to VLANs, ingress and egress ports, and VLAN tags.

Tagging and untagging

Another important concept is tagging and untagging of network packets. Tagging is the basis of 802.1Q-compliant VLANs. Every port on the GbESM can be configured as a tagged or untagged port. Each packet that exits the switch through a tagged port has some VLAN information inserted into the header, if that packet was not previously tagged. The tag information can then be used by other 802.1Q-compliant devices to make packet forwarding decisions. The VLAN information identifies the VLAN of the packet. If a port receives a packet and is not on the VLAN contained in the packet header, the switch drops the packet.

Older devices or desktop systems on a network may be tag-unaware and do not conform to the 802.1Q specification. The GbESM uses the untagged option for any port that is connected to one of these devices. Any packet leaving an untagged port does not have any tagging information inserted into it. In addition, any tags already in a packet leaving an untagged port are stripped from the packet prior to the switch forwarding the packet on.

In the next section we provide more detailed technical information on the switch and what technologies it supports.

1.3 Ports and performance features

This section lists the specifications for the ports on the switch. It also lists the performance and operational features of the GbESM.

- ▶ Ports
 - Four external copper ports for making 10/100/1000 Mbps connections to backbone, end stations, and servers.
 - Fourteen internal full-duplex gigabit ports, one connected to each of the BladeCenter blade servers.
 - Two internal full-duplex 10/100 Mbps ports for connection to the management modules. One port connects to each management module.
 - The ports can be configured for autosensing and can utilize either straight-through or crossover cables for switch-to-switch connections. The switch will detect the type of cable used and set the port accordingly. However, a cross-over cable is recommended because this cable will work with all of the modes.
- ▶ Performance and operational features of the GbESM
 - Transmission method: Store-and-forward.
 - Random-access memory (RAM) buffer: 8 MB.
 - Media access control (MAC) address learning: Automatic update, supports 28K MAC addresses.
 - Priority queues: Four priority queues per port.
 - Forwarding table age time: maximum age is 17 to 2100 seconds, default is 300 seconds.
 - 802.1D Spanning Tree support. Can be disabled on the entire switch or on a per-port basis.
 - 802.1Q Tagged virtual local area network (VLAN) support.
 - Support for 256 VLANs in total, including 128 static VLANs.
 - Link aggregation on four external ports for up to two static trunk groups or two link aggregation control protocol (LACP) 802.3ad link aggregation groups.

1.4 Switch and network management

The GbESM supports the following network management protocols and standards. Some of these are technologies that can be used to manage and monitor the switch. Others, such as Spanning Tree, are technologies the switch uses to manage the network. The network technologies listed are covered in more detail later in this paper.

- ▶ Switch monitoring and management
 - Simple network management protocol (SNMP) version 1.
 - Fully configurable either in-band or out-of-band control through SNMP-based software.
 - Flash memory for software upgrades. This can be done through trivial file transfer protocol (TFTP) which can be started through a telnet session or HTTP session.
 - Supports password-enabled Web-based management and a telnet remote console.

- Built-in SNMP management:
 - Bridge management information base (MIB) (RFC 1493)
 - MIB-II (RFC 1213)
- ▶ Network management
 - 802.1P/Q MIB (RFC 2674)
 - Interface MIB (RFC 2233)
 - Mini-RMON MIB (RFC 1757) - four groups. The remote monitoring (RMON) specification defines the counters for the receive functions only. However, the switch provides counters for both receive and transmit functions.
 - Spanning Tree Protocol (STP) for creation of alternative backup paths and prevention of network loops.
 - TFTP support
 - Bootstrap protocol (BOOTP) support
 - Dynamic host configuration protocol (DHCP) client support

1.5 Network cables

The following cables and cable lengths are supported by the GbESM:

- ▶ 10BASE-T:
 - UTP Category 3, 4, 5 (100 meters maximum)
 - 100-ohm STP (100 meters maximum)
- ▶ 100BASE-TX:
 - UTP Category 5 (100 meters maximum)
 - EIA/TIA-568 100-ohm STP (100 meters maximum)
- ▶ 1000BASE-T:
 - UTP Category 5e (100 meters maximum)
 - EIA/TIA-568B 100-ohm STP (100 meters maximum)

1.6 Supported network standards

The following standards are supported by the GbESM. Some of these standards are explained in greater detail later in this document.

- ▶ IEEE 802.3 10BASE-T Ethernet
- ▶ IEEE 802.3u 100BASE-TX Fast Ethernet
- ▶ IEEE 802.3z Gigabit Ethernet
- ▶ IEEE 802.1Q Tagged VLAN
- ▶ IEEE 802.3ab 1000BASE-T
- ▶ IEEE 802.3x Full-duplex Flow Control
- ▶ ANSI/IEEE 802.3 NWay auto-negotiation



GbESM management and initial configuration

This chapter provides an overview of how to manage the Layer 2-7 GbE Switch Module. It introduces the different ways to establish a management session to the GbESM, including what type of physical cabling to use, and documents some of the initial switch configuration options. The supported firmware on the GbESM used while writing this Redpaper was v20.0. The management module firmware version was 53c. To determine the most up-to-date versions available, open a Web session to the management module, then click **Monitors** → **Firmware VPD** to get the current firmware levels on all devices installed in the BladeCenter chassis.

2.1 GbESM management through the management module

The BladeCenter management module has one external port you can use for management. You can only manage the management module via this external port. Initially, the four external Ethernet ports on the GbESM are disabled. Therefore, you must configure the switch via the external port on the management module.

At the time of the writing of this paper, only one management module was supported in the BladeCenter chassis. The support for a second redundant management module is planned for the future. The GbESM has one internal 100 Mbps Ethernet connection for each management module. On the management module, this internal connection is labeled Eth1. The internal connection provides the Ethernet connectivity between the management module and the switch modules for management purposes. All the modules in the BladeCenter chassis (up to four) share Eth1 on the management module. In order to have IP connectivity between the management module and the GbESM, configure the management module external interface Eth0 and the internal interface Eth1, with IP addresses in the same subnet as the GbESM.

Note: By default, the four external Ethernet ports on the GbESM are disabled. Therefore, you must initially use the 10/100 Mbps Ethernet port on the management module to configure the switch.

We recommend connecting the external 10/100 Mbps Ethernet port on the management module to a network dedicated to management (shown in Figure 2-1). We also recommend accessing the Ethernet switch modules via the 10/100 Mbps port on the management module. Ideally, the management module, management station, switch modules, and DHCP server should connect to the same management subnet.

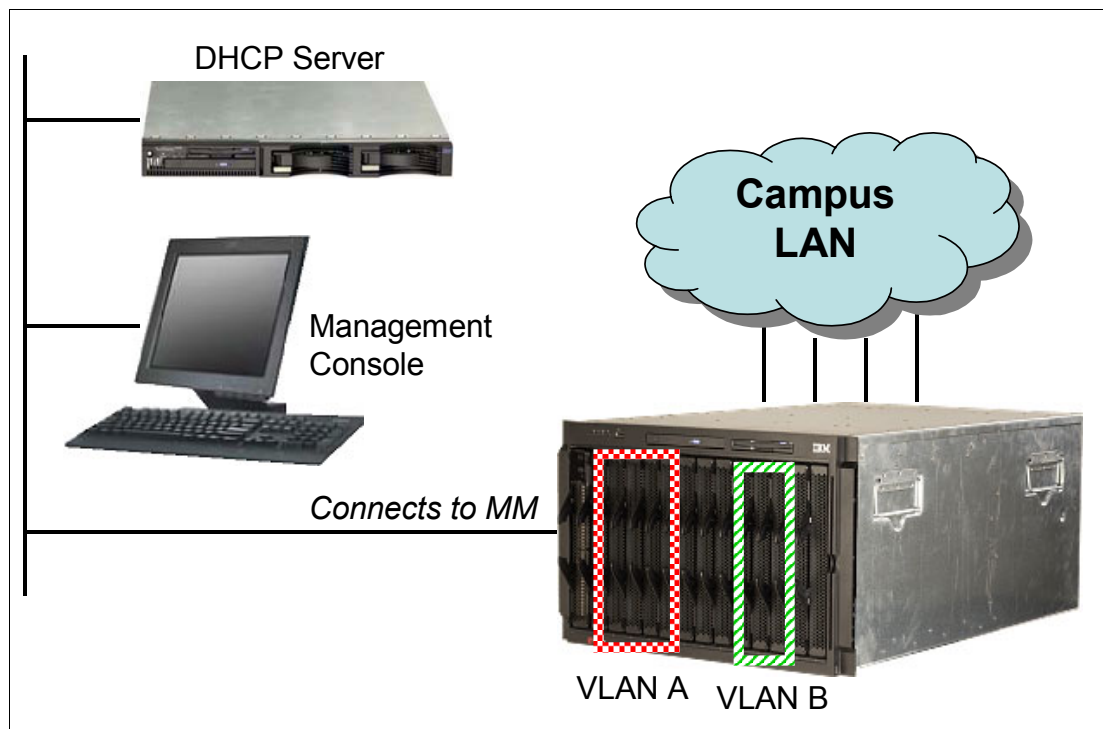


Figure 2-1 A typical management network

There are two primary reasons for this design:

- ▶ Separating the production network and the management network improves data security.
- ▶ If a problem occurs on the production network, the management module and switch modules remain accessible.

Note: Configuration changes are not effective (do not change the operation of the switch) until the “Apply” command is issued. This can be done from the command line or by pressing the Apply button near the top of the browser display. It is not necessary to issue the “Save” command to implement configuration changes. It is, in fact, good practice to *not* issue the save command until one is sure that the changes are functioning as expected. If changes cause an unanticipated problem, the “revert flash” command will back them out, as will a re-boot of the switch. Once the changes are Saved, it is necessary to manually back them out.

2.1.1 Establishing a physical connection to the management module

The only way to access the management module is through the external 10/100 Mbps Ethernet port on the front of the module. To establish the physical connection to the management module use one of the following methods:

- ▶ Use a Category 3, 4, 5, or higher unshielded twisted pair (UTP) straight through cable to connect the Ethernet port on the management module to a switch in a network that has an accessible management station.
- ▶ Use a Category 3, 4, 5, or higher cross-over cable to connect a management station (PC, laptop, and so forth) directly to the external Ethernet port of the management module.

Note: The 10/100 external Ethernet port on the management module uses a media dependent interface (MDI). This standard describes the interface for unshielded twisted pair (UTP) cable. For two devices to communicate, the transmit wires of one device must connect to the receive wires of the other device. You can accomplish this by using a cross-over cable or a port that implements the cross-over function internally (MDI-X port).

2.1.2 Using the management module Web interface to initially configure GbESM

Once you establish the physical connection to the management module, configure the management station with an available IP address in the same subnet as the management module. By default, the subnet is 192.168.70.0/24.

You have two primary methods to manage the management module:

- ▶ HTTP Web interface
- ▶ IBM Director

We used the management module Web interface to demonstrate the initial switch configuration.

Note: The Web browser you use must support Java, JavaScript 1.2, and a minimum screen resolution of 800x600 with 256 colors.

Follow these steps to establish a management session with the management module, and to configure the initial recommended switch settings:

1. Open a Web browser and connect to the management module using the configured IP address. The default IP address for the management module external interface is 192.168.70.125.

Note: The default IP address for the internal interface is 192.168.70.126.

2. Enter the userid and password. The defaults are USERID and PASSWORD (case-sensitive with a zero in the place of the letter O). Click **OK**.
3. At the initial screen, click **Continue** to access the management session.

Configuring the GbESM IP address

When you first install a GbESM in the BladeCenter unit, you must use the management module to configure the initial IP address. The management interface on the GbESM does not support the DHCP or BOOTP protocols. Once an IP address is assigned to the GbESM, you can use a GbESM management session to configure the other settings on the switch. By default, the GbESM IP address information is set as described in Table 1-1 on page 4. If you have changed the default management module IP address information to match your network you must also configure the GbESM IP address information to an address on the same subnet. Once you configure the GbESM module IP address, you can initiate a management session to the GbESM through the management module. To configure an IP address, subnet mask, and default gateway on a GbESM follow these steps:

1. From the management module menu, select **Switch Tasks** → **Management**. You will see a window similar to Figure 2-2.
2. Select **Switch Module 1**.

Note: The GbESMs are shipped with factory IP addresses. The factory IP addresses for the GbESMs in slot 1 and slot 2 are 10.90.90.91 and 10.90.90.92, respectively. These IP addresses are different from the default IP address that gets assigned by the management module. The IP address configuration pushed to the switch by the management module will take precedence.

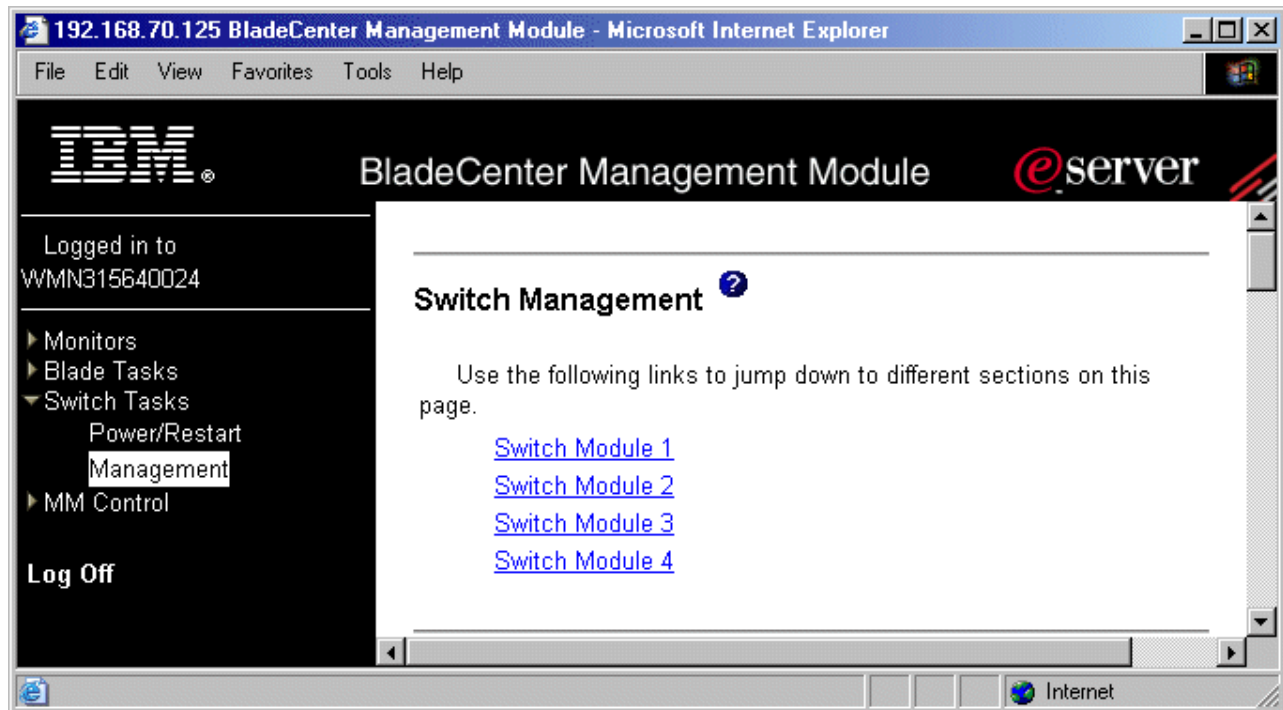


Figure 2-2 Switch Management

3. As shown in Figure 2-3 on page 16, complete the IP Address, Subnet Mask and Default Gateway fields. As stated previously, the management module will automatically assign IP address information based on the slot the GbESM is installed in.

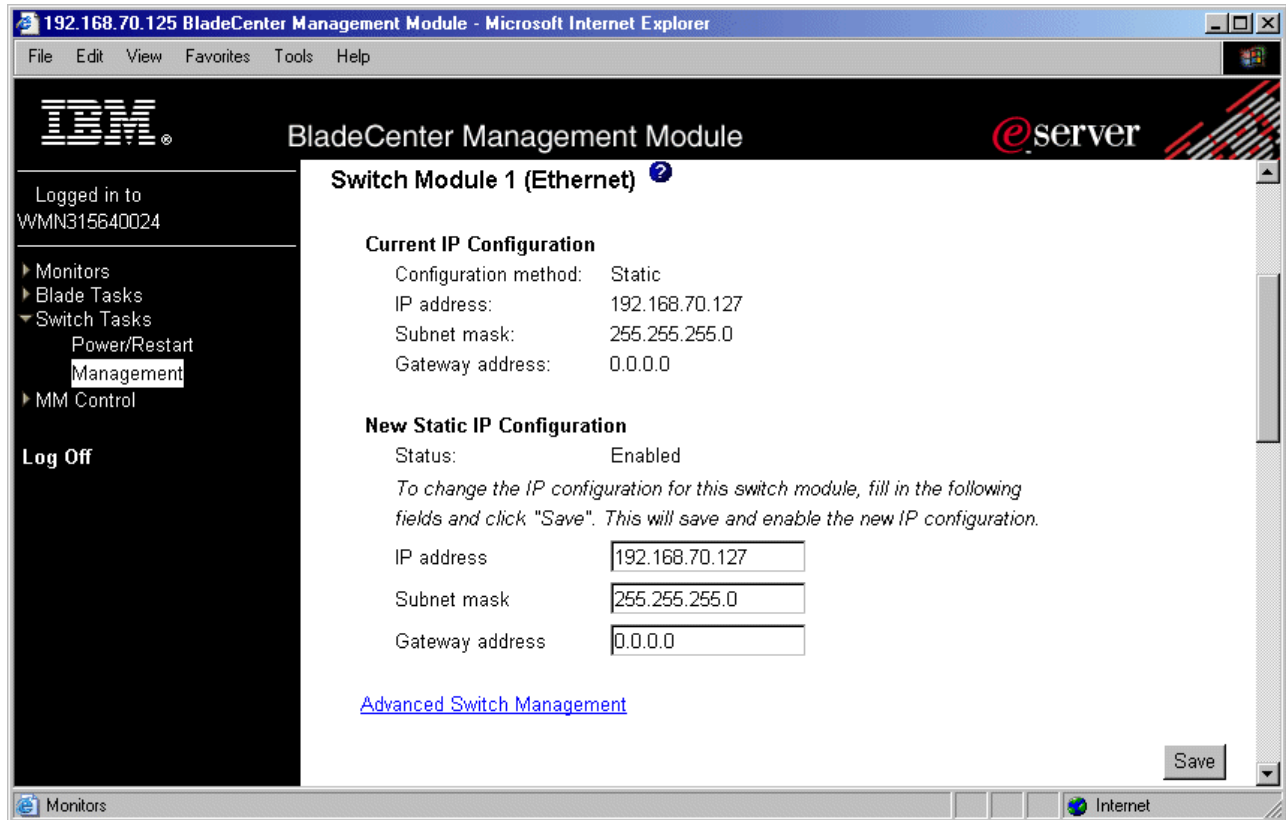


Figure 2-3 Switch module IP configuration

4. Click **Save**. The Current IP Configuration section of the panel displays the new IP address.
5. Click **Advanced Management** and select **Enabled** from the drop-down box next to "Preserve new IP configuration on all resets." Then click **Save**. Enabling this setting ensures that if the switch needs to be reset back to factory defaults, the management module will reassign the current IP address of the switch back to it.
6. To verify the IP connectivity between the management module and the GbESM, click **Advanced Switch Management** (Figure 2-4 on page 17). Then click **Send Ping Requests**. Click **Cancel** to return to the previous menu.

Enabling external Ethernet ports

By default, the four external ports on the GbESM are disabled. Enable the external ports on the GbESM to allow Ethernet connectivity to the external infrastructure. To enable the external ports, follow these steps:

1. From the Advanced Switch Management panel, click **Advanced Setup**.
2. Select **Enabled** from the list box for the option to enable "External ports" (Figure 2-4 on page 17).

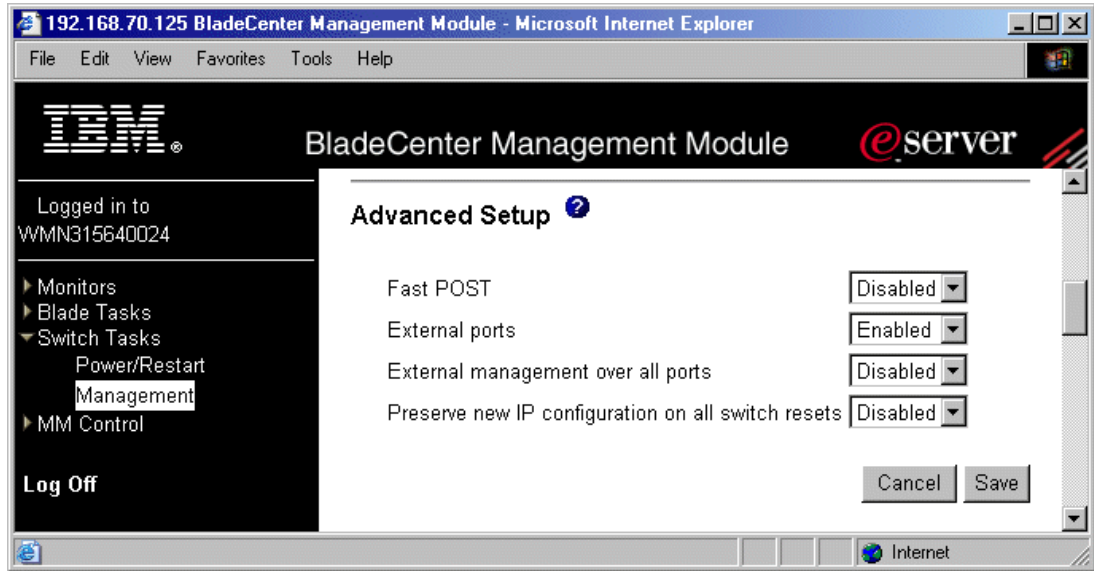


Figure 2-4 Switch Advanced Setup

3. Click **Save**.

Important: The IBM @server BladeCenter ships with the external GbESM ports disabled. You must use the management module to enable them.

Enabling external management

We recommend that you manage the GbESMs through the management module; however, if you want in-band management access to the four external ports on the GbESM or if you configure the GbESM to send SNMP traps, then you should enable external management over all external ports. If external management is not enabled, you must establish a GbESM management session through the management module.

1. From the Advanced Setup window (Figure 2-4), select **Enabled** from the “External management over all ports” list box.
2. Click **Save**.

Enabling option preserve IP address

In order to ensure that the GbESM IP address configured in the management module is saved even after a switch factory reset, you need to enable this option.

1. From the Advanced Setup window, select **Enabled** from the “Preserve new IP configuration on all switch resets” list box.
2. Click **Save**.

Note: Repeat the GbESM configuration process for switch module two if a second switch is installed.

Configuring the switch IP address

You can change the GbESM IP address on the management interface using either a telnet or a Web session directly connected to the GbESM, but you should not do so. If you change the IP address of the switch outside the management module Web interface, you will lose your connection to the GbESM and you must reconnect to the switch after the IP address change.

In addition, the management module must update its NVRAM before it can reconnect to the switch. The management module performs this update at set intervals, so it can take up to ten minutes. In order to avoid this situation, you should change the IP address of the switch using a management session to the management module. This immediately updates the management module NVRAM and you can reconnect to the GbESM Web session using the new IP address. There are no restrictions on updating the IP addresses of interfaces created on the public VLAN.

You can create additional IP interfaces on the switch that are accessible from your public network. The only restriction is each interface must be on a different subnet. Also, unlike the management interface 128, interface 1 will support BOOTP for setting its IP address information.

Resetting to factory defaults

If needed, the GbESM can also be set to factory defaults in two different ways. The preferred method is to use the MM Web interface to reset the switch. Figure 2-5 shows the Restore Factory Defaults screen. To reset the switch to factory defaults using the MM:

1. Open a session of the MM in the Web interface.
2. Click **I/O Module Tasks** → **Management** in the left frame.
3. Click the bay that has the switch you wish to reset.
4. Click **Advanced Management**.
5. Click **Restore Factory Defaults**.
6. Click the **Restore Defaults** button. This will restore the defaults to the switch and reset it.

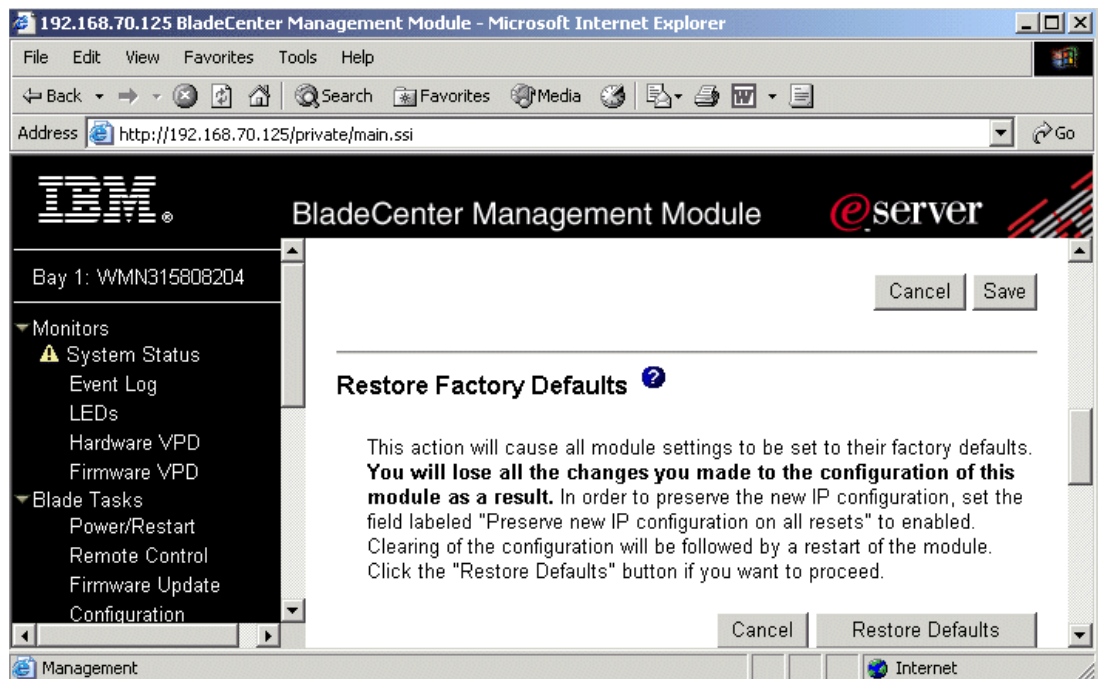
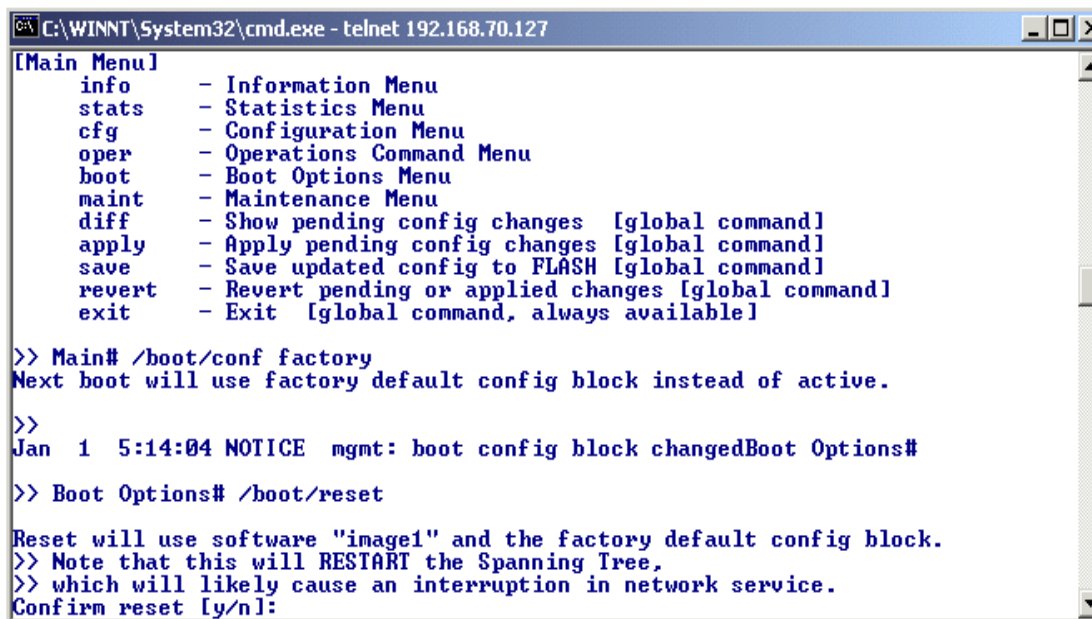


Figure 2-5 Restore factory defaults

Note: When using either the Management Module interface or the CLI to reset the switch, if the option to “Preserve IP configuration on all resets” is not enabled under advanced management, the switch will lose the IP address settings on the management interface and must be reconfigured via the management module.

The switch can also be reset to factory defaults using the CLI. Figure 2-6 shows the commands necessary. To reset the switch using the CLI:

1. Open a telnet session to the switch and log in.
2. Enter **/boot/conf factory** at the prompt to make the factory configuration the active configuration.
3. Enter **/boot/reset** to reset the switch and load the factory configuration.
4. Answer the prompt to confirm the reset. Since the switch is being set to factory defaults, if you have a telnet session open to an interface with an IP address other than the default, you will lose connectivity to the switch.



```
C:\WINNT\System32\cmd.exe - telnet 192.168.70.127
[Main Menu]
  info      - Information Menu
  stats     - Statistics Menu
  cfg       - Configuration Menu
  oper      - Operations Command Menu
  boot      - Boot Options Menu
  maint     - Maintenance Menu
  diff      - Show pending config changes [global command]
  apply     - Apply pending config changes [global command]
  save      - Save updated config to FLASH [global command]
  revert    - Revert pending or applied changes [global command]
  exit      - Exit [global command, always available]

>> Main# /boot/conf factory
Next boot will use factory default config block instead of active.
>>
Jan 1 5:14:04 NOTICE mgmt: boot config block changedBoot Options#
>> Boot Options# /boot/reset

Reset will use software "image1" and the factory default config block.
>> Note that this will RESTART the Spanning Tree,
>> which will likely cause an interruption in network service.
Confirm reset [y/n]:
```

Figure 2-6 Restore factory defaults: Command line

2.2 GbESM management using the GbESM Web interface

After you complete the initial configuration for the GbESM through the management module, establish a management session to the GbESM. Use the GbESM management interface for all of the other switch configuration. The following chapters of this paper document some of the advanced configuration options.

You can manage the GbESM by using any one of the following methods:

- ▶ BladeCenter management module Web utility
- ▶ GbESM Web session, used interchangeably with Browser-Based Interface or BBI
- ▶ GbESM command line telnet session or CLI

The management module Web utility is used to initially configure the GbESM or to launch a Web or telnet management session to the GbESM. We next use a GbESM Web session to demonstrate some additional basic switch configuration examples.

Note: The CLI for the GbESM is more powerful and flexible than the Web-based configuration tool. Basic configuration of the switch is covered in both the BBI and CLI. However more advanced items such as uploading a new OS image are only covered using the CLI. In addition, the how-tos in Chapters 5 and 6 will be covered using the CLI.

2.2.1 Basic GbESM configuration examples

This section provides examples of some of the basic recommended switch configurations. Use the following steps to establish a management session to the GbESM:

1. From the management module Web utility select **Switch Tasks** → **Management**.
2. Select the appropriate GbESM.
3. Click **Advanced Switch Management**.
4. Click **Start Telnet/Web Session**. You will see a window similar to Figure 2-7.

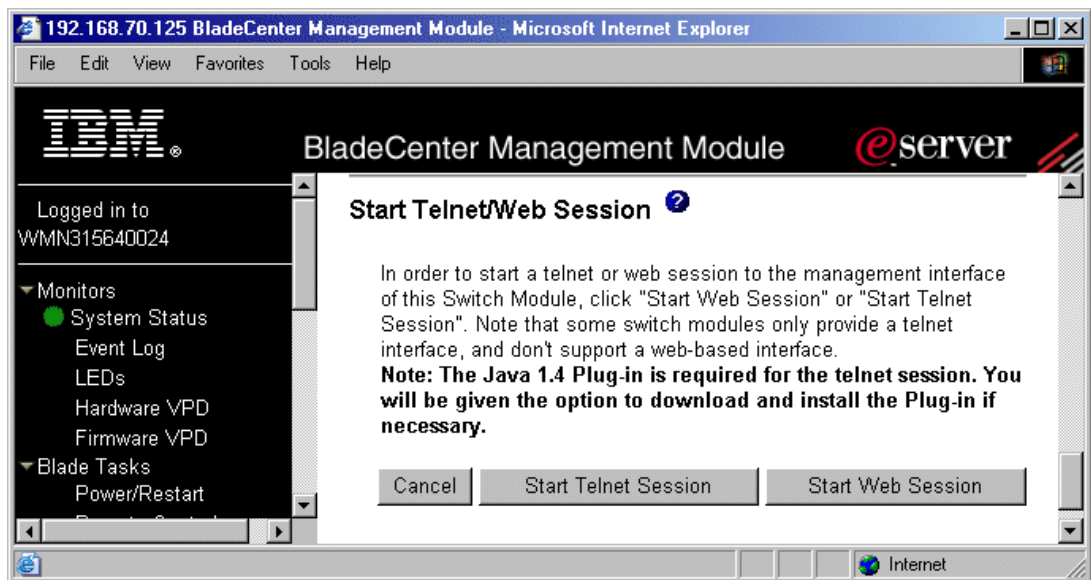


Figure 2-7 Starting a Web session

5. Click **Start Web Session** to open the Browser-based Interface (or BBI) for the GbESM.
6. Enter the userid and password for the GbESM. The defaults are admin and admin (and they are case-sensitive). Click **OK**.

Note: We recommend that you manage the GbESM through the Ethernet port of the management module; however, if desired, you can connect directly to the GbESM if you enable external management over all ports via the management module Web interface. You must also create an interface on the default VLAN to manage the switch directly. You can then open a Web browser and point the browser to the IP address of the GbESM to access the GbESM Web interface directly. See the *Alteon OS 20.0 Application Guide, Part 215654-A* (Nortel) for further information.

During the writing of the redpaper, we discovered some non-standard conventions used in the BBI that may confuse some readers, so we offer the following tips for using it.

Tip: Here are some tips for using the BBI.

- ▶ The folder icons in the tree in the left frame all expand, but there is no '+' beside them to indicate that they do so, nor is there a '-' to indicate that the folder trees are expanded.
- ▶ To expand a folder in the tree, click on the folder, not the words next to it. If the caption next to the folder is a hyperlink, it opens its own page, it does not expand the folder tree.
- ▶ Clicking on words that are not hyperlinks has no effect. You must click on a folder, an icon (at lower levels of the tree) or words that are a hyperlink.

The three buttons across the top frame - **Configure**, **Statistics**, and **Dashboard** set the context of the page displayed in the main frame. For example, to configure the switch IP address which is found under **Switch** → **General**, click **Configure** to set the context, then click **Switch** → **General**. If you are already at a screen you wish to configure but the Configure context has not been set, you must set the Configure context, then reload that screen.

2.2.2 Configuring the GbESM via the Browser-Based Interface

We now take a brief look at the Browser-Based Interface (BBI) on the GbESM. Everything that can be done here can also be done in the CLI. More emphasis will be placed on configuring the switch using the CLI rather than using the BBI.

The Switch Information panel displays the MAC address of the switch, as well as the firmware and hardware versions. Use the following steps to configure the system and contact information:

1. From the GbESM Web interface, click the folder icon next to **Nortel Networks Layer 2-7 Gbe Switch** in the left-hand frame.
2. Click the folder icon next to **Switch** in the left-hand frame.
3. Click the **CONFIGURE** button at the top of the page.
4. Click the icon next to **General** in the drop-down list under **Switch**. On a window similar to Figure 2-8 on page 22, you will see options such as IP Address and Network Mask fields that can be configured on this page. Other options on this page include date/time settings, syslog settings (if you have a syslog server), and SNMP settings.

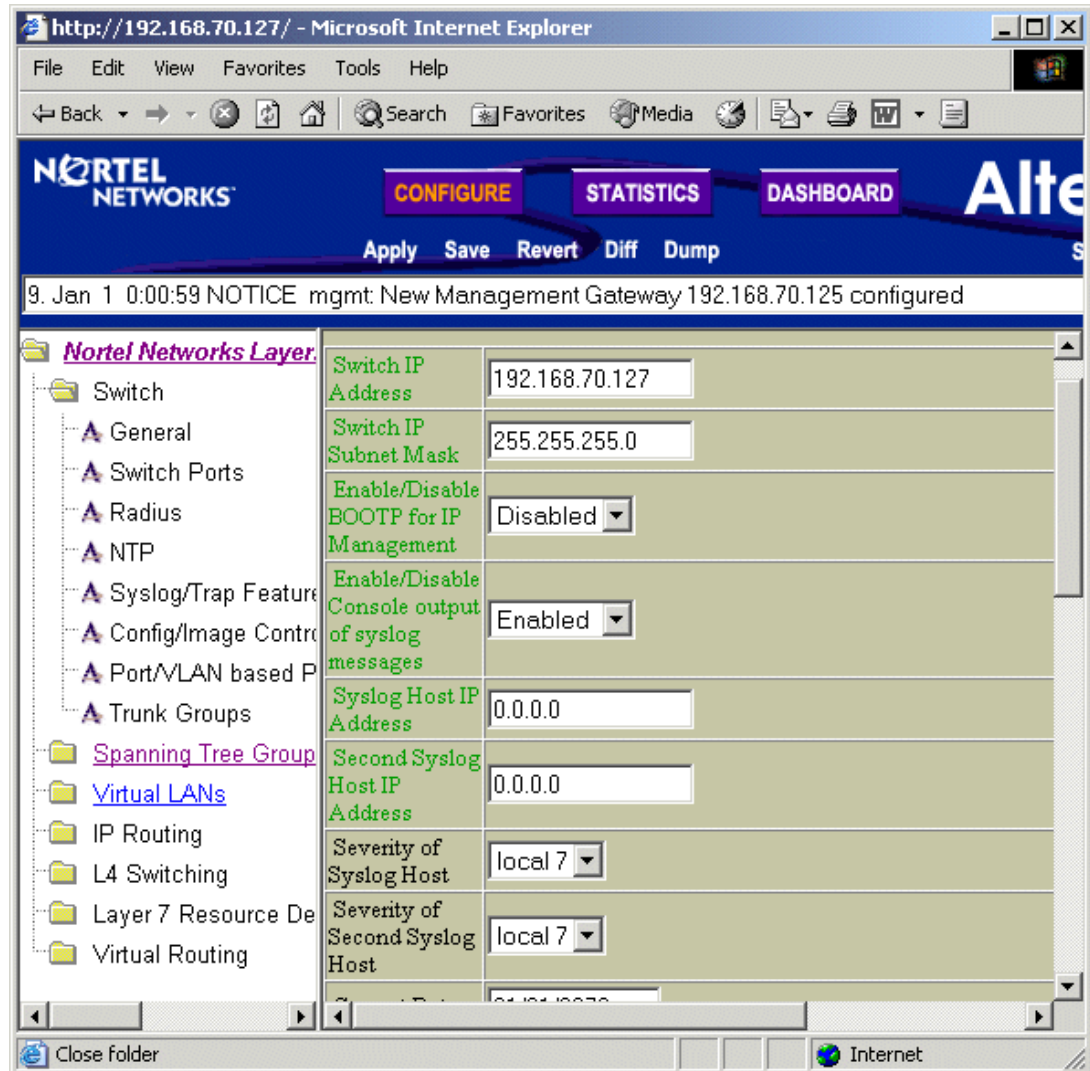


Figure 2-8 Switch information

5. Browse through some of the other links in the left-hand frame to get more familiar with where the configuration options for the switch are located.
6. If you have made any changes to the switch and wish to save them, click **Apply** to apply the changes to the current running config.
7. Click **Save** to save the changes to NVRAM.

2.2.3 Configuring the switch via the CLI

The CLI is much more flexible for configuring the switch than the BBI. It is scriptable, requires less overhead to run, and since it is a telnet session, it can be run from any OS, graphical or not. The main screen is shown in Figure 2-9 on page 23. The **stats** menu gives statistics about the switch. The **cfg** menu contains all of the configuration options for the switch. The **oper** menu contains all of the operational commands. Some of these commands can change the state of the switch, but these changes only apply until the next reboot and are not permanent. The **boot** menu contains the commands to control the booting of the switch - which image to boot from, which config to boot, and the **gting** and **ptimg** commands for getting and putting firmware files to the switch. The **maint** menu contains all of the commands

for maintenance of the switch. The commands to manipulate the arp cache and forwarding database are here, as well as the commands to obtain dumps of the current state of the switch for tech support. The rest of the options on the main menu—**diff**, **apply**, **save**, **revert**, and **exit**—are all global commands that will work anywhere on the switch. You can see from Figure 2-9 what each of the commands will do. The **help** command is also global and lists all the global commands.

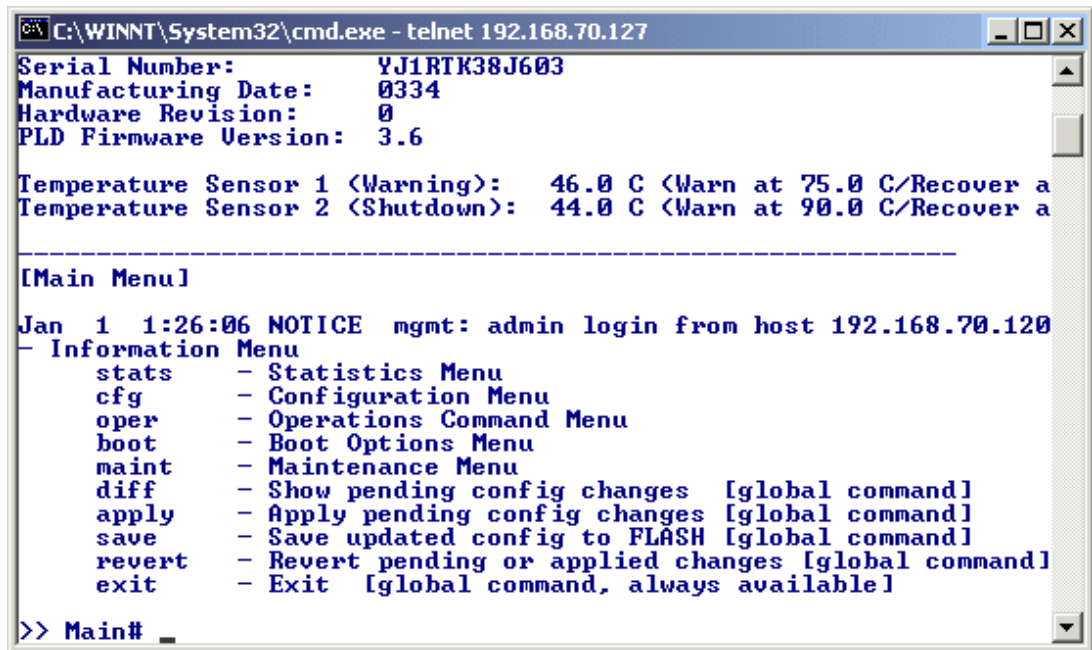


Figure 2-9 CLI Main Menu

To help you get familiar with using the CLI, we describe how to configure and enable an IP interface on the public network, we configure Link Aggregation (Trunking), and we also update the firmware on the switch. As you get more involved with managing the GbESM, an excellent reference for using the CLI is the *Alteon OS 20.0 Application Guide*, Part 215654-A (Nortel). It contains an in-depth listing of each of the commands in the CLI and their function.

However, before we begin, here are a few notes on using the CLI. Commands are entered in the format of a directory listing. For example, to find the current state of port EXT4 on the switch, you would enter:

```
/oper/port EXT4/cur
```

where the port menu is a sub-menu of the oper menu, and EXT 4 is a port on the port menu. Each port will have a sub-menu in turn. cur is an option on that port's menu. To find out the current IP settings for each interface configured on the switch, you would enter:

```
/info/13/ip
```

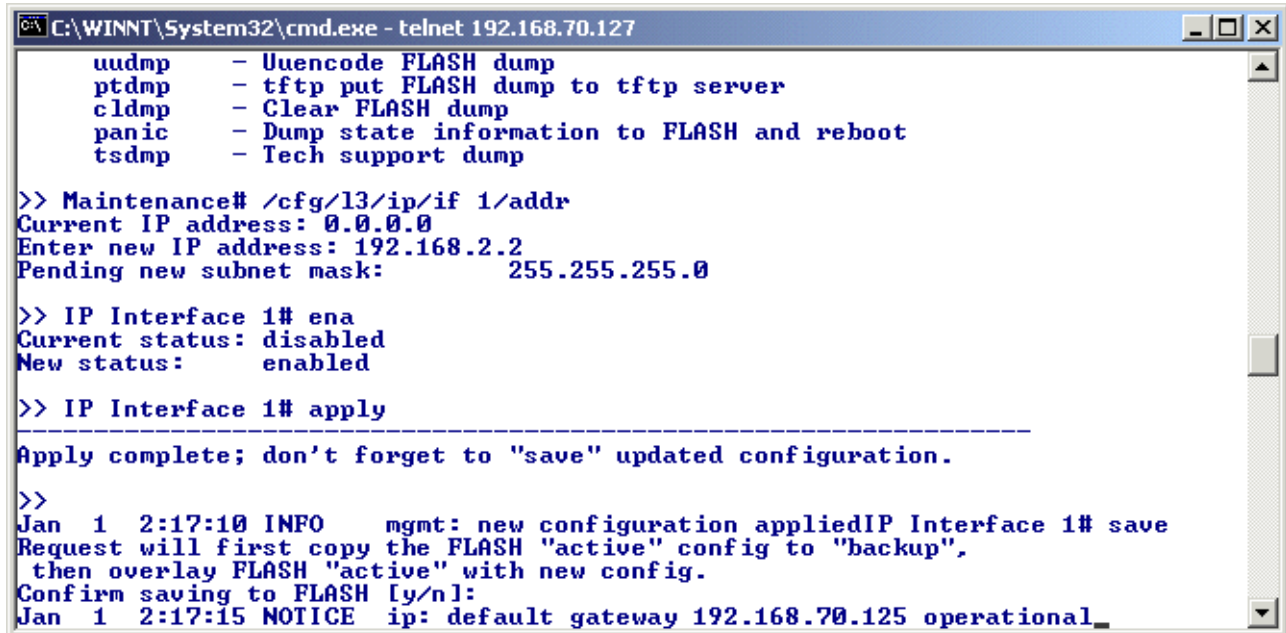
Commands can also be abbreviated, so long as multiple choices on any one menu do not fit the same abbreviation. For example, the command `/info/13/ip` could be shortened to `/info/13/i` and it yields the same result either way. However, the command `/c/13/i` generates an ambiguous command error because **ip** and **if** are both valid commands on the `/config/13` menu. Experiment to find out which commands can be shortened and which cannot. Also, when you save a config from the switch, it gets saved as a list of the commands to recreate that config in their abbreviated form. An excellent way to learn more about the commands is to save a config and look at the file. For the examples in this Redpaper we use

the fully qualified commands. At the end of the examples, Example 2-1 on page 27 lists the configuration file that results from the commands you enter.

Configuring an IP interface

The first example is configuring and enabling an IP interface on the GbESM. The GbESM supports up to 128 interfaces, with 128 already being used as the management interface. Our example will use interface 1. Figure 2-10 shows configuration of the interface. For simplicity, the figure shows the command entered all on one line. However, the example has been broken into several different steps so that you can see each menu and better understand how the CLI functions.

1. Enter `/cfg/13/ip/if` to start the interface configuration.
2. Enter `1` as the interface you wish to configure.
3. Enter `addr` to configure the IP address of the interface and enter the interface IP Address. The subnet mask is auto filled based on the class of the IP address you enter. If you wish to change it, go to step 4; otherwise, skip to step 5.
4. Enter `mask` to configure the subnet mask of the interface and enter the subnet mask.
5. Enter `vlan` to put the interface on another VLAN other than the default VLAN 1. If you do not wish to do so, skip this step.
6. Enter `ena` to enable the interface.
7. Enter `apply` to apply the changes and activate the new interface.
8. Enter `save` to save the changes to NVRAM and answer `y` to overwrite the active config with the new config. The interface is now active.



```
C:\WINNT\System32\cmd.exe - telnet 192.168.70.127
uudmp - Uuencode FLASH dump
ptdmp - tftp put FLASH dump to tftp server
cldmp - Clear FLASH dump
panic - Dump state information to FLASH and reboot
tsdmp - Tech support dump

>> Maintenance# /cfg/13/ip/if 1/addr
Current IP address: 0.0.0.0
Enter new IP address: 192.168.2.2
Pending new subnet mask: 255.255.255.0

>> IP Interface 1# ena
Current status: disabled
New status: enabled

>> IP Interface 1# apply
-----
Apply complete; don't forget to "save" updated configuration.

>>
Jan 1 2:17:10 INFO mgmt: new configuration appliedIP Interface 1# save
Request will first copy the FLASH "active" config to "backup",
then overlay FLASH "active" with new config.
Confirm saving to FLASH [y/n]:
Jan 1 2:17:15 NOTICE ip: default gateway 192.168.70.125 operational_
```

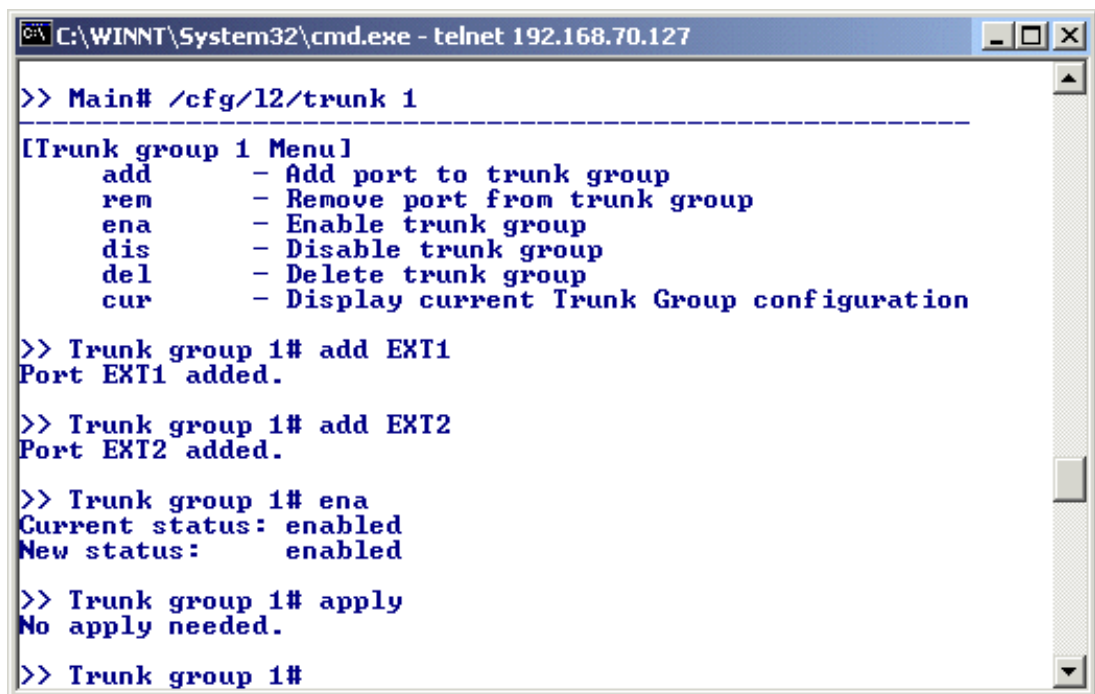
Figure 2-10 Setting IP address

Configuring link aggregation

The GbESM can also perform link aggregation or trunking. This feature enables the switch to treat the external physical links (up to 4) as a single virtual link. Inherent in the trunking is load balancing and fault tolerance. If one of the links in the trunk fails, the traffic from that link is simply shifted to the other links in the trunk. Figure 2-11 on page 25 shows the configuration

of a trunk up to the apply command. It again uses the single command line for brevity. The directions are given as several steps so you can see more of the CLI and its menus.

1. Enter `/cfg` at the main menu.
2. Enter `12` at the `cfg` menu.
3. Enter `trunk` at the `l2` menu, then enter the trunk number - either 1 or 2.
4. Enter `add <port>` to add a port to the trunk where `<port>` is EXT1, EXT2, EXT3 or EXT4.
5. Repeat Step 4 for all the ports you wish to add to the trunk.
6. Enter `ena` to enable the new trunk group.
7. Enter `apply` to apply the changes and activate the new trunk.
8. Enter `save` to save the changes to NVRAM and answer `y` to overwrite the active config with the new config. The trunk is now active.



```
>> Main# /cfg/l2/trunk 1
-----
[Trunk group 1 Menu]
  add   - Add port to trunk group
  rem   - Remove port from trunk group
  ena   - Enable trunk group
  dis   - Disable trunk group
  del   - Delete trunk group
  cur   - Display current Trunk Group configuration

>> Trunk group 1# add EXT1
Port EXT1 added.

>> Trunk group 1# add EXT2
Port EXT2 added.

>> Trunk group 1# ena
Current status: enabled
New status:     enabled

>> Trunk group 1# apply
No apply needed.

>> Trunk group 1#
```

Figure 2-11 Configuring link aggregation

Upgrading the firmware

To upgrade the firmware on the GbESM, you must use the TFTP transfer from a telnet session and the CLI. There is not an option on the Web session to upgrade the firmware.

Figure 2-12 on page 26 shows the process to load a new OS image file onto the switch.

Important: Before updating the firmware, save any configuration changes to the NVRAM of the GbESM. From the telnet session enter **Apply**, then press **Enter**. Type **Save** and press **Enter**. Answer **y** to the prompt that asks to confirm saving to flash. Answer **y** to the prompt that asks if you want to change the boot to the active config block. After the switch is updated, run the command `/boot/conf active` to set the boot back to the current configuration.

The firmware for the GbESM is contained in two files. One is a boot image file, the other is the OS image file. Use the following steps to upgrade the firmware on the GbESM via the telnet session:

1. Type `/boot/gtimg`.
2. Enter where the new image file will be placed. We are upgrading the boot image file first, so enter `boot`. That is the location for the boot image file.
3. Enter the IP address of the TFTP server.
4. Enter the fully qualified pathname for the boot image file on the TFTP server.
5. The switch will report the current version of the boot kernel on the switch and ask if you wish to replace it with new file. If you wish to continue, enter `y`.
6. Once the download is finished, go back to Step 1 and repeat the process for the OS image file. In step 2, enter `image1` or `image2` as the location to store the new image file.
7. If the download location is the same as the location for the currently loaded OS image, the switch will warn you that a failed download could result in an inoperative switch. If the download location is different from the location of the currently loaded OS image, the image file will download. After the download is finished the switch will prompt whether you wish to use the old location or the new location. Figure 2-12 shows a successful download of the OS image to `image2`. It also shows the information on the download that was just completed and the prompt as to whether the user wishes to change to `boot image2`.
8. Type `/boot/reset` to reset the switch and boot with the new firmware files.

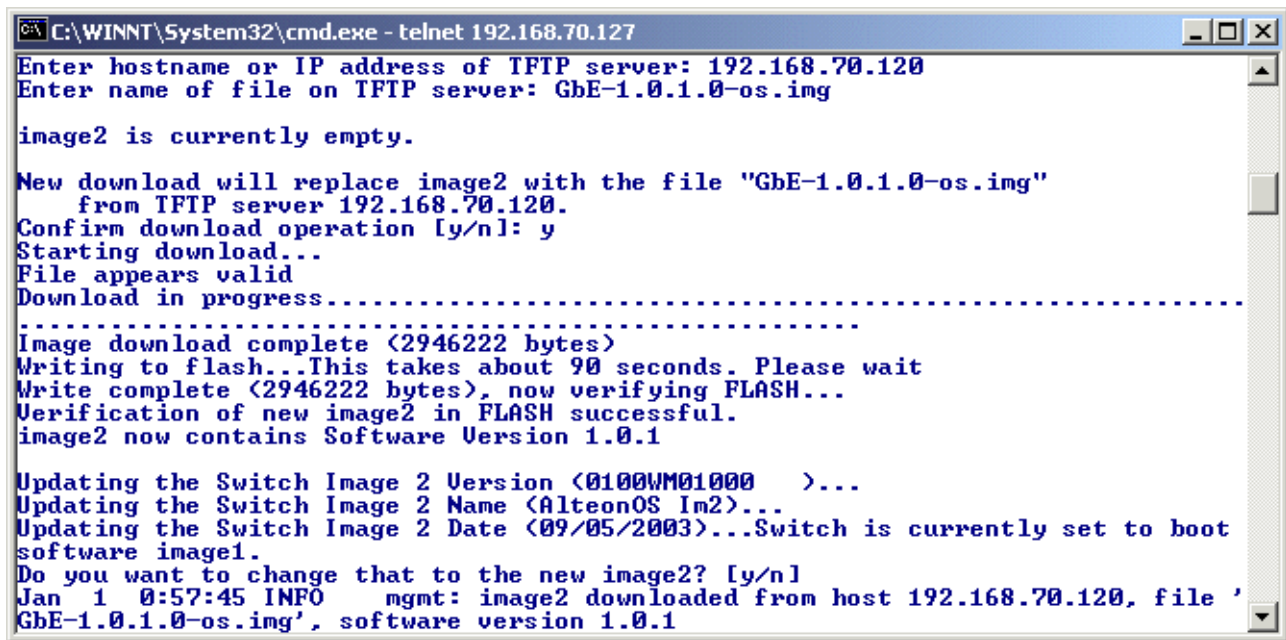


Figure 2-12 OS image download

Capturing the current configuration

There are a few ways to capture the current configuration in the CLI. The first is to use a TFTP server to push the configuration file to the server. However, in some text editors the resulting file is a single long line of text. Though this method requires a TFTP server running on the network, it does work with any telnet client. To capture the configuration by pushing a file to a TFTP server:

1. Enter `/cfg/ptcfg` at the command line.

2. Enter the IP address of the TFTP server.
3. Enter the filename you wish to save the file to.

A second way to capture the current configuration does not require a TFTP server and will result in a more people-friendly file with each command on its own line. However, this method requires a terminal emulator that can capture text. This example uses a telnet session configured in Hyperterminal to capture the text. The commands on the switch are the same for any software, but the steps to set the software to capture the text may be different. If your terminal emulator does not support this, you will have to use the TFTP method. To capture the configuration using text capture in HyperTerminal:

1. Click **Transfer** → **Capture Text** in the HyperTerminal menus and enter the filename to capture to.
2. Enter `/cfg/dump` at the switch command line to dump the configuration.
3. Click **Transfer** → **Capture Text** → **Stop** in the HyperTerminal menus to stop the text capture.
4. Open the file you captured. The first line will read `>>Information# dump` and must be removed. The last line that reads `>>Information#` must also be removed. The edited configuration file from our commands entered earlier is listed in Example 2-1.

Example 2-1 Example Configuration File

```
script start "Nortel Networks Layer2-7 GbE Switch Module" 4 /**** DO NOT EDIT THIS LINE!
/* Configuration dump taken 0:17:40 Thu Jan 1, 2070
/* Version 20.0.1, Base MAC address 00:0c:f8:2a:6f:00
/c/12/trunk 1
    dis
    add EXT1
    add EXT2
/c/13/if 1
    ena
    addr 192.168.2.2
/
script end /**** DO NOT EDIT THIS LINE!

>> Configuration# /cfg/dump
script start "Nortel Networks Layer2-7 GbE Switch Module" 4 /**** DO NOT EDIT THIS LINE!
/* Configuration dump taken 0:27:53 Thu Jan 1, 2070
/* Version 20.0.1, Base MAC address 00:0c:f8:2a:6f:00
/c/12/trunk 1
    dis
    add EXT1
    add EXT2
/c/13/if 1
    ena
    addr 192.168.2.2
/
script end /**** DO NOT EDIT THIS LINE!
```

Configuring user accounts

This section describes the user accounts on the switch.

The seven user accounts listed in Table 2-1 on page 28 are the default accounts on the GbESM.

Table 2-1 Description of default user accounts

User account	Description/Tasks performed	Password
User	Can view switch statistics but cannot make changes.	user
SLB Operator	Can manage content servers and configure options on the SLB menus, but not filters.	slboper
Layer 4 Operator	Reserved for future use.	l4oper
Operator	The Operator manages all functions of the switch. In addition to SLB Operator functions, the Operator can reset ports or the entire switch.	oper
SLB Administrator	The SLB Administrator configures and manages content servers and other Internet services and their loads.	slbadmin
Layer 4 Administrator	In addition to SLB Administrator functions, the Layer 4 Administrator can configure all parameters on the SLB menus, including filters and bandwidth management.	l4admin
Administrator	The super-user Administrator has complete access to all menus, information, and configuration commands on the switch.	admin

There is no mechanism on the GbESM for adding users directly to the switch. Instead, the GbESM supports the RADIUS (Remote Authentication Dial-in User Service) method to authenticate and authorize remote administrators for managing the switch. Any additional users you wish to add to the switch first need to be added to the RADIUS server. The switch must also be configured to use the RADIUS server. Then the new user simply logs into the switch. Users only need to be defined at the server and can access any switch that is configured to use that server. RADIUS is not required to manage the switch. The administrator can change the passwords of any of the default users to secure the switch. RADIUS is only required if you wish to add additional users to the switch.

RADIUS is based on a client/server model. The Remote Access Server (RAS)—the switch—is a client to the backend database server. A remote user (the remote administrator) interacts only with the RAS, not the backend server and database. The GbESM, acting as the RADIUS client, communicates to the RADIUS server to authenticate and authorize a remote administrator.

How RADIUS authentication works

1. Remote administrator connects to the switch and provides user name and password.
2. Using Authentication/Authorization protocol, the switch sends request to the authentication server.
3. Authentication server checks the request against the user ID database.
4. Using RADIUS protocol, the authentication server instructs the switch to grant or deny administrative access.

For more information on RADIUS and how to configure the switch to access a RADIUS server, reference the *Alteon OS 20.0 Application Guide, Part 215654-A* (Nortel).

When you access the switch using one of these accounts, the authentication methods differ slightly for the CLI and Web interfaces. In the case of the CLI, all you need to enter is a password. The switch knows what access level you have and will show the appropriate menu based on the password you have entered. To access the switch as user, you would enter the

password for the `user` account. As `user`, you are presented with a much more limited menu than if you were to access the switch as `admin`. Figure 2-13 and Figure 2-14 show the different main menus for `admin` and `user`. As you can see, `user` has a much more limited set of commands available than `admin`.

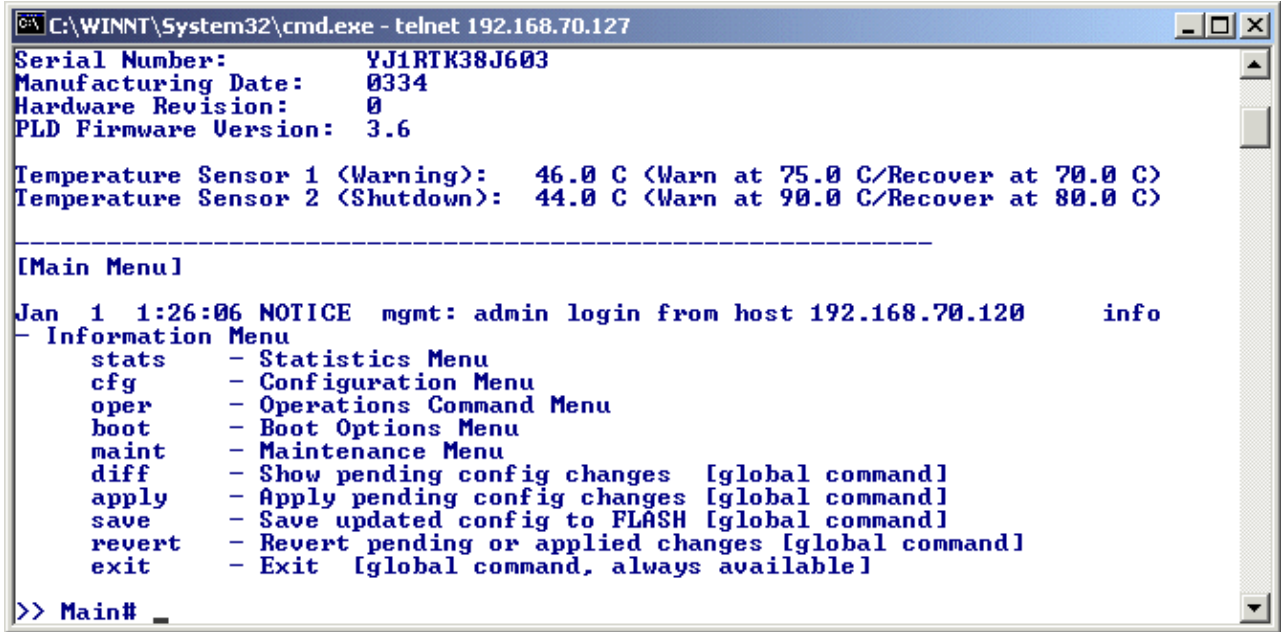


Figure 2-13 Main menu for user: admin

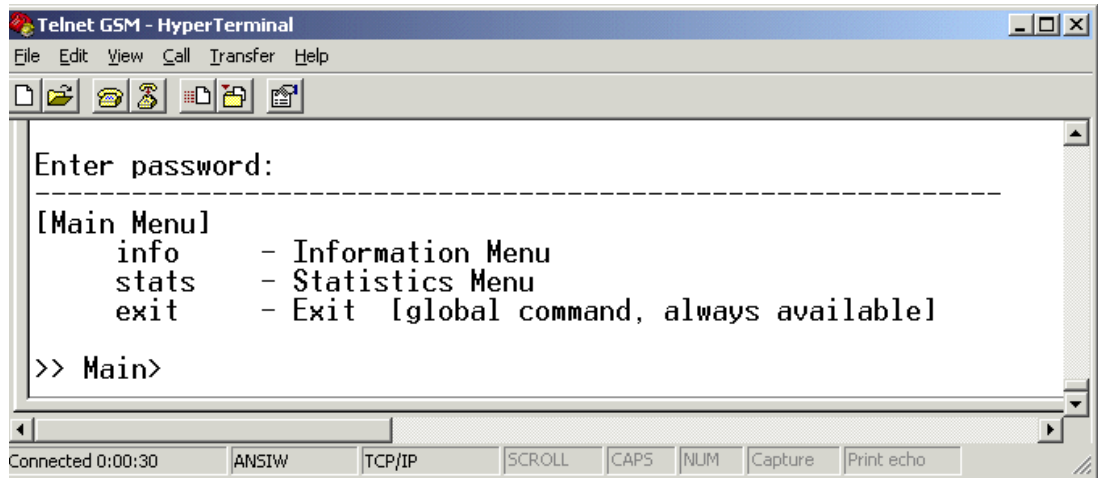


Figure 2-14 Main menu for user: user

When you access the switch through the Web interface, you are prompted for a username and a password. For all the default users on the switch, the username and password are the same by default. If you wish to change any of the passwords, you must be logged in as `admin`. To change a password enter the command:

```
/cfg/sys/access/user
```

This will bring up the menu shown in Figure 2-15 on page 30. From this menu you can change the passwords of all the default users on the switch.

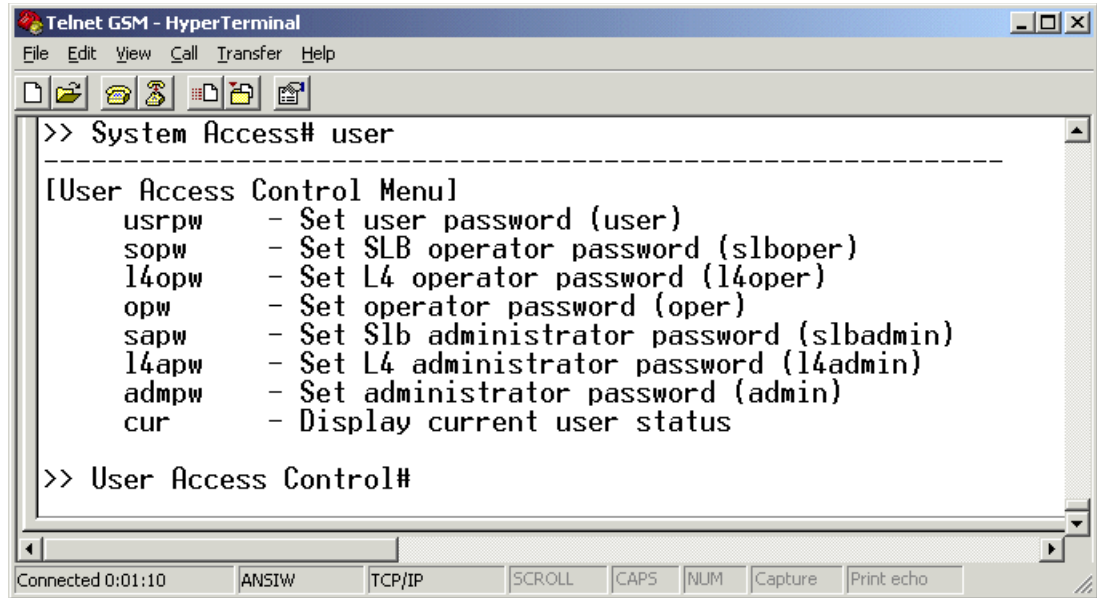


Figure 2-15 User password control



Layer 2-7 GbE Switch Module for IBM @server BladeCenter functionality

In this chapter, we introduce how the Layer 2-7 GbE Switch Module for IBM @server BladeCenter operates. For this purpose the OSI reference model is used.

Before reading this section, you should have a basic understanding of the Ethernet protocol and its vocabulary (for example, MAC address, frame, shared media, collision, segment, CRC error).

Important: The intent of this Redpaper is to introduce networking terms and concepts to the users and administrators of the IBM @server BladeCenter product. Our explanations of these terms and concepts are meant to give the non-networking professional an overall view of networking, not substitute for in-depth training on networking fundamentals. The examples illustrated in this paper are meant to provide real “working” examples of networking configurations where an IBM @server BladeCenter is deployed, but of course many other configurations are possible.

Tip: For more in-depth information about the features discussed in this chapter refer to the Application Guide and the Command Reference.

For general networking knowledge, we recommend the Cisco Internetworking Technology Handbook available at:

http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/index.htm

The OSI reference model

The OSI reference model is the primary architecture model for networks. It describes how data and network information are communicated from applications on one computer, through the network media, to an application on another computer. The OSI reference model categorizes the various processes needed in a communications session into seven distinct functional layers. The layers are organized as shown in Figure 3-1 based on the natural sequence of events that occur during a communication session.

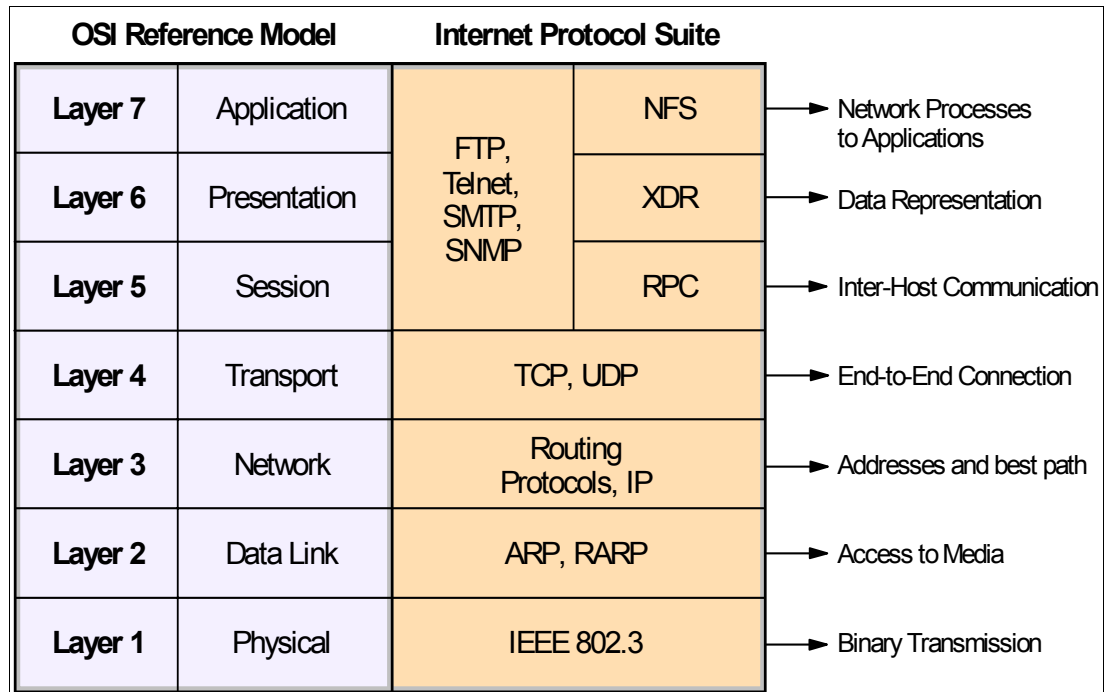


Figure 3-1 The OSI reference model with examples of related protocols

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter has functionality related to all of the seven layers. This means that the Layer 2-7 GbE Switch Module for IBM @server BladeCenter not only could be used to consolidate the Layer 2/3 LAN switches that sit behind the servers, but also SSL appliances, server load balancers, cache appliances, application health checking, and others.

Infrastructure would now be much simpler to deploy, administer, and troubleshoot. Furthermore, by integrating the Layer 2-7 switching functionality into the BladeCenter, many of the benefits that such switches deliver become an integral part of the BladeCenter value proposition, for example, improved security, better application performance, SSL encryption/decryption, server load balancing, cache redirection, application health checking, on-demand computing, and so forth.

This results in a platform with higher application availability, increased application performance, better scalability, simplified management, and advanced security. Furthermore, it simplifies the data center topology and reduces the number of discreet devices, which means customers spend less in acquiring, integrating, managing, and maintaining their infrastructure—all of which results in a significantly lower Total Cost of Ownership!

Figure 3-2 show where the GbESM BladeCenter is positioned in the OSI model.

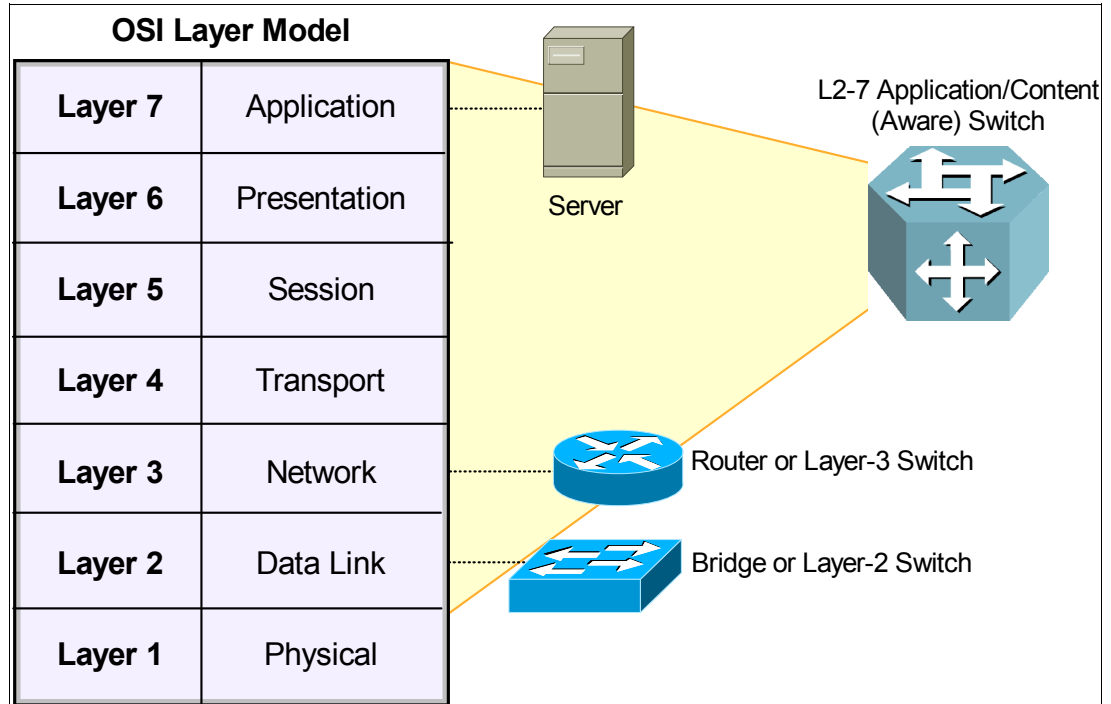


Figure 3-2 Positioning of the GbESM in the OSI reference model

3.1 Layer 1: Physical layer

Layer 1 of the OSI reference model is called the *physical layer*. The layer is responsible for transmission of the bit stream. Data frames from Layer 2 enter the layer and are transmitted one bit at a time to the media. This layer is also responsible for reception of incoming streams of data, also one bit at a time. These streams are then sent to the data link layer.

Hubs are network devices that belong in the physical layer.

3.2 Layer 2: Data link layer

The *data link layer* has two responsibilities: transmit and receive. For transmitting to take place, data is packed into frames. A *frame* contains enough information to ensure that the data can be successfully sent across a LAN to the final destination. The layer uses its own MAC address to accomplish this.

Bridging

An *Ethernet bridge* is a two-port device that forwards traffic by analyzing the MAC addresses of frames passing through it. A bridge builds and maintains an internal *forwarding table* by inspecting the source MAC addresses of the frames entering its ports. The contents of the forwarding table determine whether frames entering one port should be forwarded to the other port, or discarded.

When a bridge powers on, the forwarding table is empty initially. After the first frame from the network arrives, the bridge *learns* the MAC address of the *source* machine and its location (that is, what port it is connected to). The bridge stores these two pieces of information in the table and searches for the MAC address of the *destination* machine. Because the destination

machine has not transmitted any frames yet, the bridge does not know where it is. Therefore, the bridge sends the frame to the other port. The process of populating a bridge forwarding table is called *source learning*.

As other machines transmit frames, the bridge learns more source MAC addresses and their associated port locations. With this information, the bridge will discard a frame (for example, not forward it to the other port) if its forwarding table indicates that the destination MAC address is on the same port as the source MAC address. By making decisions to forward or discard frames, a bridge acts as a filter to limit the amount of network traffic passed from one port to the other port.

Isolating traffic into two segments also isolates collisions in the shared media. When two or more machines in one segment transmit simultaneously, a collision occurs. After a collision event, all machines in the segment must back off (that is, stop transmitting) for a random period of time. As more devices connect to one segment of shared media, there is a greater probability of multiple collisions occurring. Any device that detects contiguous collisions must increase its backoff time exponentially.

Therefore, when a bridge splits one segment into two smaller segments, two benefits emerge. First, by filtering data frames, it reduces the probability of a collision taking place. Second, by isolating collision traffic, a bridge prevents collisions in one segment from disrupting communication in the other segment. In other words, Ethernet bridges separate the media into two *collision domains* to improve traffic flow.

Bridges have limitations, however. Since they typically contain two ports, bridged Ethernet networks are not very scalable. For example, to create three collision domains, we require two bridges. Four collision domains require three bridges. Within each collision domain, multiple devices still need to share the Ethernet media.

Switches

Ethernet switches are functionally equivalent to bridges. However, switches have more than two ports and contain special hardware, such as application-specific integrated circuits (ASICs), to provide much higher performance than bridges. The ability to connect to several segments makes a switch much more versatile than a bridge. We can create many small collision domains using one switch, further improving on the benefits of bridging. In fact, by connecting only one device to each switch port, we virtually eliminate the drawbacks of shared media. Every collision domain now contains only two transmitters:

1. The attached device, which could be a server, workstation, printer, another switch, and so forth.
2. The switch itself, which forwards frames from other ports.

Aging time

The switch forwarding table, like the one in a bridge, must contain up-to-date information about MAC addresses and their port locations. Without accurate entries, a switch might forward frames to the wrong port, to all ports, or not forward them at all. Source learning by itself is useful for adding new MAC addresses to the forwarding table only when devices transmit frames. If a machine is disconnected from the network or moved to another location, the switch needs to remove the old information. The key to this process is the *aging time*, which determines how long a MAC address is stored in the forwarding table after the device stops transmitting frames. The default value for the Layer 2-7 GbE Switch Module for IBM @server BladeCenter is 300 seconds. In other words, if the switch does not detect any frames with a specific source MAC address after five minutes, it deletes the MAC address entry from the forwarding table. This will force the switch to relearn the source MAC address when the machine transmits another frame.

Increasing the aging time allows MAC addresses to stay in the table longer. This is beneficial for devices that rarely move around in the network or those that do not transmit frames within every 300 seconds. For example, some printers are very “quiet;” their primary task is to receive data for printing. The disadvantage of a long aging time is the potential for incorrect switch forwarding decisions. The switch will forward frames to the wrong port if a machine is relocated to another part of the network, but does not transmit any frames to update the switch.

Decreasing the aging time is useful when devices connect to different parts of the network on a regular basis (for example, laptops). If the switch receives a frame with a destination MAC address not stored in its forwarding table, then it *floods* the frame out all ports (except the port where the frame originated from.) This enables the relocated device to receive traffic even before it has transmitted frames. When the relocated device does transmit a frame, the switch relearns its MAC address, so flooding is not necessary. The disadvantage of a short aging time is excessive flooding. Because the forwarding table loses information too quickly, the switch frequently must send frames to all ports, similar to the operation of an Ethernet hub. This defeats the purpose of switching.

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter also supports static entries in the forwarding table. The aging time has no effect on static MAC addresses. In rare instances, we might want to statically configure a MAC address and map it to a port. This is primarily for security reasons, and ensures that a machine can receive traffic on the assigned port only. If someone moves the machine to another port, it cannot communicate through the switch.

Frame forwarding modes

Switches generally operate in one of two modes: *cut-through* or *store-and-forward*. In cut-through mode, a switch immediately forwards a frame and begins transmission after receiving the destination MAC address. This occurs while the original frame is still being received by the ingress port. The switch cannot discard the frame even if it has a CRC error, since most of the frame has already been transmitted to the destination. Cut-through operation does not work across ports that have different data rates. For example, a frame received at 10 Mb/s on one port cannot be forwarded in cut-through mode to a port at 100 Mb/s.

To forward traffic between different data rates, a switch must buffer the entire frame first. This is the store-and-forward mode of operation. Since the switch reads a complete frame before transmission, it checks for a CRC error and discards the frame if it detects an error. Store-and-forward introduces latency (delay) into the switching process, but it prevents any errors from propagating to other ports. Also, newer standards (such as IEEE 802.1Q) require that switches insert additional fields into Ethernet frames and recalculate the frame check sequence (FCS) to prevent false CRC errors. For these reasons, the Layer 2-7 GbE Switch Module for IBM @server BladeCenter operates in store-and-forward mode only.

Spanning Tree Protocol

Spanning Tree Protocol as defined by the IEEE 802.1D (STP bridge protocol) enables a network device to understand its connections to other devices. STP is a Layer 2 link management protocol that provides path redundancy while preventing undesirable loops in the network.

For a Layer 2 Ethernet network to function properly, only one active path can exist between any two stations. STP operation is transparent to end stations, which cannot detect whether they are connected to a single LAN segment or a switched LAN of multiple segments. The Layer 2-7 GbE Switch Module for IBM @server BladeCenter can be configured to use STP

(the IEEE 802.1D bridge protocol) on all VLANs. By default, STP is enabled on all ports. If needed, you can disable (and enable) STP on a per-VLAN basis.

STP is not needed on the internal ports because the internal connections within BladeCenter create a hub-and-spoke physical topology without loops, with the Layer 2-7 GbE Switch Module for IBM @server BladeCenter at the hub.

Important: There is no need to run STP on the external ports if all external ports are part of the same Link Aggregation trunk group, or if external ports in separate trunk groups are also in different VLANs. The external ports do not forward Layer 2 packets between links within a trunk group or between VLANs.

The only time that STP would need to be enabled on the external ports is if two external ports are part of the same VLAN and are not part of the same Link Aggregation trunk group. Generally, there is no need for this to be true.

Note that this does not mean that the external network is not running STP. In fact, in general, it probably is. It just means that the external port generally operates as an end system and not an L2 switch, and therefore doesn't participate in the STP process.

Background on STP

The spanning tree algorithm enables switches and transparent bridges to dynamically discover a loop-free network (tree) and provide a single physical path between any two stations attached to the network (spanning). Although this standard applies to both transparent bridges and ethernet switches, we will be referencing Ethernet switches in this section.

The IEEE 802.1d standard defines how the spanning tree will work in a switching environment. If there were loops in the network topology, there would be network configurations where:

- ▶ Frames endlessly circulate within the network
- ▶ Communication is prevented because a bridge may incorrectly assert that a source and destination address are on the same side of the bridge

Layer 2 LAN ports send and receive STP frames at regular intervals. Network devices do not forward these frames, but use the frames to construct a loop-free path.

STP defines a tree with a root bridge and a loop-free path from the root to all network devices in the Layer 2 network. STP forces redundant data paths into a standby (blocked) state. If a network segment in the spanning tree fails and a redundant path exists, the STP algorithm recalculates the spanning tree topology and activates the standby path. When two Layer 2 LAN ports on a network device are part of a loop, the STP port priority and port path cost setting determine which port is put in the forwarding state and which port is put in the blocking state. The STP port priority value represents the location of a port in the network topology and how well located it is to pass traffic. The STP port path cost value represents media speed.

Understanding bridge protocol data units

To participate in the spanning tree protocol, each switch will initially assume it is the root bridge and will transmit a Hello BPDU (bridge protocol data unit) on each of its ports. This message will be sent every Hello time. *Hello time* is one of the spanning tree configuration parameters that can be specified for each switch during the switch configuration. This Hello BPDU will have the following characteristics:

1. The source address will be the address of the transmitting switch.

2. The destination address will be '800143000000 '.
3. The source and destination SAPs will be '42 '.
4. The Root ID field will contain the ID of the transmitting switch.
5. The Bridge ID field will contain the ID of the transmitting switch.
6. The Path Cost field will contain 0.
7. It will be sent out by the switch on all its ports.

Each Hello BPDU sent out on a switch port will be received by all the other switches that are connected to the LAN attached to that port. BPDUs are not broadcast frames. They are sent to switches, and selectively updated and forwarded to other switches. Each switch uses the information received in the Hello BPDUs to determine the root bridge, the designated bridges, and the designated ports within each designated bridge. To do this, each switch will continue transmitting a Hello BPDU on each of its ports until it receives a better Hello BPDU than the one it is transmitting on that port. The better Hello BPDU will be determined based on the following information contained in the Hello BPDU (listed in order of their significance):

1. The lowest root ID
2. The lowest path cost to the switch
3. The lowest transmitting switch ID
4. The lowest port ID

As soon as a switch receives such a Hello BPDU on a port, it will stop transmitting any further Hello BPDUs on that port and will use the information received in the better Hello BPDU to transmit a new Hello BPDU on all its other ports. The new Hello BPDU will have the following characteristics:

1. The source address will be the address of this switch.
2. The destination address will be '800143000000 '.
3. The source and destination SAP will be '42 '.
4. The Root ID field will contain the root ID received in the better Hello BPDU.
5. The Bridge ID field will contain the ID of this switch.
6. The Path Cost field will be the sum of the path cost received in the better Hello BPDU plus the path cost defined for the switch port on which the better Hello BPDU was received.
7. The new Hello BPDU will be sent out by the switch on all its ports except the port on which the better Hello BPDU was received.

This process will be repeated by all the switches until:

1. There is one switch (root bridge) remaining that is still transmitting its original Hello BPDU.
2. One switch (designated bridge) on each LAN is transmitting the Hello BPDU based on the Hello BPDU received from the root bridge.

On the designated switch, the port on which the best Hello BPDU is received is the root port and all the ports to which the Hello BPDUs are transmitted are the designated ports.

Note: There may be some ports on the designated bridge over which the switch will not be transmitting Hello BPDUs due to the fact that the received BPDUs on those ports are better than the one this switch would be able to transmit (but they are not better than the Hello BPDU received on its root port). Once the root and designated bridges have been elected, the root ports and the designated ports will be put in forwarding state and all the other ports will be put in blocking state.

Switches and network topology changes

In Figure 3-3, Switch A is elected as the root bridge because the bridge priority of all the network devices is set to the default (32768) and Switch A has the lowest MAC address. However, due to traffic patterns, number of forwarding ports, or link types, Switch A might not be the ideal root bridge. By increasing the priority (lowering the numerical value) of the ideal network device so that it becomes the root bridge, one forces an STP recalculation to form a new spanning tree topology with the ideal network device as the root.

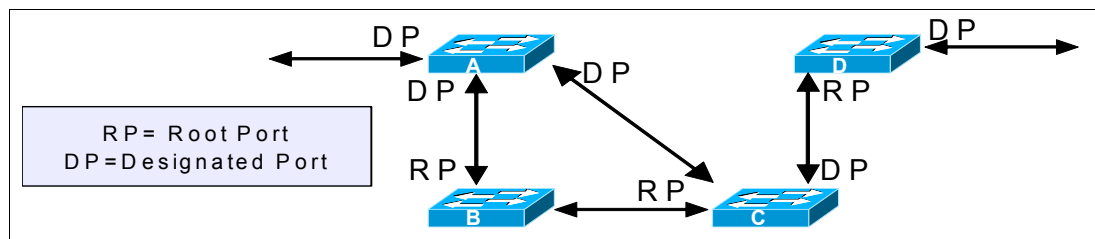


Figure 3-3 Election of root bridge

When the spanning tree topology is calculated based on default parameters, the path between source and destination end stations in a switched network might not be ideal. For instance, connecting higher-speed links to a port that has a higher number than the current root port can cause a root-port change. The goal is to make the fastest link the root port. For example, assume that one port on Switch B is a 1Gbps, and another port on Switch B, at 100 Mbps, is the root port. Network traffic might be more efficient over the 1Gbps link. By changing the STP port priority on the 1Gbps port to a higher priority (lower numerical value) than the root port, the 1Gbps becomes the new root port.

STP port states

Propagation delays can occur when protocol information passes through a switched LAN. As a result, topology changes can take place at different times and at different places in a switched network. When a Layer 2 LAN port transitions directly from nonparticipation in the spanning tree topology to the forwarding state, it can create temporary data loops. Ports must wait for new topology information to propagate through the switched LAN before starting to forward frames. They must allow the frame lifetime to expire for frames that have been forwarded using the old topology.

Each Layer 2 LAN port using STP exists in one of the following five states:

1. **Blocking** – The Layer 2 LAN port does not participate in frame forwarding.
2. **Listening** – The first transitional state after the blocking state, when STP determines that the Layer 2 LAN port should participate in frame forwarding.
3. **Learning** – The Layer 2 LAN port prepares to participate in frame forwarding.
4. **Forwarding** – The Layer 2 LAN port forwards frames.
5. **Disabled** – The Layer 2 LAN port does not participate in STP and is not forwarding frames.

Figure 3-4 on page 39 illustrates how a Layer 2 LAN port moves through the five states

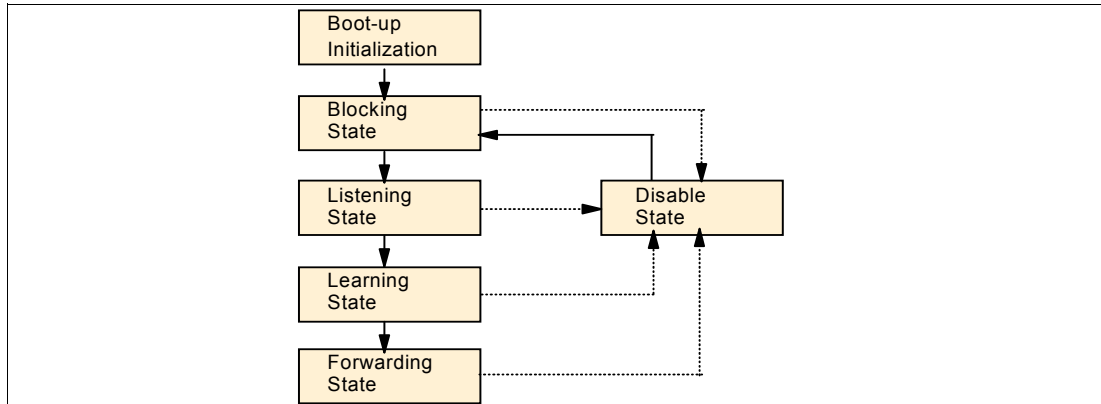


Figure 3-4 STP port states

When one enables STP, every port in the switch, VLAN, and network goes through the blocking state and the transitory states of listening and learning at power up. If properly configured, each Layer 2 LAN port stabilizes to the forwarding or blocking state. When the STP algorithm places a Layer 2 LAN port in the forwarding state, the following process occurs:

1. The Layer 2 LAN port is put into the listening state while it waits for protocol information that suggests it should go to the blocking state.
2. The Layer 2 LAN port waits for the forward delay timer to expire, moves the Layer 2 LAN port to the learning state, and resets the forward delay timer.
3. In the learning state, the Layer 2 LAN port continues to block frame forwarding as it learns end station location information for the forwarding database.
4. The Layer 2 LAN port waits for the forward delay timer to expire and then moves the Layer 2 LAN port to the forwarding state, where both learning and frame forwarding are enabled.

Multiple Spanning Tree (IEEE 802.1s)

MISTP (802.1s) is an IEEE standard which allows several VLANs to be mapped to a reduced number of spanning-tree instances. This is possible since most networks do not need more than a few logical topologies. Each instance handles multiple VLANs that have the same Layer 2 topology. The Layer 2-7 GbE Switch Module for IBM @server BladeCenter does not support 802.1s at the time of writing this book. However, the GbESM can interoperate in a Cisco environment that is using PVST+. Per VLAN Spanning Tree Plus (PVST+) provides the same functionality as PVST using 802.1Q trunking technology rather than ISL. PVST+ is an enhancement to the 802.1Q specification.

Virtual LANs

This section provides an overview of virtual LANs (IEEE 802.1Q, VLANs)

A virtual local area network (VLAN) is a logical association of network devices based on certain policies or rules defined in a switch. VLANs provide additional flexibility in designing networks since we are no longer constrained by the physical location of machines. To understand VLANs, we need to discuss how broadcast traffic affects a switch.

Recall that a switch reads the destination MAC address of a frame to make a forwarding decision. If the destination address is a known MAC address in the forwarding table, then the switch transmits the frame on the corresponding port. If the MAC address is not in the table, flooding occurs. Frames containing a destination address of an individual device are called *unicast* frames. In contrast, *broadcast* frames contain a special destination address that all

devices recognize. Every device, upon receiving a frame with a broadcast destination address, must process the frame, whether or not it was intended for that device.

When a switch receives a broadcast frame from one of its ports, it transmits a copy of the frame to all other active ports. Broadcast traffic is necessary for the operation of certain protocols and applications. Unfortunately, when too much broadcast traffic exists in the network, performance degrades. As the number of broadcast frames increases, machines must spend more CPU time to process them. To combat this problem, VLANs were created to isolate broadcasts to a subset of ports in the switch.

Whereas switches and bridges keep Ethernet collisions within domains, VLANs isolate traffic (both unicast and broadcast) within *broadcast domains*. In the Layer 2-7 GbE Switch Module for IBM @server BladeCenter, all ports initially belong to the same VLAN, also known as the default VLAN. Once we create a new VLAN with ports assigned to it, the Layer 2-7 GbE Switch Module for IBM @server BladeCenter prevents traffic in one VLAN from interfering with traffic in the other VLAN. This is equivalent to having two physical networks with no connection between them. For example, a blade server in one VLAN cannot communicate with a blade server in another VLAN unless routing is enabled on the Layer 2-7 GbE Switch Module for IBM @server BladeCenter or on an external device that connects the two VLANs together.

The problem with traditional VLANs

The VLANs described thus far assume that all machines connect to a single switch. If we expand our network by connecting an additional switch, it cannot extend or enlarge the VLANs from the first switch. For example, in Figure 3-5, we connect port 1 on Switch B to port 6 on Switch A. Both ports belong in VLAN 30, so traffic in that VLAN flows freely between switches. However, VLAN 10 and VLAN 20 are isolated in each switch. Devices in VLAN 10 or VLAN 20 of Switch A cannot communicate with devices in VLAN 10 or VLAN 20 of Switch B.

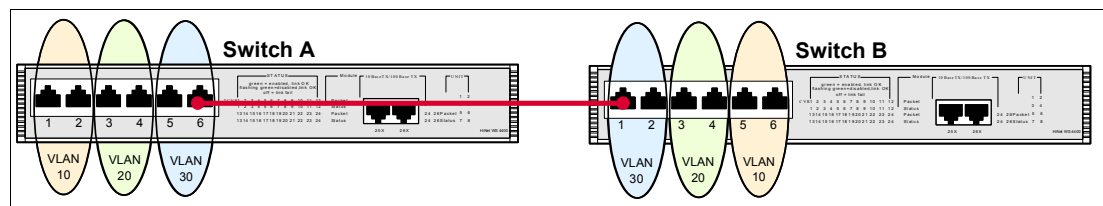


Figure 3-5 Both VLANs 10 and 20 are isolated in each switch; only VLAN 30 spans both switches

Alternatively, if we do not create any VLANs in Switch B (that is, all ports are in the default VLAN), then all devices attached to Switch B would be grouped into VLAN 30 in the first switch (Figure 3-6). Again, Switch B cannot access VLAN 10 and VLAN 20 in Switch A. In both cases, we can only connect one VLAN in Switch A to one VLAN in Switch B.

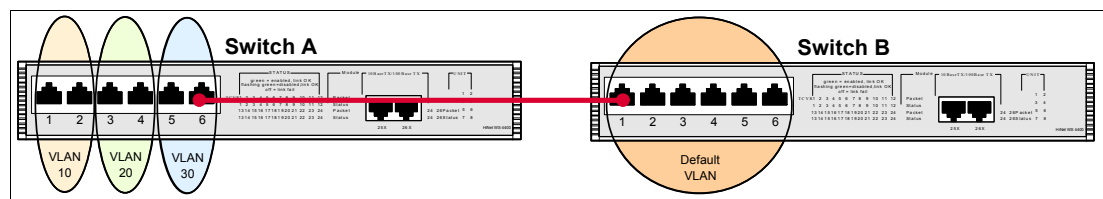


Figure 3-6 Switch B can communicate with Switch A in VLAN 3 only

To solve this problem, we require a mechanism to satisfy the following requirements:

1. Allow traffic from multiple VLANs in one switch to cross a single link to another switch.

2. Identify the VLAN that each frame belongs in *before* it crosses the link so the remote switch knows where to forward the traffic.

Figure 3-7 shows these two ideas at work. Port 6 on Switch A and port 1 on Switch B are no longer in VLAN 30. Instead, they are ports designated for carrying traffic to and from different VLANs. The frames that travel across the link have VLAN identification so they can be forwarded to the appropriate ports by the remote switch.

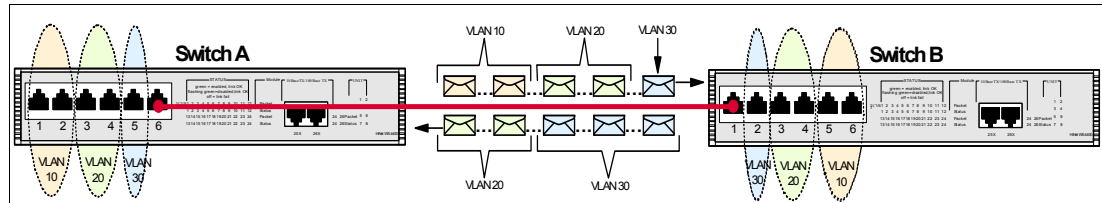


Figure 3-7 Transporting three VLANs across a single link

IEEE 802.1Q is a standard protocol that enables VLANs to span multiple Ethernet switches. Prior to 802.1Q, manufacturers used proprietary protocols to forward VLAN traffic between their switches. Unfortunately, this meant that we could not mix different brands of switches and still maintain VLAN consistency across the network. Also, with 802.1Q, we still have the flexibility to connect with non-802.1Q (legacy) switches.

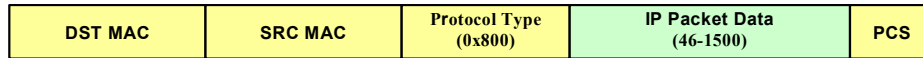
To tag or not to tag

One of the key concepts of 802.1Q is the tag header. An Ethernet frame that contains a tag header immediately following the source MAC address field is called a *tagged* frame (see Figure 3-8). Switches (such as the Layer 2-7 GbE Switch Module for IBM @server BladeCenter, and also some network interface cards) that support 802.1Q read the contents of the tag to determine in which VLAN the frame belongs. Likewise, an *untagged* frame is an Ethernet frame that does not contain a tag header. These frames do not contain a VLAN identifier. For each VLAN, switch ports are either in tagging mode or untagging mode, depending on the type of device connected.

If we set a switch port to tagging mode for a particular VLAN, then every frame from that VLAN will be transmitted with a tag header. If we set a switch port to untagging mode for a particular VLAN, then every frame from that VLAN will be transmitted without a tag header. The tagging/untagging parameter only affects frames *transmitted* by the switch. It does not affect frames received by the switch. In other words, the tagging/untagging function affects egress (outgoing) traffic only. If the tagging parameter is turned on, untagged frames received will be assigned to the PVID of the ingress port.

With the tag header, switches are able to preserve VLAN traffic as it travels from the source to the destination. VLAN-tagged frames are never modified by switches. Once a tag is created at the source, all 802.1Q switches preserve the tag to ensure proper VLAN isolation. A switch removes the tag only when the frame must be sent to a device that only accepts untagged Ethernet frames (for example, PCs, servers, printers, and non-802.1Q switches). In general, we enable tagging on both ends of a link that supports 802.1Q, thereby creating a VLAN *trunk* connection. The trunk can carry as few as one VLAN, or as many as 128 VLANs (in the Layer 2-7 GbE Switch Module for IBM @server BladeCenter).

Ethernet Frame Encapsulation of IP traffic -



Ethernet Frame with VLAN Encapsulation of IP traffic -

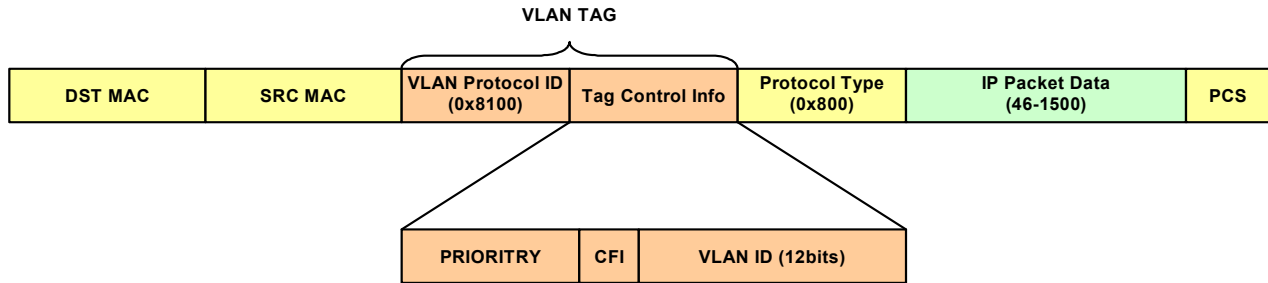


Figure 3-8 Difference between a tagged and an untagged Ethernet frame

Link Aggregation

Link Aggregation (802.3ad) is a method of combining physical network links into a single logical link for increased bandwidth. With Link Aggregation we are able to increase the capacity and availability of the communications channel between devices (both switches and end stations) using existing Fast Ethernet and Gigabit Ethernet technology. Two or more Gigabit Ethernet connections are combined in order to increase the bandwidth capability and to create resilient and redundant links. A set of multiple parallel physical links between two devices is grouped together to form a single logical link. Link Aggregation also provides load balancing where the processing and communications activity is distributed across several links in a trunk so that no single link is overwhelmed. By taking multiple LAN connections (as shown in Figure 3-9) and treating them as a unified, aggregated link, we can achieve practical benefits in many applications.

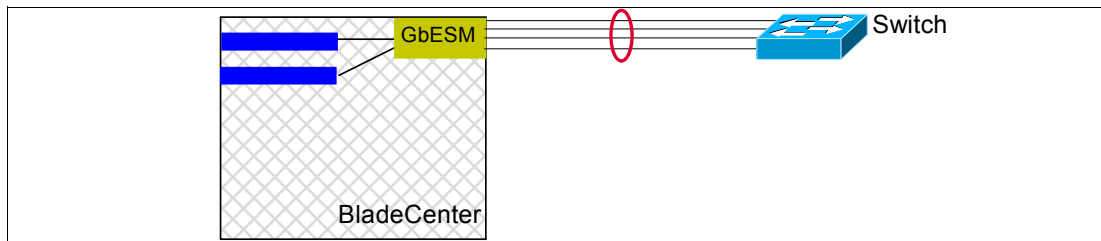


Figure 3-9 Link aggregation

Link Aggregation provides the following important benefits:

1. Increased bandwidth - The capacity of multiple links is combined into one logical link.
2. Increased availability - The failure or replacement of a single link within a Link Aggregation Group need not cause failure from the perspective of a MAC Client.
3. Linearly incremental bandwidth - Bandwidth can be increased in unit multiples as opposed to the order-of-magnitude increase available through Physical Layer technology options (10 Mb/s, 100 Mb/s, 1000 Mb/s, and so forth).
4. Load sharing - MAC Client traffic may be distributed across multiple links.

5. Deterministic behavior - Depending on the selection algorithm chosen, the configuration can be made to resolve deterministically; that is, the resulting aggregation can be made independent of the order in which events occur, and can be completely determined by the capabilities of the individual links and their physical connectivity.
6. Low risk of duplication or disordering of frames - During both steady-state operation and link (re-) configuration, there is a high probability that frames are neither duplicated nor disordered.
7. Support of existing IEEE 802.3 MAC Clients (frames transmitted are ordinary MAC frames) - No change is required to existing higher-layer protocols or applications to use Link Aggregation.
8. Backwards compatibility with aggregation-unaware devices - Links that cannot take part in Link Aggregation (either because of their inherent capabilities, management configuration, or the capabilities of the devices to which they attach) operate as normal, individual IEEE 802.3 links.
9. Accommodation of differing capabilities and constraints - Devices with differing hardware and software constraints on Link Aggregation are, to the extent possible, accommodated.
10. No change to the IEEE 802.3 frame format - Link Aggregation neither adds to, nor changes the contents of frames exchanged between MAC Clients.
11. Network management support - The standard specifies appropriate management objects for configuration, monitoring, and control of Link Aggregation.

3.3 Layer 3: Network layer

The *network layer* defines end-to-end delivery of packets. To accomplish this, it defines logical addressing so that any endpoint can be identified. It also defines how routing works and how routes are learned so that packets can be delivered.

IP addresses

To be able to identify a host on the Internet, each host is assigned an address, the IP address, or Internet address. When the host is attached to more than one network, it is called multi-homed and it has one IP address for each network interface. The IP address consists of a pair of numbers:

IP address = <network number><host number>

The network number part of the IP address is centrally administered by the Internet Network Information Center (the InterNIC) and is unique throughout the Internet. IP addresses are 32-bit numbers usually represented in a dotted decimal form (as the decimal representation of four 8-bit values concatenated with dots). For example 128.2.7.9 is an IP address with 128.2 being the network number and 7.9 being the host number.

IP addresses are used by the IP protocol to identify uniquely a host on the internet. IP datagrams (the basic data packets exchanged between hosts) are transmitted by some physical network attached to the host, and each IP datagram contains a source IP address and a destination IP address. To send a datagram to a certain IP destination, the target IP address must be translated or mapped to a physical MAC address. This may require transmissions on the network to find out the destination's physical network address (for example, on LANs the Address Resolution Protocol is used to translate IP addresses to physical MAC addresses).

The first bits of the IP address specify how the rest of the address should be separated into its network and host part.

There are five classes of IP addresses as shown in Figure 3-10.

Class	Valid Network Numbers
A	1.0.0.0 to 126.0.0.0
B	128.1.0.0 to 191.254.0.0
C	192.0.1.0 to 223.255.254.0
D	224.0.0.0 to 239.255.255.254
E	240.0.0.0 to 255.255.255.255

Figure 3-10 Valid Network Numbers

Class D addresses are used for multicast purposes and class E addresses are reserved by IETF for its own research. Class A, B, and C are addresses used on the Internet.

Private IP addressing

When IP addresses that aren't connected to the Internet are needed, they can be pulled from a set of IP networks called private Internets (RFC 1918). The RFC defines a set of networks, identified in Figure 3-11, that will never be assigned to any organization as registered network numbers.

Class	Private Network Numbers
A	10.0.0.0 to 10.255.255.255
B	172.16.0.0 to 172.31.255.255
C	192.168.0.0 to 192.168.255.255

Figure 3-11 Private network numbers

Network Address Translation

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter redirects incoming client packets to chosen server blades using Network Address Translation (NAT). NAT is an Internet standard that allows the use of one set of IP addresses for internal traffic and a second set of addresses for external traffic. NAT serves two main purposes:

1. It provides a type of firewall by hiding internal IP addresses. This increases network security.
2. It enables a company to use more internal IP addresses. Since they're used internally only, there's no possibility of conflict with public IP addresses used by other companies and organizations.

In the following NAT example, a company has configured its Layer 2-7 GbE Switch Module for IBM @server BladeCenter internal network with private IP addresses. A private network is one that is isolated from the global Internet and is, therefore, free from the usual restrictions requiring the use of registered, globally unique IP addresses. With NAT, private networks are not required to remain isolated. NAT capabilities within the switch allow internal, private

network IP addresses to be translated to valid, publicly advertised IP addresses and back again.

NAT executes two related functions: one for the external client-side switch port, and one for the internal, server-side switch port. The client-side function translates incoming requests for the publicly advertised server IP address to the server's internal private network address. The function of the server-side switch port reverses the process, translating the server's private address information to a valid public address.

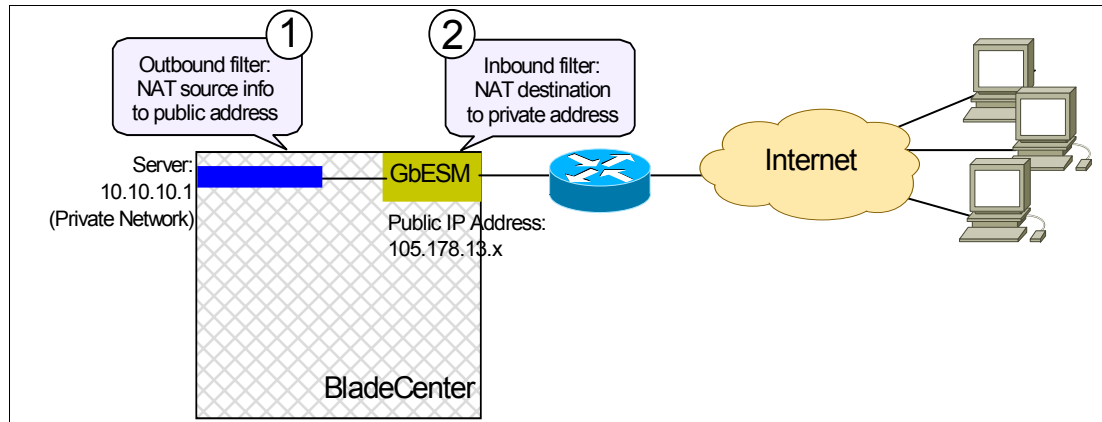


Figure 3-12 Network Address Translation

When traffic is received by the Layer 2-7 GbE Switch Module for IBM @server BladeCenter destined for a VIP, the switch modifies the Destination IP Address, Destination MAC address, and L2/L3 checksum information and forwards the packet to a server blade that provides the desired service using its unique IP address. Returning traffic, from the server to the client, is modified to place the VIP in the Source IP address field. The example in Figure 3-12 shows only a single server blade, but there could be many that the Layer 2-7 GbE Switch Module for IBM @server BladeCenter could send requests arriving on the VIP to, using private addresses such as 10.10.10.2, 10.10.10.3,.... 10.10.10.x.

Routing

Routing is the act of moving information across an inter-network from a source to a destination. Along the way, at least one intermediate node typically is encountered. Routing is often contrasted with bridging, which might seem, to the casual observer, to accomplish precisely the same thing. The primary difference between the two is that *bridging* occurs at Layer 2 (the link layer) of the OSI reference model, whereas *routing* occurs at Layer 3 (the network layer). This distinction provides routing and bridging with different information to use in the process of moving information from source to destination, so the two functions accomplish their tasks in different ways.

Routing involves two basic activities: determining optimal routing paths and transporting information groups (typically called packets) through an inter-network.

Routing protocols use metrics to evaluate what path will be the best for a packet to travel. A *metric* is a standard of measurement, such as path bandwidth, that is used by routing algorithms to determine the optimal path to a destination. To aid the process of path determination, routing algorithms initialize and maintain routing tables, which contain route information. Route information varies depending on the routing algorithm used.

Routing algorithms fill routing tables with a variety of information. Destination or the next hop associations tell a router that a particular destination can be reached optimally by sending the packet to a particular router representing the “next hop” on the way to the final destination.

When a router receives an incoming packet, it checks the destination address and attempts to associate this address with a next hop.

Routing tables also can contain other information, such as data about the desirability of a path. Routers compare metrics to determine optimal routes, and these metrics differ depending on the design of the routing algorithm used. A variety of common metrics are introduced and described later in this chapter.

Routing tables also can contain other information, such as data about the desirability of a path. Routers compare metrics to determine optimal routes, and these metrics differ depending on the design of the routing algorithm used.

Routers communicate with one another and maintain their routing tables through the transmission of a variety of messages. The routing update message is one such message that generally consists of all or a portion of a routing table. By analyzing routing updates from all other routers, a router can build a detailed picture of network topology. A link-state advertisement, another example of a message sent between routers, informs other routers of the state of the sender's links. Link information also can be used to build a complete picture of network topology to enable routers to determine optimal routes to network destinations.

Routed protocols are transported by routers across an inter-network. In general, routed protocols in this context also are referred to as network protocols. IP is an example of a routed protocol. These network protocols perform a variety of functions required for communication between user applications in source and destination devices, and these functions can differ widely among protocol suites.

The layer 3 routing table of the Layer 2-7 GbE Switch Module for IBM @server BladeCenter may be populated by using static routes and dynamic routing protocols. The Layer 2-7 GbE Switch Module for IBM @server BladeCenter supports the following dynamic routing protocols:

1. RIP
2. OSPF
3. BGP

RIP

The Routing Information Protocol (RIP) is a distance-vector protocol that uses hop count as its metric. A hop count of 4 indicates that a packet must traverse four routes to get to the destination. RIP is widely used for routing traffic in the global Internet and is an interior gateway protocol (IGP), which means that it performs routing within a single autonomous system.

RIP is designed for networks that are much smaller than the large networks we have in most places today. The following are some of the shortcomings of RIPv1 supported in GbESM:

- ▶ No authentication of routing updates from other devices.
- ▶ Only classfull routing updates wastes many valid IP addresses.
- ▶ Slow convergence due to updates only every 30 seconds.
- ▶ The limit of 15 hops does not scale in larger networks.
- ▶ The hop count does not differ on the speed of the links between network devices.

OSPF

OSPF was created because RIP was increasingly unable to serve large, heterogeneous inter-networks. OSPF has two primary characteristics:

- ▶ The protocol is open, which means that its specification is in the public domain. The OSPF specification is published as Request For Comments (RFC) 1247.
- ▶ OSPF is based on the Shortest Path First (SPF) algorithm, which sometimes is referred to as the Dijkstra algorithm, named for the person credited with its creation.

OSPF is a link-state routing protocol that calls for the sending of link-state advertisements (LSAs) to all other routers within the same hierarchical area. Information on attached interfaces, metrics used, and other variables is included in OSPF LSAs. As OSPF routers accumulate link-state information, they use the SPF algorithm to calculate the shortest path to each node.

Unlike RIP, OSPF can operate within a hierarchy. The largest entity within the hierarchy is the autonomous system (AS), which is a collection of networks under a common administration that share a common routing strategy. OSPF is an interior gateway (IGP) routing protocol, although it is capable of receiving routes from and sending routes to other autonomous systems. An AS can be divided into a number of areas, which are groups of contiguous networks and attached hosts. Routers with multiple interfaces can participate in multiple areas. These routers, which are called area border routers, maintain separate topological databases for each area.

A *topological database* is essentially an overall picture of networks in relationship to routers. The topological database contains the collection of LSAs received from all routers in the same area. Because routers within the same area share the same information, they have identical topological databases. The term *domain* sometimes is used to describe a portion of the network in which all routers have identical topological databases. Domain is frequently used interchangeably with AS. An area's topology is invisible to entities outside the area. By keeping area topologies separate, OSPF passes less routing traffic than it would if the AS were not partitioned. Area partitioning creates two different types of OSPF routing, depending on whether the source and destination are in the same or different areas. Intra-area routing occurs when the source and destination are in the same area; inter-area routing occurs when they are in different areas.

An OSPF *backbone*, as illustrated in Figure 3-13 on page 48, is responsible for distributing routing information between areas. It consists of all area border routers, networks not wholly contained in any area, and their attached routers.

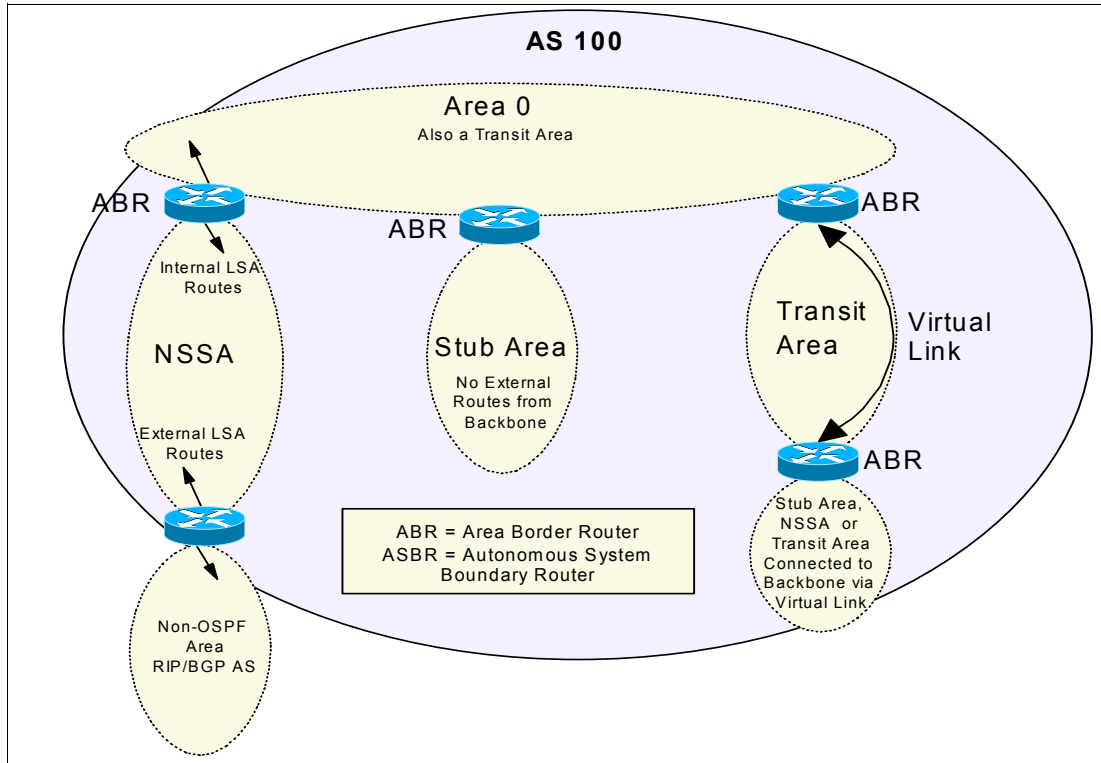


Figure 3-13 The different types of OSPF areas

Areas can be defined in such a way that the backbone is not contiguous. In this case, backbone connectivity must be restored through virtual links. *Virtual links* are configured between any backbone routers that share a link to a non-backbone area and function as if they were direct links.

AS border routers running OSPF learn about exterior routes through exterior gateway protocols (EGPs), such as Exterior Gateway Protocol (EGP) or Border Gateway Protocol (BGP), or through configuration information.

BGP-4

As with any routing protocol, BGP maintains routing tables, transmits routing updates, and bases routing decisions on routing metrics. The primary function of a BGP system is to exchange network reachability information, including information about the list of autonomous system paths, with other BGP systems, as shown in Figure 3-14 on page 49. This information can be used to construct a graph of autonomous system connectivity from which routing loops can be pruned and with which autonomous system-level policy decisions can be enforced.

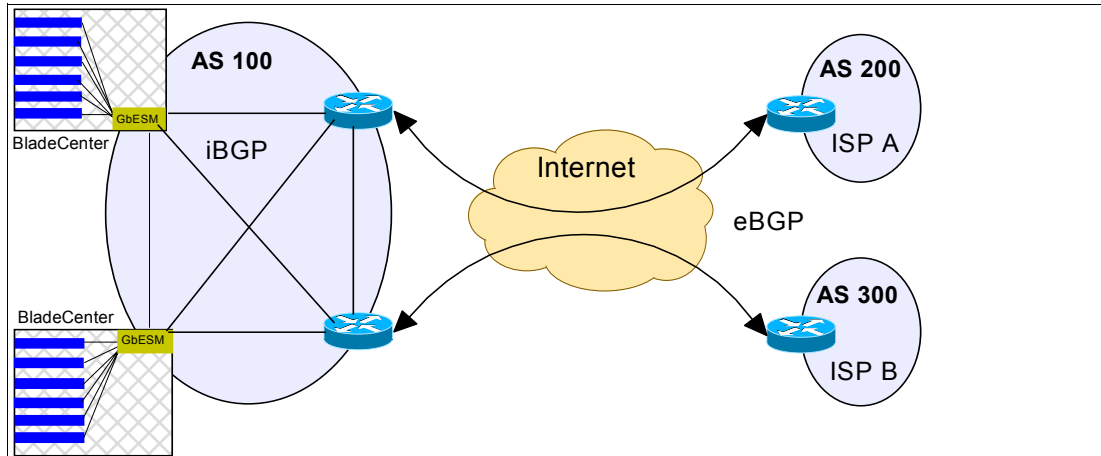


Figure 3-14 Using BGP to peer to two different ISPs

Each BGP router maintains a routing table that lists all feasible paths to a particular network. The router does not refresh the routing table, however. Instead, routing information received from peer routers is retained until an incremental update is received. BGP devices exchange routing information upon initial data exchange and after incremental updates. When a router first connects to the network, BGP routers exchange their entire BGP routing tables. Similarly, when the routing table changes, the routers send the portion of the routing table that has changed. BGP routers do not send regularly scheduled routing updates, and BGP routing updates advertise only the optimal path to a network. BGP uses a single routing metric to determine the best path to a given network. This metric consists of an arbitrary unit number that specifies the degree of preference of a particular link. The BGP metric typically is assigned to each link by the network administrator. The value assigned to a link can be based on any number of criteria, including the number of autonomous systems through which the path passes, stability, speed, delay, or cost.

Virtual Router Redundancy Protocol

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter supports high-availability network topologies through an enhanced implementation of the Virtual Router Redundancy Protocol (VRRP).

In a high-availability network topology, no device can create a single point-of-failure for the network or force a single point-of-failure to any other part of the network. To achieve this usually requires redundancy for all vital network components. With service availability becoming a major concern on the Internet, service providers are increasingly deploying Internet traffic control devices, such as application switches, in redundant configurations.

Hot-Standby redundancy

Traditionally, redundancy configurations have been various forms of hot standby, where one switch is active and the other is in a standby mode. One way to achieve hot standby with the Layer 2-7 GbE Switch Module for IBM @server BladeCenter is shown in Figure 3-15 on page 50.

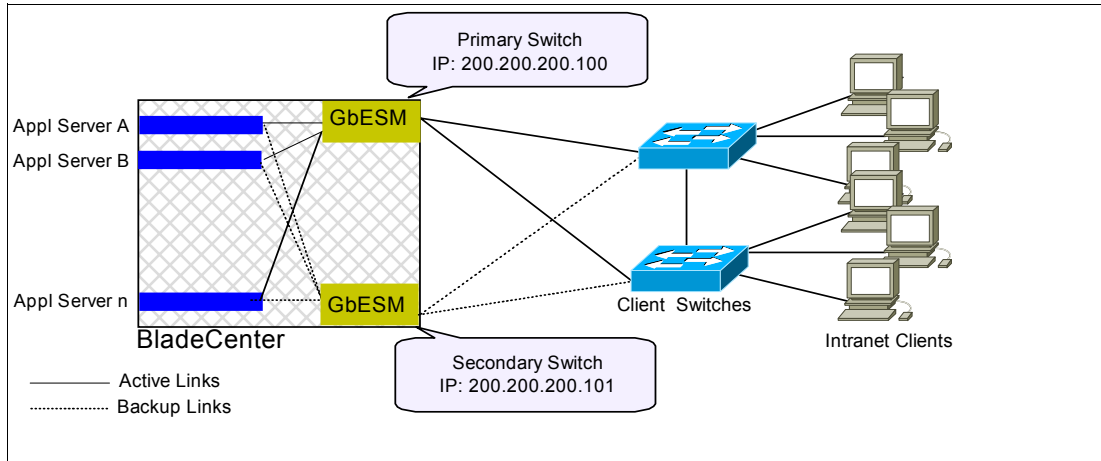


Figure 3-15 Hot-Standby redundancy

In this case, both Layer 2-7 GbE Switch Modules for IBM @server BladeCenter are providing routed access to the network for all server blades. The server blades are configured so that their default gateway is 200.200.200.100 and their secondary default gateway is 200.200.200.101. Should the primary Layer 2-7 GbE Switch Module for IBM @server BladeCenter fail, the server blades will eventually switch to using the secondary one, although the failover may take a few seconds.

Active-Standby redundancy

Another way to provide hot standby redundancy is by using VRRP between the Layer 2-7 GbE Switch Modules for IBM @server BladeCenter.

VRRP enables very fast failover for redundant router configurations within a LAN, providing alternate router paths for a host to eliminate single points-of-failure within a network. Each participating VRRP-capable routing device is configured with the same virtual router IP address and ID number. The virtual router IP address is configured as the default gateway on each server blade. No backup default gateway needs to be configured.

As shown in Figure 3-16 on page 51, one of the Layer 2-7 GbE Switch Modules for IBM @server BladeCenter is elected as the master, based on a number of priority criteria, and assumes control of the shared virtual router IP address. If the master fails, the backup Layer 2-7 GbE Switch Modules for IBM @server BladeCenter will take control of the virtual router IP address and actively process traffic addressed to it.

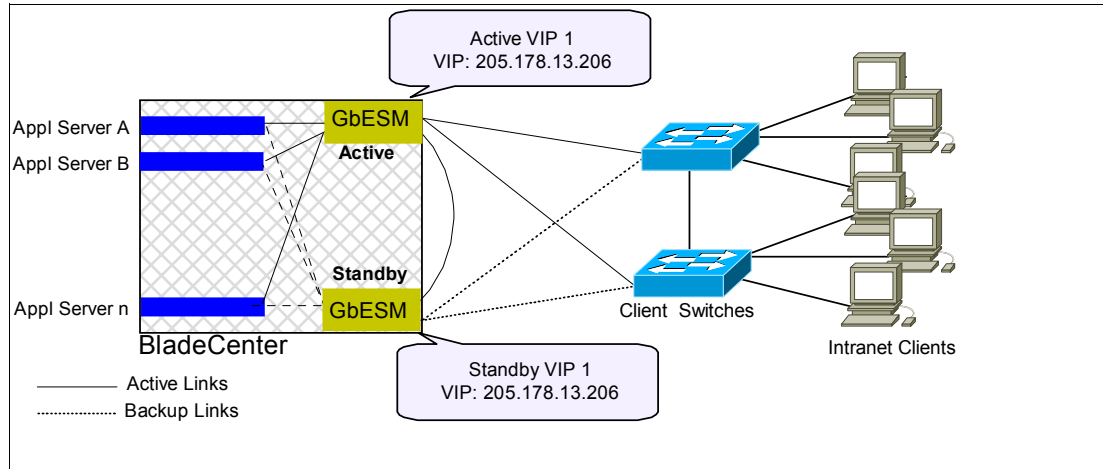


Figure 3-16 Active-Standby redundancy example

VRRP requires a healthy Layer 2 connection between all devices that are part of the same virtual router group, in this case the two Layer 2-7 GbE Switch Modules in the same BladeCenter. This connection is usually made via a switched connection over the external network. The exception to this is if the network devices are pure routers and cannot be used to create a Layer 2 connection between GbESMs. In that case, a direct physical link needs to be made between the GbESMs.

Virtual Routers and VRRP are important because there may be times when the Layer 2-7 GbE Switch Module for IBM @server BladeCenter is used as a Layer2/3 device for the purpose of achieving isolation between the external and internal Ethernet networks. Using VRRP, the Layer 3 functions of the Layer 2-7 GbE Switch Module for IBM @server BladeCenter can be made redundant without having to implement dynamic routing of any kind.

Active-Active redundancy

Although Active-Standby with VRRP increases site availability by removing single points-of-failure, and allows for fast failover and simple server blade configuration, it is not enough to meet most needs. Increasingly, end users view Active-Standby as an inefficient use of network resources because one functional application switch sits by idly until a failure calls it into action. Customers now demand that vendor equipment support redundant configurations where all devices can process traffic when they are healthy, increasing site throughput and decreasing user response times when no device has failed.

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter has extended VRRP to include virtual servers, allowing full active/active redundancy between its Layer 4 switches. In an active-active configuration, shown in Figure 3-17 on page 52, both switches can actively process traffic for given IP routing interfaces or load balancing virtual servers (VIPs).

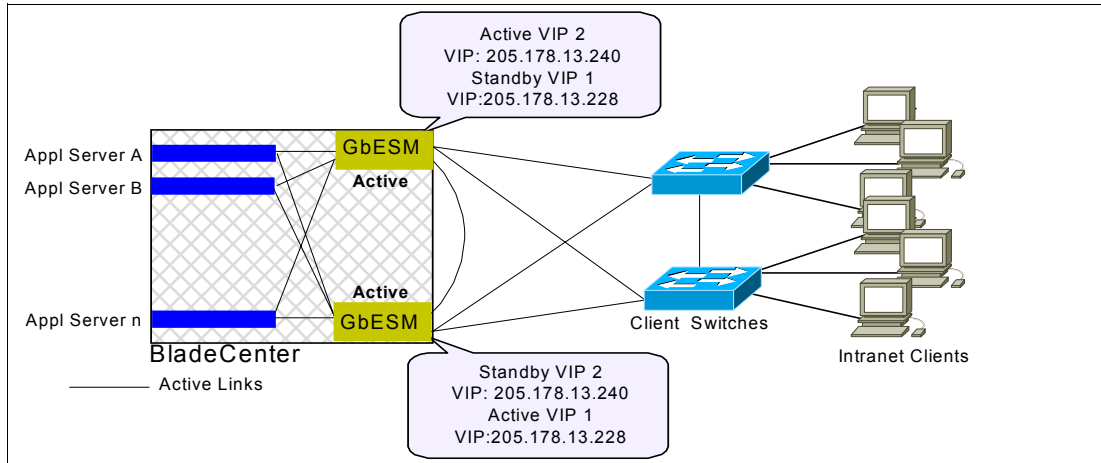


Figure 3-17 Active-Active redundancy

In this example, one switch is the master for VIP 1, but is the backup for VIP 2. The other switch is the master for VIP 2, but is the backup for VIP 1. Therefore, both switches are actively processing traffic simultaneously.

Active/active operation is very important because it:

- ▶ Justifies the investment in two Layer 2-7 GbE Switch Modules per BladeCenter.
- ▶ Increases the maximum effective I/O of the BladeCenter from 4 Gbps to 8 Gbps.
- ▶ Doubles the effective L4-7 switching horsepower within a BladeCenter, which can be useful if the Layer 2-7 GbE Switch Modules for IBM @server BladeCenter are configured to implement a lot of complex Layer 7 pattern matches, filters, features, and policies.

There are two ways to set up the services associated with the VIPs.

1. Both switches support active traffic but are configured so that they do not simultaneously support the same service. Each switch is active for its own set of services, such as IP routing interfaces or load-balancing virtual server IP addresses, and acts as a standby for other services on the other switch. If either switch fails, the remaining switch takes over processing for all services. The backup switch may forward Layer 2 and Layer 3 traffic, as appropriate.
2. The switches provide redundancy for each other, with both active at the same time for the same services. In other words, both VIP1 and VIP2 represent a set of real servers providing the same service.

Active-Active redundancy across two BladeCenters

The Inter-Chassis Redundancy Link (ICRL) feature allows you to design and build a massively scalable computing grid comprised of multiple BladeCenters. ICRL allows service pools to extend across chassis, and automates service fail-over upon the failure of server blades, Layer 2-7 GbE Switch Modules for IBM @server BladeCenter, or an entire chassis.

Multiple ICRL configurations are supported. Two or more BladeCenter chassis with one Layer 2-7 GbE Switch Module for IBM @server BladeCenter each can be configured for Active-Standby redundancy, where one chassis is processing the Virtual Server traffic. If you have multiple VIPs, you can utilize an Active-Active configuration, as shown in Figure 3-18 on page 53, to provide load sharing across both chassis.

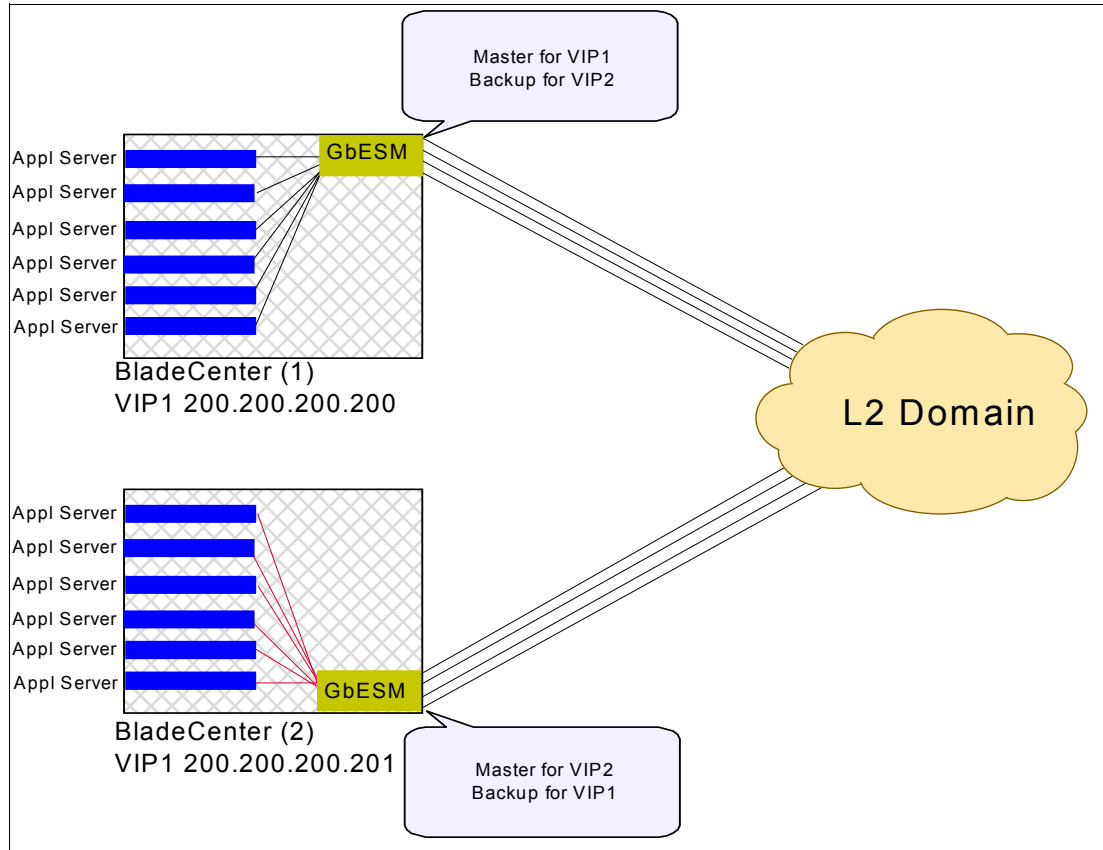


Figure 3-18 High availability across multiple chassis

Figure 3-18 demonstrates a case where redundancy is achieved with two BladeCenters and a Layer 2-7 GbE Switch Module for IBM @server BladeCenter module in each BladeCenter. Each Layer 2-7 GbE Switch Module for IBM @server BladeCenter is *Active* for a set of VIPs within its BladeCenter and *Standby* for the VIPs that the other one supports. Whether the Active VIPs in each BladeCenter represent the same service or not is transparent to the Layer 2-7 GbE Switch Modules for IBM @server BladeCenter and is an implementation option. Independent of whether the Active VIPs in each BladeCenter represent the same service, application images and application data/content need to be accessible within or from each BladeCenter so that they can take over operation for the VIPs they are acting as Standby for. When all devices are healthy, each BladeCenter has 4 Gbps of aggregate I/O throughput, for an overall throughput of 8 Gbps. (Note that more BladeCenters could have been used to achieve nx4 Gbps throughput).

To further scale your computing grid, you can install a second Layer 2-7 GbE Switch Module for IBM @server BladeCenter to each of your BladeCenters and configure the Virtual Servers for which you wish to provide failover. This provides high availability within the chassis as well as across multiple chassis. This topology can be scaled further by adding additional BladeCenters (equipped with Layer 2-7 GbE Switch Module for IBM @server BladeCenter) to the design, as shown in Figure 3-19 on page 54.

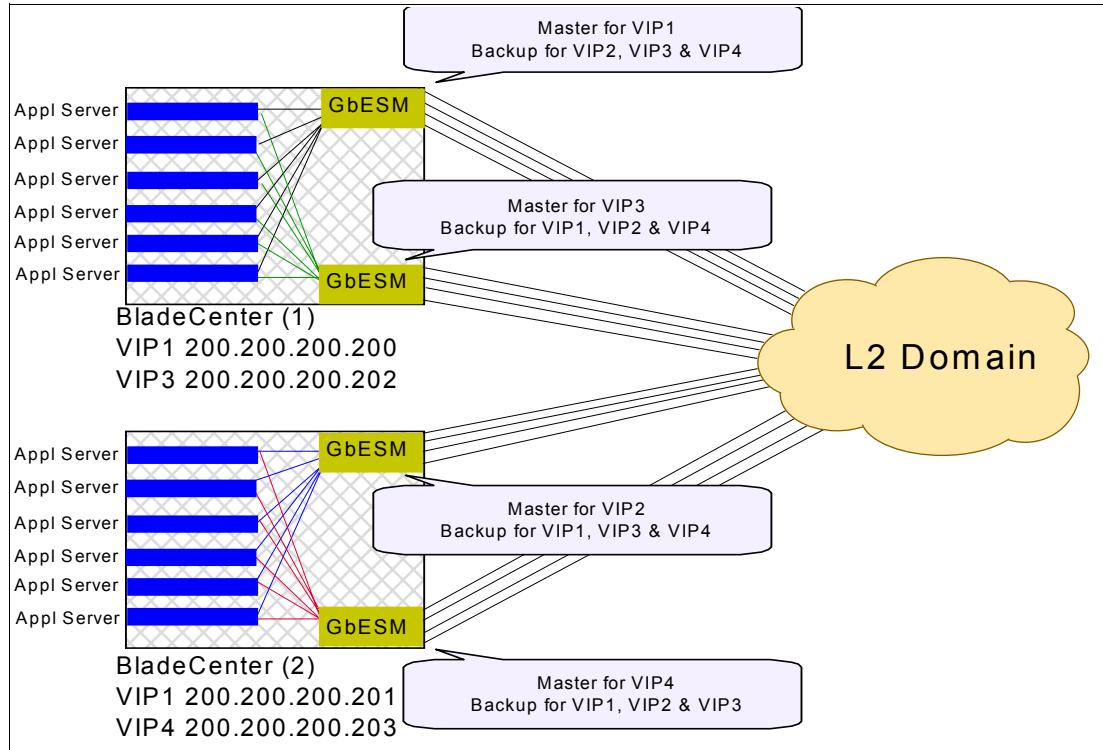


Figure 3-19 High availability within the chassis as well across multiple chassis

In Figure 3-19, when everything is healthy, each Layer 2-7 GbE Switch Module for IBM @server BladeCenter in each BladeCenter is actively supporting one of the Virtual Services. Each of the Layer 2-7 GbE Switch Modules for IBM @server BladeCenter is configured to be a backup for every service, providing quadruple redundancy.

The configuration provides 16 Gbps I/O throughput.

If each Layer 2-7 GbE Switch Module for IBM @server BladeCenter is connected to a different access switch or router, then no single failure will cause a BladeCenter to be isolated. However, if a Layer 2-7 GbE Switch Module for IBM @server BladeCenter fails or an access switch/router fails, overall I/O throughput will decrease by 4 Gbps.

3.4 Layer 4: Transport layer

The *transport layer* is responsible for end-to-end integrity of transmissions. The layer detects packets that are damaged or lost in transit and can automatically generate a retransmit request.

The layer also has the responsibility of resequencing packets that arrive out of order. This can be caused by the packets taking different paths through the network or due to retransmitting packets that have been damaged during transit.

The layer uses the Transmitting Control Protocol (TCP) and the User Datagram Protocol (UDP) for the end-to-end transport of packets. TCP is a connection-oriented protocol, while the UDP is a connectionless protocol.

The life of a TCP connection

Layer 4 switching operates on units of TCP connections, as shown in Figure 3-20 on page 55. You can't spray packets that are part of the same TCP connection across multiple servers. That means that a Layer 4 switch needs to be TCP-aware in the sense that it needs to be able to:

- ▶ Identify the beginning of a new TCP connection
- ▶ Assign that connection to a real server
- ▶ Make sure that all ensuing packets related to that TCP connection continue to be sent to the same real server

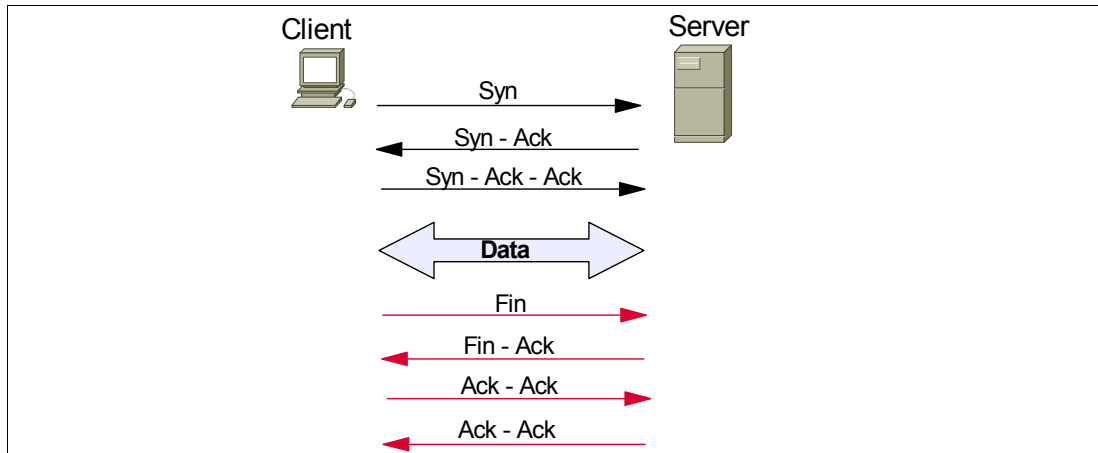


Figure 3-20 The life of a TCP connection

TCP connections get started when a client sends a special TCP packet called a SYN packet (identified by setting a special bit in the TCP header) to the desired destination to initiate the creation of a connection. If the server is ready, willing, and able to respond, it will reply with a SYN-ACK packet that contains some parameters that need to be agreed to and used consistently by both sides of the connection. The client then sends a SYN-ACK-ACK to the server, often with some actual data or a content request included (or immediately sends a request packet after sending the SYN-ACK-ACK), to kick off the real conversation part of the connection.

When the data transfer related to the TCP connection is complete from the client's point of view, it will send a FIN packet to the server and start a final process of handshaking to tear down the connection. Alternatively, if the server hasn't heard from the client in a while, it will time out the connections. Figure 3-20 shows the life of a TCP connection.

Delayed binding

With L4 switching, TCP connection dynamics change a little bit. L4 switching implements the concept of "Virtual Services" that are indexed by "Virtual IP Addresses" (or VIPs). Let's use a specific example to explain how this works, as shown in Figure 3-21 on page 56. Consider a service whose host name in a DNS sense is "A.com." Normally, there would be a server set up somewhere with that host name and IP address 100.2.2.2 (for example). With L4 switching, a Layer 2-7 GbE Switch Module for IBM @server BladeCenter takes ownership if the address 100.2.2.2 is a VIP. The GbESM has multiple (in this example, two) "real" server blades behind it capable of delivering the service A.com, whose addresses can be anything – they are only of local significance. In this case, let's call them 10.1 and 10.2.

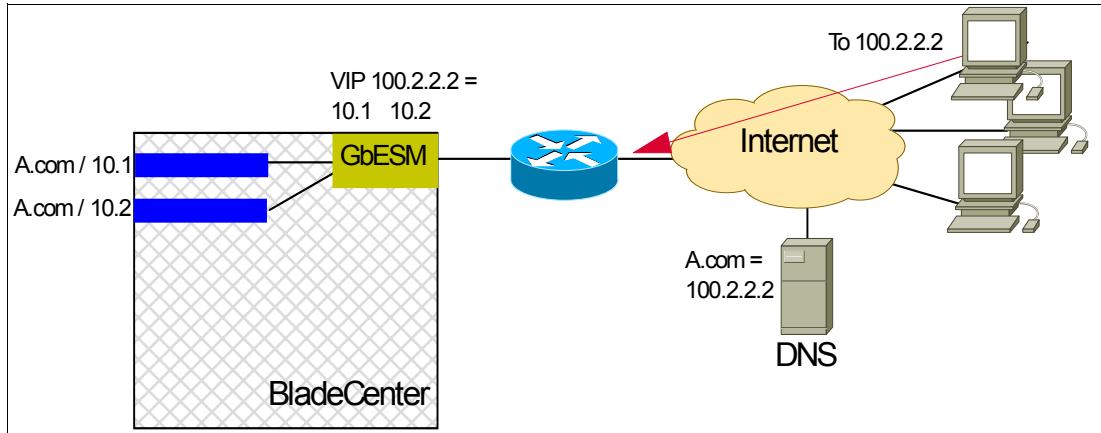


Figure 3-21 The life of an L4 switched TCP connection

1. When the GbESM receives a packet from the user, it inspects the packet to determine what physical server to send it to. If the packet belongs to an existing connection or user, the packet is sent to the server already assigned to that connection.
2. If the packet does not belong to an existing connection, but instead is a SYN packet from a user desiring to establish a new connection, the GbESM determines which physical server to send the request to, based on some knowledge of the health and load of the candidate server blades, and some policy for making such a decision.
3. Once one of the servers is chosen, the scheduler uses a re-direction mechanism to dispatch the request to the chosen server blade. The request must arrive to that server blade addressed to its IP address and MAC address.
4. At this time, the GbESM creates some state information for this connection. The GbESM maintains a binding table, which reflects the present association of users (or connections, or sessions) to server blades.
5. The chosen server blade receives the packet, which has its own IP address in the destination field; thus to it this looks like a perfectly normal TCP connection request.
6. If the chosen application server is ready and able, it answers the user with a TCP SYN-ACK. When received by the user, the TCP SYN-ACK must have the VIP address in the Source IP Address field. This is achieved by having the GbESM intercept the returned packet to translate the source IP address of the packet back into the Virtual IP address.
7. What would happen if the user's computer received a packet that had the chosen server blade's IP address in the source field instead of the virtual server's address? The user's computer would reject this frame with a TCP RESET – it would notice the imbalance between the destination IP address that it sent the TCP SYN request to, and the source IP address of the responding server for the TCP SYN-ACK.
8. The user's computer sends an ACK for the TCP SYN-ACK, and completes the 3-way handshake. The GbESM, again, is the recipient of this frame because of the destination IP address. The GbE Switch Module inspects the packet, determines it is associated with an existing session (based on the user's IP address at a minimum) and sends the packet to the same chosen server blade. What would happen if the Layer 2-7 GbE Switch Module for IBM @server BladeCenter did not maintain session state, and it sent the ACK frame to a different server blade? The "new" server blade would receive a SYN-ACK-ACK frame associated with a session that was never started with a SYN, as far as it is concerned. It would reply back with a TCP RESET.

9. Every packet coming from the user to the load-balanced set of server blades must pass through the Layer 2-7 GbE Switch Module for IBM @server BladeCenter so that it can inspect the frames and perform the appropriate load balancing mechanism.

In executing this process, the GbE Switch Module must execute a couple of critical functions:

1. Deciding the “best available” server.
2. Modifying packets to and from the target server to ensure the proper connectivity throughout the life of the TCP connection.

Some of the options for these two critical functions are described in the following sections.

Metrics for real server groups

Metrics are used for selecting which server blade in a group will receive the next client connection. There are several options for metrics. When considering what metric or combination of metrics to use, you should remember that perfectly equal load balancing between server blades in a Virtual Service Pool is generally difficult to achieve, but is also generally irrelevant. Usually, server blades deliver application services very well when lightly loaded and as the load increases, until a saturation point is reached, representing “a knee in the load versus latency curve,” at which point serious problems can occur, including server failure. The main purpose of load balancing is to maintain application availability at whatever level of performance the healthy resources allow. What this generally means in practice is to keep all servers working at operating points below the knee of the curve.

Minimum misses

For L4 switching, the client source IP address and real server blade IP address are used. All requests from a specific client are sent to the same server. This metric is useful for applications where client information must be retained on the server between sessions. With this metric, server load becomes most evenly balanced as the number of active clients with different source or destination addresses increases. When selecting a server, the switch calculates a value for each available real server based on the relevant IP address information. The server with highest value is assigned the connection. The *minmisses* metric attempts to minimize the disruption of persistency when servers are removed from service. This metric should be used only when persistence is desired. It is a weak form of persistency and is not sufficient to uniquely identify a client.

Hash

The *hash* metric uses IP address information in the client request to select a server. For SLB, the client source IP address is used. All requests from a specific client will be sent to the same server. This option is useful for applications where client information must be retained between sessions.

Note: The hash metric provides more distributed load balancing than minimum misses at any given instant. It should be used if the statistical load balancing achieved using minmisses is not as optimal as desired. If the load balancing statistics with minmisses indicate that one server is processing significantly more requests over time than other servers, consider using the hash metric.

Least connections

With the *leastconns* metric, the number of connections currently open on each real server is measured in real time. The server with the fewest current connections is considered to be the best choice for the next client connection request. This option is the most self-regulating, with the fastest servers typically getting the most connections over time.

Round robin

With the *roundrobin* metric, new connections are issued to each server in turn; that is, the first real server in the group gets the first connection, the second real server gets the next connection, followed by the third real server, and so on. When all the real servers in this group have received at least one connection, the issuing process starts over with the first real server.

Response time

The *response* metric uses real server response time to assign sessions to servers. The response time between the servers and the switch is used as the weighting factor. The switch monitors and records the amount of time it takes for each real server to reply to a health check to adjust the real server weights. The weights are adjusted so they are inversely proportional to a moving average of response time. In such a scenario, a server with half the response time of another server will receive a weight twice as large.

Bandwidth

The *bandwidth* metric uses real server octet counts to assign sessions to a server. The switch monitors the number of octets sent between the server and the switch. Then, the real server weights are adjusted so they are inversely proportional to the number of octets that the real server processes during the last interval. Servers that process more octets are considered to have less available bandwidth than servers that have processed fewer octets. For example, the server that processes half the amount of octets over the last interval receives twice the weight of the other servers. The higher the bandwidth used, the smaller the weight assigned to the server. Based on this weighting, the subsequent requests go to the server with the highest amount of free bandwidth. These weights are automatically assigned. The bandwidth metric requires identical servers with identical connections.

Weights for real servers

Weights can be assigned to each real server. These weights can bias load balancing to give the fastest real servers a larger share of connections. Weight is specified as a number from 1 to 48. Each increment increases the number of connections the real server gets. By default, each real server is given a weight setting of 1. A setting of 10 would assign the server roughly 10 times the number of connections as a server with a weight of 1.

Note: Weights are not applied when using the hash or minmisses metrics.

Connection time-outs for real servers

In some cases, open TCP/IP sessions might not be closed properly (for example, the switch receives the SYN for the session, but no FIN is sent). If a session is inactive for 10 minutes (the default), it is removed from the session table in the switch.

Maximum connections for real servers

You can set the number of open connections each real server is allowed to handle for SLB. Values average from approximately 500 HTTP connections for slower servers to 1500 for quicker, multiprocessor servers. The appropriate value also depends on the duration of each session and how much CPU capacity is occupied by processing each session. Connections that use a lot of Java or CGI scripts for forms or searches require more server resources and thus a lower *maxcon* limit. You may wish to use a performance benchmark tool to determine how many connections your real servers can handle.

When a server reaches its *maxcon* limit, the switch no longer sends new connections to the server. When the server drops back below the *maxcon* limit, new sessions are again allowed.

Using delayed binding to prevent DoS attacks

The delayed binding feature on the switch prevents SYN Denial-of-Service (DoS) attacks on the server. DoS occurs when the server or switch is denied servicing the client because it is saturated with invalid traffic.

Typically, a three-way handshake occurs before a client connects to a server. The client sends out a synchronization (SYN) request to the server. The server allocates an area and sets aside some resource (such as memory) to process the client requests, and acknowledges the client by sending a SYN ACK. The client then acknowledges the SYN ACK by sending an acknowledgement (ACK) back to the server, thus completing the three-way handshake. If the client does not acknowledge the server's SYN ACK with a data request (REQ) and, instead, sends another SYN request, the server gets saturated with SYN requests. As a result, all of the server's resources are consumed and it can no longer service legitimate client requests.

In the GbE Switch Module, SYN attack detection is enabled by default, whenever delayed binding is enabled. GbESM SYN attack detection:

- ▶ Provides a way to track half-open connections
- ▶ Activates a trap notifying that the configured threshold is exceeded
- ▶ Monitors DoS attacks and proactively signals alarm
- ▶ Provides enhanced security
- ▶ Improves visibility and protection for DoS attacks
- ▶ Times out half-open sessions which appear in the session table as the result of a DoS attack.

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter does not allocate a session until a valid SYN ACK is received from the client or the three-way handshake is complete.

Once the GbESM receives a valid ACK or DATA REQ from the client, it sends a SYN request to the server on behalf of the client, waits for the server to respond with a SYN ACK, and then forwards the client's DATA REQ to the server. Basically, the GbESM delays binding the client session to the server until the proper handshakes are complete. Thus, with delayed binding, two independent TCP connections span a session: one from the client to the Layer 2-7 GbE Switch Module for IBM @server BladeCenter, and the second from the switch to the selected server. The switch temporarily terminates each TCP connection until content has been received, thus preventing the server from being inundated with SYN requests.

Mapping ports

A Layer 2-7 GbE Switch Module for IBM @server BladeCenter allows you to hide the identity of a port for security by mapping a virtual server port to a different real server port.

Mapping is required when administrators choose to execute their real server processes on different ports than the well-known TCP/UDP ports. Otherwise, virtual server ports are mapped directly to real server ports by default and require no mapping configuration.

To take advantage of multi-CPU or multi-process servers, Layer 2-7 GbE Switch Module for IBM @server BladeCenter can be configured to map a single virtual port to multiple real ports. This capability allows the site managers, for example, to differentiate users of a service by using multiple service ports to process client requests.

In Figure 3-22 on page 60, four real servers are used to support a single service (HTTP). Clients access this service through a virtual server with IP address 192.168.2.100 on virtual port 80.

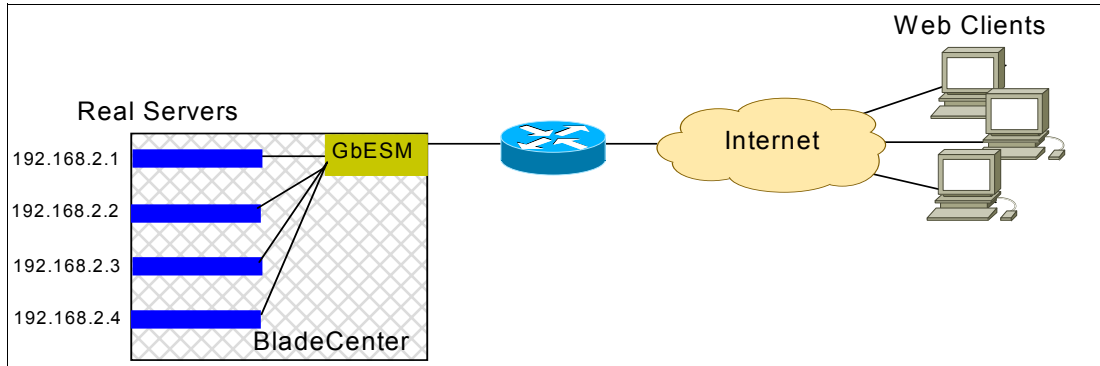


Figure 3-22 Basic Virtual Port to Real Port Mapping Configuration

Since each real server uses two ports (8001 and 8002, as shown in Figure 3-23) for HTTP services, the logical real servers are:

- 192.168.2.1/8001
- 192.168.2.1/8002
- 192.168.2.2/8001
- 192.168.2.2/8002
- 192.168.2.3/8001
- 192.168.2.3/8002
- 192.168.2.4/8001
- 192.168.2.4/8002

Domain name	Virtual server IP address	Ports activated	Port mapping	Real server IP address
www.A.com	192.168.2.100	80 (HTTP)	8001 (rport 1) 8002 (rport 2)	192.168.2.1 (RIP 1) 192.168.2.2 (RIP 2) 192.168.2.3 (RIP 3) 192.168.2.4 (RIP 4)

Figure 3-23 Virtual server ports

Layer 4 Filtering

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter can be used to filter traffic according to various IP address, protocol, and Layer 4 port criteria.

TCP health checks

The GbESM monitors the health of servers and applications by sending Layer 4 connection requests (TCP SYN packets) for each load-balanced TCP service to each server in the server group on a regular basis. The rate at which these connection requests are sent is a user-configurable parameter. These connection requests identify both failed servers and failed services on a healthy server. When a connection request succeeds, the session switch quickly closes the connection by sending a TCP FIN (finished) packet.

Note: TCP health check is a default health check after you have configured the switch for a particular service.

TCP rate limiting

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter allows you to prevent a client or a group of clients from claiming all the TCP resources on the servers. This is done by monitoring the rate of incoming TCP connection requests to a virtual IP address and limiting the client requests with a known set of IP addresses. The TCP rate limit is defined as the

maximum number of TCP connection requests within a configured time window. The switch monitors the number of new TCP connections and when it exceeds the configured limit, any new TCP connection request is blocked. When this occurs, the client is said to be held down. The client is held down for a specified period of time, after which new TCP connection requests from the client are allowed to pass through again.

The main goals of L4-7 switching is to prevent server resources from being overloaded. Often, it is not the total number of connections or active sessions or pending requests that overloads a server, but the arrival rate of new connections and requests. One way the TCP rate limiting feature can be used is to set the client group to all clients and set a maximum arrival rate. If loads exceed that, the server will not have new requests sent to it until the arrival rate falls back below the set threshold.

3.5 Layer 5: Session layer

The *session layer* is responsible for setting up, managing, and tearing down sessions between presentation layer entities. One application's data is kept separated from any other application's data by the layer.

The layer also controls whether the sessions that are established use half-duplex or full-duplex mode. Half-duplex mode means that only one device can transmit at any time, while full-duplex mode means that two devices can transmit simultaneously.

For a full-duplex connection to have flow control, both devices must support a common protocol, such as IEEE 802.3x. In this scheme, a congested device sends a special MAC Control PAUSE frame to instruct the other machine to stop transmission for a specific period of time. The GbE Switch Module supports 802.3x, which is automatically enabled on all internal ports. Since the blade servers also support 802.3x, we know that full-duplex flow control exists on internal connections. However, any external device connected to the GbE Switch Module must also support 802.3x to ensure that PAUSE frames are generated and interpreted correctly. For example, if the GbESM has a full-duplex link to another switch which does not use 802.3x, then either the GbESM or the external switch will discard frames during congestion.

3.6 Layer 6: Presentation layer

The *presentation layer* presents the data to the application layer. The layer acts as a translator and ensures that data transferred from the application layer of one system can be read by the application layer of another host.

3.7 Layer 7: Application layer

The *application layer* provides the interface between user applications and the network's services. When a client asks for a Web page on a server, the application of the client sends a request to the appropriate Layer 7 protocol. The protocol initiates a communication session that traverses down the OSI stack as shown in Figure 3-24 on page 62.

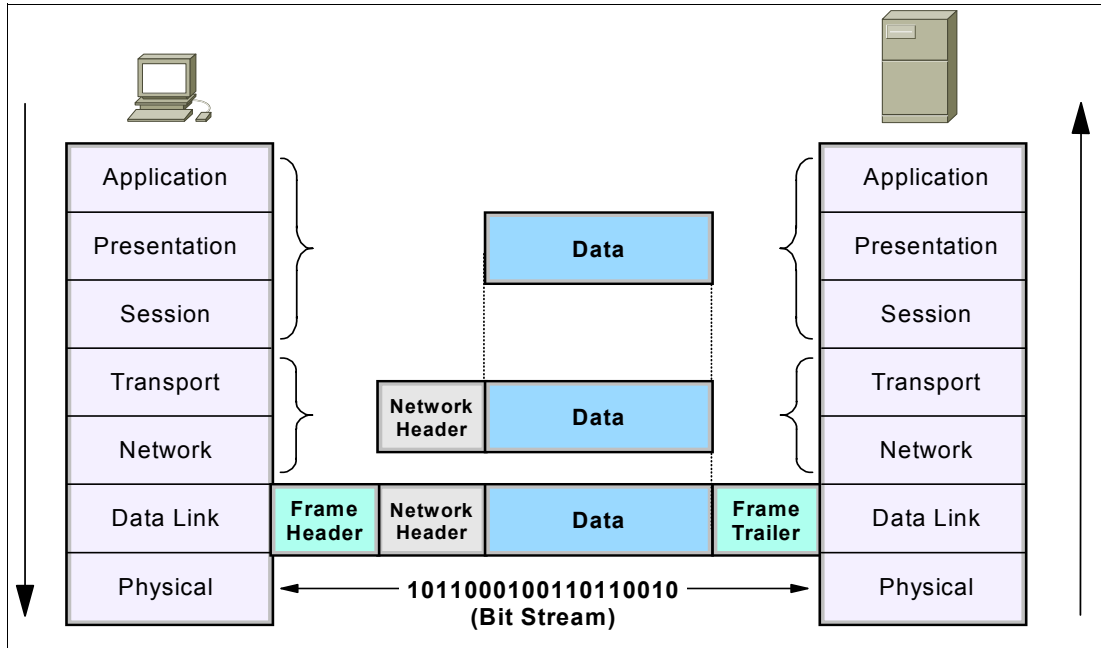


Figure 3-24 The OSI stack

This section discusses the Layer 7 capabilities of the Layer 2-7 GbE Switch Module for IBM @server BladeCenter, with most of the examples based on HTTP environments.

Some of the benefits that can be achieved by deploying Layer 7 switching are:

- ▶ Web sites may have so much content associated with their domain name that it needs to be split up across multiple file systems. One could allow each Web server access to each file system by cross-mounting all the file systems, but this gets unwieldy as the number of file systems gets larger and if it routinely changes. Another approach is to assign access to portions of the directory space to certain Web server clusters, but still advertise the site under one domain name, such as www.ibm.com@. The GbESM, as a front end to this site, must be able to inspect the URL request (including filename and path-name) and send the requests for www.ibm.com/marketing/ to one server, /research/ to another server, /admin/ to yet another server, and so forth.
- ▶ There may be a marketing advantage to advertising a single domain name for a particular service, but supporting specific and separate communities behind that single domain name. In that way, the number and type of servers supporting each community can be tuned to their needs, and, if different levels of service are warranted between the communities, traffic can be managed to achieve that.
- ▶ It may be useful to have servers that generate dynamic content optimized for the computation/processing function, and servers that provide relatively static data (like images, HTML text, and so forth) optimized for fast retrieval from disk storage. URL awareness allows the GbESM to look for requests with dynamic server page calls or CGI script executions, versus static Web page element requests. Dynamic requests would be sent to the application-optimized servers and the static requests would be sent to the storage-optimized servers.

Layer 7 load balancing

Tip: This section does not include all possible scenarios for layer 7 load balancing. For more information about other layer 7 load balancing scenarios refer to the *Alteon OS 20.0 Application Guide*.

URL-based server load balancing

URL-based server load balancing allows you to optimize resource access and server performance. Content dispersion can be optimized by making load-balancing decisions on the entire path and filename of each URL. URL requests are load balanced among multiple servers matching the URL, according to the load balancing metric configured for the real server group (leastConns is the default).

In Figure 3-25, the following criteria are specified for content load balancing:

- ▶ Requests with the string “/images” in the URL are sent to real server group 1.
- ▶ Requests with “.bin” in the URL are forwarded to real server group 2.
- ▶ Requests with URLs starting with “/product:” are sent to real server group 1 and 3.
- ▶ Requests containing URLs with anything else are sent to real server group 1, 2, and 3. These servers have been defined with the “any” string.

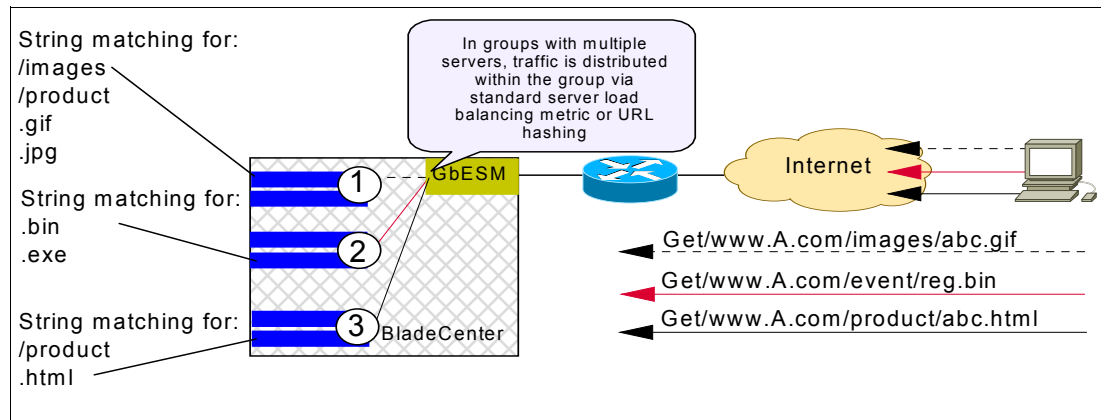


Figure 3-25 URL-based server load balancing

Cookie-based preferential load balancing

Cookies can be used to provide preferential services for customers, ensuring that certain users are offered better access to resources than other users when site resources are scarce. For example, a Web server could authenticate a user via a password and then set cookies to identify them as “Gold,” “Silver,” or “Bronze” customers. Using cookies, you can distinguish individuals or groups of users and place them into groups or communities that get redirected to better resources and receive better services than all other users.

Cookie-based preferential services enable the following support:

- ▶ Redirect higher priority users to a larger server or server group.
- ▶ Identify a user group and redirect them to a particular server.
- ▶ Serve content based on user identity.
- ▶ Prioritize access to scarce resources on a Web site.

- ▶ Provide better services to repeat customers, based on access count. Clients that receive preferential service can be distinguished from other users by one of the following methods:
 - Individual user: A specific individual user could be distinguished by IP address, login authentication, or permanent HTTP cookie.
 - User communities: Some set of users, such as “Premium Users” for service providers who pay higher membership fees than “Normal Users” could be identified by source address range, login authentication, or permanent HTTP cookie.
 - Applications: Users could be identified by the specific application they are using. For example, priority can be given to HTTPS traffic that is performing credit card transactions versus HTTP browsing traffic.
 - Content: Users could be identified by the specific content they are accessing.

Based on one or more of these criteria, you can load balance requests to different server groups.

Browser-smart load balancing

Sometimes some browser types can't handle certain data types, image formats, and the like. When that is the case, it may be useful to set up specific servers to service specific browser types. HTTP requests can be directed to different servers based on browser type by inspecting the “User-Agent” header as shown in Example 3-1.

Example 3-1 Inspecting “User-Agent” header

```
GET /products/2224/ HTTP/1.0
User-agent: Mozilla/3.0
Accept: text/html, image/gif, image/jpeg
```

DNS load balancing

The Internet name registry has become so large that a single server cannot keep track of all the entries. This is resolved by splitting the registry and saving it on different servers. If you have large DNS server farms, the GbESM allows you to load balance traffic based on DNS names. To load balance DNS names, the host name is extracted from the query, processed by the regular expressions engine, and the request is sent to the appropriate real server as shown in Figure 3-26.

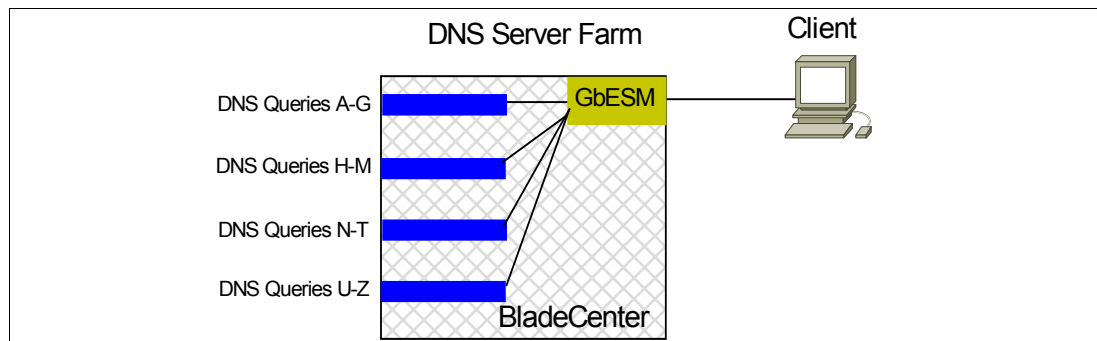


Figure 3-26 DNS server farm load balancing

The figure shows a DNS server farm load balancing DNS queries based on DNS names. Requests with DNS names beginning with A through G are sent to Server 1; DNS names beginning with H through M are sent to Server 2; DNS names beginning with N through T are sent to Server 3; DNS names beginning with U through Z are sent to Server 4.

Layer 7-based persistence

In a typical server load balancing environment, traffic comes from various client networks across the Internet to the virtual server IP address on the GbE Switch Module. The switch then load balances this traffic among the available real servers.

In any authenticated Web-based application, it is necessary to provide a persistent connection between a client and the content server to which it is connected. Because HTTP does not carry any state information for these applications, it is important for the browser to be mapped to the same real server for each HTTP request until the transaction is complete. This ensures that the client traffic is not load balanced mid-session to a different real server, forcing the user to restart the entire transaction.

Persistence-based server load balancing enables the network administrator to configure the network to redirect requests from a client to the same real server that initially handled the request. Persistence is an important consideration for administrators of e-commerce Web sites, where a server may have data associated with a specific user that is not dynamically shared with other servers at the site. In the GbE Switch Module, persistence can be based on the following characteristics: source IP address, cookies, and Secure Sockets Layer (SSL) session ID.

Tip: This section does not include in-depth information about persistence. For more information, refer to the *Alteon OS 20.0 Application Guide*.

Source IP address

Until recently, the only way to achieve TCP/IP session persistence was to use the source IP address as the key identifier. There are two major conditions which cause problems when session persistence is based on a packet's IP source address:

1. Many clients sharing the same source IP address (proxied clients). Proxied clients appear to the switch as a single source IP address and do not take advantage of server load balancing on the switch. When many individual clients behind a firewall use the same proxied source IP address, requests are directed to the same server, without the benefit of load balancing the traffic across multiple servers. Persistence is supported without the capability of effectively distributing traffic load. Also, persistence is broken if you have multiple proxy servers behind the switch performing server load balancing. The switch changes the client's address to different proxy addresses as attempts are made to load balance client requests
2. Single clients sharing a pool of source IP addresses. When individual clients share a pool of source IP addresses, persistence for any given request cannot be assured. Although each source IP address is directed to a specific server, the source IP address itself is randomly selected, thereby making it impossible to predict which server will receive the request. Server load balancing is supported, but without persistence for any given client.

SSL session ID-based persistence

SSL is a set of protocols built on top of TCP/IP that allows an application server and client to communicate over an encrypted HTTP session, providing authentication, non-repudiation, and security. The SSL protocol handshake is performed using clear (unencrypted) text. The content data is then encrypted (using an algorithm exchanged during the handshake) prior to being transmitted.

Using the SSL session ID, the switch forwards the client request to the same real server to which it was bound during the last session. Because SSL protocol allows many TCP connections to use the same session ID from the same client to a server, key exchange needs to be done only when the session ID expires. This reduces server overhead and

provides a mechanism, even when the client IP address changes, to send all sessions to the same real server, as shown in Figure 3-27.

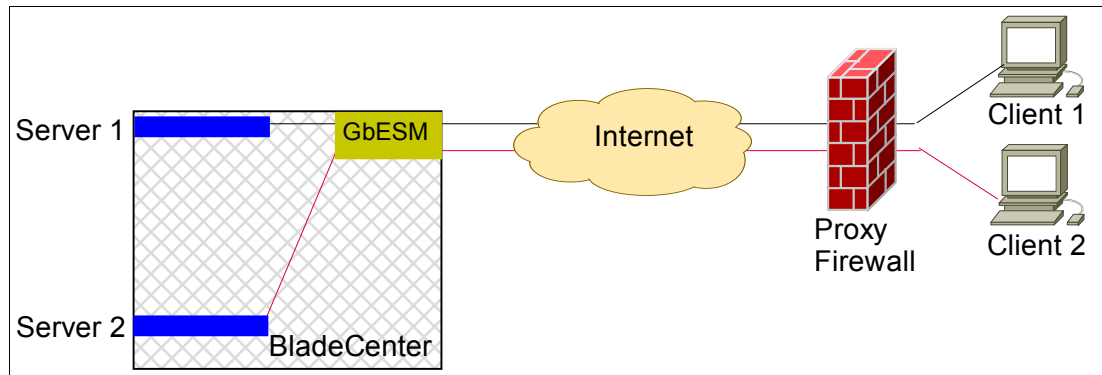


Figure 3-27 SSL session ID-based persistence

1. An SSL Hello handshake occurs between Client 1 and Server 1 via the GbE Switch Module.
2. An SSL session ID is assigned to Client 1 by Server 1.
3. The switch records the SSL session ID.
4. The switch selects a real server based on the existing server load balancing settings. As a result, subsequent connections from Client 2 with the same SSL session ID are directed to Server 1.
5. Client 2 appears to the Layer 2-7 GbE Switch Module for IBM @server BladeCenter to have the same source IP address as Client 1 because they share the same proxy firewall. However, the switch does not automatically direct Client 2 traffic to Server 1 based on the source IP address. Instead, an SSL session ID for the new traffic is assigned. Based on server load balancing settings, the connection from Client 2 is spliced to Server 2. As a result, subsequent connections from Client 2 with the same SSL session ID are directed to Server 2.

“Shopping cart” persistency and cookie-based persistence

Cookies are a mechanism for maintaining state between clients and servers. When the server receives a client request, the server issues a cookie, or token, to the client, which the client then sends to the server on all subsequent requests. Using cookies, the server does not require authentication, the client IP address, or any other time-consuming mechanism to determine that the user is the same user that sent the original request.

In the simplest case, the cookie may be just a “customer ID” assigned to the user. It may be a token of trust, allowing the user to skip authentication while his or her cookie is valid. It may also be a key that associates the user with additional state data that is kept on the server, such as a shopping cart and its contents. In a more complex application, the cookie may be encoded so that it actually contains more data than just a single key or an identification number. The cookie may contain the user’s preferences for a site that allows their pages to be customized.

The cookie-based persistence feature solves the proxy server problem and gives better load distribution at the server site. In the switch, cookies are used to route client traffic back to the same physical server to maintain session persistence.

Figure 3-28 on page 67 shows the different steps in cookie-based persistence.

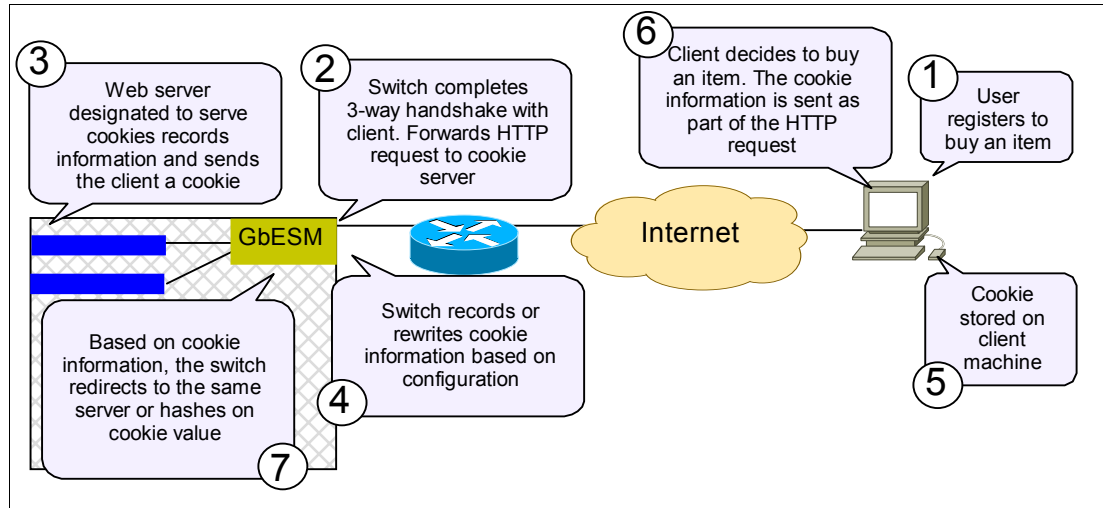


Figure 3-28 Cookie-based persistence

Health checking

Layer 2-7 GbE Switch Module for IBM @server BladeCenter running server load balancing monitors the servers in the real server group and the load-balanced applications running on them. If a switch detects that a server or application has failed, it will not direct any new connection requests to that server. When a service fails, the switch can remove the individual service from the load balancing algorithm without affecting other services provided by that server.

Tip: This section does not include in-depth information about health checking. For more information, refer to the *Alteon OS 20.0 Application Guide*.

Application-specific health checks

Many times, an application process can fail without the knowledge of the TCP or IP processes. In that case, a ping test or even a TCP connection test can succeed, even though the server is not able to service application requests. To detect such failures, a higher-level test, actually exercising the application logic, is required.

Over the years, the software that the Layer 2-7 GbE Switch Module for IBM @server BladeCenter is based on has been used in a lot of different environments for a lot of different purposes. The GbESM uses a suite of application-level tests, some pre-defined and others with user/field definable parameters, that can be used to exercise application logic.

The presently available pre-defined tests are the following:

- ▶ HTTP Health Checks
- ▶ UDP-Based DNS Health Checks
- ▶ FTP Server Health Checks
- ▶ POP3 Server Health Checks
- ▶ SMTP Server Health Checks
- ▶ IMAP Server Health Checks
- ▶ NNTP Server Health Checks
- ▶ RADIUS Server Health Checks
- ▶ HTTPS/SSL Server Health Checks
- ▶ WAP Gateway Health Checks
- ▶ LDAP Health Checks
- ▶ ARP Health Checks

Script-based health checks

Although the Layer 2-7 GbE Switch Module for IBM @server BladeCenter has a lot of “canned” health checks for testing a lot of environments, the number of applications and environments the switch can be used in is far greater than the full extent of these tests. Therefore, a mechanism has been created to allow user/field creation of specific tests related to specific applications. These “send/expect” structured health checks dynamically verify application and content availability using scripts. These scripts execute a sequence of tests, all of which are oriented towards sending a particular request and checking for delivery of the expected response.

Layer 7 filtering

The GbESM allows you to secure your switch from virus attacks by configuring the switch with a list of potential offending string patterns (HTTP URL requests). The switch examines the HTTP content of the incoming client request for the matching string pattern. If the matching virus pattern is found, then the packet is dropped, as shown in Figure 3-29, and a reset frame is sent to the offending client. SYSLOG messages and SNMP traps are generated to warn operators of a possible attack.

A layer 7 deny filter is basically a deny filter except that the deny action is delayed until HTTP content is examined to see if the packet should be denied.

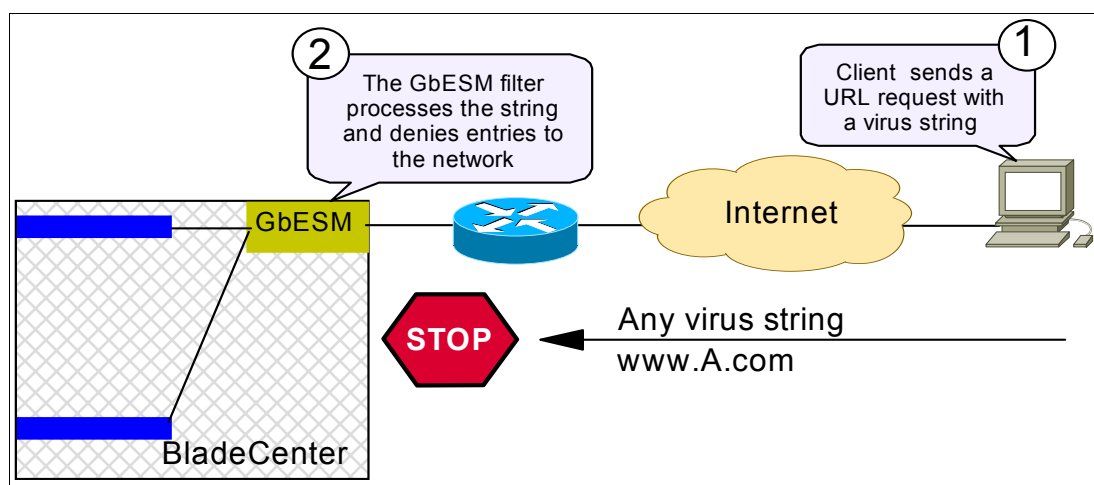


Figure 3-29 Layer 7 deny filter

Application redirection

Application redirection improves network bandwidth and provides unique network solutions. Filters can be created to redirect traffic to cache and application servers, improving speed of repeated client access to common Web or application content and freeing valuable network bandwidth.

Most of the information downloaded from the Internet is not unique, as clients will often access the Web page many times for additional information or to explore other links. Duplicate information also gets requested as the components that make up Internet data at a particular Web site (pictures, buttons, frames, text and so on) are reloaded from page to page. When you consider this scenario, as illustrated in Figure 3-30 on page 69, it becomes apparent that redundant requests can consume a considerable amount of your available bandwidth to the Internet.

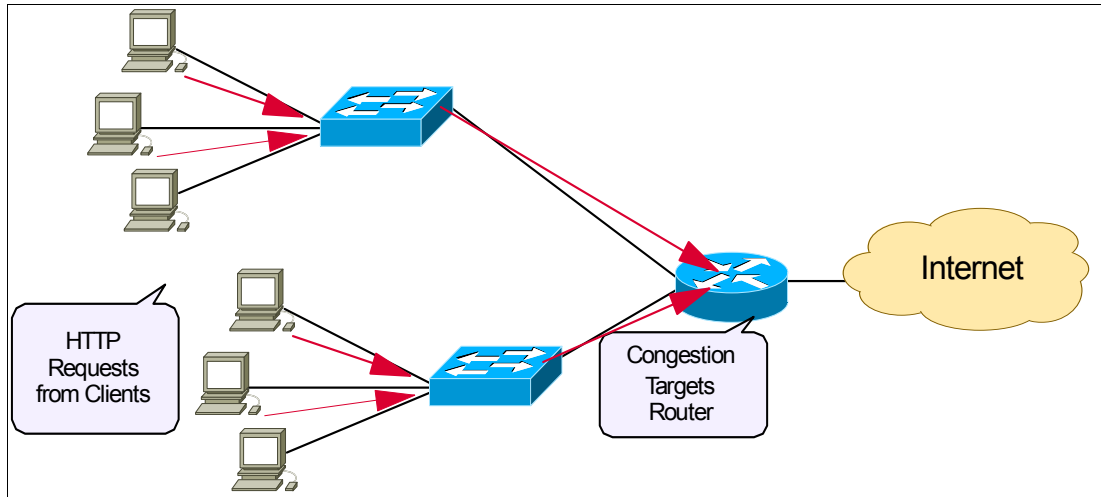


Figure 3-30 Traditional network without cache redirection

Application redirection can help reduce the traffic congestion during peak loads. When application redirection filters are properly configured for the GbESM, outbound client requests for Internet data are intercepted and redirected to a group of application or cache servers on your network. The servers duplicate and store inbound Internet data that has been requested by your clients. If the servers recognize a client's outbound request as one that can be filled with cached information, the servers supply the information rather than send the request across the Internet, as shown in Figure 3-31. In addition to increasing the efficiency of your network, accessing locally cached information can be much faster than requesting the same information across the Internet.

The network needs a solution that addresses the following key concerns:

- ▶ The solution must be readily scalable.
- ▶ The administrator should not need to reconfigure all the clients' browsers to use proxy servers.

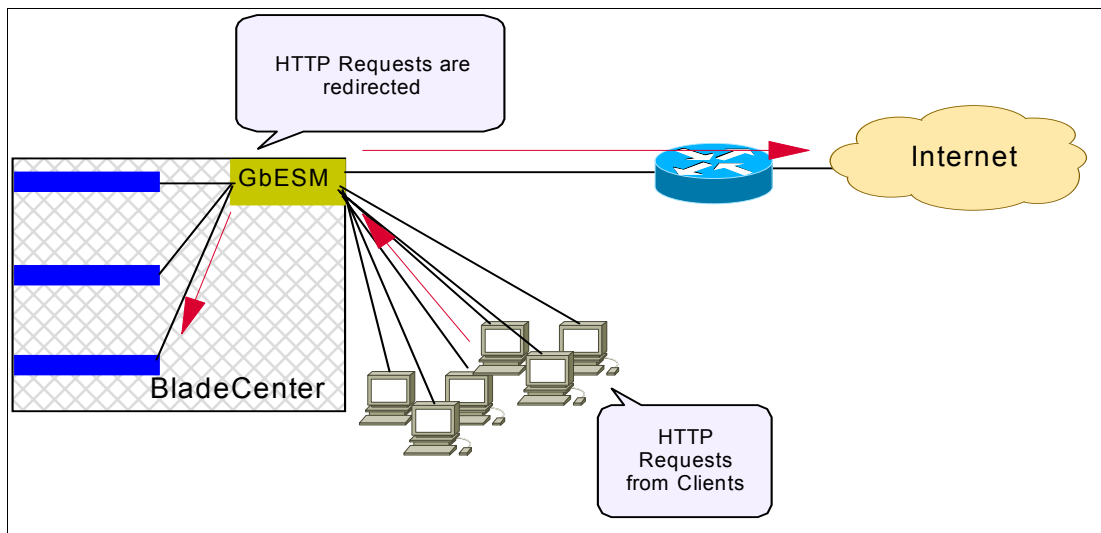


Figure 3-31 Network with cache redirection

If you have more clients than switch ports, connect the clients to a layer 2 switch.

Adding a Layer 2-7 GbE Switch Module for IBM @server BladeCenter as shown in Figure 3-31 addresses these issues:

- ▶ Cache servers can be added or removed dynamically without interrupting services.
- ▶ Performance is improved by balancing the cached request load across multiple servers.
- ▶ More servers can be added at any time to increase processing power.
- ▶ The proxy is transparent to the client.
- ▶ Frames that are not associated with HTTP requests are normally passed to the router.



Layer 2 configuration guidelines

In this chapter, we present a few guidelines for configuring Layer 2-related parameters on the Layer 2-7 GbE Switch Module. Particular focus is placed on the parameters that are considered routine and essential to basic switch configuration and operation. These include:

- ▶ Multi-port trunks (port aggregation groups)
- ▶ VLANs
- ▶ VLAN Trunks
- ▶ Spanning Tree Groups

4.1 Configuring multi-port trunks

Trunk groups are useful for connecting a Layer 2-7 GbE Switch Module to third-party devices that support link aggregation, such as Cisco routers and switches with EtherChannel technology (not ISL trunking technology). A *trunk*¹ is a group of ports that act together, combining their bandwidth to create a single, larger port. Up to 2 trunk groups can be configured on the Layer 2-7 GbE Switch Module, with the following restrictions:

- ▶ Any external (EXT) switch port can belong to no more than one trunk group.
- ▶ Up to four EXT ports can belong to the same trunk group.
- ▶ Best performance is achieved when all ports in a trunk are configured for the same speed.
- ▶ Trunking of other devices must interoperate with Cisco EtherChannel technology. Nortel Networks trunk group technology is compatible with these devices when they are configured manually.

When using port trunk groups between a Layer 2-7 GbE Switch Module and a remote gigabit Ethernet switch, you can create a virtual link between the switches operating up to 4 Gigabit per second, depending on how many physical ports are combined.

Statistical load distribution

Out-bound network traffic is statistically distributed among the ports in a trunk group. The Layer 2-7 GbE Switch Module uses the Layer 2 MAC address information present in each transmitted frame for determining load distribution. There are no configuration parameters associated with this distribution function.

Each packet's particular combination of source and destination MAC addresses results in selecting one line in the trunk group for data transmission. If there are enough Layer 2 devices feeding the trunk lines, then traffic distribution becomes relatively even. In cases where most of the traffic is being sent to the same remote device, such as a router, the traffic distribution will be influenced by the source MAC address.

Since each trunk group is comprised of multiple physical links, the trunk group is inherently fault tolerant. As long as one connection between the switches is available, the trunk remains active. Statistical load balancing is maintained whenever a port in a trunk group is lost or returned to service.

Before you configure trunks

When you create and enable a trunk, the trunk members (switch ports) take on certain settings necessary for correct operation of the trunking feature.

Before you configure your trunk, you must consider these settings, along with specific configuration rules, as follows:

1. Read and follow the configuration rules provided in the next section.
2. Determine which switch ports (up to four) are to become trunk members (the specific ports making up the trunk). Ensure that the chosen switch ports are set to Enabled, using the `/cfg/port` command. Trunk member ports must have the same VLAN configuration.
3. Consider how the existing spanning tree will react to the new trunk configuration.
4. Consider how existing VLANs will be affected by the addition of a trunk.

¹ Should not be confused with a VLAN Trunk. VLAN Trunking is discussed later in this chapter.

Trunk group configuration rules

The trunking feature operates according to specific configuration rules. When creating trunks, consider the following rules that determine how a trunk group reacts in any network topology:

- ▶ Trunking is only applied to the 4 external (EXT) ports of the Layer 2-7 GbE Switch Module.
- ▶ All trunks must originate from one device, and lead to one physical or logical destination device². For example, you cannot combine a link from BladeCenter Switch 1 and a link from BladeCenter Switch 2 into one trunk group. Any physical switch port can belong to only one trunk group.
- ▶ Trunking from third-party devices must comply with Cisco EtherChannel technology.
- ▶ All trunk member ports on the Layer 2-7 GbE Switch Module must be assigned to the same VLAN configuration before the trunk can be enabled.
- ▶ If you change the VLAN settings of any trunk member, you cannot apply the change until you change the VLAN settings of all trunk members.
- ▶ When an active port is configured in a trunk, the port becomes a trunk member when you enable the trunk using the `/cfg/12/trunk/ena` command. The spanning tree parameters for the port then change to reflect the new trunk settings.
- ▶ All trunk members must be in the same Spanning Tree Group (STG) and can belong to only one Spanning Tree Group. However, if all ports are tagged, then all trunk ports can belong to multiple STGs.
- ▶ If you change the spanning tree participation of any trunk member to Enabled or Disabled, the spanning tree participation of all members of that trunk changes similarly.
- ▶ When a trunk is enabled, the trunk spanning tree participation setting takes precedence over that of any trunk member.
- ▶ You cannot configure a trunk member as a monitor port in a port-mirroring configuration.
- ▶ Trunks cannot be monitored by a monitor port; however, trunk members can be monitored.

4.1.1 Port trunking example

In this section we describe the steps to configure the switches shown in Figure 4-1 on page 74, in which three ports are trunked between two switches.

² Nortel's Split Multi-link Trunking (SMLT) allows two physical switch devices to appear as a single logical device. See Chapter 6 for a more detailed configuration discussion.

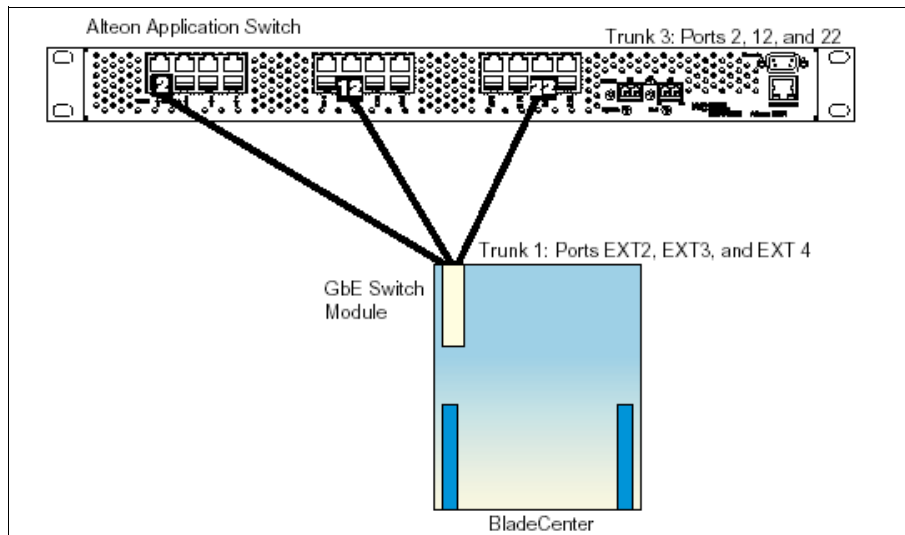


Figure 4-1 Port trunking example

Prior to configuring each switch in this example, you must connect to the appropriate switch's Command Line Interface (CLI) as the administrator.

Note: For details about accessing and using any of the menu commands described in this example, see the *Alteon OS 20.0 Command Reference*.

1. On the Layer 2-7 GbE Switch Module:
 - a. Define a trunk group. (See Example 4-1)
 - b. Apply and verify the configuration.
Examine the resulting information. If any settings are incorrect, make appropriate changes.
 - c. Save your new configuration changes.
2. Repeat the process on the opposite switch per the manufacturer's guidelines.
3. Connect the switch ports that will be involved in the trunk group. A cross-over cable is recommended for this connection.

Note: If a third-party device supporting link aggregation is used (such as Cisco routers and switches with EtherChannel technology), trunk groups on the third-party device should be configured manually. Connection problems will arise when automatic trunk group negotiation, such as Cisco's PAgP or IEEE LACP, is enabled on the third-party device.

Example 4-1 Defining a trunk group

```
>> # /cfg/12/trunk 1 (Select trunk group 1)
>> Trunk group 1# add EXT2 (Add port EXT2 to trunk group 1)
>> Trunk group 1# add EXT3 (Add port EXT3 to trunk group 1)
>> Trunk group 1# add EXT4 (Add port EXT4 to trunk group 1)
>> Trunk group 1# ena (Enable trunk group 1)
>> Trunk group 1# apply (Make your changes active)
>> Trunk group 1# cur (View current trunking configuration)
>> Trunk group 1# save (Save for restore after reboot)
```

4. Examine the trunking information on each switch.

>> /info/12/trunk (View trunking information)

Information about each port in each configured trunk group will be displayed. Make sure that trunk groups consist of the expected ports and that each port is in the expected state.

Trunk configuration menu options (/cfg/12/trunk)

Table 4-1 lists the configuration commands for configuring trunks.

Table 4-1 Trunk configuration command syntax and usage

Command syntax	Purpose
add <port alias or number (EXT1-EXT4)>	Adds a physical port to the current trunk group.
rem <port alias or number (EXT1-EXT4)>	Removes a physical port from the current trunk group.
ena	Enables the current trunk group.
dis	Turns the current trunk group off.
del	Removes the current trunk group configuration.
cur	Displays current trunk group parameters.

Note: Either port numbers or port aliases may be used. Internal port numbers 1 through 14 correspond to blade bays 1 through 14. Port numbers 17 through 20 correspond to external ports 1 through 4 (EXT1-EXT4).

Factory default setting for trunks

The factory default setting is for no trunk groups to be active. Each of the four external ports is configured as an individual port. Connecting two or more ports to the same Layer 2 switch device will create a potential Layer 2 loop situation.

4.2 Configuring IEEE 802.1Q VLANs

Setting up virtual LANs (VLANs) is a way to segment networks to increase network flexibility without changing the physical network topology. With network segmentation, each switch port connects to a segment that is a single broadcast domain. When a switch port is configured to be a member of a VLAN, it is added to a group of ports (workgroup) that belong to one broadcast domain. Within the IBM BladeCenter chassis, VLANs can be used to partition selected blades as groups within their own domain within the chassis.

Ports are grouped into broadcast domains by assigning them to the same VLAN. Frames received in one VLAN can only be forwarded within that VLAN, and multicast, broadcast, and unknown unicast frames are flooded only to ports in the same VLAN.

When we enable tagging on a switch port, the switch preserves the 802.1Q VLAN tag as frames exit the port. Conversely, when we disable tagging, the switch removes the VLAN tag. A tag is never created at a port during egress. This implies that tag headers are created somewhere else in the process. Indeed, when a normal, untagged Ethernet frame arrives at a switch port, the switch must insert a tag into the frame to identify in which VLAN the frame belongs. In the Layer 2-7 GbE Switch Module, each internal and external port has one (and only one) Port VLAN Identifier (PVID) associated with it. The PVID determines the VLAN identifier (VID) that is inserted into the tag of an originally untagged frame for use within the

switch. The PVID of the port that receives the untagged frame will determine the VID inserted into that frame as it is received.

The PVID comes into play only for untagged frames as they enter a port. The PVID has no bearing on egress frames. By using the PVID to create the VID in the 802.1Q tag header, a switch guarantees that all frames are tagged as they enter its ports. In other words, the PVID is an ingress parameter for a port. In contrast, tagging/untagging is an egress parameter. These two concepts are critical to understanding how 802.1Q VLANs work, and how to configure the Layer 2-7 GbE Switch Module. Although not obvious at first, PVIDs are always involved in the frame forwarding operation. Even in factory-default state, all ports are assigned a PVID of 1, meaning that all untagged frames are tagged with a VID of 1 at ingress. Also, by default, all ports are set to untagging, meaning that all frames are untagged at egress.

4.2.1 VLANs and port VLAN ID numbers

This section discusses VLAN configuration and use.

VLAN numbers

The Layer 2-7 GbE Switch Module supports up to 128 VLANs per switch. Even though the maximum number of VLANs supported at any given time is 128, each can be identified with any number between 1 and 4095. VLAN 1 is the default VLAN for the external ports and the internal blade ports. VLAN 4095 is reserved and used by the interface to the internal BladeCenter management modules.

Creating VLANs using only the PVID

When we change the PVID of a port to a value other than 1, the firmware checks whether a VLAN already exists with a matching VID. If the VLAN does not exist, the firmware automatically creates a new VLAN with a matching VID (that is, if we select a PVID of 20, then a VLAN with a VID of 20 is created).

Viewing VLANs

The VLAN Information is displayed in the VLAN information menu, as shown in Figure 4-2. VLAN information can be viewed by issuing the `/info/12/vlan` command. The VLAN number, VLAN name, and member ports are displayed along with the status (enabled or disabled) and whether or not jumbo frames are allowed on this VLAN.

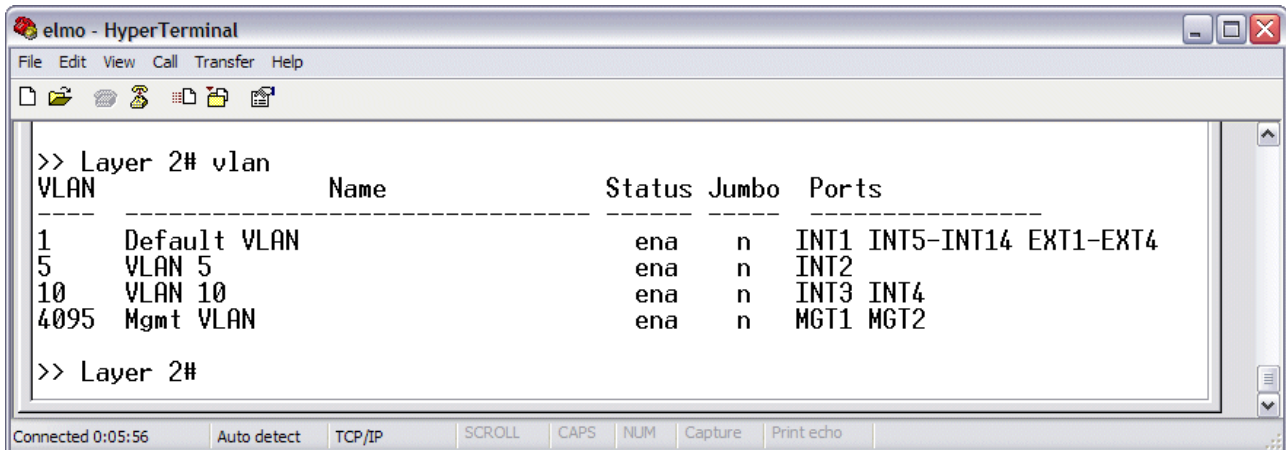


Figure 4-2 VLAN information menu

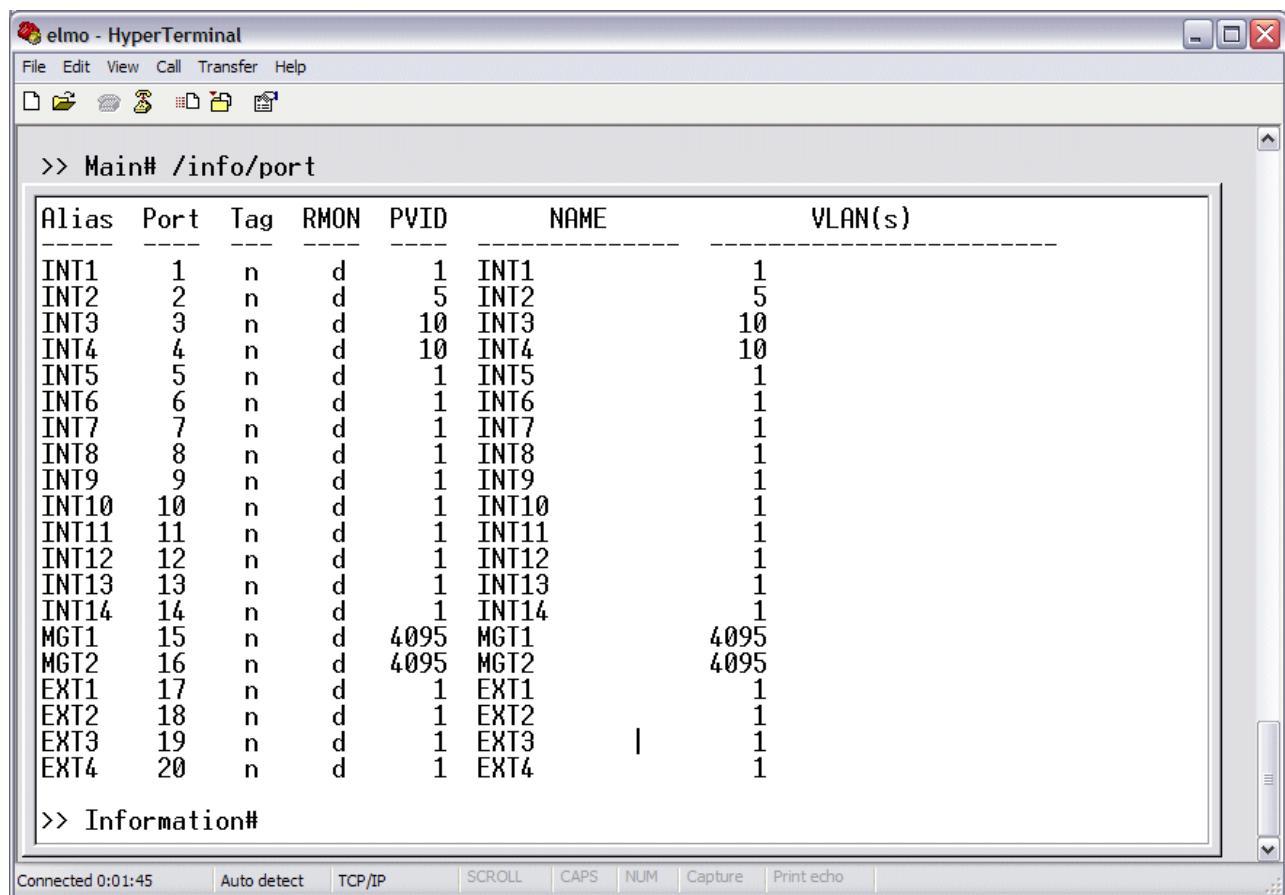
PVID numbers

As discussed previously, each port in the switch has a configurable default VLAN number, known as its PVID. This places all non-management ports on the same VLAN initially, although each port's PVID is configurable to any VLAN number between 1 and 4094.

The default configuration settings for Layer 2-7 GbE Switch Module have all ports set as untagged members of VLAN 1 with all ports configured as PVID = 1. In the default configuration example shown in Figure 4-5 on page 78, all incoming packets are assigned to VLAN 1 by the default port VLAN identifier (PVID =1).

Viewing and configuring PVIDs

You can view PVIDs via the `/info/12/port` command, as shown in Figure 4-3. The information screen lists each of the ports, along with the PVID assigned. The same screen shows the configured 802.1q VLANs and whether or not tagging is enabled.



```
>> Main# /info/port
```

Alias	Port	Tag	RMON	PVID	NAME	VLAN(s)
INT1	1	n	d	1	INT1	1
INT2	2	n	d	5	INT2	5
INT3	3	n	d	10	INT3	10
INT4	4	n	d	10	INT4	10
INT5	5	n	d	1	INT5	1
INT6	6	n	d	1	INT6	1
INT7	7	n	d	1	INT7	1
INT8	8	n	d	1	INT8	1
INT9	9	n	d	1	INT9	1
INT10	10	n	d	1	INT10	1
INT11	11	n	d	1	INT11	1
INT12	12	n	d	1	INT12	1
INT13	13	n	d	1	INT13	1
INT14	14	n	d	1	INT14	1
MGT1	15	n	d	4095	MGT1	4095
MGT2	16	n	d	4095	MGT2	4095
EXT1	17	n	d	1	EXT1	1
EXT2	18	n	d	1	EXT2	1
EXT3	19	n	d	1	EXT3	1
EXT4	20	n	d	1	EXT4	1

```
>> Information#
```

Figure 4-3 VLAN port information

The PVID can be changed for an individual port by the configuration command, as shown in Figure 4-4 on page 78. For example, to change the PVID of INT7 from 1 to 10, the command `/cfg/12/vlan 10/add int7` will initiate the change.

```

elmo - HyperTerminal
File Edit View Call Transfer Help
VLAN
-----
1   Default VLAN      ena   n   INT1 INT5-INT14 EXT1-EXT4
5   VLAN 5            ena   n   INT2
10  VLAN 10           ena   n   INT3 INT4
4095 Mgmt VLAN        ena   n   MGT1 MGT2
>> Layer 2#

```

Figure 4-4 Changing port PVID

Important: Each port on the switch can belong to one or more VLANs, and each VLAN can have any number of switch ports in its membership. Any port that belongs to multiple VLANs, however, must have VLAN tagging enabled. This will result in all egress packets having an 802.1q tag.

Any untagged frames (those with no VLAN specified) are classified with the sending port's PVID.

You can choose to partition the Server Blades into two or more PVID groups by assigning PVIDs on a per port basis. There is no requirement to enable tagging if PVIDs are the only means for assigning VLANs. Blades that are in separate PVIDs are assigned to separate Layer 2 broadcast domains.

In Figure 4-5, untagged incoming packets are assigned directly to VLAN 1 (PVID = 1). Port 5 is a member of VLAN 1, and transmits the frame as untagged.

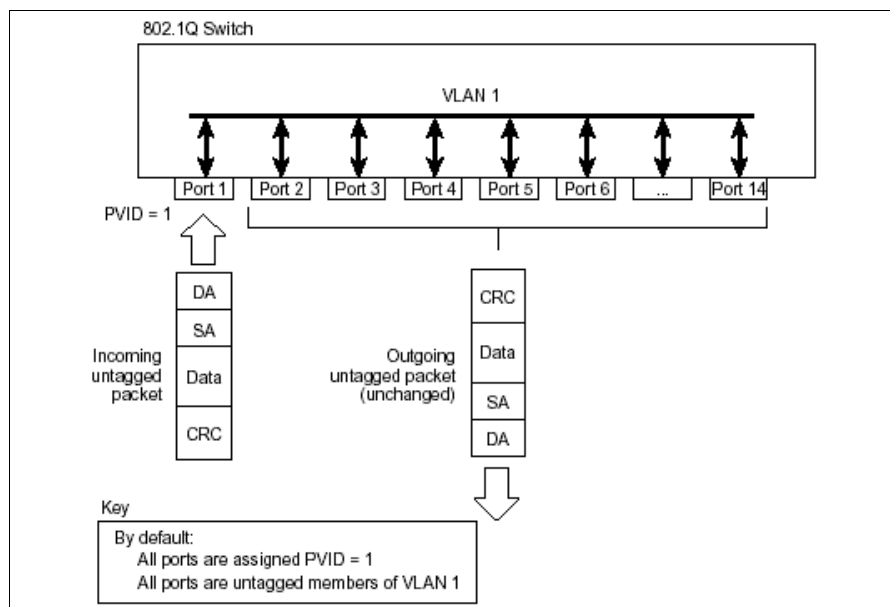


Figure 4-5 VLAN example with all ports on PVID=1

IEEE 802.1Q VLAN Tagging

Although they appear similar, VLAN Identifiers (VIDs) are not the same as Port VLAN Identifiers (PVIDs). Whereas one PVID is fixed to each port (that is, all untagged ingress traffic is associated with one VLAN), a port can potentially transmit and receive tagged frames with different VIDs. Once a frame is tagged, we no longer refer to a PVID and there is no requirement that the VID and PVID match. Instead, we refer to a tagged frame's VID. Any VLAN in an 802.1Q network is also referenced by its VID, a number which ranges from 1 to 4094. Therefore, a tagged frame with a specific VID is associated with the VLAN having the same VID. The distinction between VIDs and PVIDs allows us to configure 802.1Q VLANs in a flexible manner.

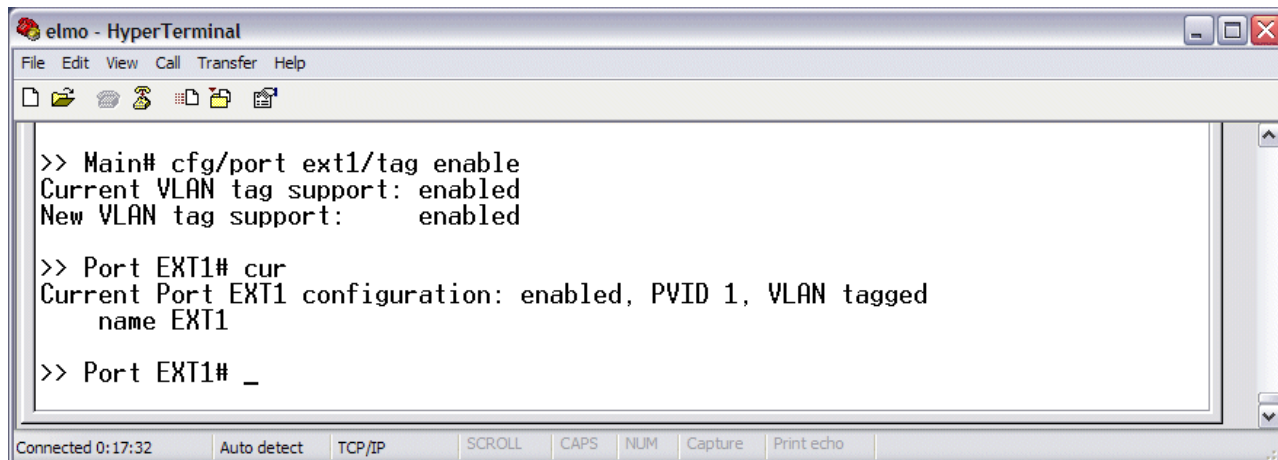
The Layer 2-7 GbE Switch Module software supports 802.1Q VLAN tagging, providing standards-based VLAN support for Ethernet systems. Tagging places the VLAN identifier in the frame header, allowing each port to belong to multiple VLANs. When you add a port to multiple VLANs, you must also enable tagging on that port. Since tagging fundamentally changes the format of frames transmitted on a tagged port, you must carefully plan network designs to prevent tagged frames from being transmitted to devices that do not support 802.1Q VLAN tags, or devices where tagging is not enabled. For example, when tagging is enabled on the internal ports of a BladeCenter, the Server Blades must be configured separately to handle the VLAN tagged packets.

Important terms used with the 802.1Q tagging feature are:

- ▶ VLAN identifier (VID): The 12-bit portion of the VLAN tag in the frame header that identifies an explicit VLAN.
- ▶ Port VLAN identifier (PVID): A classification mechanism that associates a port with a specific VLAN. For example, a port with a PVID of 3 (PVID =3) assigns all untagged frames received on this port to VLAN 3.
- ▶ Tagged frame: A frame that carries VLAN tagging information in the header. This VLAN tagging information is a 32-bit field (VLAN tag) in the frame header that identifies the frame as belonging to a specific VLAN. Untagged frames are marked (tagged) with this classification as they leave the switch through a port that is configured as a tagged port.
- ▶ Untagged frame: A frame that does not carry any VLAN tagging information in the frame header.

Enabling tagging for a port

Tagging is enabled on a per port basis. When tagging is enabled, a VLAN tag is required for all outbound frames. Frames will be tagged with either the PVID or with the VID that was appended by an external device (for example, a server, router, or external switch). Tagging is enabled by the `/cfg/port <port alias or number>/tag enable` command, as shown in Figure 4-6 on page 80.



```
elmo - HyperTerminal
File Edit View Call Transfer Help
[Icons]
>> Main# cfg/port ext1/tag enable
Current VLAN tag support: enabled
New VLAN tag support:    enabled

>> Port EXT1# cur
Current Port EXT1 configuration: enabled, PVID 1, VLAN tagged
name EXT1

>> Port EXT1# _

Connected 0:17:32  Auto detect  TCP/IP  SCROLL  CAPS  NUM  Capture  Print echo
```

Figure 4-6 VLAN tag enable

4.2.2 VLAN topologies and design considerations

By default, the Layer 2-7 GbE Switch Module software is configured so that management ports, INT15 and INT16, are configured to PVID 4095. The user is not allowed to change this VLAN or disable the management ports.

If you are configuring a Spanning Tree Group, note that each Spanning Tree Groups 2-16 may contain only one VLAN. This is discussed in more detail later in this section.

VLAN configuration rules with mirroring

VLANs operate according to specific configuration rules. When creating VLANs, consider the following rules that determine how the configured VLAN reacts in any network topology:

- ▶ All ports involved in trunking and port mirroring must have the same VLAN configuration. If a port is on a trunk with a mirroring port, the VLAN configuration cannot be changed.
- ▶ All ports that are involved in port mirroring must have memberships in the same VLANs. If a port is configured for port mirroring, the port's VLAN membership cannot be changed.

The network example shown in Figure 4-7 on page 81 includes three separate VLANs.

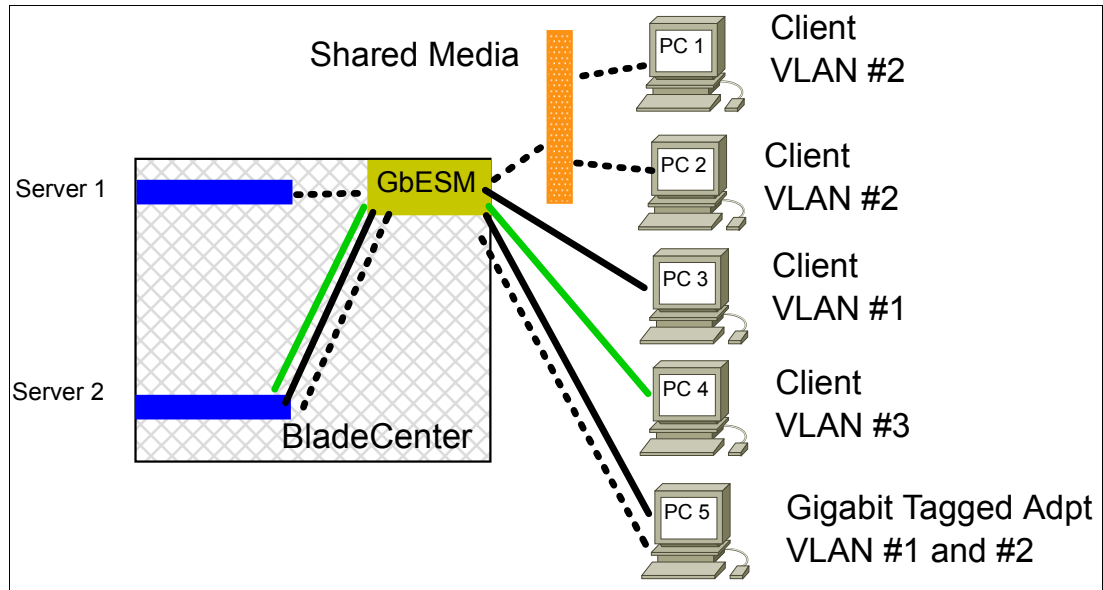


Figure 4-7 Multiple VLAN configuration

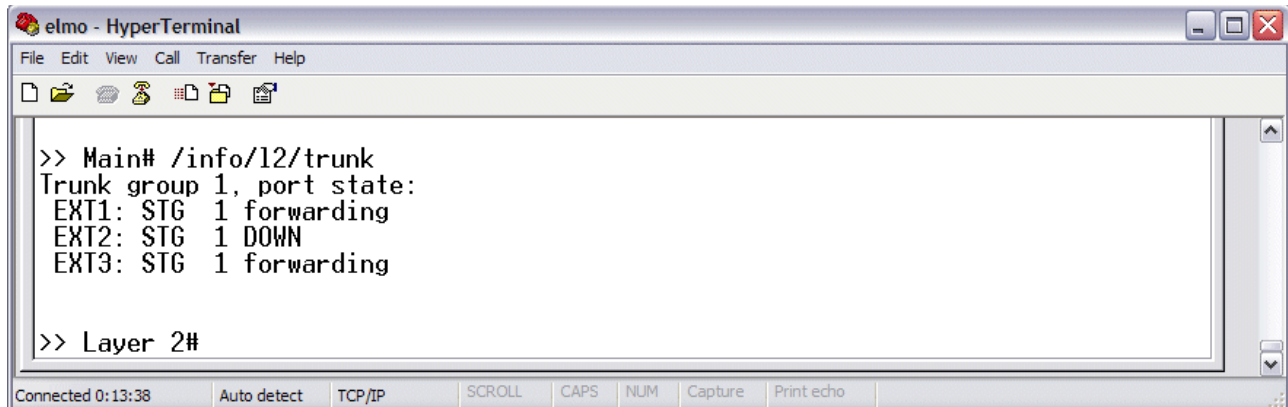
In the pictured configuration, the components have the following characteristics:

- ▶ GbE Switch Module - This switch is configured for three VLANs that represent three different IP subnets. Three servers and five clients are attached to the switch.
- ▶ Server #1 - Server 1 is part of VLAN 3 and only has presence in one IP subnet. The port that the VLAN is attached to is configured only for VLAN 3, so VLAN tagging is off.
- ▶ Server #2 - Server 2 is a high-use server and needs to be accessed from all VLANs and IP subnets. The server has a VLAN-tagging enabled driver with VLAN tagging turned on. The server is attached to one of the switch's internal Gigabit Ethernet ports that is configured for VLANs 1, 2, and 3. Tagging is turned on. Because of the VLAN tagging capabilities of both the server and the switch, the server is able to communicate on all three IP subnets in this network. Broadcast separation between all three VLANs and subnets, however, is maintained.
- ▶ PCs #1 and #2 - These PCs are attached to a shared media hub that is then connected to the switch. They belong to VLAN 2 and are logically in the same IP subnet as Server 2 and PC5. Tagging is not enabled on their switch port.
- ▶ PC #3 - A member of VLAN 1, this PC can only communicate with Server 2 and PC5.
- ▶ PC #4 - A member of VLAN 3, this PC can only communicate with Server 1 and Server 2.
- ▶ PC #5 - A member of both VLAN 1 and VLAN 2, this PC has VLAN-tagging Gigabit Ethernet adapter installed. It can communicate with Server #2 via VLAN 1, and to PC #1 and PC #2 via VLAN 2. The switch port to which it is connected is configured for both VLAN 1 and VLAN 2 and has tagging enabled.

Creating VLAN trunks

A VLAN Trunk is created on a switch-to-switch link whenever traffic from multiple tagged VLANs shares the same interswitch cable or link aggregation group. This is implemented by defining one or more external (EXT) ports or multi-link trunks to be members on two or more VLANs. Likewise, the opposite switch ports must also be configured to be members of the same VLANs.

An example VLAN Trunk configuration is shown in Figure 2-11 on page 25. The current status of the trunks can be displayed using the `/info/L2/trunk` command, as shown in Figure 4-8.



```
elmo - HyperTerminal
File Edit View Call Transfer Help
>> Main# /info/l2/trunk
Trunk group 1, port state:
EXT1: STG 1 forwarding
EXT2: STG 1 DOWN
EXT3: STG 1 forwarding
>> Layer 2#
Connected 0:13:38 Auto detect TCP/IP SCROLL CAPS NUM Capture Print echo
```

Figure 4-8 Port trunk information example

VLAN configuration

`/cfg/12/vlan <VLAN number>`

The commands in this menu configure VLAN attributes, change the status of the VLAN, delete the VLAN, and change the port membership of the VLAN. For more information on configuring VLANs, see *Command Reference Alteon OS 20.0: L2-7 GbE Switch Module for IBM eServer BladeCenter*.

By default, the VLAN menu option is disabled except VLAN 1, which is enabled all the time. Ports INT1-INT14 and ports EXT1-EXT4 are in VLAN 1 by default.

VLAN configuration menu options (`/cfg/12/vlan`)

Command syntax and usage:

name - Assigns a name to the VLAN or changes the existing name. The default VLAN name is the first one.

stg <Spanning Tree Group index [1-16]> - Assigns a VLAN to a Spanning Tree Group.

add <port alias or number (1-20)> - Adds port(s) or trunk group(s) to the VLAN membership.

rem <port alias or number (1-20)> - Removes port(s) or trunk group(s) from this VLAN.

def <list of port numbers> - Defines which ports are members of this VLAN. Every port must be a member of at least one VLAN. By default, port INT1-INT14 and EXT1-EXT4 are in VLAN 1.

jumbo - Define support for jumbo frames (enable/disable).

ena - Enables this VLAN.

dis - Disables this VLAN without removing it from the configuration.

del - Deletes this VLAN.

cur - Displays the current VLAN configuration.

Note: All ports must belong to at least one VLAN. Any port which is removed from a VLAN and which is not a member of any other VLAN is automatically added to default VLAN #1. You cannot remove a port from VLAN #1 if the port has no membership in any other VLAN. Also, you cannot add a port to more than one VLAN unless the port has VLAN tagging turned on.

4.3 Configuring the Spanning Tree Protocol

The Layer 2-7 GbE Switch Module supports the IEEE 802.1d Spanning Tree Protocol (STP). STP is used to prevent loops in the network topology. There are 16 spanning tree groups (STGs) that can be configured on the switch (STG 16 is reserved for management). VLANs are mapped to separate STGs to provide separate spanning trees. STG is enabled on all EXT ports by default.

Note: When VRRP is used for active/active redundancy, STG must be enabled.

Spanning Tree configuration menu (/cfg/12/stg)

Command syntax and usage:

brg - Displays the Bridge Spanning Tree Menu.

port <port alias or number (1-20)> - Displays the Spanning Tree Port Menu.

add <VLAN number (1-4095)> - Associates a VLAN with a spanning tree and requires an external VLAN ID as a parameter.

remove <VLAN number (1-4095)> - Breaks the association between a VLAN and a spanning tree and requires an external VLAN ID as a parameter.

clear - Removes all VLANs from a spanning tree.

on - Globally enables Spanning Tree Protocol.

off - Globally disables Spanning Tree Protocol.

default - Restores a spanning tree instance to its default configuration.

cur - Displays current Spanning Tree Protocol parameters.

Bridge Spanning Tree configuration (/cfg/12/stg/brg)

Spanning tree bridge parameters affect the global STG operation of the switch. STG bridge parameters include:

- ▶ Bridge priority
- ▶ Bridge hello time
- ▶ Bridge maximum age
- ▶ Forwarding delay
- ▶ Bridge aging time

When configuring STG bridge parameters, the following formulas must be used:

- ▶ $2*(fwd-1) > mxage$
- ▶ $2*(hello+1) < mxage$

Bridge Spanning Tree menu options (/cfg/12/stg/brg)

Command syntax and usage:

prior <new bridge priority (0-65535)> Configures the bridge priority. The bridge priority parameter controls which bridge on the network is the STG root bridge. To make this switch the root bridge, configure the bridge priority lower than all other switches and bridges on your network. The lower the value, the higher the bridge priority. The range is 0 to 65535, and the default is 32768.

hello <new bridge hello time (1-10 secs)> Configures the bridge hello time. The hello time specifies how often the root bridge transmits a configuration bridge protocol data unit (BPDU). Any bridge that is not the root bridge uses the root bridge hello value. The range is 1 to 10 seconds, and the default is 2 seconds.

mxage <new bridge max age (6-40 secs)> Configures the bridge maximum age. The maximum age parameter specifies the maximum time the bridge waits without receiving a configuration bridge protocol data unit before it re-configures the STG network. The range is 6 to 40 seconds, and the default is 20 seconds.

fwd <new bridge Forward Delay (4-30 secs)> Configures the bridge forward delay parameter. The forward delay parameter specifies the amount of time that a bridge port has to wait before it changes from the listening state to the learning state and from the learning state to the forwarding state. The range is 4 to 30 seconds, and the default is 15 seconds.

aging <new bridge Aging Time (1-65535 secs, 0 to disable)> Configures the forwarding database aging time. The aging time specifies the amount of time the bridge waits without receiving a packet from a station before removing the station from the forwarding database. The range is 1 to 65535 seconds, and the default is 300 seconds. To disable aging, set this parameter to 0.

current Displays the current bridge STG parameters.

Spanning Tree port parameters are used to modify STG operation on an individual port basis. STG port parameters include:

- ▶ Port priority
- ▶ Port path cost

The port option of STG is turned on by default.

Spanning Tree port menu (/cfg/12/stg/port)

Command syntax and usage:

prior <new port Priority (0-255)> Configures the port priority. The port priority helps determine which bridge port becomes the designated port. In a network topology that has multiple bridge ports connected to a single segment, the port with the lowest port priority becomes the designated port for the segment. The range is 0 to 255, and the default is 128.

cost <new port Path Cost (1-65535, 0 for default)> Configures the port path cost. The port path cost is used to help determine the designated port for a segment. Generally speaking, the faster the port, the lower the path cost. The range is 1 to 65535. The default is 10 for 100 Mbps ports, and 1 for Gigabit ports. A value of 0 indicates that the default cost will be computed for an auto-negotiated link speed.

on - Enables STG on the port.

off - Disables STG on the port.

cur - Displays the current STG port parameters.



Deploying the Layer 2-7 GbE Switch Module for IBM @server BladeCenter in a Cisco environment

This chapter provides information on deploying the Layer 2-7 GbE Switch Module for IBM @server BladeCenter (GbESM) in an infrastructure based on Cisco Systems switches. Specific GbESM and Cisco configurations are covered for a number of possible scenarios.

Note: The scope of this book is to provide Layer 2 connectivity between the Layer 2-7 GbE Switch Module for IBM @server BladeCenter and Cisco devices covering a number of possible scenarios. For more in-depth configuration examples configuring Layer 3-7 features, refer to the *Alteon OS 20.0 Application Guide*, Part 215654-A (Nortel) and *Alteon OS 20.0 Command Reference*, Part 215655-A (Nortel).

5.1 Introduction

The Layer 2-7 GbE Switch Module for IBM @server BladeCenter is a Layer 2-7 aware switching device. It provides the necessary features and functionality to attach the BladeCenter components to virtually any standards-based network infrastructure, including those based on Cisco Systems equipment.

While the GbESM and Cisco share a common set of specifications with which they are compatible (this was covered in Chapter 1), it is important to note that there are some features and standards that are fairly commonly deployed in Cisco infrastructures that are not found in the GbESM. Cisco enhancements and standards that are commonly encountered in a Cisco infrastructure, but that are not supported by the current GbESM, include the following:

- ▶ BackboneFast - To aid in the rapid convergence of layer 2 networks
- ▶ UplinkFast - To aid in the rapid convergence of layer 2 networks
- ▶ UDLD - UniDirectional Link Detection, to reduce the possibility of STP loops
- ▶ CDP - Cisco Discovery Protocol, to aid in management and troubleshooting
- ▶ ISL - VLAN Trunking (Cisco Proprietary)
- ▶ PAgP - Port Aggregation Protocol (Cisco EtherChannel)
- ▶ VTP - VLAN Trunking Protocol, similar to GVRP
- ▶ DTP - Dynamic Trunking Protocol, to auto-negotiate trunk type and state
- ▶ VMPS - VLAN Management Policy Server (only on some Cisco switches)

Note: All of these protocols and capabilities can still be turned on and leveraged within the core network.

5.2 Architecture summary

A complete discussion of various network architectures is beyond the scope of this document. Instead, this section provides a few brief remarks on the subject.

5.2.1 Datacenter networks introduction

While what constitutes a datacenter and its associated architecture is open to discussion, it is reasonably expected that many BladeCenter deployments will occur in some form of datacenter environment. It can also be safely assumed that datacenter networks are fairly unique environments from other parts of the enterprise network, and can be more demanding in their requirements.

High availability, performance, and security tend to be far more critical, with elements of firewalls, content caching, load balancing, and security playing a much larger role than in other areas. With this stated, this document does not attempt to discuss all of the various elements that may constitute a datacenter, but rather focuses only on attaching the BladeCenter to Cisco switches.

Note: This document should not be considered a substitute for an Architectural Reference document, and the examples provided should not be employed without first understanding the specific needs of your particular environment.

For those seeking more details and information on Cisco's definition of datacenter architectures, along with recommendations, visit:

http://www.cisco.com/en/US/netso1/ns110/ns53/net_solution_home.html

5.2.2 Common Cisco components

As previously mentioned, the datacenter is a rather unique environment, with demands that can far exceed those of other areas of the enterprise network. With this in mind, it is recommended that certain highly robust and scalable designs and platforms be used in the datacenter.

With regard to design, Cisco traditionally proposes a three level architecture when referring to enterprise environments. This architecture promotes both robustness and scalability, and is made up of an access layer, a distribution layer, and a core layer. When working specifically with datacenter architectures, a common current approach is more often a two layer design as shown in Figure 5-1.

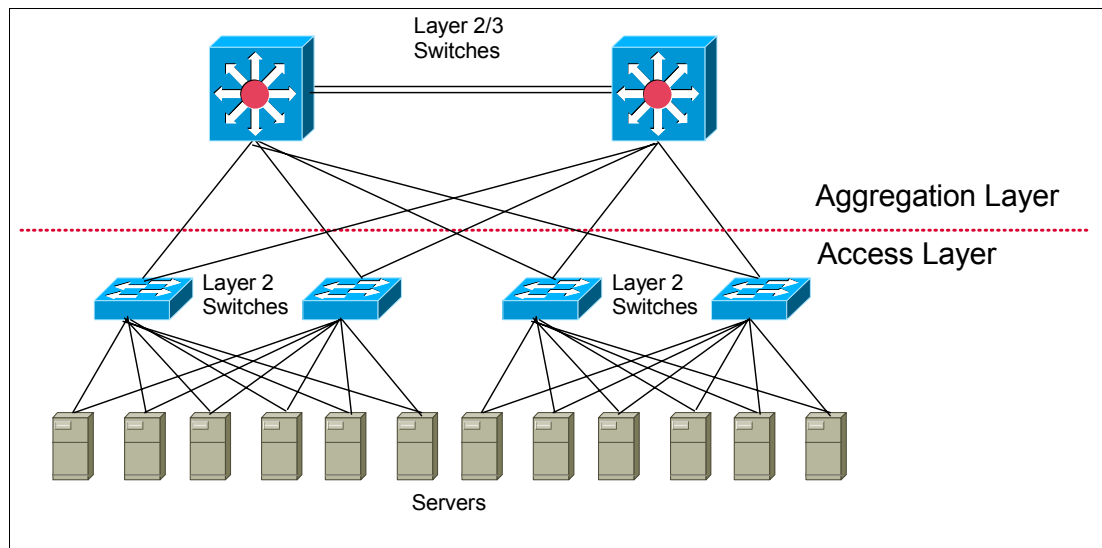


Figure 5-1 Two-level datacenter design

The two layers that make up this design are:

- ▶ An access layer (to attach servers). Sometimes referred to as the front-end layer, this is primarily a Layer 2 (OSI) network.
- ▶ An aggregation layer (to tie the components of the datacenter to each other and into the rest of the enterprise network via the enterprise “core”). In the aforementioned three layer design, this aggregation layer would be referred to as the distribution layer. One could also look at this as being the core layer of a datacenter, where elements of Layer 2 and Layer 3 (OSI) come together.

With a two layer design such as this in mind, the BladeCenter can be placed in this type of architecture at one of two points, as follows.

- ▶ Attaching the GbESMs of the BladeCenter to the access layer switches, as shown in Figure 5-2 on page 88. This in essence adds a third layer to the two layer architecture.

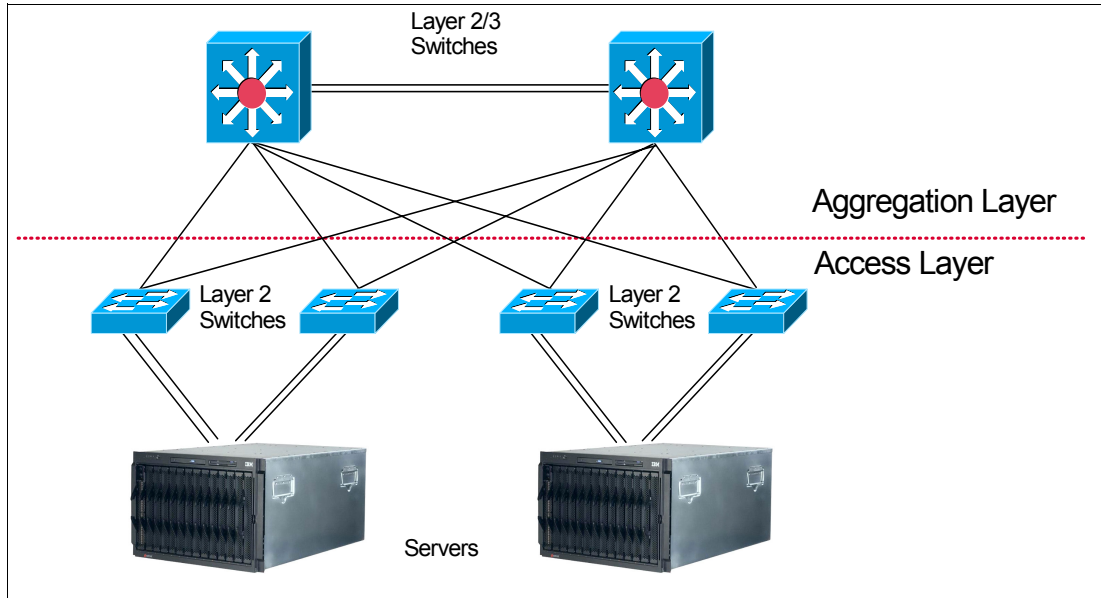


Figure 5-2 Attaching the GbESM to the access layer switches

- ▶ Attaching the GbESMs of the BladeCenter directly to the aggregation layer switches, as shown in Figure 5-3. This maintains the two layer approach, with the GbESMs essentially becoming the access layer.

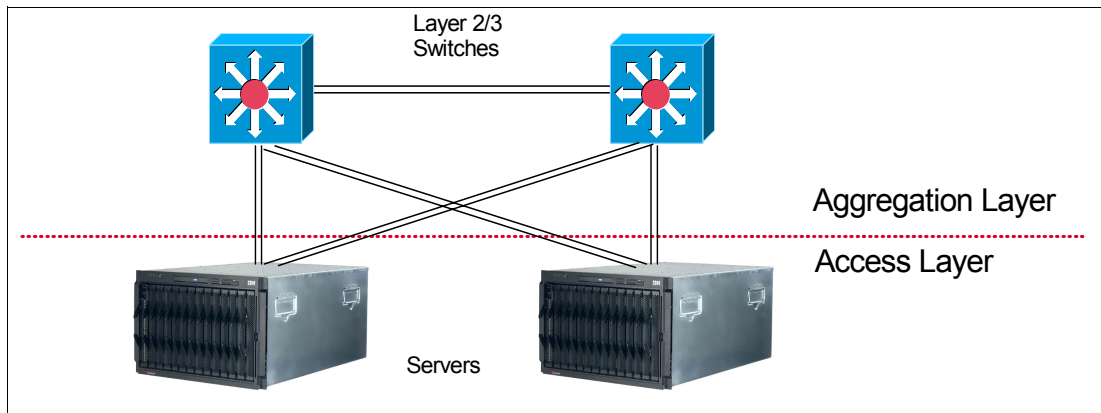


Figure 5-3 Attaching the GbESM's of the BladeCenter directly to the aggregation layer

Since each of these scenarios has its own advantages, architecture selection must be based on specific customer requirements. When GbESMs are attached to the access layer, the network architecture provides more flexibility in handling unexpected growth. When GbESMs directly interface to the aggregation layer, the network benefits from a reduced number of network levels.

Attaching to the access layer tends to be strictly a Layer 2 (OSI) connection, while attaching to the aggregation level may be at Layer 2 or Layer 3 (OSI). While direct Layer 3 connectivity is beyond the scope of this document, and all configuration examples given assume Layer 2, it is important to consider the pros and cons of attaching to the network at Layer 2 or Layer 3. An example of this might be that connecting at Layer 2 can lead to complications with the different implementations of STP employed by the GbESM and the Cisco equipment, while connecting at Layer 3 can lead to inefficient use of IP subnet addresses and limit growth potential.

With regard to platforms, the platform of choice at the aggregation level is the Cisco Catalyst 6500 series of switches. This modular platform offers high performance and availability, and operation in multiple OSI layers, as well as providing investment protection. Beyond switching and routing, the 6500 also currently offers various service modules (such as Firewall and Intrusion Detection) that allow this platform to serve many roles in the datacenter, as well as the role of the aggregation element. Figure 5-4 shows positioning of the Cisco Catalyst Switches.

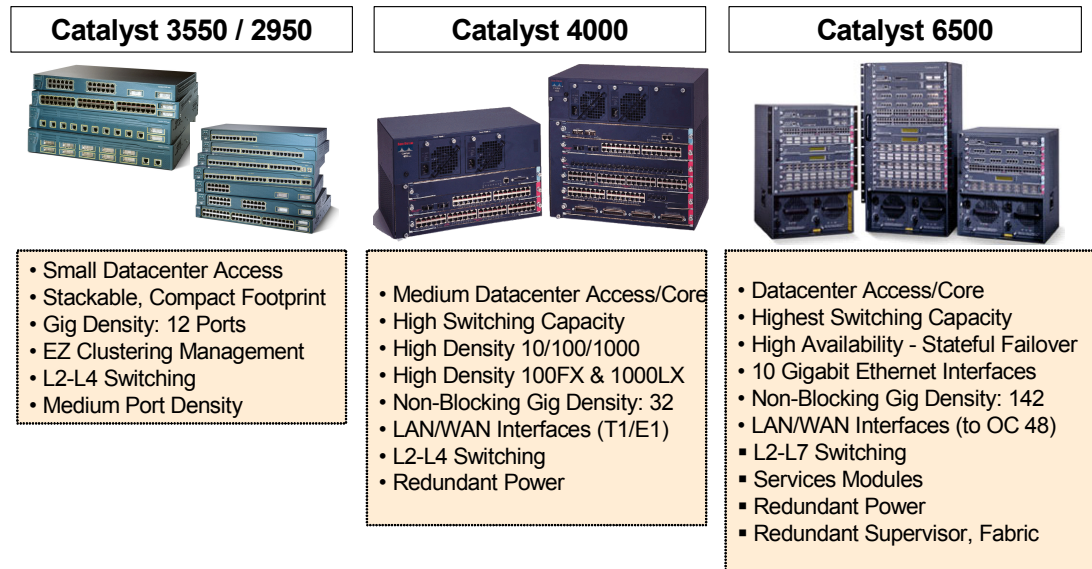


Figure 5-4 Positioning of Cisco Catalyst Switches

Note: In Q3 2003 Cisco released the Catalyst 3750 series of switches. These switches are positioned between the Catalyst 3550 series and the Catalyst 4000 series. Since little information about the 3750 series was available when this book was written, it is not included in Figure 5-4.

Our testing was performed utilizing the @server BladeCenter GbESM and two Cisco Catalyst 4006 switches running CatOS and two Cisco Catalyst 4006s running native IOS.

Note: It is not the intent of this chapter to promote the use of the 4006 for datacenter deployment.

5.3 Guidelines, rules, and comments

Before getting into the specifics of some of the combinations of configurations available in this environment, it is necessary to discuss some basics of configuration and operation as implemented in our lab during the creation of this document.

All configurations were tested with the following equipment and code revisions:

- ▶ IBM BladeCenter 8677
 - Three single CPU Blade servers with Windows® 2000 Server with SP4
 - One Management module

- Two Ethernet Switch Modules (GbESM)
 - GbESM: L2-7 GbE Switch Module
 - Software Version 20.0.1
 - PLD Firmware version: 3.6
- ▶ Two Cisco Catalyst 4006s running Native IOS
 - WS-X4515 Catalyst 4500 Supervisor IV, 2 ports GBIC
 - Native IOS Version 12.1(13)EW
 - ROM Version 12.1(12r)EW
 - WS-X4232-GB-RJ 2 ports Gig uplinks and 32 ports 10/100BaseT
- ▶ Two Cisco Catalyst 4006s running CatOS
 - WS-X4013 Catalyst 4000 Supervisor II, 2 ports GBIC
 - CatOS Version 7.6(3)
 - ROM Version 5.4(1)
 - WS-X4232-GB-RJ 2 ports Gig uplinks and 32 ports 10/100BaseT

Also note that available features and command syntax can be different with different versions of code. This document was prepared utilizing the features and syntax from the aforementioned revisions of code, and as such, might vary from other revisions.

5.3.1 Cleaning and resetting systems

All configuration and testing was performed on clean systems.

If working in a production network, be sure to understand the consequences of any commands issued. Failure to completely understand the operation of commands can lead to network-down conditions.

Important: Performing the operations described in this chapter will result in all configuration data being lost on the mentioned devices, which in turn *will* lead to network downtime if performed on production systems. The commands are presented here only to indicate preparation performed prior to commencement of lab testing.

To reset the Catalyst switches to factory defaults, you need access to the switch console using either a physical console or a Telnet connection. You will also need the console/enable passwords.

CatOS - Catalyst 4006

This procedure applies to Catalyst 4000/5000/6000 switches running CatOS. These switches store the configuration in NVRAM. The configuration is saved in the NVRAM automatically wherever users enter the commands in enable mode.

In the case of the Cisco Catalyst 4006 running CatOS, a **clear config all** command was issued, followed by a **reset** command as shown in Example 5-1.

Example 5-1 Cleaning and resetting CatOS

```
Cat4k> (enable) clear config all
This command will clear all configuration in NVRAM.
This command will cause ifIndex to be reassigned on the next system startup.
Do you want to continue (y/n) [n]? y
```

```
.....  
.....  
.....  
System configuration cleared.  
Console> (enable)
```

When the configuration is erased, you do not need to **reload** the switch. The configuration takes effect immediately and the switch returns to the factory default configuration.

Native IOS - Catalyst 4006

The running configuration and the startup configuration are the two configuration files on switches running Cisco IOS. The running configuration is stored in the RAM and the startup configuration is stored in the NVRAM. When you make a configuration change to a switch, the change becomes part of the running configuration. The change does not automatically become part of the startup configuration, which is used each time the switch restarts. The configuration changes can be saved into NVRAM (startup configuration) by issuing the **write memory** command. If you do not save your changes to NVRAM, they are lost when the switch reloads.

On Catalyst 4006 running Cisco IOS (Native Mode), the VLAN information is stored in a separate file called *vlan.dat* file. To reset these switches to factory defaults, you need to delete the *startup-config* and the *vlan.dat* file stored in *const_nvram*.

The switch can be reset to factory defaults by issuing the **erase cat4000_flash:** command followed by the command **write erase**. For this to take effect you have to finish with the **reload** command. This is shown in Example 5-2.

Note: Depending on the platform you use, the *vlan.dat* file is placed in different locations: On Catalyst 4000s, issue the **erase cat4000_flash:** command. On Catalyst 2950/3550/2900XL/3500XLs, issue the delete **flash:vlan.dat** command.

Example 5-2 Cleaning and resetting Native IOS

```
Cisco-1-IOS#erase cat4000_flash:  
Erasing the cat4000_flash filesystem will remove all files! Continue? [confirm]  
[OK]  
Erase of cat4000_flash: complete  
Cisco-1-IOS#write erase  
Erasing the nvram filesystem will remove all configuration files! Continue? [confirm]  
[OK]  
Erase of nvram: complete  
Cisco-1-IOS#  
00:03:11: %SYS-7-NV_BLOCK_INIT: Initalized the geometry of nvram  
Cisco-1-IOS#reload  
  
System configuration has been modified. Save? [yes/no]: n  
Proceed with reload? [confirm]  
  
00:03:30: %SYS-5-RELOAD: Reload requested
```

Cleaning and resetting the GbESM

In the case of the GbESM, the module was restored to factory defaults via the command line interface. The command **boot/conf factory** tells the switch to use the factory setting during the next boot. The **reset** command restarts the switch, as shown in Example 5-3.

Example 5-3 Cleaning and resetting GbESM

```
>> Main# boot/conf factory
Next boot will use factory default config block instead of active.
>>
Jan 1 0:06:28 NOTICE mgmt: boot config block changedBoot Options#

>> Boot Options# reset

Reset will use software "image1" and the factory default config block.
>> Note that this will RESTART the Spanning Tree,
>> which will likely cause an interruption in network service.
Confirm reset [y/n]:y
```

5.3.2 Rules for attaching the eServer BladeCenter to a Cisco infrastructure

This section contains information on things to consider when attaching the BladeCenter GbESM to a Cisco infrastructure. It is highly recommended that you review this entire section prior to any initial configuration changes.

General guidelines and comments

These are some comments and recommendations that are general in nature, and not part of a specific technology covered elsewhere.

Cable selection

Selection of the cable type (cross-over or straight-through) to use between the GbESM and a Cisco switch can vary, depending whether you are using auto-negotiation or not on the GbESM. When a port on the GbESM is configured for auto-negotiation, it will automatically try to determine the best speed, duplex, and MDI/MDI-X configuration. The auto MDI/MDI-X usually permits any cable (straight-through or cross-over) to be used, as the GbESM can adjust for either cable. When you change the GbESM from Auto to a fixed speed/duplex, this also disables the auto MDI/MDI-X feature. When this happens, a cross-over cable *must* be used to connect a GbESM to a Cisco switch.

Tip: It is recommended that you always use a cross-over cable between the GbESM and a Cisco switch, to ensure that no matter how the speed/duplex is configured, the link will continue to operate.

Speed/duplex selection

The decision to allow a port to auto-negotiate its speed and duplex, or to force it to a set value, is a subject of frequent debate. Testing in the lab has shown that GbESM can correctly negotiate the link when attaching to Cisco switches. With that said, conventional wisdom indicates to always hard-code critical links within a network. During the lab testing we found problems when we did not hard-code the GbESM side. This was especially true in scenarios where static link aggregation was used. We therefore recommend that you hard-code the speed and duplex between the GbESM, while using the default auto-negotiation on the Cisco switches. In the case of the examples presented in this chapter, this will always be set to 100/full on the GbESM and Auto on the Cisco side.

Important: In production environments, it is strongly recommended that 1000BaseT connections be used (available on all Cisco platforms suitable for datacenter environments) to ensure the best possible throughput. Copper-based 1000BaseTs are only supported up to 100 meters.

If the cable distance exceeds 100 meters, fiber must be used. 1000Mbps fiber converters could then be implemented between the BladeCenter and the Core network. In high availability solutions, the fiber converters should have redundant power supplies.

Use of the term trunk

There are a number of terms and acronyms used in the industry that have proven to be the source of much confusion. One such term is the word *trunk* or *trunking*. This term has been used to describe a number of technologies, most commonly the act of bundling links together to increase performance and reliability, and the act of carrying multiple VLANs on a single connection.

Unless otherwise stated, in this chapter, the following definitions apply:

- ▶ *Trunk* or *trunking* - The act of carrying multiple VLANs on a single connection (the connection might be a single link, or a group of links aggregated together to form a Link Aggregation Group). The IEEE specification for VLAN trunking is 802.1.Q.
- ▶ *Aggregation* or *link aggregation* - The act of bundling multiple physical links into one logical link, for the purposes of increasing throughput or offering increased reliability, or both. Link Aggregation is often referred to as EtherChannel in the Cisco world. The IEEE specification for link aggregation is 802.3ad (now part of 802.3 2002).

Note: There are some places in the command line to the GbESM where what is defined as *aggregation* above is shown in the menus as *trunk*. In those cases, we display the menu option and commands as they exist in the command line (for example, Port Trunking), but this still refers to aggregation as defined in this document.

Use of the term Native VLAN

The term *Native VLAN* is used throughout this chapter to describe a single designated untagged VLAN in an 802.1Q trunk. The 802.1Q specification does not define this term, but the concept of untagged VLANs on a trunk is defined within the specification. Cisco has adopted this term to describe a VLAN that provides, among other things, backward compatibility with a device that might not understand 802.1Q, such that at least some communications can take place across this link. In Cisco networks, the Native VLAN is most often VLAN 1, but this can be changed to any VLAN. Any device connecting to a Cisco switch via an 802.1Q trunk should define at least one untagged VLAN, and it must match the untagged VLAN on the Cisco switch (for example, they both define the untagged VLAN as VLAN 1). Note that both the GbESM and all Cisco switches default to VLAN 1 as being the untagged VLAN. More comments about the operation and configuration of the Native VLAN can be found elsewhere within this chapter.

Guidelines and comments - VLANs and trunking

On the Cisco external switch, always configure the trunks connecting to a GbESM as 802.1Q. The GbESM does not support Cisco ISL (Inter-Switch Link).

On the Cisco external switch, always configure the connections connecting to an GbESM as a *trunk* port, rather than an *access* port. This can help to prevent unintentional Spanning Tree loops.

On the Cisco external switch, always use the *nonegotiate* option with connections going to an GbESM, to ensure a trunk can be established. The GbESM does not support Cisco DTP (Dynamic Trunking Protocol) to determine trunk type.

It is recommended that the Native 802.1Q VLAN be VLAN 1 (the default for both the GbESM and all Cisco switches).

Regardless of what number is assigned to the Native VLAN, it must be the same on both sides of the trunk (most commonly both sides call the Native VLAN, VLAN 1). Having different Native VLANs on either side of the trunks can lead to unexpected and undesired operation of the network brought on by Spanning Tree loops.

Make sure that the trunk is *carrying* the native VLAN (usually VLAN 1). There is some control available to prevent a given trunk from carrying a given VLAN. If using this feature, do not block the Native VLAN. By default, both the GbESM and Cisco switches carry VLAN 1 on every trunk.

Guidelines and comments - Spanning Tree

It is recommended that you always leave STP disabled on the BladeCenter GbESM, but the Cisco switches should have STP enabled. The main reason for disabling the STP on the GbESM is that having it enabled could, by a misconfiguration or a cabling error, result in a network loop.

On the Cisco external switch, do not use a link configured as *Access* to connect to the GbESM. Always use a *Trunk* link. Forcing the link to be a trunk link can reduce the likelihood of Spanning Tree loops.

When deploying in an environment where a GbESM will be directly attaching to multiple Cisco switches in a common layer 2 cloud, one has the choice of letting the GbESM and the connecting switches decide which redundant links to block, or manually controlling what links get blocked. It is recommended that you set the path costs on the links between the GbESM and the Cisco switch such that one can be assured, if all links are operational, that a specific set of links will go into blocking.

The choice of path cost, and thus path selection, in your production environment is up to your network administrator, and will depend on factors such as the network architecture, the location of the root switch, the distance of the root switch to the GbESMs, and various port cost settings in between.

The recommendations given about Spanning Tree configuration are based on the following scenario:

Network description

As shown in Figure 5-5 on page 95, a single GbESM with dual aggregated links, each aggregated link going to separate Cisco switches, each Cisco switch joined to each other via a single link (simulating a layer 2 network beyond the switches).

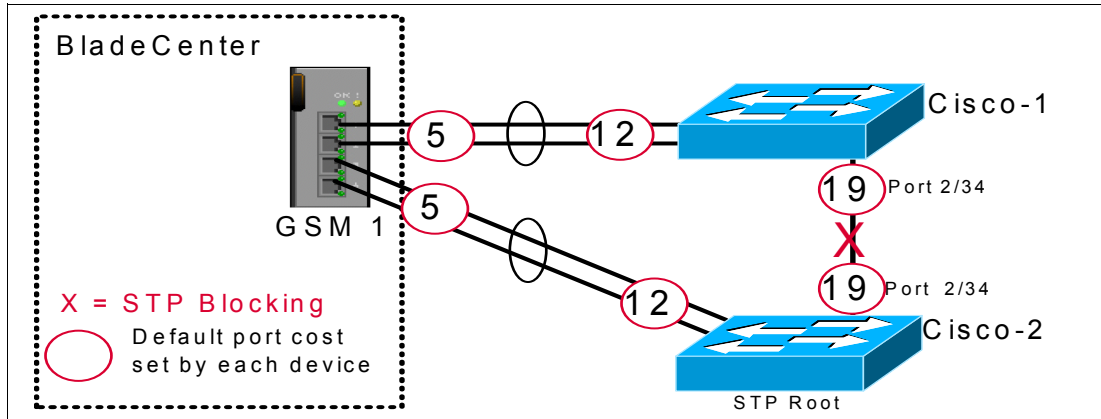


Figure 5-5 Default Port cost for each interface

In the examples presented in this chapter, where multiple Cisco switches are used to provide redundancy, the Cisco-2 switch is always forced to be the root (chosen for consistency of results for the examples, and not because it is a good design). Having the root directly attached to the GbESM is not necessarily recommended in redundant configurations, as flow patterns can become less than obvious.

Based on the default port cost when the GbESM and both Cisco switches are running Spanning Tree, we do not have a desirable traffic pattern, as shown in Example 5-4. We can see that all traffic is going through the GbESM and the port between Cisco-1 and Cisco-2 are in blocking.

Example 5-4 Default spanning tree in solution with STP enabled on all devices.

Cisco-1-IOS#sh spanning-tree root

Vlan	Root ID	Root Cost	Hello Time	Max Age	Fwd Dly	Root Port
VLAN0001	8193 0004.dda6.0e00	17	2	20	15	Po1
VLAN0005	8197 0004.dda6.0e00	12	2	20	15	Po1
VLAN0010	8202 0004.dda6.0e00	12	2	20	15	Po1

Note: The example shown here uses 100Mbps links. Using 1000Mbps links would give lower cost than shown in this example.

We can also see that VLAN 1 has a higher root cost than the path to the roots for VLAN 5 and 10. This is the direct result of joining a Cisco switch running PVST+ to a Mono Spanning Tree switch. The STP for VLAN 1 is taking into account the GbESM link costs, $12+5=17$ (as STP is common on VLAN 1 for both PVST+ and Mono Spanning Tree devices). The other VLANs see the link from Cisco-1 through the GbESM to Cisco-2 as a straight connection through to Cisco-2, cost 12 (as BPDUs for VLANs other than VLAN 1 are essentially tunneled through the GbESM as multicast packets for that VLAN), and are choosing to forward traffic through the GbESM.

While this would be the case if all devices were running either Cisco's PVST+ or Mono Spanning Tree, it is not the case in this mixed environment.

It should be noted that as long as the root is not one of the directly attached switches to the GbESM, this flow will be unlikely to occur (as other path costs will come into play that will more than likely offset this undesired behavior). Even so, it helps to be aware of the possibility

of this occurring, as only traffic destined for the BladeCenter should be forwarded to the GbESM to ensure maximum throughput is available for the blade servers.

For this example, a simple solution is to set the port cost on the po1 interface of Cisco-1 such that it becomes an undesirable path, as shown in Example 5-5 (anything higher than 19 will work) for all VLANs. The following example demonstrates this change:

Example 5-5 Setting higher portcost on Po1 interface

```
Cisco-1-IOS#Conf ter
Enter configuration commands, one per line. End with CNTL/Z.
Cisco-1-IOS(config)#int po1
Cisco-1-IOS(config-if)#spanning-tree cost 100
Cisco-1-IOS(config-if)#end
Cisco-1-IOS#sh span root
```

Vlan	Root ID	Cost	Time	Age	Dly	Root Port
VLAN0001	8193 0004.dda6.0e00	19	2	20	15	Fa2/34
VLAN0005	8197 0004.dda6.0e00	19	2	20	15	Fa2/34
VLAN0010	8202 0004.dda6.0e00	19	2	20	15	Fa2/34

Figure 5-6 shows that with change to the port cost, all VLANs are now choosing the link between Cisco-1 and Cisco-2 as the path to the root (Cisco-2).

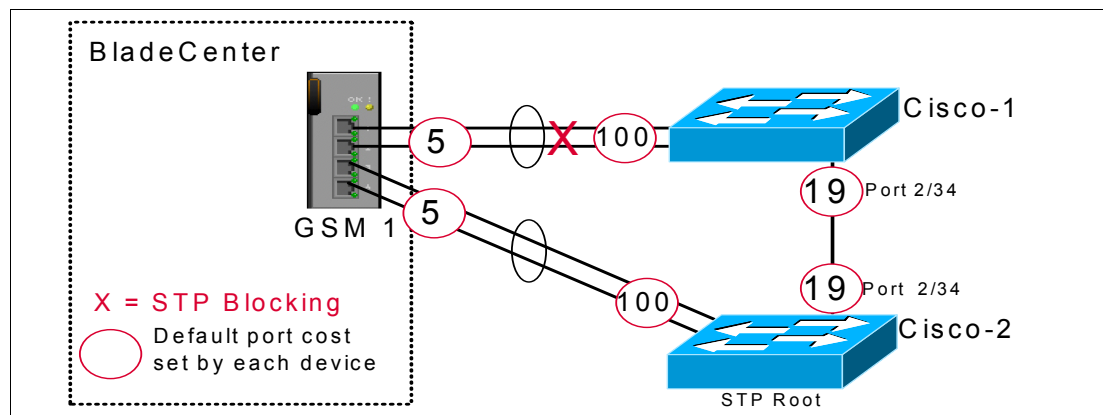


Figure 5-6 Adjusting port costs to block traffic through the GbESM

5.4 Preliminary information on configuration examples

Before we discuss specific configuration examples, it is necessary to discuss some of the basics for all configurations in this chapter.

Some comments on the examples offered

Within the examples provided are syntax and screen captures of commands to complete the desired task. It is possible that a production switch might have configuration commands already in place that conflict with these commands. It is the responsibility of the person configuring the Cisco switch and the GbESM to fully comprehend any changes and their consequences. Failure to fully understand the commands can lead to network-down conditions.

Examples showing the use of link aggregation are always shown using an even number of sequential links. Both the BladeCenter and most Cisco products support using an odd number of links in the aggregation, as well as non-sequential links.

The examples provided assume a layer 2 network exists, and that the reader is attempting to connect the BladeCenter to this layer 2 network. Where appropriate, comments on ports being blocked via STP are included. If the network behind the initial switches is a layer 3 network (routed), the comments on STP blocking will probably be incorrect, depending on your specific environment. It is important for the administrator to understand these situations.

The examples provided do not go into network architecture design, rather they only cover the specifics of interfacing the BladeCenter into a Cisco infrastructure with certain characteristics. It is assumed that the administrator understands the need for and ramifications of a proper network design (see 5.2, "Architecture summary" on page 86 for a brief introduction to datacenter architectures).

Examples that show use of link aggregation are only in reference to layer 2 link aggregation. Some Cisco switches support layer 3 aggregation, but their description and use are not discussed in these examples.

There are various options to control the way traffic is load balanced over any aggregated links. This section assumes that default load balancing is in use.

General rules and comments for configuring the GbESM

While there are many possible ways to create a desired configuration, care must always be taken to understand the consequences of any such configuration.

The GbESM supports configuration via BBI (browser-based interface) and a telnet interface. This chapter only discusses the telnet interface. The telnet interface is accessible by telnet session to the IP address of the GbESM to be configured, and logging in. The default user ID for the telnet session is as follows:

User ID: **admin** (All lowercase)

Recommended sequence of configuration

The following are the basic steps we followed in producing most of the examples presented in this paper:

1. Shut down or un-cable the links to be configured (Table 5-1 on page 99).
2. Configure the GbESM.
 - d. Configure speed and duplex of links.
 - e. Configure any desired VLANs and VLAN trunking options.
 - f. Configure any desired aggregation links.
 - g. Save the configuration to NVRAM.
3. Configure the Cisco equipment.
 - a. Configure desired speed and duplex of links.
 - b. Configure any desired VLANs and VLAN trunking options.
 - c. Configure any desired aggregation links.
 - d. Depending on the Cisco switch, save the config to NVRAM.
4. Re-enable or re-cable the links that were disabled in step 1 (Table 5-2 on page 99).
5. Confirm the desired operation of the configuration.

Base configuration options common to all examples

The following configuration options are common to all of the examples. These are only for demonstration purposes in the examples, and more than likely *will not* be duplicated in your particular environment.

- ▶ All configurations have three VLANs configured: VLAN 1, VLAN 5, and VLAN 10.
- ▶ All configurations assume that VLAN 1 is the native VLAN (Native VLAN is untagged).
- ▶ All configurations assume that all VLANs will be carried on all trunks. It is possible to limit which VLANs are carried on a given trunk, but this is not presented in this chapter.
- ▶ All configurations force one of the Cisco switches to be the Spanning Tree root for all VLANs.

Important: There is a high probability that any existing network will already have a desired switch configured as the root. It is very important that you understand the proper selection of the root bridge, and that the GbESM not be allowed to become the root bridge. Allowing the GbESM to become the root bridge can result in sub-optimal data flow within the layer 2 network.

- ▶ The following blade servers internal to the BladeCenter are placed in the specified VLANs during the configuration stage of each example:
 - BladeServer 1: VLAN 1
 - BladeServer 2: VLAN 5
 - BladeServer 3: VLAN 10

Summary of disconnect procedure, to be performed for each example

When performing initial configurations or making changes to existing configurations that might have an impact on Spanning Tree (such as changing link aggregation), it is recommended that you leave connections un-cabled, or shut down, prior to making the configuration changes. This will reduce the likelihood of any temporary Spanning Tree loops and possible network-down conditions that might result in the process of adding or changing configurations.

Table 5-1 on page 99 shows three basic options to disable the connection.

1. Option 1: Disable the GbESM interface
2. Option 2: Disable the Cisco interface
3. Option 3: Unplug the cable(s)

Choose the one best suited to your situation. For example, if you will not be physically at the equipment while you are performing the configuration, physically disconnecting the cables is not your best option.

Table 5-1 Pre-configuration step: Disable the links being configured

Description and comments	Action
Option 1 - Disable the GbESM interface. Repeat for any other Ext interfaces involved in the configuration.	Perform the following from the telnet interface: >> Main# cfg/port Enter port (INT1-14, MGT1-2, EXT1-4): ext1 >> Port EXT1# dis Current status: enabled New status: disabled >> Port EXT1# apply
Option 2a - Disable the Cisco interface (CatOS). Repeat for any other desired interfaces.	Perform the following from the enable mode: To disable a single port: set port disable 2/3 To disable a range of ports from 2/3 to 2/6: set port disable 2/3-6
Option 2b - Disable the Cisco interface (IOS). Repeat for any other desired interfaces.	Perform the following from the enable mode: To disable a single port: config t inter FastEthernet 2/3 shutdown end To disable a range of ports from FastEthernet 2/3 - FastEthernet 2/6: config t inter range FastEthernet 2/3 -6 shutdown end
Option 3 - Pull the cable(s).	Unplug the cable(s) from either the GbESM or the Cisco external switch of the link(s) as necessary.

Summary of reconnect procedure, to be performed for each example

Table 5-2 includes the steps performed once the configuration of both sides of the connection is complete. It should be the reverse of whatever procedure was followed from Table 5-1.

Table 5-2 Post configuration step: Reconnecting the devices

Description and comments	Action
Option 1 - Re-enable the GbESM interface. Repeat for any other Ext interfaces involved in the configuration.	Perform the following from the telnet interface: >> Main# cfg/port Enter port (INT1-14, MGT1-2, EXT1-4): ext1 >> Port EXT1# ena Current status: disabled New status: enabled >> Port EXT1# apply
Option 2a - Enable the Cisco interface (CatOS). Repeat for any other desired interfaces.	Perform the following from the enable mode: To enable a single port: set port enable 2/3 To enable a range of ports from 2/3 to 2/6: set port enable 2/3-6

Description and comments	Action
Option 2b - Enable the Cisco interface (IOS). Repeat for any other desired interfaces.	Perform the following from the enable mode: To enable a single port: <pre> config t inter FastEthernet 2/3 no shutdown end </pre> To enable a range of ports from FastEthernet 2/3 - FastEthernet 2/6: <pre> config t inter range FastEthernet 2/3 - 6 no shutdown end </pre>
Option 3 - Plug the cable(s) in to the port(s).	Plug in the cable(s) from either the GbESM or the Cisco external switch of the link(s) as necessary.

GbESM base configuration

As already noted, the preparation of examples for this chapter included wiping out each device and installing an initial base configuration prior to performing the procedures given each example.

For the GbESMs there should be no changes necessary after the Factory Reset. Be aware that if this were a brand new @server BladeCenter, you would need to connect to the BladeCenter's Management Module (default 192.168.70.125) and enable the external interfaces of the GbESMs (the default is Disabled) at least once. This can be done by logging in to the Management Module (ID= *USERID* and Password = *PASSWORD*) and clicking **I/O Module Tasks** → **Management** → **Bay X** (where **X** is the GbESM to manage), click **Advanced Management** → **Advanced Setup** and change External ports to Enabled (see Figure 5-7).

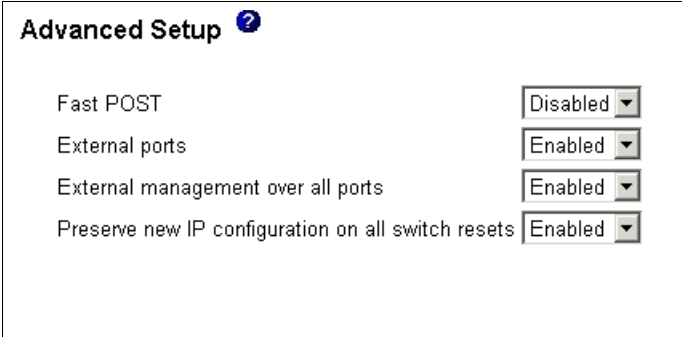


Figure 5-7 Enabling the external ports for the first time via the Management Module

After the change is done, click **Save**.

Cat 4006 CatOS-based switch configuration

As with the GbESM- and IOS-based switches, the CatOS-based switch was wiped out and a base configuration applied, prior to each example. After the switch was reset to the factory default, the following configuration was applied to the CatOS-based switch to simulate

portions of a pre-existing Cisco network. As already mentioned, this is only for the examples shown here, and more than likely will vary from your production network.

Several things to note about this base configuration:

- ▶ Module 2 is a 4232-GB-RJ, with two ports of GBIC based GigE and 34 ports of 10/100BaseT.
- ▶ Port 2/20 has been configured as a test point to test access to VLAN 1.
- ▶ Port 2/15 has been configured as a test point to test access to VLAN 5.
- ▶ Port 2/10 has been configured as a test point to test access to VLAN 10.
- ▶ Port 2/3 through 2/6 will be configured as desired in each section. This will include setting up the trunk link to the GbESM, as well as any desired aggregation.
- ▶ In examples with more than one Cisco switch, port 2/34 on switch Cisco-1 will be connected to port 2/34 on Cisco-2 to simulate a Layer 2 cloud behind the two Cisco switches (this will help to show any Spanning Tree blocking that might occur with a given configuration).
- ▶ In the examples with more than one Cisco switch, the switch named Cisco-2 will be forced to become the root switch (for consistency purposes only). This can be done by running the **set spantree priority YYYYY X** command such that YYYYY is lower on Cisco-2 than it is on Cisco-1 for each of the three sample VLANs (VLAN in the command shown here marked as X).
- ▶ Telnet and Enable passwords are all set to `cisco` (lowercase).
- ▶ The configuration in Example 5-6 is for a CatOS-based switch acting as Cisco-1 (VLAN 1 management IP address set to 192.168.70.202).
- ▶ For a Cisco-2 configuration of a CatOS-based switch, the VLAN 1 management IP address would be set to 192.168.70.203, and the name of the switch would be changed accordingly (Cisco-2-CatOS).

Example 5-6 CatOS configuration

```
#version 7.6(3)
!
!
#system web interface version(s)
set password $2$4CZy$7im5fQA35UD.7hR5us/U0.
set enablepass $2$NMK/$VnHGG4PHSDFPFG28d7IpP0
set logout 0
!
#system
set system name Cisco-1-CatOS
!
!
#vtp
set vtp domain IBM
set vtp mode transparent
set vlan 5,10
!
#ip
# For Cisco-1 deployment
set interface sc0 1 192.168.70.202/255.255.255.0 192.168.70.255
!
set interface s10 down
set interface me1 down
!
#vlan                                <VlanId>
```

```

set spantree priority 24576 1
set spantree priority 24576 5
set spantree priority 24576 10
!
#module 2 : 34-port 10/100/1000 Ethernet
set vlan 5 2/15
!
set vlan 10 2/10
!
!
!
set trunk 2/34 nonegotiate dot1q
set spantree portfast 2/10,2/15,2/20 enable

```

Cat 4006 IOS-based switch configuration

As with the GbESM, each switch in each example was wiped out and a base configuration installed prior to the commencement of each example. In the case of IOS-based switches, after the switch was reset to the factory default, the following configuration was applied (to simulate portions of a pre-existing Cisco network). As already mentioned, this was only for the examples shown here, and more than likely will vary from your production network.

Several things to note about this base configuration:

- ▶ Module 2 is a 4232-GB-RJ, with two ports of GBIC based GigE and 34 ports of 10/100BaseT.
- ▶ Port 2/20 has been configured as a test point to test access to VLAN 1.
- ▶ Port 2/15 has been configured as a test point to test access to VLAN 5.
- ▶ Port 2/10 has been configured as a test point to test access to VLAN 10.
- ▶ Port 2/3 through 2/6 will be configured as desired in each section. This will include setting up the trunk link to the GbESM, as well as any desired aggregation.
- ▶ In examples with more than one Cisco switch, port 2/34 on switch Cisco-1 will be connected to port 2/34 on Cisco-2 to simulate a Layer 2 cloud behind the two Cisco switches (this will help to show any Spanning Tree blocking that might occur with a given configuration)
- ▶ In the examples with more than one Cisco switch, the switch named Cisco-2 will be forced to become the root switch (for consistency purposes only). This can be done by running the **spanning-tree vlan X priority YYYYY** command such that YYYYY is lower on Cisco-2 than it is on Cisco-1 for each of the three sample VLANs.
- ▶ Telnet and Enable passwords are all set to `cisco` (lowercase).
- ▶ The following configuration example is for an IOS based switch acting as Cisco-1 (VLAN 1 management IP address set to 192.168.70.200).
- ▶ For a Cisco-2 configuration of an IOS based switch, the VLAN 1 management IP address would be set to 192.168.70.201, and the name of the switch would be changed accordingly (Cisco-2-`IOS`).

Example 5-7 IOS-based configuration

```

!
hostname Cisco-1-IOS
!
enable secret 5 $1$QHTt$SvAEWBZAtQUvWj.uzH07Y1
enable password Cisco
!
ip subnet-zero

```

```

no ip domain-lookup
!
vtp domain IBM
vtp mode transparent

vlan 5
  name VLAN5
!
vlan 10
  name VLAN10
!
spanning-tree mode pvst
spanning-tree extend system-id
spanning-tree vlan 1 priority 8192
spanning-tree vlan 5 priority 8192
spanning-tree vlan 10 priority 8192
!
!
interface FastEthernet2/10
  description VLAN 10 Test point
  switchport
  switchport access vlan 10
  switchport mode access
  no ip address
  spanning-tree portfast
!
interface FastEthernet2/15
  description VLAN 5 Test point
  switchport
  switchport access vlan 5
  switchport mode access
  no ip address
  spanning-tree portfast
!
interface FastEthernet2/20
  description VLAN 1 Test point
  switchport
  switchport mode access
  no ip address
  spanning-tree portfast
!
interface FastEthernet2/34
  description Trunk port to rest of layer 2 network
  switchport
  switchport trunk encapsulation dot1q
  switchport mode trunk
  switchport nonegotiate
!
!
interface Vlan1
  ip address 192.168.70.200 255.255.255.0
  no shut
!
ip classless
!
!
line con 0
  exec-timeout 0 0
line vty 0 4

```

```
exec-timeout 0 0
password cisco
login
!
end
```

5.5 Configuration examples

This section contains the actual configuration examples for installing and verifying the IBM @server BladeCenter Layer 2-7 GbE Switch Module in a layer 2 Cisco infrastructure environment.

Note: All examples given are based on layer 2 network connections. While it is possible (and desirable in some environments) to connect the layer 2 interfaces of the GbESMs directly to routed interfaces (layer 3) of Cisco switches, that discussion is beyond the scope of this document and is not included here.

Note: While some of the examples provided show a number of ways to achieve layer 2 redundancy, they do not take into account some of the possible issues that may be encountered with Spanning Tree in complex layer 2 networks. With this in mind, the examples should only be viewed as a possibility of what can be done, and not what is necessarily desirable in any specific environment. While Spanning Tree can be a valuable tool to control loops, it should not be used as a substitute for good designs that minimize possible loops but still offer the desired redundancy.

5.5.1 Single GbESM, single link to a single Cisco switch

In this example (Figure 5-8), we discuss a very basic configuration that includes a single GbESM, a single Cisco switch, and a single link between the two. This configuration offers minimal performance and redundancy, and might be used for initial installation and testing of an IBM @server BladeCenter, or in an environment that does not require maximum performance or redundancy.

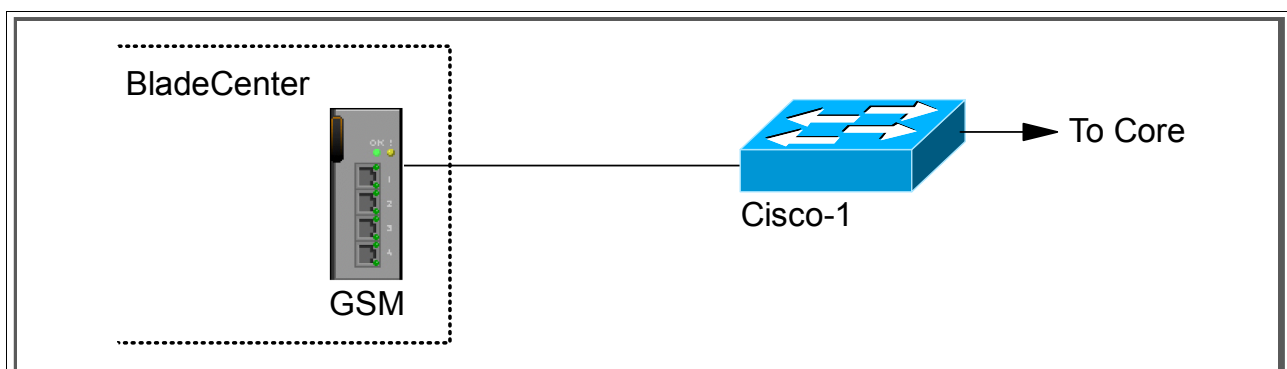


Figure 5-8 Single GbESM, Single Cisco Switch, Single link

Step 1: Take down the link

It is always advisable to disable the link prior to making any configuration changes. Review Table 5-1 on page 99 for procedures.

Step 2: Configure the GbESM side of the link

This section walks you through the sequence of actions required to configure the GbESM for this example (refer to Table 5-3 on page 105).

The following assumptions have been made for this example:

- ▶ The user is already logged into the GbESM with *root* level access.
- ▶ Port EXT1 on the GbESM in Switch Module Bay 1 is being used as the link between IBM and Cisco.
- ▶ Commands are being performed in the sequence shown.
- ▶ The GbESM is starting from a default config per the example on page 100.
- ▶ Cisco switches being used are 10/100Mbps based and we will be setting the GbESM port to 10/100Mbps. This means a cross-over cable *must* be used for the link between the GbESM and the Cisco switch.

Table 5-3 Configuring the GbESM

Description and comments	Actions to perform via CLI interface to GbESM
<p>Step 1 - Configure Internal Ports</p> <p>Turn off STP. Place the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer <i>y</i> to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. This corresponds to BladeServer bay 3. Answer <i>y</i> to the prompt to confirm changing the PVID on port INT2.</p> <p>All other PVIDs should be set for 1 with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 2- Configure Tagging on External Port</p> <p>This enables tagging on the external port so that it can be member of VLANS 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1.</p> <p>Enter apply to apply the new configuration</p>
<p>Step 3- Configure the External Port.</p> <p>This step sets the speed, duplex and VLAN setting for the external port. This will allow external port 1 to carry traffic for VLANS 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p> <p>As already noted, it will be necessary to use a cross-over cable on the link between the GbESM and the Cisco switch because turning off auto-negotiation also turns off the auto selection of MDI/MDI-X.</p>	<p>Enter the command: <code>cfg/port ext1/gig/speed 100</code> to set port speed to 100Mbps.</p> <p>Enter the command: <code>cfg/port ext1/gig/mode full</code> to set duplex settings to full.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10.</p> <p>Enter apply to apply the new configuration</p>

Description and comments	Actions to perform via CLI interface to GbESM
Step 4- Save GbESM config to NVRAM Failure to perform this step will result in all changes to the GbESM being lost if the @server BladeCenter is powered off or the GbESM is otherwise restarted.	Save the changes, <code>save</code>

Step 3: Configure the Cisco switch

The following assumptions have been made for this example (Table 5-4):

- ▶ VLANs 1, 5, and 10 already exist in the Cisco configuration and will be part of the 802.1Q trunk to the GbESM.
- ▶ The user is already logged into the switch and the switch is in enable mode.
- ▶ The lowest available compatible port is being used:
 - For the CatOS switch being used in this example, port 2/3 is being used.
 - For the IOS switch being used in this example, port FastEthernet 2/3 is being used.
- ▶ Commands are being performed in the sequence shown.
- ▶ The switches are starting from default configs as shown in Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS).
- ▶ Cisco switches being used are 10/100Mbps based.

If the Cisco switches used had been 1000-based, the configuration would have been different than in this example.

Attention: During the testing we found that hardcoding the CatOS and IOS speed prevented us from getting the link aggregation working. The problem was solved by using the `auto` setting on the ports of the Cisco equipment. This is the default setting.

Table 5-4 Configuring the Cisco equipment

Description and comments	CatOS-based switch	IOS-based switch
Step 1: Configure speed and enable the ports. Use default setting (auto.)	(does not apply)	(does not apply)
Step 2: Configure 802.1Q trunking Forces link to become an 802.1Q VLAN trunk.	<code>set trunk 2/1 dot1q nonegotiate</code>	<pre> config t interface FastEthernet 2/3 switchport trunk encap dot1q switchport mode trunk switchport nonegotiate end </pre>
Step 3: Save config to NVRAM. Only necessary on IOS based switches.	(does not apply)	<code>write mem</code>

Step 4: Reconnect the devices

This is the final step to bring the connection into full operation. This is the reverse of whatever procedure was used in Step 1. See Table 5-2 on page 99 for details on how to re-establish the links.

Step 5: Verify the configuration

This section provides options for verifying the correct and desired operation.

Verify correct operation on the BladeCenter GbESM

Verify the configurations on the GbESM look as follows:

As in Example 5-8, verify the port state by typing the **info/link** command. The link of EXT1 should be up and the Speed/Duplex setting should be 100/Full.

Example 5-8 Verifying link status, speed, and duplex settings

```
>> Main# info/link
-----
Alias  Port  Speed  Duplex  Flow Ctrl  Link
-----  --TX-----RX--  -----
INT1   1     1000   full    yes     yes     up
INT2   2     1000   full    yes     yes     up
INT3   3     1000   full    yes     yes     up
INT4   4     1000   full    yes     yes     down
INT5   5     1000   full    yes     yes     down
INT6   6     1000   full    yes     yes     down
INT7   7     1000   full    yes     yes     down
INT8   8     1000   full    yes     yes     down
INT9   9     1000   full    yes     yes     down
INT10  10    1000   full    yes     yes     down
INT11  11    1000   full    yes     yes     down
INT12  12    1000   full    yes     yes     down
INT13  13    1000   full    yes     yes     down
INT14  14    1000   full    yes     yes     down
MGT1   15    100    full    yes     yes     up
MGT2   16    100    full    yes     yes     disabled
EXT1   17    100    full    no      no      up
EXT2   18    any    any     yes     yes     down
EXT3   19    any    any     yes     yes     down
EXT4   20    any    any     yes     yes     down

>> Information#
```

As in Example 5-9, verify VLAN configurations by the **info/port** command. Make sure that INT2 belongs to VLAN 5 and INT3 belongs to VLAN 10. EXT1 should have y in the tag and should have VLAN 1, 5, and 10 associated to its port.

Example 5-9 Shows the VLAN and Tag information

```
>>>> Main# info/port
Alias  Port  Tag  RMON  PVID  NAME  VLAN(s)
-----  --
INT1   1     n    d     1     INT1  1
INT2   2     n    d     5     INT2  5
INT3   3     n    d     10    INT3  10
INT4   4     n    d     1     INT4  1
INT5   5     n    d     1     INT5  1
INT6   6     n    d     1     INT6  1
INT7   7     n    d     1     INT7  1
```

```

INT8    8    n    d    1    INT8    1
INT9    9    n    d    1    INT9    1
INT10   10   n    d    1    INT10   1
INT11   11   n    d    1    INT11   1
INT12   12   n    d    1    INT12   1
INT13   13   n    d    1    INT13   1
INT14   14   n    d    1    INT14   1
MGT1    15   n    d    4095 MGT1    4095
MGT2    16   n    d    4095 MGT2    4095
EXT1    17   y    d    1    EXT1    1    5    10
EXT2    18   n    d    1    EXT2    1
EXT3    19   n    d    1    EXT3    1
EXT4    20   n    d    1    EXT4    1

```

>> Information#

As in Example 5-10, verify that Spanning Tree is turned off for all ports by using the **info/12/stg** command. It is also possible to see that INT1, INT2, INT3 and EXT1 are in FORWARDING state.

Example 5-10 Shows Spanning Tree status

>> Main# info/12/stg

Spanning Tree Group 1: Off, FDB aging timer 300

Port	Priority	Cost	State	Designated Bridge	Des Port
INT1	0	0	FORWARDING	*	
INT2	0	0	FORWARDING	*	
INT3	0	0	FORWARDING	*	
INT4	0	0	DISABLED	*	
INT5	0	0	DISABLED	*	
INT6	0	0	DISABLED	*	
INT7	0	0	DISABLED	*	
INT8	0	0	DISABLED	*	
INT9	0	0	DISABLED	*	
INT10	0	0	DISABLED	*	
INT11	0	0	DISABLED	*	
INT12	0	0	DISABLED	*	
INT13	0	0	DISABLED	*	
INT14	0	0	DISABLED	*	
EXT1	0	0	FORWARDING	*	
EXT2	0	0	DISABLED	*	
EXT3	0	0	DISABLED	*	
EXT4	0	0	DISABLED	*	

* = STP turned off for this port.

Verify correct operation on the Cisco external switch

Table 5-5 shows some commands you can use to verify the desired configuration and operation of the Cisco equipment.

Table 5-5 Verifying the configuration and operation of the Cisco external switch of the connection

Description and comments	CatOS-based switch	IOS-based switch
Review running config for desired statements.	show config Review for the following: <ul style="list-style-type: none"> • set trunk 2/1 nonegotiate dot1q 1-1005,1025-4094 	show run Review for the following: <ul style="list-style-type: none"> • interface FastEthernet2/3 • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate
Show speed and duplex.	show port status 2/1 Should show the following: <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100 	show int FastEthernet 2/3 status Should show the following: <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100
Show trunking status. Link must be up before a trunk will come up.	show port trunk 2/1 Should show the following: <ul style="list-style-type: none"> • Mode = nonegotiate • Encapsulation = dot1q • Status = Trunking • Native VLAN = 1 	show int FastEthernet 2/3 trunk Should show the following: <ul style="list-style-type: none"> • Mode = on • Encapsulation =802.1q • Status = Trunking • Native VLAN = 1
Ping the GbESM. Where x.x.x.x is the IP address of the GbESM (must be in same VLAN as subnet being pinged).	ping x.x.x.x Based on the Cisco-1 being at 192.168.70.202 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.	ping x.x.x.x Based on the Cisco-1 being at 192.168.70.200 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.
Attach a device to VLAN 5 and 10 on the Cisco switch and verify a ping to one of the blade servers in that same VLAN works as desired.	For VLAN 5, attach a device to port 2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port 2/10 and attempt to ping across to the IP address on the blade server in slot 3 (note that IP subnets must match for ping to work).	For VLAN 5, attach a device to port fa0/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port fa0/10 and attempt to ping across to the IP address on the blade server in slot 3 (note that IP subnets must match for ping to work).

5.5.2 Single GbESM, single link to two Cisco switches

Figure 5-9 on page 110 shows a little more robust configuration, with a single GbESM connecting to two different Cisco switches. It still offers minimal performance but with increased redundancy should an uplink or one of the external switches fail. It does not offer any redundancy in the event of a GbESM failure.

Based on our specific configuration for this example (layer 2 network, Cisco-2 is root and port cost on both Cisco links toward the GbESM set to 100), Spanning Tree will block the connection between Cisco-1 and the GbESM at the Cisco external switch until such time as the link from the GbESM to Cisco-2 (or Cisco-2 itself) goes down.

Important: The choice of root and port cost settings were only made *for this example*, and may be a poor choice in a production network (certainly placing the root switch up against the GbESM in a datacenter environment would not be very common). It is very important that any time a GbESM is connected in a redundant fashion, the location of the root switch and various port costs in the network be accounted for to ensure correct data flow (we need to prevent non-BladeCenter traffic from flowing through the GbESM).

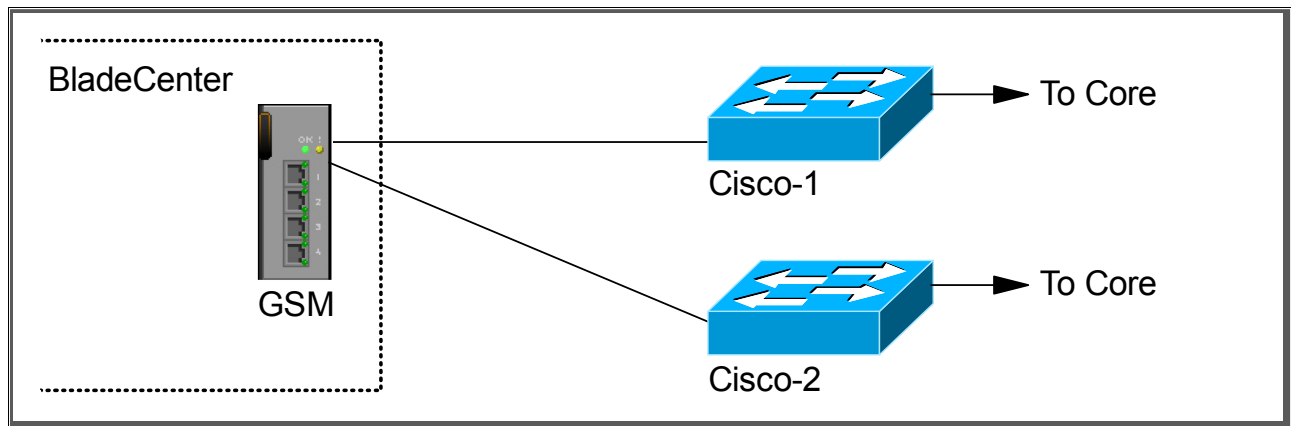


Figure 5-9 Single GbESM with single links to two different Cisco switches

Step 1: Take down the links

It is always advisable to disable the links prior to making any configuration changes. See Table 5-1 on page 99 for procedures.

Step 2: Configure the GbESM side of the link

Table 5-6 on page 111 walks you through the sequence of actions required to configure the GbESM for this example.

The following assumptions have been made for this example:

- ▶ The user is already logged into the GbESM with *root* level access.
- ▶ Port EXT1 on the GbESM in Switch Module Bay 1 is being used as the link between the @server BladeCenter and Cisco-1.
- ▶ Port EXT2 on the GbESM in Switch Module Bay 1 is being used as the link between the @server BladeCenter and Cisco-2.
- ▶ Commands are being performed in the sequence shown.
- ▶ The GbESM is starting from a default config per the example on page 100.
- ▶ Cisco switches being used are 10/100Mbps based and we will be setting the GbESM port to 10/100Mbps full duplex. This means a cross-over cable *must* be used for the link between the GbESM and the Cisco switch.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Table 5-6 Configuring the GbESM

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 1 - Configure Internal Ports</p> <p>Turn off STP. Then place the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. This corresponds to BladeServer bay 3. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>All other PVIDs should be set for 1 with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 2- Configure Tagging on External Ports</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1.</p> <p>Enter the command: <code>/cfg/port EXT2/tag e</code> to enable tagging on external port 2.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 3- Configure the External Ports</p> <p>This step sets the speed, duplex, and VLAN setting for the external ports. This will allow external ports 1 and 2 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p> <p>As already noted, it will be necessary to use a cross-over cable on the link between the GbESM and the Cisco switch, as tuning off auto-negotiation also turns off the auto selection of MDI/MDI-X.</p>	<p>Enter the command: <code>cfg/port ext1/gig/speed 100</code> to set port speed to 100Mbps for external port 1. Repeat for external port 2</p> <p>Enter the command: <code>cfg/port ext1/gig/mode full</code> to set duplex settings to full for external port 1. Repeat for external port 2</p> <p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external port 2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external port 2.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 4- Save GbESM config to NVRAM</p> <p>Failure to perform this step will result in all changes to the GbESM being lost if the @server BladeCenter is powered off or the GbESM is otherwise restarted.</p>	<p>Save the changes, save.</p>

Step 3: Configure the Cisco switches

The following assumptions have been made for this example (refer to Table 5-7):

- ▶ VLANs 1, 5, and 10 already exist in the Cisco configuration and will be part of the 802.1Q trunk to the GbESM.
- ▶ The user is already logged into the switch and the switch is in enable mode.
- ▶ The lowest available compatible port is being used:
 - For the CatOS switches being used in this example, port 2/3 is being used.
 - For the IOS switches being used in this example, port FastEthernet 2/3 are being used.

- ▶ Commands are being performed in the sequence shown.
- ▶ The switches are starting from a default config per Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS).
- ▶ Cisco switches being used are 10/100Mbps based.

If the Cisco switches used had been 1000 based, the configuration would have been different than in this example

Attention: During the testing we found that hardcoding the CatOS and IOS speed prevented us from getting the link aggregation working. The problem was solved by using the *auto* setting on the ports of the Cisco equipment. This is the default setting.

Table 5-7 Configuring the Cisco equipment

Description and comments	CatOS-based switch	IOS-based switch
Step 1: Configure speed and enable the ports. Use default setting (auto).	(does not apply)	(does not apply)
Step 2: Configure 802.1Q trunking Forces link to become an 802.1Q VLAN trunk. Need to perform on both Cisco-1 and Cisco-2.	<code>set trunk 2/1 dot1q nonegotiate</code>	<code>config t interface FastEthernet 2/3 switchport trunk encap dot1q switchport mode trunk switchport nonegotiate</code>
Step 3: Configure Spanning Tree port cost. Need to perform on both Cisco-1 and Cisco-2. Setting the port cost higher than default helps to prevent VLAN traffic between Cisco-1 and Cisco-2, other than VLAN 1, from being switched through the GbESM. This results in a more predictable flow. For more information, review the section titled: "Guidelines and comments - Spanning Tree" on page 94	<code>set spantree portcost 2/1 100</code>	<code>spanning-tree cost 100 end</code>
Step 4: Save config to NVRAM. Only necessary on IOS-based switches. Need to perform on both Cisco-1 and Cisco-2.	(does not apply)	<code>write mem</code>

Step 4: Reconnect the devices

This is the final step to bring the connection into full operation. This will be the reverse of whatever procedure was used in Step 1. See Table 5-2 on page 99 for details on how to re-establish the links.

Step 5: Verify the configuration

This section provides options for verifying the correct and desired operation.

Certain elements of this verification depend on the existence of a layer 2 network behind the two Cisco switches. This was simulated for this example by tying the two switches together

with an 802.1Q trunk link. (See Example 5-6 on page 101 (CatOS) and Example 5-7 on page 102 (IOS) to see how this link was configured).

Verify correct operation on the BladeCenter GbESM

Verify the configurations on the GbESM look as follows.

As in Example 5-11, verify the port state by typing the `info/link` command. The link of EXT1, and EXT2 should be Up and the Speed/Duplex setting should be 100/Full.

Example 5-11 Verifying Link Status, Speed and Duplex settings

```
>> Main# info/link
```

Alias	Port	Speed	Duplex	Flow Ctrl		Link
-----	----	-----	-----	--TX--	---RX---	-----
INT1	1	1000	full	yes	yes	up
INT2	2	1000	full	yes	yes	up
INT3	3	1000	full	yes	yes	up
INT4	4	1000	full	yes	yes	down
INT5	5	1000	full	yes	yes	down
INT6	6	1000	full	yes	yes	down
INT7	7	1000	full	yes	yes	down
INT8	8	1000	full	yes	yes	down
INT9	9	1000	full	yes	yes	down
INT10	10	1000	full	yes	yes	down
INT11	11	1000	full	yes	yes	down
INT12	12	1000	full	yes	yes	down
INT13	13	1000	full	yes	yes	down
INT14	14	1000	full	yes	yes	down
MGT1	15	100	full	yes	yes	up
MGT2	16	100	full	yes	yes	disabled
EXT1	17	100	full	no	no	up
EXT2	18	100	full	no	no	up
EXT3	19	any	any	yes	yes	down
EXT4	20	any	any	yes	yes	down

```
>> Information#
```

As in Example 5-12, verify VLAN configurations by the `info/port` command. Make sure that INT2 belongs to VLAN 5 and INT3 belongs to VLAN 10. EXT1 and EXT2 should have y in the tag and should have VLAN 1, 5, and 10 associated to its ports.

Example 5-12 Shows the VLAN and tag information

```
>> Main# info/port
```

Alias	Port	Tag	RMON	PVID	NAME	VLAN(s)
-----	----	---	----	----	-----	-----
INT1	1	n	d	1	INT1	1
INT2	2	n	d	5	INT2	5
INT3	3	n	d	10	INT3	10
INT4	4	n	d	1	INT4	1
INT5	5	n	d	1	INT5	1
INT6	6	n	d	1	INT6	1
INT7	7	n	d	1	INT7	1
INT8	8	n	d	1	INT8	1
INT9	9	n	d	1	INT9	1
INT10	10	n	d	1	INT10	1
INT11	11	n	d	1	INT11	1
INT12	12	n	d	1	INT12	1

```

INT13 13 n d 1 INT13 1
INT14 14 n d 1 INT14 1
MGT1 15 n d 4095 MGT1 4095
MGT2 16 n d 4095 MGT2 4095
EXT1 17 y d 1 EXT1 1 5 10
EXT2 18 y d 1 EXT2 1 5 10
EXT3 19 n d 1 EXT3 1
EXT4 20 n d 1 EXT4 1

```

>> Information#

As in Example 5-13, verify that Spanning Tree is turned off for all ports by using the **info/12/stg** command. It is also possible to see that INT1, INT2, INT3, EXT1 and EXT 2 are in FORWARDING state. Based on the configurations in this example (Cisco-2 is root switch, and setting of path costs) these ports should show FORWARDING. The Cisco external switch of the EXT1 link should show blocking for that connection.

Example 5-13 Show Spanning Tree status

>> Main# info/12/stg

Spanning Tree Group 1: Off, FDB aging timer 300

Port	Priority	Cost	State	Designated Bridge	Des Port
INT1	0	0	FORWARDING *		
INT2	0	0	FORWARDING *		
INT3	0	0	FORWARDING *		
INT4	0	0	DISABLED *		
INT5	0	0	DISABLED *		
INT6	0	0	DISABLED *		
INT7	0	0	DISABLED *		
INT8	0	0	DISABLED *		
INT9	0	0	DISABLED *		
INT10	0	0	DISABLED *		
INT11	0	0	DISABLED *		
INT12	0	0	DISABLED *		
INT13	0	0	DISABLED *		
INT14	0	0	DISABLED *		
EXT1	0	0	FORWARDING *		
EXT2	0	0	FORWARDING *		
EXT3	0	0	DISABLED *		
EXT4	0	0	DISABLED *		

* = STP turned off for this port

Verify correct operation on the Cisco external switch

Table 5-8 on page 115 includes some commands you can use to verify the desired configuration and operation of the Cisco equipment.

Table 5-8 Verifying the configuration

Description and comments	CatOS-based switch	IOS-based switch
Review running config for desired statements.	show config Review for the following: <ul style="list-style-type: none"> • set trunk 2/1 nonegotiate dot1q 1-1005,1025-4094 • set spantree portcost 2/1 100 	show run Review for the following on interface FastEthernet 2/3 <ul style="list-style-type: none"> • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate • spanning-tree cost 100
Show speed and duplex.	show port status 2/1 Should show the following: <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100 	show int FastEthernet 2/3 status Should show the following: <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100
Show trunking status. Link must be up before a trunk will come up.	show port trunk 2/1 Should show the following: <ul style="list-style-type: none"> • Mode = nonegotiate • Encapsulation = dot1q • Status = Trunking • Native VLAN = 1 	show int FastEthernet 2/3 trunk Should show the following: <ul style="list-style-type: none"> • Mode = on • Encapsulation =802.1q • Status = Trunking • Native VLAN = 1
Show Spanning Tree status. In this configuration Cisco-2 should show forwarding and Cisco-1 blocking	show spantree 2/1 State should show forwarding for connection on Cisco-2 and blocking for connection between Cisco-1 and GbESM for all VLANs.	show spanning int FastEthernet 2/3 State should show forwarding for connection on Cisco-2 and blocking for connection between Cisco-1 and GbESM for all VLANs
Ping the GbESM. Where x.x.x.x is the IP address of the GbESM (must be in same VLAN as subnet being pinged).	ping x.x.x.x Based on the Cisco-1 being at 192.168.70.202 and Cisco-2 at 192.168.70.203 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.	ping x.x.x.x Based on the Cisco-1 being at 192.168.70.200 and Cisco-2 at 192.168.70.201 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.
Attach a device to VLAN 5 and 10 on the Cisco switch and verify a ping to one of the blade servers in that same VLAN, works as desired	For VLAN 5, attach a device to port 2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port 2/10 and attempt to ping across to the IP address on the blade server in slots 3 (note that IP subnets must match for ping to work).	For VLAN 5, attach a device to port fa2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port fa2/10 and attempt to ping across to the IP address on the blade server in slots 3 (note that IP subnets must match for ping to work).

5.5.3 Single GbESM, four port static aggregation to a single Cisco switch

This example (Figure 5-10) shows a single GbESM using all four external ports dynamically aggregated into a single pipe to a single Cisco switch. It produces the maximum performance from a single GbESM and also offers redundancy in the event of a link failure. It does not offer any redundancy in the event of a switch failure (either on the part of the GbESM or the part of the Cisco switch). This configuration is suitable for environments that require maximum throughput with a single GbESM, but that are not very concerned with switch redundancy.

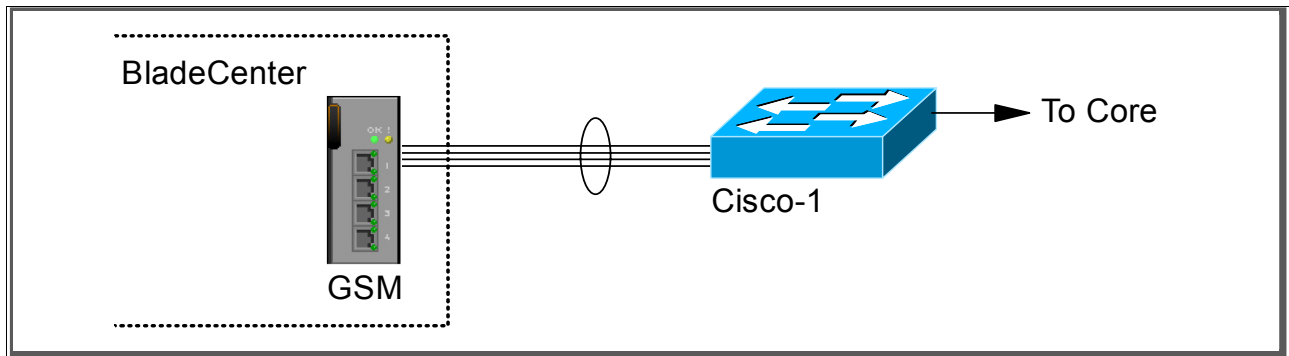


Figure 5-10 Single GbESM with four port static aggregation into a single Cisco switch

Step 1: Take down the links

It is always advisable to disable the links prior to making any configuration changes. See Table 5-1 on page 99 for procedures.

Step 2: Configure the IBM side of the link

Table 5-9 on page 117 walks you through the actions required to configure the GbESM for this example.

The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with *root* level access.
- ▶ Commands are being performed in the sequence shown.
- ▶ The GbESM is starting from a default config per the “GbESM base configuration” on page 100.
- ▶ Cisco switches being used for this lab are 10/100Mbps and we set the GbESM ports to 100/100 full duplex. This means a cross-over cable *must* be used for the links between the GbESM and the Cisco switch.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Table 5-9 Configuring the GbESM

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 1: Configure internal ports.</p> <p>Turn off STP. Then place the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. This corresponds to BladeServer bay 3. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>All other PVIDs should be set for 1 with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 2: Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 3: Configure the external ports.</p> <p>This step sets the speed, duplex and VLAN setting for the external ports. This will allow external ports 1 and 2 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p> <p>As already noted, it will be necessary to use a cross-over cable on the link between the GbESM and the Cisco switch, as turning off auto-negotiation also turns off the auto selection of MDI/MDI-X.</p>	<p>Enter the command: <code>cfg/port ext1/gig/speed 100</code> to set portspeed to 100Mbps for external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>cfg/port ext1/gig/mode full</code> to set duplex settings to full for external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2, 3, and 4.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 4- Define a trunk group.</p>	<p>Enter the command: <code>/cfg/12/trunk 1/add ext1</code> to add ports to the trunk group. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>enable</code> to enable trunk group.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 5: Save GbESM config to NVRAM.</p> <p>Failure to perform this step will result in all changes to the GbESM being lost if the @server BladeCenter is powered off or the GbESM is otherwise restarted.</p>	<p>Save the changes, save.</p>

Step 3: Configure the Cisco switch

The following assumptions have been made for this example (refer to Table 5-10):

- ▶ VLANs 1, 5, and 10 already exist in the Cisco configuration and will be part of the 802.1Q trunk to the GbESM.
- ▶ The user is already logged into the switch and the switch is in enable mode.

- ▶ The lowest available compatible ports are being used:
 - For the CatOS switch used in this example, ports 2/3 through 2/6 are used.
 - For the IOS switch used in this example, ports FastEthernet 2/3 through 2/6 are used.
- ▶ Commands are performed in the sequence shown.
- ▶ The switches are starting from a default config per Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS).
- ▶ Cisco switches used for this lab are 10/100Mbps.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Attention: During the testing we found that hardcoding the CatOS and IOS speed prevented us from getting the link aggregation working. The problem was solved by using the *auto* setting on the ports of the Cisco equipment. This is the default setting.

Table 5-10 Configuring the Cisco equipment

Description and comments	CatOS-based switch	IOS-based switch
Step 1: Configure speed and enable the ports. Use default setting (auto).	(does not apply)	(does not apply)
Step 2: Configure 802.1Q trunking. Forces the link to become an 802.1Q VLAN trunk.	<code>set trunk 2/3 nonegotiate</code> <code>set trunk 2/4 nonegotiate</code> <code>set trunk 2/5 nonegotiate</code> <code>set trunk 2/6 nonegotiate</code>	<code>Configure t</code> <code>int range FastEthernet 2/3 - 6</code> <code>switchport trunk encap dot1q</code> <code>switchport mode trunk</code> <code>switchport nonegotiate</code> Note that the range option on IOS is only available in newer revisions of code. If the range option is not available, you will need to repeat steps 2 through 3 for each interface.
Step 3: Configure static link aggregation.	<code>set port channel 2/3-6 mode on</code>	<code>channel-group 1 mode on</code> <code>end</code>
Step 4: Save config to NVRAM. Only necessary on IOS-based switches.	(does not apply)	<code>write mem</code>

Step 4: Reconnect the devices

This is the final step to bring the connection into full operation. This will be the reverse of whatever procedure was used in Step 1. See Table 5-2 on page 99 for details on how to re-establish the links.

Step 5: Verify the configuration

This section provides options for verifying the correct and desired operation.

Verify correct operation on the BladeCenter GbESM

Verify that the configurations on the GbESM look as follows.

As shown in Example 5-14, verify the port state by typing the `info/link` command. The links of EXT1, EXT2, EXT3, and EXT4 should be Up and the Speed/Duplex setting should be 100/Full.

Example 5-14 Verifying link status, speed and duplex settings

```
>> Main# info/link
-----
Alias  Port  Speed  Duplex  Flow Ctrl  Link
-----  ---  -----  -
--TX--RX--
INT1   1     1000   full    yes       yes       up
INT2   2     1000   full    yes       yes       up
INT3   3     1000   full    yes       yes       up
INT4   4     1000   full    yes       yes       down
INT5   5     1000   full    yes       yes       down
INT6   6     1000   full    yes       yes       down
INT7   7     1000   full    yes       yes       down
INT8   8     1000   full    yes       yes       down
INT9   9     1000   full    yes       yes       down
INT10  10    1000   full    yes       yes       down
INT11  11    1000   full    yes       yes       down
INT12  12    1000   full    yes       yes       down
INT13  13    1000   full    yes       yes       down
INT14  14    1000   full    yes       yes       down
MGT1   15    100    full    yes       yes       up
MGT2   16    100    full    yes       yes       disabled
EXT1   17    100    full    no        no        up
EXT2   18    100    full    no        no        up
EXT3   19    100    full    no        no        up
EXT4   20    100    full    no        no        up

>> Information#
```

As shown in Example 5-15, verify VLAN configurations by issuing the `info/port` command. Make sure that INT2 belongs to VLAN 5 and INT3 belongs to VLAN 10. EXT1, EXT2, EXT3 and EXT4 should have y in the tag and should have VLAN 1, 5, and 10 associated to its ports.

Example 5-15 Shows the VLAN and Tag information

```
>> Main# info/port
Alias  Port  Tag  RMON  PVID  NAME  VLAN(s)
-----  ---  ---  ---  ---  -
INT1   1     n    d     1     INT1  1
INT2   2     n    d     5     INT2  5
INT3   3     n    d     10    INT3  10
INT4   4     n    d     1     INT4  1
INT5   5     n    d     1     INT5  1
INT6   6     n    d     1     INT6  1
INT7   7     n    d     1     INT7  1
INT8   8     n    d     1     INT8  1
INT9   9     n    d     1     INT9  1
INT10  10    n    d     1     INT10 1
INT11  11    n    d     1     INT11 1
INT12  12    n    d     1     INT12 1
INT13  13    n    d     1     INT13 1
INT14  14    n    d     1     INT14 1
MGT1   15    n    d     4095  MGT1  4095
```

```

MGT2 16 n d 4095 MGT2 4095
EXT1 17 y d 1 EXT1 1 5 10
EXT2 18 y d 1 EXT2 1 5 10
EXT3 19 y d 1 EXT3 1 5 10
EXT4 20 y d 1 EXT4 1 5 10

```

>> Information#

As shown in Example 5-16, verify that Spanning Tree is turned off for all ports by using the **info/12/stg** command. It is also possible to see that INT1, INT2, INT3, EXT1, EXT2, EXT3 and EXT4 are in FORWARDING state.

Example 5-16 Show Spanning Tree status

>> Main# info/12/stg

Spanning Tree Group 1: Off, FDB aging timer 300

Port	Priority	Cost	State	Designated Bridge	Des Port
INT1	0	0	FORWARDING *		
INT2	0	0	FORWARDING *		
INT3	0	0	FORWARDING *		
INT4	0	0	DISABLED *		
INT5	0	0	DISABLED *		
INT6	0	0	DISABLED *		
INT7	0	0	DISABLED *		
INT8	0	0	DISABLED *		
INT9	0	0	DISABLED *		
INT10	0	0	DISABLED *		
INT11	0	0	DISABLED *		
INT12	0	0	DISABLED *		
INT13	0	0	DISABLED *		
INT14	0	0	DISABLED *		
EXT1	0	0	FORWARDING *		
EXT2	0	0	FORWARDING *		
EXT3	0	0	FORWARDING *		
EXT4	0	0	FORWARDING *		

* = STP turned off for this port.

As shown in Example 5-17 on page 120, verify that trunk group 1 is forwarding for EXT1 through EXT4 by using the **info/12/trunk** command.

Example 5-17 Verify that ports EXT1 through EXT4 are forwarding

>> Main# info/12/trunk

Trunk group 1, port state:

```

EXT1: STG 1 forwarding
EXT2: STG 1 forwarding
EXT3: STG 1 forwarding
EXT4: STG 1 forwarding

```

As shown in Example 5-18, you can also verify that the aggregation is operational by generating traffic across the link using the command **stats/port ext1/if** and monitoring the counters for the four EXT ports. Depending on how much traffic is being generated, and from how many sources, the numbers will vary across the ports. For this example, 10000, 1400

byte pings were being sent from the Cisco switch to the IP address of the GbESM. In this case, most traffic is using EXT1.

Example 5-18 Traffic traversing the links

```
>> Main# stats/port ext1/if
```

```
-----  
Interface statistics for port EXT1:  
          ifHCIn Counters      ifHCOut Counters  
Octets:                30428                0  
UcastPkts:              0                0  
BroadcastPkts:         0                0  
MulticastPkts:         428                0  
Discards:              0                0  
Errors:                0                0
```

```
>> Main# stats/port ext2/if
```

```
-----  
Interface statistics for port EXT2:  
          ifHCIn Counters      ifHCOut Counters  
Octets:                943                0  
UcastPkts:              0                0  
BroadcastPkts:         1                0  
MulticastPkts:         3                0  
Discards:              0                0  
Errors:                0                0
```

```
>> Main# stats/port ext3/if
```

```
-----  
Interface statistics for port EXT3:  
          ifHCIn Counters      ifHCOut Counters  
Octets:                1236                0  
UcastPkts:              0                0  
BroadcastPkts:         1                0  
MulticastPkts:         4                0  
Discards:              0                0  
Errors:                0                0
```

```
>> Main# stats/port ext4/if
```

```
-----  
Interface statistics for port EXT4:  
          ifHCIn Counters      ifHCOut Counters  
Octets:                1593                0  
UcastPkts:              0                0  
BroadcastPkts:         0                0  
MulticastPkts:         7                0  
Discards:              0                0  
Errors:                0                0
```

If one were to pull the cable for EXT1, the traffic would switch over to a different EXT port, usually with no loss of packets (usually under 1 second).

As shown in Example 5-19, with the same traffic flowing on the link, the cable to EXT1 is removed. The result is that the traffic previously carried on EXT1 is now shared over the 3 remaining ports (it could have gone to any of the available EXT ports in the aggregation). The counters in the example were not cleared, so they include the numbers generated in Example 5-18 on page 121.

Example 5-19 Traffic traversing the links when EXT1 is physically removed

Jan 1 0:47:16 NOTICE system: link down on port EXT1

>> Main# stats/port ext1/if

Interface statistics for port EXT1:
 ifHCIn Counters ifHCOut Counters
Octets: 34277 0
UcastPkts: 0 0
BroadcastPkts: 0 0
MulticastPkts: 553 0
Discards: 0 0
Errors: 0 0

>> Main# stats/port ext2/if

Interface statistics for port EXT2:
 ifHCIn Counters ifHCOut Counters
Octets: 33172 809
UcastPkts: 0 0
BroadcastPkts: 1 5
MulticastPkts: 454 0
Discards: 0 0
Errors: 0 0

>> Main# stats/port ext3/if

Interface statistics for port EXT3:
 ifHCIn Counters ifHCOut Counters
Octets: 2701 809
UcastPkts: 0 0
BroadcastPkts: 1 5
MulticastPkts: 9 0
Discards: 0 0
Errors: 0 0

>> Main# stats/port ext4/if

Interface statistics for port EXT4:
 ifHCIn Counters ifHCOut Counters
Octets: 2829 809
UcastPkts: 0 0
BroadcastPkts: 1 5
MulticastPkts: 11 0
Discards: 0 0
Errors: 0 0

Verify correct operation on the Cisco external switch

Table 5-11 on page 123 provides some commands you can use to verify the desired configuration and operation of the Cisco equipment.

Table 5-11 Verifying the configuration and operation of the Cisco external switch of the connection

Description and comments	CatOS-based switch	IOS-based switch
Review running config for desired statements.	<p>show config</p> <p>Review for the following:</p> <ul style="list-style-type: none"> • set trunk 2/3 negotiate dot1q 1-1005,1025-4094 • set trunk 2/4 negotiate dot1q 1-1005,1025-4094 • set trunk 2/5 negotiate dot1q 1-1005,1025-4094 • set trunk 2/6 negotiate dot1q 1-1005,1025-4094 • set port channel 2/3-6 mode on 	<p>show run</p> <p>Review for the following on interface Port-channel1:</p> <ul style="list-style-type: none"> • switchport • switchport trunk encapsulation dot1q • switchport mode trunk • switchport negotiate <p>Note that the values in Port-channel1 may not show up if the aggregation has never come up since first being configured.</p> <p>Review for the following on int fastethernet 2/3 through fastethernet 2/6:</p> <ul style="list-style-type: none"> • switchport trunk encapsulation dot1q • switchport mode trunk • switchport negotiate • channel-group 1 mode on
Show speed and duplex.	<p>Run the following command on each interface, 2/3 through 2/6:</p> <p>show port status 2/3</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100 	<p>Run the following command on each interface, fastethernet 2/3 through 2/6:</p> <p>show int fastethernet 2/3 status</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100
Show trunking status. Aggregation must be up before the trunk will come up.	<p>Run the following command on each interface, 2/3 through 2/6:</p> <p>show port trunk 2/3</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = negotiate • Encapsulation = dot1q • Status = Trunking • Native VLAN = 1 	<p>show int Port-channel 1 trunk</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = on • Encapsulation =802.1q • Status = Trunking • Native VLAN = 1
Review status of the aggregated link.	<p>show port channel</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • Channel Mode = on <p>show channel traffic</p> <p>Should show how the links in the channel are being utilized.</p>	<p>show etherchannel summary</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Protocol = - (- = no protocol (static)) • Ports fa2/3 - 6 = (P) (P) = part of an aggregation group
Ping the GbESM. Where x.x.x.x is the IP address of the GbESM (must be in same VLAN as subnet being pinged).	<p>ping x.x.x.x</p> <p>Based on the Cisco-1 being at 192.168.70.202 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.</p>	<p>ping x.x.x.x</p> <p>Based on the Cisco-1 being at 192.168.70.200 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.</p>

Description and comments	CatOS-based switch	IOS-based switch
Attach a device to VLAN 5 and 10 on the Cisco switch and verify a ping to one of the blade servers in that same VLAN works as desired.	For VLAN 5, attach a device to port 2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port 2/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).	For VLAN 5, attach a device to port fa0/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port fa0/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).

5.5.4 Single GbESM, dual port static aggregation to two Cisco switches

This example (Figure 5-11) offers a compromise between performance and redundancy. It makes use of a single GbESM with two ports aggregated to one Cisco switch, and the remaining two ports aggregated to a second Cisco switch. While suitable for those seeking a compromise between performance and redundancy, it suffers from no protection in the event of a GbESM failure.

Based on our specific configuration for this example (layer 2 network, Cisco-2 is root and port cost on both Cisco links toward the GbESM set to 100) Spanning Tree will block the connection between Cisco-1 and the GbESM, at the Cisco external switch until such time as the aggregation from the GbESM to Cisco-2 (or Cisco-2 itself) goes down. The choice of root and port cost settings in this example was only made for this example, and may be a poor choice in a production network. (Placing the root switch up against the GbESM in a datacenter environment would not be very common.) It is very important that any time an GbESM is connected in a redundant fashion, the location of the root switch and various port costs in the network be accounted for to ensure correct data flow (need to prevent non-BladeCenter traffic from flowing through the GbESM).

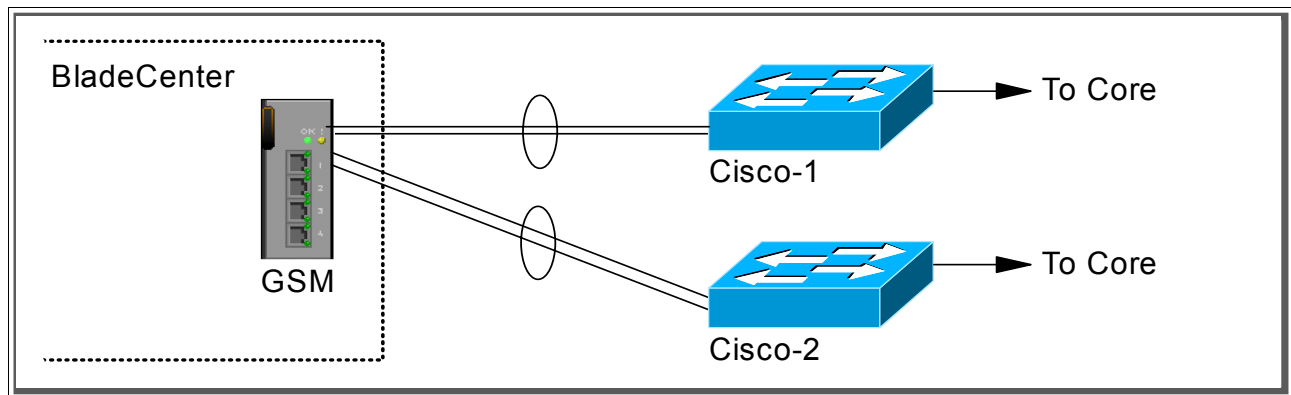


Figure 5-11 Single GbESM with two 2 port static aggregated links to two different Cisco switches

Step 1: Take down the links

It is always advisable to disable the links prior to making any configuration changes. See Table 5-1 on page 99 for procedures.

Step 2: Configure the GbESM side of the link

Table 5-12 on page 125 walks you through the actions required to configure the GbESM for this example.

The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with *root* level access.
- ▶ Commands are performed in the sequence shown.
- ▶ The GbESM is starting from a default config per “GbESM base configuration” on page 100.
- ▶ Cisco switches used for this lab are 10/100Mbps and we set the GbESM ports to 100 full duplex. This means a cross-over cable *must* be used for the links between the GbESM and the Cisco switch.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Table 5-12 Configuring the GbESM

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 1 - Configure internal ports.</p> <p>Turn off STP. Then place the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. This corresponds to BladeServer bay 3. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>All other PVIDs should be set to 1 with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 2- Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter apply to apply the new configuration</p>
<p>Step 3- Configure the external ports.</p> <p>This step sets the speed, duplex and VLAN setting for the external ports. This will allow external port 1 and 2 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p> <p>As already noted, it will be necessary to use a cross-over cable on the link between the GbESM and the Cisco switch, as turning off auto-negotiation also turns off the auto selection of MDI/MDI-X.</p>	<p>Enter the command: <code>cfg/port ext1/gig/speed 100</code> to set port speed to 100Mbps for external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>cfg/port ext1/gig/mode full</code> to set duplex settings to full for external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2, 3, and 4.</p> <p>Enter apply to apply the new configuration.</p>

Description and comments	Actions to perform via Web interface to GbESM
Step 4- Define trunk groups.	Enter the command: <code>/cfg/12/trunk 1/add ext1</code> to add ports to the trunk group. Repeat for external port 2. Enter the command: <code>enable</code> to enable trunk group. Enter the command: <code>/cfg/12/trunk 2/add ext3</code> to add ports to the trunk group. Repeat for external port 4. Enter the command: <code>enable</code> to enable trunk group. Enter <code>apply</code> to apply the new configuration
Step 5- Save GbESM config to NVRAM. Failure to perform this step will result in all changes to the GbESM being lost if the @server BladeCenter is powered off or the GbESM is otherwise restarted.	Save the changes, <code>save</code>

Step 3: Configure the Cisco switches

The following assumptions have been made for this example (refer to Table 5-13 on page 126):

- ▶ VLANs 1, 5, and 10 already exist in the Cisco configuration and will be part of the 802.1Q trunk to the GbESM.
- ▶ The user is already logged into the switch and the switch is in enable mode.
- ▶ The lowest available compatible ports are used:
 - For the CatOS switches used in this example, ports 2/3 through 2/4 are used.
 - For the IOS switch used in this example, ports FastEthernet 2/3 through 2/4 are used.
- ▶ Commands are performed in the sequence shown.
- ▶ The switches are starting from a default config per Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS).
- ▶ Cisco switches used for this lab are 10/100Mbps.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Attention: During the testing we found that hardcoding the CatOS and IOS speed prevented us from getting the link aggregation working. The problem was solved by using the `auto` setting on the ports of the Cisco equipment. This is the default setting.

Table 5-13 Configuring the Cisco equipment

Description and comments	CatOS-based switch	IOS-based switch
Step 1 - Configure speed and enable the ports. Use default setting (auto).	(does not apply)	(does not apply)

Description and comments	CatOS-based switch	IOS-based switch
<p>Step 2- Configure 802.1Q trunking.</p> <p>Forces link to become an 802.1Q VLAN trunk.</p>	<pre>set trunk 2/3 nonegotiate set trunk 2/4 nonegotiate</pre>	<pre>Configure t int range FastEthernet 2/3 - 4 switchport trunk encap dot1q switchport mode trunk switchport nonegotiate</pre> <p>Note that the range option on IOS is only available in newer revisions of code. If the range option is not available you will need to repeat step 2 through 3 for each interface.</p>
<p>Step 3- Configure Spanning Tree port cost on individual ports.</p> <p>This must be performed on both Cisco-1 and Cisco-2.</p> <p>Setting the port cost higher than default helps to prevent VLAN traffic between Cisco-1 and Cisco-2, other than VLAN 1, from being switched through the GbESM. This results in a more optimal flow.</p>	<pre>set spantree portcost 2/3-4 100</pre> <p>Note that for this particular design, leaving the links at their default cost will produce split traffic flows (see “Guidelines and comments - Spanning Tree” on page 94 for details on this behavior)</p>	<pre>spanning-tree cost 100</pre> <p>Note that for this particular design, leaving the links at their default cost will produce split traffic flows (see “Guidelines and comments - Spanning Tree” on page 94 for details on this behavior)</p>
<p>Step 4- Configure static link aggregation.</p> <p>Must be performed on both Cisco-1 and Cisco-2 switch.</p>	<pre>set port channel 2/3-4 mode on</pre>	<pre>channel-group 1 mode on exit</pre>
<p>Step 5- Set port cost on the aggregated link.</p> <p>Must be performed on both Cisco-1 and Cisco-2 switch.</p>	<p>(does not apply)</p> <p>CatOS uses the values from the individual ports to meet this requirement.</p>	<pre>int port-channel 1 spanning-tree cost 100 end</pre> <p>See Step 3 for details. In this case, set the cost for the whole aggregation as well as the individual ports.</p>
<p>Step 6- Save config to NVRAM.</p> <p>Only necessary on IOS-based switches.</p> <p>Must be performed on both Cisco-1 and Cisco-2 switch.</p>	<p>(does not apply)</p>	<pre>write mem</pre>

Step 4: Reconnect the devices

This is the final step to bring the connection into full operation. This will be the reverse of whatever procedure was used in Step 1. See Table 5-2 on page 99 for details on how to re-establish the links.

Step 5: Verify the configuration

This section provides options for verifying the correct and desired operation.

Certain elements of this verification depend on the existence of a layer 2 network behind the two Cisco switches. This was simulated for this example by tying the two switches together with an 802.1Q trunk link (see Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS) for details on how this link was configured).

Verify correct operation on the BladeCenter GbESM

Verify that the configurations on the GbESM look as follows:

As shown in Example 5-20, verify the port state by typing the **info/link** command. The link of EXT1, EXT2, EXT3 and EXT4 should be Up and the Speed/Duplex setting should be 100/Full.

Example 5-20 Verify link status, speed and duplex settings

```
>> Main# info/link
-----
Alias  Port  Speed  Duplex  Flow Ctrl  Link
-----  ---  -----  -
--TX--RX--
INT1    1    1000    full    yes    yes    up
INT2    2    1000    full    yes    yes    up
INT3    3    1000    full    yes    yes    up
INT4    4    1000    full    yes    yes    down
INT5    5    1000    full    yes    yes    down
INT6    6    1000    full    yes    yes    down
INT7    7    1000    full    yes    yes    down
INT8    8    1000    full    yes    yes    down
INT9    9    1000    full    yes    yes    down
INT10   10   1000    full    yes    yes    down
INT11   11   1000    full    yes    yes    down
INT12   12   1000    full    yes    yes    down
INT13   13   1000    full    yes    yes    down
INT14   14   1000    full    yes    yes    down
MGT1    15    100    full    yes    yes    up
MGT2    16    100    full    yes    yes    disabled
EXT1    17    100    full    no     no     up
EXT2    18    100    full    no     no     up
EXT3    19    100    full    no     no     up
EXT4    20    100    full    no     no     up
>> Information#
```

As shown in Example 5-21, verify VLAN configurations by the **info/port** command. Make sure that INT2 belongs to VLAN 5 and INT3 belongs to VLAN 10. EXT1, EXT2, EXT3 and EXT4 should have y in the Tag and should have VLAN 1, 5, and 10 associated to its ports.

Example 5-21 VLAN and tag information

```
>> Main# info/port
Alias  Port  Tag  RMON  PVID  NAME  VLAN(s)
-----  ---  ---  ---  ---  -
INT1    1    n    d     1    INT1    1
INT2    2    n    d     5    INT2    5
INT3    3    n    d    10    INT3    10
INT4    4    n    d     1    INT4    1
INT5    5    n    d     1    INT5    1
INT6    6    n    d     1    INT6    1
INT7    7    n    d     1    INT7    1
INT8    8    n    d     1    INT8    1
INT9    9    n    d     1    INT9    1
INT10   10   n    d     1    INT10   1
```

```

INT11 11 n d 1 INT11 1
INT12 12 n d 1 INT12 1
INT13 13 n d 1 INT13 1
INT14 14 n d 1 INT14 1
MGT1 15 n d 4095 MGT1 4095
MGT2 16 n d 4095 MGT2 4095
EXT1 17 y d 1 EXT1 1 5 10
EXT2 18 y d 1 EXT2 1 5 10
EXT3 19 y d 1 EXT3 1 5 10
EXT4 20 y d 1 EXT4 1 5 10
>> Information#

```

As shown in Example 5-22, verify that Spanning Tree is turned off for all ports by using the `info/12/stg` command. It is also possible to see that INT1, INT2, INT3, EXT1, EXT2, EXT3 and EXT4 are in FORWARDING state.

Example 5-22 Spanning Tree status

```

>> Main# info/12/stg
-----
Spanning Tree Group 1: Off, FDB aging timer 300

Port   Priority   Cost   State   Designated Bridge   Des Port
-----
INT1   0          0      FORWARDING *
INT2   0          0      FORWARDING *
INT3   0          0      FORWARDING *
INT4   0          0      DISABLED *
INT5   0          0      DISABLED *
INT6   0          0      DISABLED *
INT7   0          0      DISABLED *
INT8   0          0      DISABLED *
INT9   0          0      DISABLED *
INT10  0          0      DISABLED *
INT11  0          0      DISABLED *
INT12  0          0      DISABLED *
INT13  0          0      DISABLED *
INT14  0          0      DISABLED *
EXT1   0          0      FORWARDING *
EXT2   0          0      FORWARDING *
EXT3   0          0      FORWARDING *
EXT4   0          0      FORWARDING *
* = STP turned off for this port.

```

As shown in Example 5-23, verify that trunk group 1 and trunk group 2 are forwarding for EXT1 through EXT4 by using the `info/12/trunk` command.

Example 5-23 Ensure that port EXT1 through EXT4 are forwarding

```

>> Main# info/12/trunk
Trunk group 1, port state:
EXT1: STG 1 forwarding
EXT2: STG 1 forwarding

Trunk group 2, port state:
EXT3: STG 1 forwarding
EXT4: STG 1 forwarding

```

Verify correct operation on the Cisco external switch

Table 5-14 includes some commands you can use to verify the desired configuration and operation of the Cisco equipment.

Table 5-14 Verifying the configuration and operation of the Cisco external switch

Description and comments	CatOS-based switch	IOS-based switch
<p>Review running config for desired statements.</p> <p>Perform on both Cisco-1 and Cisco-2.</p>	<p>show config</p> <p>Review for the following:</p> <ul style="list-style-type: none"> • set trunk 2/3 nonegotiate dot1q 1-1005,1025-4094 • set trunk 2/4 nonegotiate dot1q 1-1005,1025-4094 • set trunk 2/34 nonegotiate dot1q 1-1005,1025-4094 • set spantree portcost 2/3-4 100 • set port channel 2/3-4 mode on 	<p>show run</p> <p>Review for the following on interface Port-channel1:</p> <ul style="list-style-type: none"> • switchport • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate • spanning-tree cost 100 <p>Note that the values in Port-channel1 may not show up if the aggregation has never come up since first being configured.</p> <p>Review for the following on int fastethernet 2/3 through fastethernet 2/4:</p> <ul style="list-style-type: none"> • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate • spanning-tree cost 100 • channel-group 1 mode on
<p>Show speed and duplex.</p> <p>Perform on both Cisco-1 and Cisco-2.</p>	<p>Issue the following command on each interface, 2/3 and 2/4:</p> <p>show port status 2/3</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100 	<p>Issue the following command on both interfaces, fa2/3 and fa2/4:</p> <p>show int fastethernet 2/3 status</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100
<p>Show trunking status.</p> <p>Aggregation must be up before the trunk will come up.</p> <p>Perform on both Cisco-1 and Cisco-2.</p>	<p>Issue the following command on each interface, 2/3 and 2/4:</p> <p>show port trunk 2/3</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = nonegotiate • Encapsulation = dot1q • Status = Trunking • Native VLAN = 1 	<p>show int Port-channel 1 trunk</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = on • Encapsulation =802.1q • Status = Trunking • Native VLAN = 1
<p>Review status of the aggregated link.</p> <p>Perform on both Cisco-1 and Cisco-2.</p>	<p>show port channel</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • Channel Mode = on <p>show channel traffic</p> <p>Should show how the links in the channel are being utilized.</p>	<p>show etherchannel summary</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Protocol = - (- = no protocol (static)). • Ports fa2/3 -4 = (P) (P) = part of an aggregation group.

Description and comments	CatOS-based switch	IOS-based switch
<p>Show Spanning Tree status.</p> <p>Note that values shown here are very specific to the example configurations, and may not reflect production information.</p> <p>Perform on both Cisco-1 and Cisco-2.</p>	<p>show spantree 2/3 show spantree 2/4</p> <p>Based on the configuration and cost setting in this example:</p> <ul style="list-style-type: none"> • Cisco 1 switch should show all three VLANs Blocking for these two interfaces. • Cisco-2 switch (root) should show all three VLANs Forwarding for these two interfaces. • Interface 2/34 should also show forwarding for all three VLANs. 	<p>show spanning int fastethernet 2/3 show spanning int fastethernet 2/4 show spanning int po 1</p> <p>Based on the configuration and cost settings in this example:</p> <ul style="list-style-type: none"> • Cisco-1 switch should show BLK (Blocking) for all three VLANs, for each of these interfaces. • Cisco-2 switch (root) should show FWD (Forwarding) for all three VLANs, for each of these interfaces. • Interface fastethernet 2/34 should show forwarding for all three VLANs for both Cisco-1 and Cisco-2.
<p>Ping the GbESM.</p> <p>Where x.x.x.x is the IP address of the GbESM (must be in same VLAN as subnet being pinged).</p>	<p>ping x.x.x.x</p> <p>Based on the Cisco-1 being at 192.168.70.202 and Cisco-2 at 192.168.70.203 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.</p>	<p>ping x.x.x.x</p> <p>Based on the Cisco-1 being at 192.168.70.200 and Cisco-2 at 192.168.70.201 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.</p>
<p>Attach a device to VLAN 5 and 10 on the Cisco switch and verify a ping to one of the blade servers in that same VLAN works as desired.</p>	<p>For VLAN 5, attach a device to port 2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work).</p> <p>For VLAN 10, attach a device to port 2/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).</p>	<p>For VLAN 5, attach a device to port fa0/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work).</p> <p>For VLAN 10, attach a device to port fa0/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).</p>
<p>Test redundancy.</p>	<p>Experiment with shutting down the top two or bottom two ports for the GbESM to force Spanning Tree to make the other path active. Verify with ping tests after Spanning Tree has stabilized.</p>	<p>Experiment with shutting down the top two or bottom two ports for the GbESM to force Spanning Tree to make the other path active. Verify with ping tests after Spanning Tree has stabilized.</p>

5.5.5 Dual GbESMs, each with a single link to the same Cisco switch

In this example (see Figure 5-12), we discuss a basic configuration that includes dual GbESMs, each with a single link to a single Cisco switch. This configuration offers minimal performance and limited redundancy and might be used for initial installation and testing of a dual BladeCenter GbESM, or in an environment that does not require maximum performance. It does offer limited redundancy in the form of two GbESMs, but depends on the operating systems installed on the blade servers to perform redundancy in the event of a GbESM failure.

One thing to note here is that the spanning tree is not an issue in this configuration because each GbESM only has one connection into the layer 2 network.

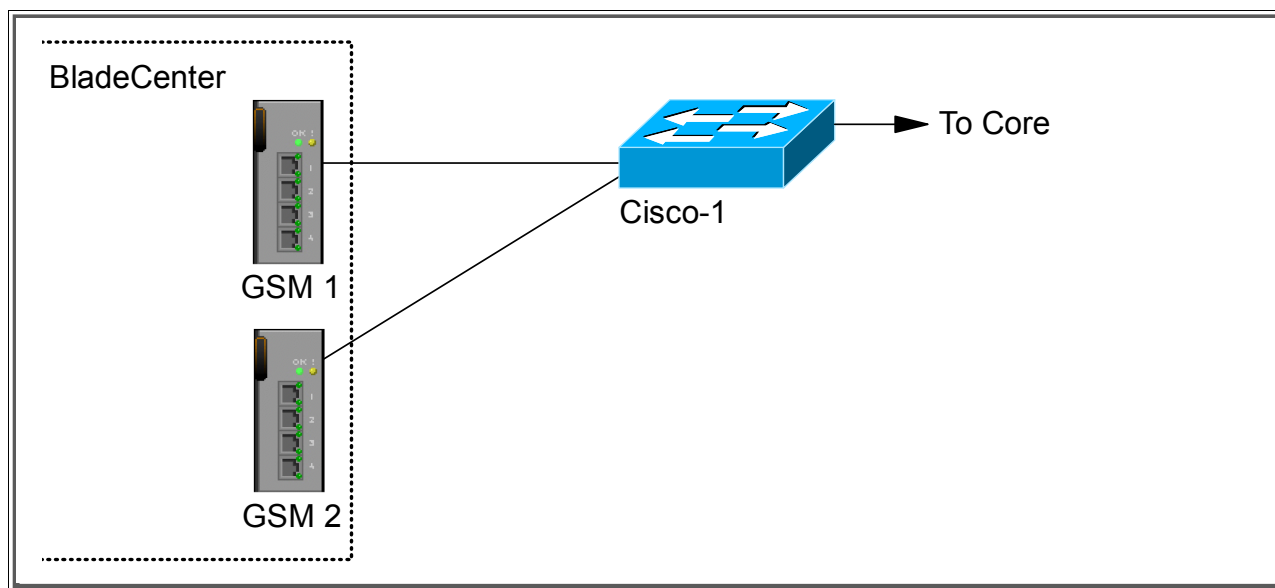


Figure 5-12 Dual GbESMs each with a single link to a single Cisco switch

For this example, refer to the configuration steps presented in 5.5.1, “Single GbESM, single link to a single Cisco switch” on page 104. The only differences are that you will have to configure both GbESMs instead of just one (follow the same procedure for both); and on the Cisco external switch, you will configure a second port (port 2/4 or FastEthernet 2/4, depending if the switch is running CatOS or IOS) (also follow the same procedure on both ports in the Cisco switch).

The verification procedure will be the same on the GbESM side, except that you will need to log in to both GbESMs and verify the desired configuration on each.

For the Cisco external switch of the connection, you can follow the same verification procedure as well, modified slightly to review each port instead of just one port.

The results should be the same for both the GbESM and the Cisco switch as they were in “Single GbESM, single link to a single Cisco switch” on page 104.

One thing to remember here is that there will be no Spanning Tree blocking in this configuration because the GbESMs are completely standalone access switches, with no internal connections to one another (except through the blade servers). It is theoretically possible to obtain and install software onto a blade server that would make it (the blade server) operate as a two port switch, in which case Spanning Tree may come into play (because now the GbESMs are being tied together both inside the IBM @server BladeCenter and external to the IBM @server BladeCenter). The discussion of this interaction is beyond the scope of this document.

5.5.6 Dual GbESMs with a single link to two different Cisco switches

The example (Figure 5-13) described in this section is almost identical to the one in 5.5.1 “Single GbESM, single link to a single Cisco switch” on page 104, with the exception of having two GbESMs and two external switches instead of one. It offers minimal performance and some redundancy. Possible uses for this configuration might be the initial installation and testing of an IBM @server BladeCenter deployment or environments that do not require maximum performance, but do require redundancy. Note that no ports will be in a Spanning

Tree blocking state in this configuration since each GbESM only has a single connection into the layer 2 network.

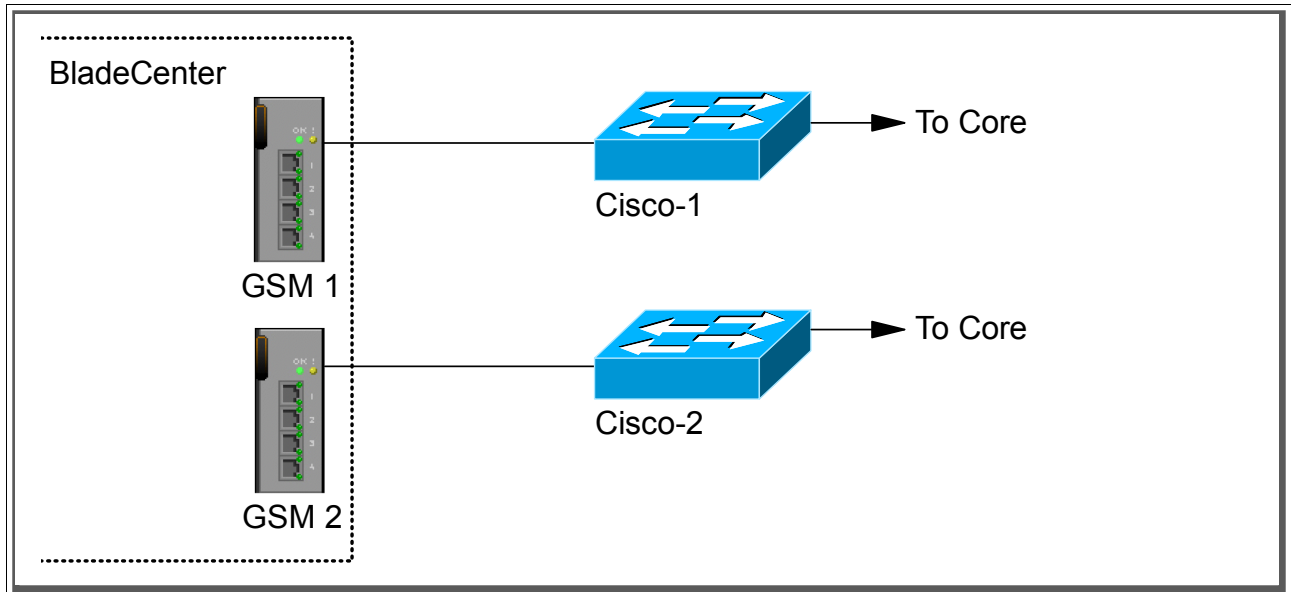


Figure 5-13 Dual GbESMs, each with a single link to two separate Cisco switches

Refer to the configuration example in 5.5.1, “Single GbESM, single link to a single Cisco switch” on page 104. The configuration steps for both the GbESM and the Cisco switches are the same, you just have to perform the set of operations twice, once for GbESM-1 and Cisco-1, and once for the GbESM-2 and Cisco-2.

The verification procedure is the same on the GbESM side, except you’ll need to log in to each GbESM and verify the desired configuration.

For the Cisco external switch of the connection, you can follow the same verification procedure as well, modified slightly to review each port instead of just one port.

For both the GbESM and the Cisco switches, the results should be the same as those described in Section 5.5.1, “Single GbESM, single link to a single Cisco switch” on page 104.

As noted for the previous example, there will be no Spanning Tree blocking in this configuration because the GbESMs are completely standalone access switches, with no internal connections to one another (except through the blade servers). It is theoretically possible to obtain and install software onto a blade server that would make it (the blade server) operate as a two port switch, in which case Spanning Tree may come in to play (since now the GbESMs are being tied together both inside the IBM @server BladeCenter and external to the IBM @server BladeCenter). The discussion of this interaction is beyond the scope of this document.

5.5.7 Dual GbESMs each with one link to separate Cisco switches

This example (see Figure 5-14) begins to show some high redundancy, but still with limited performance. It contains dual GbESMs, each cross-connected to a pair of Cisco switches. Possible uses for this configuration are in environments that are not concerned with performance but require high availability for their network connections.

Based on our specific configuration for this example (layer 2 network, Cisco-2 is root and port cost on both Cisco links toward each GbESM is set to 100), Spanning Tree will normally block the connection between Cisco-1 and the GbESMs, at the Cisco external switch. The choice of

root and port cost settings in this example were only made for this example, and may be a poor choice in a production network. (Certainly, placing the root switch up against the GbESM in a datacenter environment would not be very common.) It is very important that any time a GbESM is connected in a redundant fashion, the location of the root switch and various port costs in the network be accounted for to ensure correct data flow (that is, you need to prevent non-BladeCenter traffic from flowing through the GbESM).

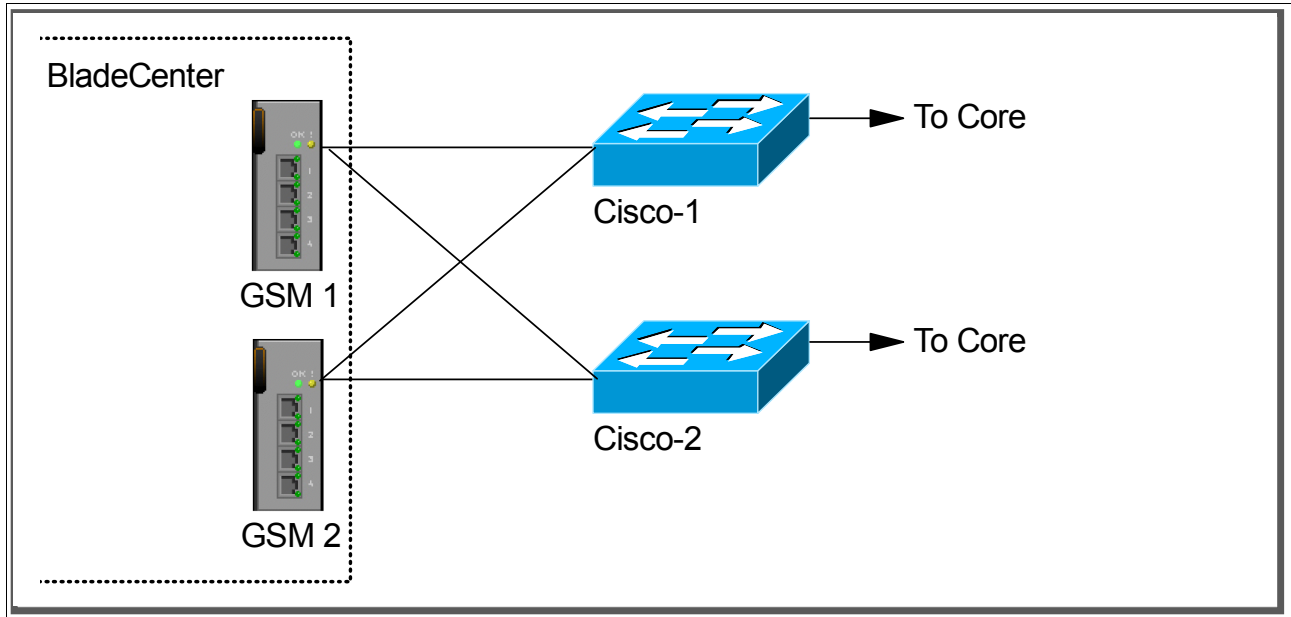


Figure 5-14 Dual GbESMs with cross connected single links to two different Cisco switches

Configuration and verification for this example is similar to 5.5.2, “Single GbESM, single link to two Cisco switches” on page 109, but contains enough differences that a whole new procedure is presented here.

Step 1: Take down the links

It is always advisable to disable the links prior to making any configuration changes. See Table 5-1 on page 99 for procedures.

Step 2: Configure the GbESM side of the link

Table 5-15 on page 135 walks you through the actions required to configure the GbESMs for this example.

The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESMs with *root* level access.
- ▶ The following cabling was used (see Figure 5-15 for a diagram):
 - GbESM-1, Ext1 goes to Cisco-1 Port 2/3 if CatOS / FastEthernet 2/3 if IOS.
 - GbESM-1, Ext2 goes to Cisco-2 Port 2/3 if CatOS / FastEthernet 2/3 if IOS.
 - GbESM-2, Ext1 goes to Cisco-1 Port 2/4 if CatOS / FastEthernet 2/4 if IOS.
 - GbESM-2, Ext2 goes to Cisco-2 Port 2/4 if CatOS / FastEthernet 2/4 if IOS.

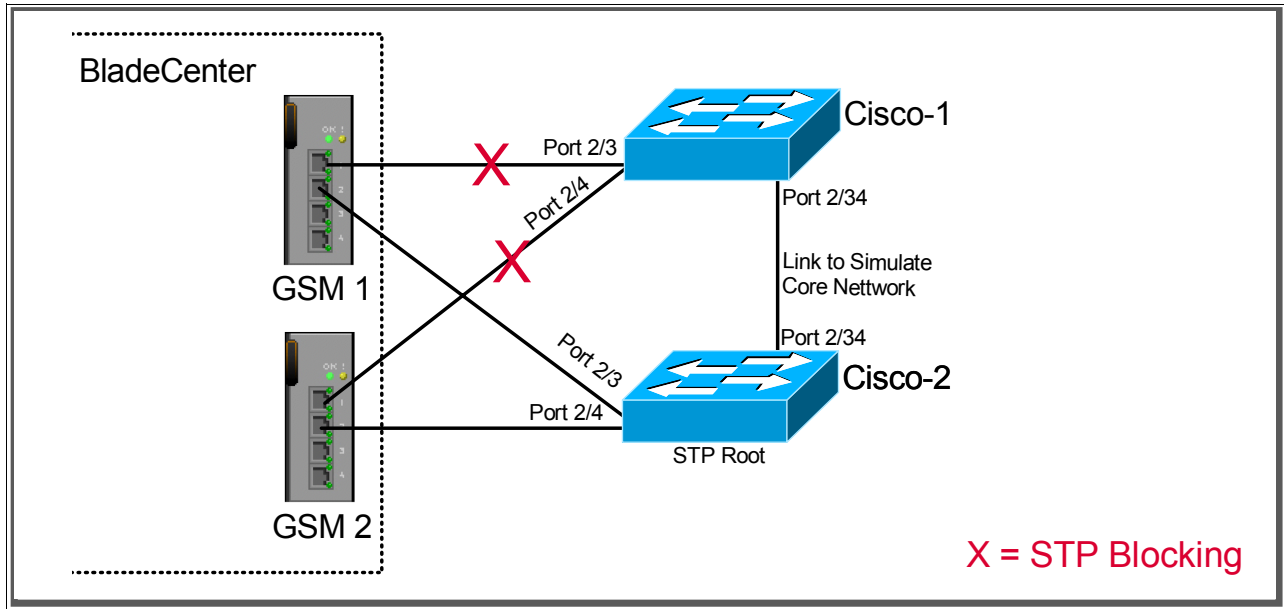


Figure 5-15 Cabling layout used in this example

- ▶ Commands are performed in the sequence shown.
- ▶ The GbESMs are starting from a default config per the “GbESM base configuration” on page 100.
- ▶ Cisco switches used are 10/100Mbps-based and we set the GbESM ports to 10/100Mbps full duplex. This means a cross-over cable *must* be used for the links between the GbESM and the Cisco switches.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Table 5-15 Configuring the GbESMs

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 1 - Configure internal ports</p> <p>Turn off STP. Then place the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged. Perform on both GbESM-1 and GbESM-2.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. This corresponds to BladeServer bay 3. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>All other PVIDs should be set to 1 with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration.</p>

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 2- Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANS 1, 5, and 10 simultaneously. Perform on both GbESM-1 and GbESM-2.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1.</p> <p>Enter the command: <code>/cfg/port EXT2/tag e</code> to enable tagging on external port 2.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 3- Configure the external ports</p> <p>This step sets the speed, duplex and VLAN setting for the external ports. This will allow external port 1 and 2 to carry traffic for VLANS 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p> <p>Use a cross-over cable on the link between the GbESM and the Cisco switch, as turning off auto-negotiation also turns off the auto selection of MDI/MDI-X.</p> <p>Perform on both GbESM-1 and GbESM-2.</p>	<p>Enter the command: <code>cfg/port ext1/gig/speed 100</code> to set port speed to 100Mbps for external port 1. Repeat for external port 2.</p> <p>Enter the command: <code>cfg/port ext1/gig/mode full</code> to set duplex settings to full for external port 1. Repeat for external port 2.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external port 2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external port 2.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 4- Save GbESM config to NVRAM.</p> <p>Failure to perform this step will result in all changes to the GbESM being lost if the @server BladeCenter is powered off or the GbESM is otherwise restarted. Perform on both GbESM-1 and GbESM-2.</p>	<p>Save the changes, save.</p>

Step 3: Configure the Cisco switches

The following assumptions have been made for this example (refer to Table 5-16):

- ▶ VLANs 1, 5, and 10 already exist in the Cisco configurations and will be part of the 802.1Q trunk to the GbESM.
- ▶ The user is already logged in to the switches and the switches are in enable mode.
- ▶ The port connections are as described in Step 2 of this example.
- ▶ Commands are performed in the sequence shown.
- ▶ The switches are starting from a default config per Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS).
- ▶ Cisco switches being used are 10/100Mbps-based.

If using switch ports that support 1000Mbps, the command to set speed will be different from what is shown in the example.

Attention: During the testing we found that hardcoding the CatOS and IOS speed prevented us from getting the link aggregation working. The problem was solved by using the *auto* setting on the ports of the Cisco equipment. This is the default setting.

Table 5-16 Configuring the Cisco equipment

Description and comments	CatOS-based switch	IOS-based switch
Step 1 - Configure speed and enable the ports. Use default setting (auto).	(does not apply)	(does not apply)
Step 2 - Configure 802.1Q trunking. Forces link to become an 802.1Q VLAN trunk.	<code>set trunk 2/3 nonegotiate</code> <code>set trunk 2/4 nonegotiate</code>	<code>Configure t</code> <code>int range FastEthernet 2/3 - 4</code> <code>switchport trunk encap dot1q</code> <code>switchport mode trunk</code> <code>switchport nonegotiate</code> Note that the range option on IOS is only available in newer revisions of code. If the range option is not available you will need to repeat step 2 through 3 for each interface.
Step 3 - Configure Spanning Tree port cost. Do this on both Cisco-1 and Cisco-2. Setting the port cost higher than default helps to prevent VLAN traffic between Cisco-1 and Cisco-2, other than VLAN 1, from being switched through the GbESM. This results in a more predictable flow.	<code>set spantree portcost 2/3 100</code> <code>set spantree portcost 2/4 100</code>	<code>spanning-tree cost 100</code> <code>end</code>
Step 4 - Save config to NVRAM. Only necessary on IOS-based switches.	(does not apply)	<code>write mem</code>

Step 4: Reconnect the devices

This is the final step to bring the connection into full operation. This will be the reverse of whatever procedure was used in Step 1. See Table 5-2 on page 99 for details on how to re-establish the links.

Step 5: Verify the configuration

This section provides options for verifying the correct and desired operation.

Certain elements of this verification depend on the existence of a layer 2 network behind the two Cisco switches. This was simulated for this example by tying the two switches together with an 802.1Q trunk link (see Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS) for details on how this link was configured).

Verify correct operation on the BladeCenter GbESM

Verify that the configurations on the GbESM look as follows.

As shown in Example 5-24, verify the port state by typing the `info/link` command. The link of EXT1, and EXT2 should be Up and the Speed/Duplex setting should be 100/Full. You should see similar results on both GbESM-1 and GbESM-2.

Example 5-24 Verify link status, speed and duplex settings

```
>> Main# info/link
-----
Alias  Port  Speed  Duplex  Flow Ctrl  Link
-----  ---  -----  -
--TX--RX--
INT1   1    1000   full    yes    yes    up
INT2   2    1000   full    yes    yes    up
INT3   3    1000   full    yes    yes    up
INT4   4    1000   full    yes    yes    down
INT5   5    1000   full    yes    yes    down
INT6   6    1000   full    yes    yes    down
INT7   7    1000   full    yes    yes    down
INT8   8    1000   full    yes    yes    down
INT9   9    1000   full    yes    yes    down
INT10  10   1000   full    yes    yes    down
INT11  11   1000   full    yes    yes    down
INT12  12   1000   full    yes    yes    down
INT13  13   1000   full    yes    yes    down
INT14  14   1000   full    yes    yes    down
MGT1   15   100    full    yes    yes    up
MGT2   16   100    full    yes    yes    disabled
EXT1   17   100    full    no     no     up
EXT2   18   100    full    no     no     up
EXT3   19   any    any     yes    yes    disabled
EXT4   20   any    any     yes    yes    disabled
>> Information#
```

As shown in Example 5-25, verify VLAN configurations using the **info/port** command. Make sure that INT2 belongs to VLAN 5 and INT3 belongs to VLAN 10. EXT1 and EXT2 should have y in the Tag and should have VLAN 1, 5, and 10 associated to its ports. You should see similar results on both GbESM-1 and GbESM-2.

Example 5-25 VLAN and tag information

```
>> Main# info/port
Alias  Port  Tag  RMON  PVID  NAME  VLAN(s)
-----  ---  ---  ---  ---  -
INT1   1    n    d     1    INT1  1
INT2   2    n    d     5    INT2  5
INT3   3    n    d    10    INT3  10
INT4   4    n    d     1    INT4  1
INT5   5    n    d     1    INT5  1
INT6   6    n    d     1    INT6  1
INT7   7    n    d     1    INT7  1
INT8   8    n    d     1    INT8  1
INT9   9    n    d     1    INT9  1
INT10  10   n    d     1    INT10 1
INT11  11   n    d     1    INT11 1
INT12  12   n    d     1    INT12 1
INT13  13   n    d     1    INT13 1
INT14  14   n    d     1    INT14 1
MGT1   15   n    d   4095  MGT1  4095
MGT2   16   n    d   4095  MGT2  4095
EXT1   17   y    d     1    EXT1  1 5 10
EXT2   18   y    d     1    EXT2  1 5 10
EXT3   19   n    d     1    EXT3  1
EXT4   20   n    d     1    EXT4  1
>> Information#
```

As shown in Example 5-26, verify that Spanning Tree is turned off for all ports by using the `info/12/stg` command. It is also possible to see that INT1, INT2, INT3, EXT1 and EXT 2 are in FORWARDING state. Based on the configurations in this example (Cisco-2 is root switch and setting of path costs) these ports should show FORWARDING. The Cisco external switch of the EXT1 link should show blocking for that connection. You should see similar results on both GbESM-1 and GbESM-2.

Example 5-26 Spanning Tree status

```
>> Main# info/12/stg
-----
Spanning Tree Group 1: Off, FDB aging timer 300

Port      Priority  Cost      State      Designated Bridge  Des Port
-----
INT1      0         0         FORWARDING *
INT2      0         0         FORWARDING *
INT3      0         0         FORWARDING *
INT4      0         0         DISABLED  *
INT5      0         0         DISABLED  *
INT6      0         0         DISABLED  *
INT7      0         0         DISABLED  *
INT8      0         0         DISABLED  *
INT9      0         0         DISABLED  *
INT10     0         0         DISABLED  *
INT11     0         0         DISABLED  *
INT12     0         0         DISABLED  *
INT13     0         0         DISABLED  *
INT14     0         0         DISABLED  *
EXT1      0         0         FORWARDING *
EXT2      0         0         FORWARDING *
EXT3      0         0         DISABLED  *
EXT4      0         0         DISABLED  *
* = STP turned off for this port.
```

Verify correct operation on the Cisco external switch

Table 5-17 includes some commands you can use to verify the desired configuration and operation of the Cisco equipment.

Table 5-17 Verifying the configuration

Description and comments	CatOS based switch	IOS based switch
Review running config for desired statements. Perform on both Cisco-1 and Cisco-2.	show config Review for the following: <ul style="list-style-type: none"> • set trunk 2/3 nonegotiate dot1q 1-1005,1025-4094 • set trunk 2/4 nonegotiate dot1q 1-1005,1025-4094 • set spantree portcost 2/3-4 100 	show run Review for the following on int FastEthernet 2/3 and 2/4: <ul style="list-style-type: none"> • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate • spanning-tree cost 100

Description and comments	CatOS based switch	IOS based switch
<p>Show speed and duplex. Perform on both Cisco-1 and Cisco-2.</p>	<p>show port status 2/3 show port status 2/4</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-Full • Speed = a-100 	<p>show int FastEthernet 2/3 status show int FastEthernet 2/4 status</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100
<p>Show trunking status. Link must be up before a trunk will come up. Perform on both Cisco-1 and Cisco-2.</p>	<p>show port trunk 2/3 show port trunk 2/4</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = nonegotiate • Encapsulation = dot1q • Status = Trunking • Native VLAN = 1 	<p>show int FastEthernet 2/3 trunk show int FastEthernet 2/4 trunk</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = on • Encapsulation =802.1q • Status = Trunking • Native VLAN = 1
<p>Show Spanning Tree status. Perform on both Cisco-1 and Cisco-2.</p>	<p>show spantree 2/3 show spantree 2/4</p> <p>State should show blocked for the links between Cisco-1 and the GbESMs for all VLANs.</p> <p>State should show forward for the links between Cisco-2 and the GbESM for all VLANS</p>	<p>show spanning int FastEthernet 2/3 show spanning int FastEthernet 2/4</p> <p>State should show blocked for the links between Cisco-1 and the GbESMs for all VLANs.</p> <p>State should show forward for the links between Cisco-2 and the GbESM for all VLANS</p>
<p>Test redundancy. Perform on both Cisco-1 and Cisco-2.</p>	<p>Experiment with redundancy by taking down Cisco-2 (which would normally be carrying all traffic) and make sure Cisco-1 takes over. Verify with ping tests after Spanning Tree has stabilized. Should be able to still ping both GbESMs from Cisco-1.</p>	<p>Experiment with redundancy by taking down Cisco-2 (which would normally be carrying all traffic) and make sure Cisco-1 takes over. Verify with ping tests after Spanning Tree has stabilized. Should be able to still ping both GbESMs from Cisco-1.</p>
<p>Ping the GbESM. Where x.x.x.x is the IP address of the GbESM (must be in same VLAN as subnet being pinged).</p>	<p>ping x.x.x.x</p> <p>Based on the Cisco-1 being at 192.168.70.202 and Cisco-2 at 192.168.70.203 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.</p>	<p>ping x.x.x.x</p> <p>Based on the Cisco-1 being at 192.168.70.200 and Cisco-2 at 192.168.70.201 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.</p>
<p>Attach a device to VLAN 5 and 10 on the Cisco switch and verify a ping to one of the blade servers in that same VLAN works as desired.</p>	<p>For VLAN 5, attach a device to port 2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work).</p> <p>For VLAN 10, attach a device to port 2/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).</p>	<p>For VLAN 5, attach a device to port fa0/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work).</p> <p>For VLAN 10, attach a device to port fa0/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).</p>

5.5.8 Dual GbESMs with 4 port static aggregation to different Cisco switches

This example (Figure 5-16) offers the maximum performance available from the @server BladeCenter, as well as fairly high redundancy, depending on the configuration of the operating systems running on the blade servers within the @server BladeCenter. It makes use of two GbESMs, each with all 4 ports aggregated into a single link, and each going to a separate Cisco switch. No ports will be in a Spanning Tree blocking state with this configuration because each GbESM only has a single (albeit aggregated) connection into the layer 2 network).

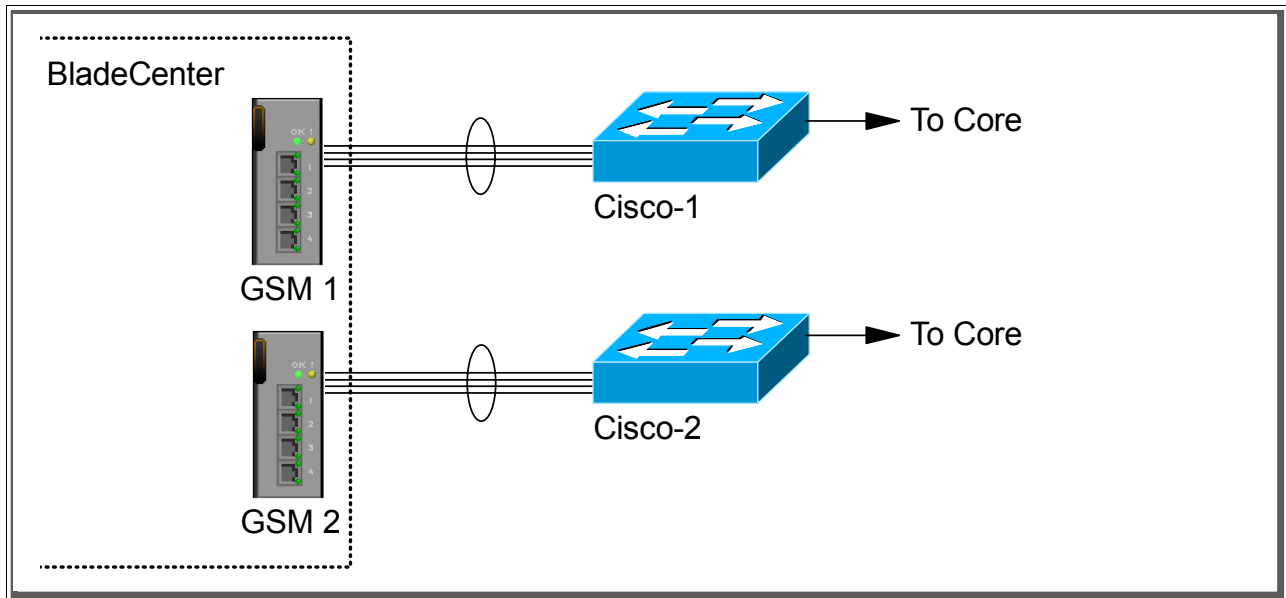


Figure 5-16 Dual GbESMs each with 4 port static aggregation to two separate Cisco switches

This example will use the same configuration and verification procedures as 5.5.3, “Single GbESM, four port static aggregation to a single Cisco switch” on page 116. As with the other dual GbESM examples, you will need to run through the configurations and verifications twice, once for the GbESM-1/Cisco-1 pair, and once for the GbESM-2/Cisco-2 pair.

5.5.9 Dual GbESMs with two static aggregation to two Cisco switches

This example (Figure 5-17) offers a good compromise between performance and high availability. It is made up of dual GbESMs, each with two, 2-port aggregated links, going to two separate Cisco switches. The 2-port aggregation provides for higher performance than a single link, and the second 2-port aggregation provides for full redundancy on any link or switch failures. Possible uses might be high availability and performance environments.

Based on our specific configuration for this example (layer 2 network, Cisco-2 is root and port cost on both Cisco links/aggregations toward the GbESM set to 100), Spanning Tree will block the connection between Cisco-1 and the GbESM, at the Cisco external switch. The choices of root and port cost settings in this example were only made for this example, and may be a poor choice in a production network. (Certainly, placing the root switch up against the GbESM in a datacenter environment would not be very common.) It is very important that any time a GbESM is connected in a redundant fashion, the location of the root switch and various port costs in the network be accounted for to ensure correct data flow (that is, you need to prevent non-BladeCenter traffic from flowing through the GbESMs).

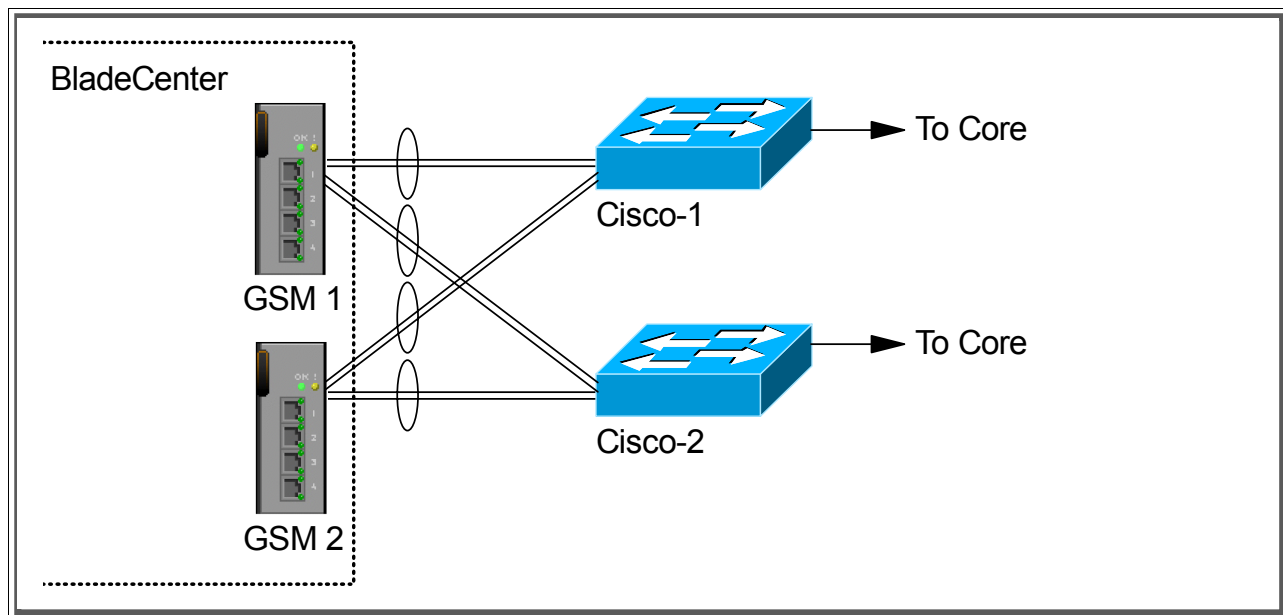


Figure 5-17 Dual GbESMs with cross connected static aggregated links to two separate Cisco switches

Step 1: Take down the links

It is always advisable to disable the links prior to making any configuration changes. See Table 5-1 on page 99 for procedures.

Step 2: Configure the IBM side of the link

Table 5-18 on page 143 walks you through the actions required to configure the GbESM for this example.

The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESMs with *root* level access.
- ▶ The following cabling was used (see Figure 5-18 on page 143 for a cabling diagram):
 - GbESM-1, Ext1 goes to Cisco-1 Port 2/3 if CatOS / FastEthernet 2/3 if IOS.
 - GbESM-1, Ext2 goes to Cisco-1 Port 2/4 if CatOS / FastEthernet 2/4 if IOS.
 - GbESM-1, Ext3 goes to Cisco-2 Port 2/3 if CatOS / FastEthernet 2/3 if IOS.
 - GbESM-1, Ext4 goes to Cisco-2 Port-2/4 if CatOS / FastEthernet 2/4 if IOS.
 - GbESM-2, Ext1 goes to Cisco-1 Port 2/5 if CatOS / FastEthernet 2/5 if IOS.
 - GbESM-2, Ext2 goes to Cisco-1 Port 2/6 if CatOS / FastEthernet 2/6 if IOS.
 - GbESM-2, Ext3 goes to Cisco-2 Port 2/5 if CatOS / FastEthernet 2/5 if IOS.
 - GbESM-2, Ext4 goes to Cisco-2 Port 2/6 if CatOS / FastEthernet 2/6 if IOS.

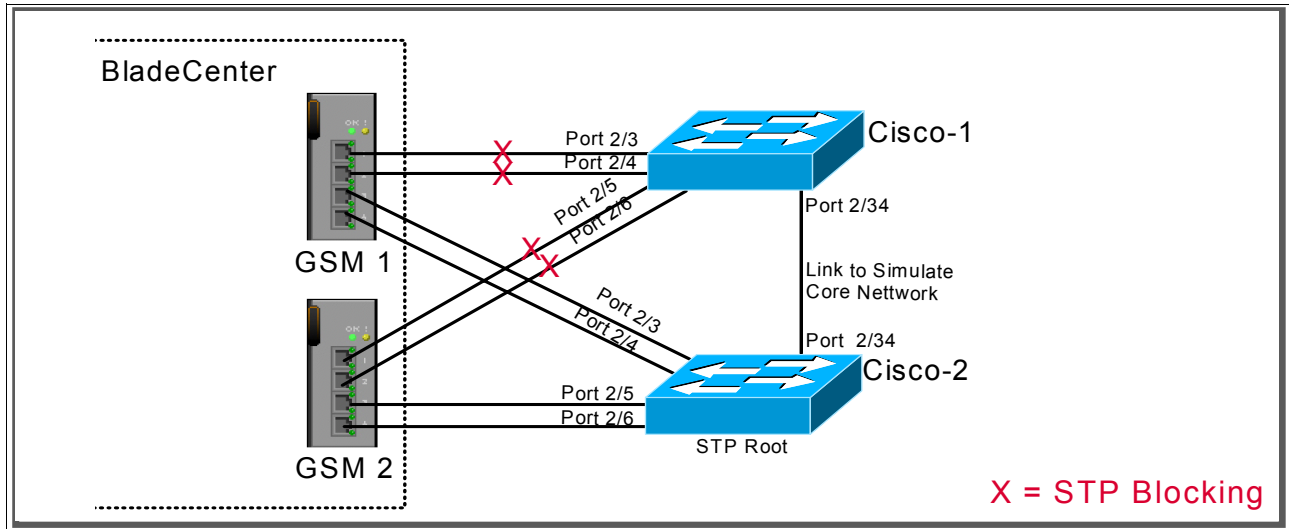


Figure 5-18 Cabling layout used in this example

- ▶ Commands are performed in the sequence shown.
- ▶ The GbESMs are starting from a default config per the “GbESM base configuration” on page 100.
- ▶ Cisco switches being used are 10/100Mbps-based and we set the GbESM port to 10/100Mbps. This means a cross-over cable *must* be used for the link between the GbESM and the Cisco switch.

If using switch ports that support 1000Mbps, the procedure to set speed will be slightly different from what is shown in the example.

Table 5-18 Configuring the GbESM

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 1 - Configure internal ports.</p> <p>Turn off STP. Then place the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged. Do this on both GbESM-1 and GbESM-2.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. This corresponds to BladeServer bay 3. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>All other PVIDs should be set for 1 with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 2 - Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously. Do this on both GbESM-1 and GbESM-2.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter apply to apply the new configuration</p>

Description and comments	Actions to perform via Web interface to GbESM
<p>Step 3- Configuring the external ports.</p> <p>This step sets the speed, duplex and VLAN setting for the external ports. This will allow external ports 1 and 2 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p> <p>As already noted, use a cross-over cable on the link between the GbESM and the Cisco switch because turning off auto-negotiation also turns off the auto selection of MDI/MDI-X. Do this on both GbESM-1 and GbESM-2.</p>	<p>Enter the command: <code>/cfg/port ext1/gig/speed 100</code> to set port speed to 100Mbps for external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/port ext1/gig/mode full</code> to set duplex settings to full for external port 1. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2, 3, and 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2, 3, and 4.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 4- Define trunk groups.</p> <p>Do this on both GbESM-1 and GbESM-2.</p>	<p>Enter the command: <code>/cfg/12/trunk 1/add ext1</code> to add ports to the trunk group. Repeat for external port 2.</p> <p>Enter the command: enable to enable trunk group.</p> <p>Enter the command: <code>/cfg/12/trunk 2/add ext3</code> to add ports to the trunk group. Repeat for external port 4.</p> <p>Enter the command: enable to enable trunk group.</p> <p>Enter apply to apply the new configuration.</p>
<p>Step 5- Save GbESM config to NVRAM.</p> <p>Failure to perform this step will result in all changes to the GbESM being lost if the @server BladeCenter is powered off or the GbESM is otherwise restarted. Do this on both GbESM-1 and GbESM-2.</p>	<p>Save the changes, save.</p>

Step 3: Configuring the Cisco switches

The following assumptions have been made for this example (refer to Table 5-19):

- ▶ VLANs 1, 5, and 10 already exist in the Cisco configuration and will be part of the 802.1Q trunk to the GbESMs.
- ▶ The user is already logged into the switches and the switches are in enable mode.
- ▶ The port connections are as described in Step 2 of this example.
- ▶ Commands are performed in the sequence shown.
- ▶ The switches are starting from a default config per Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS).
- ▶ Cisco switches being used are 10/100-based.

If using switch ports that support 1000Mbps, the command to set speed will be different from what is shown in the example.

Attention: During the testing we found that hardcoding the CatOS and IOS speed prevented us from getting the link aggregation working. The problem was solved by using the *auto* setting on the ports of the Cisco equipment. This is the default setting.

Table 5-19 Configuring the Cisco equipment

Description and comments	CatOS-based switch	IOS-based switch
<p>Step 1 - Configure speed and enable the ports.</p> <p>Use default setting (auto).</p>	(does not apply)	(does not apply)
<p>Step 2 - Configure 802.1Q trunking</p> <p>Forces link to become an 802.1Q VLAN trunk.</p> <p>Do this on both Cisco-1 and Cisco-2 switch.</p>	<pre>set trunk 2/3 nonegotiate set trunk 2/4 nonegotiate set trunk 2/5 nonegotiate set trunk 2/6 nonegotiate</pre>	<pre>Configure t int range FastEthernet 2/3 - 6 switchport trunk encap dot1q switchport mode trunk switchport nonegotiate</pre> <p>Note that the range option on IOS is only available in newer revisions of code. If the range option is not available you will need to repeat steps 2 and 3 for each interface.</p>
<p>Step 3 - Configure Spanning Tree port cost.</p> <p>Do this on both Cisco-1 and Cisco-2 switch.</p> <p>Setting the port cost higher than default performs two operations for this example:</p> <ul style="list-style-type: none"> • Forces all connections to a known Spanning Tree state (Blocking or Forwarding). • Helps to prevent VLAN traffic between Cisco-1 and Cisco-2, other than VLAN 1, from being switched through the GbESM. This forcing of the port cost helps to promote a more predictable flow. 	<p>Set the port cost on Cisco-1 and Cisco-2 to control default flow:</p> <pre>set spantree portcost 2/3-6 100</pre> <p>Note that for this particular design, leaving the links at their default cost will produce split traffic flows (see “Guidelines and comments - Spanning Tree” on page 94 for details on this behavior)</p>	<p>Set the port cost on Cisco-1 and Cisco-2 to control default flow:</p> <pre>spanning-tree cost 100 exit</pre> <p>Note that for this particular design, leaving the links at their default cost will produce split traffic flows (see “Guidelines and comments - Spanning Tree” on page 94 for details on this behavior)</p>
<p>Step 4 - Configure static link aggregation.</p> <p>Do this on both Cisco-1 and Cisco-2 switch.</p>	<p>Sets port 2/3-4 into channel group 1.</p> <pre>set port channel 2/3-4 1 set port channel 2/3-4 mode on</pre> <p>Sets port 2/5-6 into channel group 2.</p> <pre>set port channel 2/5-6 2 set port channel 2/5-6 mode on</pre>	<p>Sets FastEthernet 2/3 through 2/4 into channel group 1.</p> <pre>int range FastEthernet 2/3 - 4 channel-group 1 mode on exit</pre> <p>Sets FastEthernet 2/5 through 2/6 into channel group 2.</p> <pre>int range FastEthernet 2/5 - 6 channel-group 2 mode on exit</pre>

Description and comments	CatOS-based switch	IOS-based switch
<p>Step 5 - Set port cost on the aggregated link.</p> <p>Do this on both Cisco-1 and Cisco-2 switch.</p>	<p>(does not apply)</p> <p>CatOS uses the values from the individual ports to meet this requirement.</p>	<p>Set the channel cost on Cisco-1 and Cisco-2 to control default flow:</p> <pre>int port-channel1 spanning-tree cost 100 exit int port-channel2 spanning-tree cost 100 end</pre> <p>See Step 3 for details. In this case, set the cost for the whole aggregation as well as the individual ports.</p>
<p>Step 6 - Save config to NVRAM.</p> <p>Only necessary on IOS-based switches.</p> <p>Do this on both Cisco-1 and Cisco-2 switch</p>	<p>(does not apply)</p>	<pre>write mem</pre>

Step 4: Reconnect the devices

This is the final step to bring the connection into full operation. This will be the reverse of whatever procedure was used in Step 1. See Table 5-2 on page 99 for details on how to re-establish the links.

Step 5: Verify the configuration

This section provides options for verifying the correct and desired operation.

Certain elements of this verification depend on the existence of a layer 2 network behind the two Cisco switches. This was simulated for this example by tying the two switches together with an 802.1Q trunk link (see Example 5-7 on page 102 (IOS) and Example 5-6 on page 101 (CatOS) for details on how this link was configured).

Verify correct operation on the BladeCenter GbESM

Verify that the configurations on the GbESM look as follows:

As shown in Example 5-27, verify the port state by typing the `info/link` command. The link of EXT1, EXT2, EXT3 and EXT4 should be Up and the Speed/Duplex setting should be 100/Full.

Example 5-27 Verifying link status, speed and duplex settings

```
>> Main# info/link
-----
Alias  Port  Speed  Duplex  Flow Ctrl  Link
-----  ---  -----  -
INT1   1     1000   full    yes yes     up
INT2   2     1000   full    yes yes     up
INT3   3     1000   full    yes yes     up
INT4   4     1000   full    yes yes     down
INT5   5     1000   full    yes yes     down
INT6   6     1000   full    yes yes     down
INT7   7     1000   full    yes yes     down
INT8   8     1000   full    yes yes     down
INT9   9     1000   full    yes yes     down
```

```

INT10  10   1000  full  yes  yes  down
INT11  11   1000  full  yes  yes  down
INT12  12   1000  full  yes  yes  down
INT13  13   1000  full  yes  yes  down
INT14  14   1000  full  yes  yes  down
MGT1   15    100  full  yes  yes  up
MGT2   16    100  full  yes  yes  disabled
EXT1   17    100  full  no   no   up
EXT2   18    100  full  no   no   up
EXT3   19    100  full  no   no   up
EXT4   20    100  full  no   no   up

```

>> Information#

As shown in Example 5-28, verify VLAN configurations by the **info/port** command. Make sure that INT2 belongs to VLAN 5 and INT3 belongs to VLAN 10. EXT1, EXT2, EXT3 and EXT4 should have y in the tag and should have VLAN 1, 5, and 10 associated to its ports.

Example 5-28 VLAN and tag information

```

>> Main# info/port
Alias  Port  Tag  RMON  PVID  NAME  VLAN(s)
-----
INT1   1    n    d     1    INT1   1
INT2   2    n    d     5    INT2   5
INT3   3    n    d    10    INT3   10
INT4   4    n    d     1    INT4   1
INT5   5    n    d     1    INT5   1
INT6   6    n    d     1    INT6   1
INT7   7    n    d     1    INT7   1
INT8   8    n    d     1    INT8   1
INT9   9    n    d     1    INT9   1
INT10  10   n    d     1    INT10  1
INT11  11   n    d     1    INT11  1
INT12  12   n    d     1    INT12  1
INT13  13   n    d     1    INT13  1
INT14  14   n    d     1    INT14  1
MGT1   15   n    d   4095  MGT1   4095
MGT2   16   n    d   4095  MGT2   4095
EXT1   17   y    d     1    EXT1   1 5 10
EXT2   18   y    d     1    EXT2   1 5 10
EXT3   19   y    d     1    EXT3   1 5 10
EXT4   20   y    d     1    EXT4   1 5 10

```

>> Information#

As shown in Example 5-29, verify that Spanning Tree is turned off for all ports by using the **info/12/stg** command. It is also possible to see that INT1, INT2, INT3, EXT1, EXT2, EXT3 and EXT4 are in FORWARDING state.

Example 5-29 Spanning Tree status

```

>> Main# info/12/stg
-----
Spanning Tree Group 1: Off, FDB aging timer 300

Port  Priority  Cost  State  Designated Bridge  Des Port
-----
INT1   0          0    FORWARDING *

```

```

INT2      0      0      FORWARDING *
INT3      0      0      FORWARDING *
INT4      0      0      DISABLED *
INT5      0      0      DISABLED *
INT6      0      0      DISABLED *
INT7      0      0      DISABLED *
INT8      0      0      DISABLED *
INT9      0      0      DISABLED *
INT10     0      0      DISABLED *
INT11     0      0      DISABLED *
INT12     0      0      DISABLED *
INT13     0      0      DISABLED *
INT14     0      0      DISABLED *
EXT1      0      0      FORWARDING *
EXT2      0      0      FORWARDING *
EXT3      0      0      FORWARDING *
EXT4      0      0      FORWARDING *
* = STP turned off for this port.

```

As shown in Example 5-30, verify that trunk group 1 and trunk group 2 are forwarding for EXT1 through EXT4 by using the `info/12/trunk` command.

Example 5-30 Port EXT1 through EXT4 are forwarding

```

>> Main# info/12/trunk
Trunk group 1, port state:
EXT1: STG 1 forwarding
EXT2: STG 1 forwarding

Trunk group 2, port state:
EXT3: STG 1 forwarding
EXT4: STG 1 forwarding

```

Verify correct operation on the Cisco external switch

Table 5-20 includes some commands you can use to verify the desired configuration and operation of the Cisco equipment. Following the table we have also provided the output of some of the Cisco `show` commands.

Table 5-20 Verifying the configuration and operation of the Cisco external switch of the connection

Description and comments	CatOS-based switch	IOS-based switch
<p>Review running config for desired statements.</p> <p>Do this on both Cisco-1 and Cisco-2.</p>	<p>show config</p> <p>Review for the following:</p> <ul style="list-style-type: none"> • set port channel 2/3-4 1 • set port channel 2/5-6 2 • set trunk 2/3 nonegotiate dot1q 1-1005,1025-4094 <p>Should see a similar entry for each configured port (2/3 through 2/6).</p> <ul style="list-style-type: none"> • set spantree portcost 2/3-6 100 • set spantree portvlancost 2/3 cost 99 • set spantree portvlancost 2/4 cost 99 • set spantree portvlancost 2/5 cost 99 • set spantree portvlancost 2/6 cost 99 <p>The portvlancost settings are a result of the set spantree portcost commands</p> <ul style="list-style-type: none"> • set port channel 2/3-6 mode on <p>Notice that even though we are using 2/3-4 as an aggregation pair and 2/5-6 as an aggregation pairs, that the CatOS shows them together on the command line. They will still be used as desired but this can be confusing.</p>	<p>show run</p> <p>Review for the following on interface Port-channel1 and Port-channel2:</p> <ul style="list-style-type: none"> • switchport • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate • spanning-tree cost 100 <p>Note that the values in Port-channel1 and 2 may not show up if the aggregation has never come up since first being configured.</p> <p>Review for the following on FastEthernet 2/3 through 2/6:</p> <ul style="list-style-type: none"> • switchport trunk encapsulation dot1q • switchport mode trunk • switchport nonegotiate • spanning-tree cost 100 • channel-group 1 mode on <ul style="list-style-type: none"> • Channel-group 1 will be on FastEthernet 2/3 and 2/4 • Channel-group 2 will be on FastEthernet 2/5 and 2/6
<p>Show speed and duplex.</p>	<p>Issue the following command on each interface, 2/3 through 2/6 on both Cisco-1 and Cisco-2:</p> <p>show port status 2/3</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100 	<p>Issue the following command on FastEthernet 2/3 through 2/6 on both Cisco-1 and Cisco-2:</p> <p>show int FastEthernet 2/3 status</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Status = Connected • VLAN - Trunk • Duplex = a-full • Speed = a-100
<p>Show trunking status.</p> <p>Aggregation must be up before the trunk will come up.</p>	<p>Issue the following command on each interface, 2/3 through 2/6 on both Cisco-1 and Cisco-2:</p> <p>show port trunk 2/3</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = nonegotiate • Encapsulation = dot1q • Status = Trunking • Native VLAN = 1 	<p>Issue the following command on FastEthernet 2/3 through 2/6 on both Cisco-1 and Cisco-2:</p> <p>show int FastEthernet 2/3 trunk</p> <p>Should show the following:</p> <ul style="list-style-type: none"> • Mode = on • Encapsulation =802.1q • Status = trunk-inbndl (Po1) or (Po2) • Native VLAN = 1

Description and comments	CatOS-based switch	IOS-based switch
Review status of the aggregated link.	show port channel Should show the following: <ul style="list-style-type: none"> • Status = Connected • Channel Mode = on show channel traffic Should show how the links in the channel are being utilized.	show etherchannel summary Should show the following: <ul style="list-style-type: none"> • Protocol = - (- = no protocol (static)). • Ports fa2/3 -6 = (P) (P) = part of an aggregation group.
Show Spanning Tree status. Note that values shown here are very specific to the example configurations, and may not reflect production information.	show spantree 2/3-6 Based on the configuration and cost setting in this example: <ul style="list-style-type: none"> • Cisco 1 switch should show all three VLANs for each aggregation pair Blocking for these interfaces. • Cisco-2 switch (root) should show all three VLANs for each aggregation pair Forwarding for these interfaces. • Both Cisco-1 and Cisco-2 should show forwarding for the simulated layer 2 link (2/34). 	Issue the following command on FastEthernet 2/3 through 2/6 on both Cisco-1 and Cisco-2: show spanning int FastEthernet 2/3 show spanning int po1 Based on the configuration and cost settings in this example: <ul style="list-style-type: none"> • Cisco-1 switch should show BLK (Blocking) for all three VLANs, for each of these interfaces. • Cisco-2 switch (root) should show FWD (Forwarding) for all three VLANs, for each of these interfaces. • Both Cisco-1 and Cisco-2 should show forwarding for the simulated layer 2 link (FastEthernet 2/34).
Ping the GbESM. Where x.x.x.x is the IP address of the GbESM (must be in same VLAN as subnet being pinged).	ping x.x.x.x Based on the Cisco-1 being at 192.168.70.202 and Cisco-2 at 192.168.70.203 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.	ping x.x.x.x Based on the Cisco-1 being at 192.168.70.200 and Cisco-2 at 192.168.70.201 on VLAN 1 and the GbESM being at 192.168.70.127 on VLAN 1, should be able to ping across.
Attach a device to VLAN 5 and 10 on the Cisco switch and verify a ping to one of the blade servers in that same VLAN works as desired.	For VLAN 5, attach a device to port 2/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port 2/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).	For VLAN 5, attach a device to port fa0/15 and attempt to ping across to the IP address on the blade server in slot 2 (note that IP subnets must match for ping to work). For VLAN 10, attach a device to port fa0/10 and attempt to ping across to the IP address on the blade server in slots 3 or 4 (note that IP subnets must match for ping to work).
Test redundancy.	Experiment with redundancy by taking down Cisco-2 (which would normally be carrying all traffic) and make sure Cisco-1 takes over. Verify with ping tests after Spanning Tree has stabilized. Should be able to still ping both GbESMs from Cisco-1. Cisco-1 should also show all of its trunk ports as forwarding now.	Experiment with redundancy by taking down Cisco-2 (which would normally be carrying all traffic) and make sure Cisco-1 takes over. Verify with ping tests after Spanning Tree has stabilized. Should be able to still ping both GbESMs from Cisco-1. Cisco-1 should also show all of its trunk ports as forwarding now.

Following is the output of some of the Cisco commands we described in Table 5-20 on page 149.

Example 5-31 includes some of the show commands taken from the Cisco-1-CatOS switch. In this example we can see that STP is blocking the links pointing to the GbESM.

Example 5-31 Outputs of show commands from the Cisco-1-CatOS switch

```
Cisco-1-CatOS> (enable) show port status 2/3
Port Name          Status    Vlan      Level Duplex Speed Type
-----
2/3                connected trunk    normal a-full a-100 10/100BaseTX
=====
```

```
Cisco-1-CatOS> (enable) show port trunk 2/3
* - indicates vtp domain mismatch
# - indicates dot1q-all-tagged enabled on the port
Port    Mode          Encapsulation Status      Native vlan
-----
2/3     nonegotiate  dot1q       trunking    1

Port    Vlans allowed on trunk
-----
2/3     1-1005,1025-4094

Port    Vlans allowed and active in management domain
-----
2/3     1,5,10

Port    Vlans in spanning tree forwarding state and not pruned
-----
2/3
```

```
Cisco-1-CatOS> (enable) show port channel

Port Status    Channel Mode          Admin Ch Group Id
-----
2/3 connected on           1      802
2/4 connected on           1      802
-----
2/5 connected on           2      803
2/6 connected on           2      803

Port Device-ID                               Port-ID Platform
-----
2/3 FOX04453960(Cisco-2-CatOS)                2/3    WS-C4006
2/4 FOX04453960(Cisco-2-CatOS)                2/4    WS-C4006
-----
2/5 FOX04453960(Cisco-2-CatOS)                2/5    WS-C4006
2/6 FOX04453960(Cisco-2-CatOS)                2/6    WS-C4006
=====
```

```
Cisco-1-CatOS> (enable) show channel traffic
ChanId Port  Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
-----
802 2/3  0.00%  0.00%  43.83%  3.24%  81.74% 100.00%
802 2/4  0.00%  0.00%  56.17%  96.76%  18.26%  0.00%
-----
803 2/5  0.00%  0.00%  18.81%  32.03%  73.60% 100.00%
803 2/6  0.00%  0.00%  81.19%  67.97%  26.40%  0.00%
```

Cisco-1-CatOS> (enable)

=====

Cisco-1-CatOS> (enable) **show spantree 2/3-6**

Port	Vlan	Port-State	Cost	Prio	Portfast	Channel_id
2/3-4	1	blocking	62	32	disabled	802
2/3-4	5	blocking	62	32	disabled	802
2/3-4	10	blocking	62	32	disabled	802
2/5-6	1	blocking	62	32	disabled	803
2/5-6	5	blocking	62	32	disabled	803
2/5-6	10	blocking	62	32	disabled	803

Cisco-1-CatOS> (enable)

Example 5-32 includes some of the show commands taken from the Cisco-2-CatOS switch. In this example we can see that STP is forwarding the links pointing to the GbESM.

Example 5-32 Output of show commands from the Cisco-2-CatOS switch

Cisco-2-CatOS> (enable) **show port status 2/3**

Port	Name	Status	Vlan	Level	Duplex	Speed	Type
2/3		connected	trunk	normal	a-full	a-100	10/100BaseTX

=====

Cisco-2-CatOS> (enable) **Show port trunk 2/3**

* - indicates vtp domain mismatch

- indicates dot1q-all-tagged enabled on the port

Port	Mode	Encapsulation	Status	Native vlan
2/3	nonegotiate	dot1q	trunking	1

Port Vlans allowed on trunk

2/3	1-1005,1025-4094
-----	------------------

Port Vlans allowed and active in management domain

2/3	1,5,10
-----	--------

Port Vlans in spanning tree forwarding state and not pruned

2/3	1,5,10
-----	--------

=====

Cisco-2-CatOS> (enable) **show port channel**

Port	Status	Channel Mode	Admin Group	Ch Id
2/3	connected	on	1	802
2/4	connected	on	1	802
2/5	connected	on	2	803
2/6	connected	on	2	803

Port	Device-ID	Port-ID	Platform
------	-----------	---------	----------

```

-----
2/3 FOX05029033(Cisco-1-CatOS) 2/3 WS-C4006
2/4 FOX05029033(Cisco-1-CatOS) 2/4 WS-C4006
-----
2/5 FOX05029033(Cisco-1-CatOS) 2/5 WS-C4006
2/6 FOX05029033(Cisco-1-CatOS) 2/6 WS-C4006
-----

```

```

=====
Cisco-2-CatOS> (enable) show channel traffic
ChanId Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
-----
802 2/3 0.00% 0.00% 3.16% 44.08% 95.33% 83.54%
802 2/4 0.00% 0.00% 96.84% 55.92% 4.67% 16.46%
-----
803 2/5 0.00% 0.00% 2.01% 22.13% 88.99% 100.00%
803 2/6 0.00% 0.00% 97.99% 77.87% 11.01% 0.00%
-----

```

```

=====
Cisco-2-CatOS> (enable) show spantree 2/3-6
Port Vlan Port-State Cost Prio Portfast Channel_id
-----
2/3-4 1 forwarding 62 32 disabled 802
2/3-4 5 forwarding 62 32 disabled 802
2/3-4 10 forwarding 62 32 disabled 802
2/5-6 1 forwarding 62 32 disabled 803
2/5-6 5 forwarding 62 32 disabled 803
2/5-6 10 forwarding 62 32 disabled 803
-----

```

Example 5-33 includes some of the show commands taken from the Cisco-1-IOS switch. In this example we can see that STP is blocking the links pointing to the GbESM.

Example 5-33 outputs of show commands from the Cisco-1-IOS switch

```

Cisco-1-IOS#show interface fastEthernet 2/3 status

```

```

Port Name Status Vlan Duplex Speed Type
Fa2/3 connected trunk a-full a-100 10/100BaseTX

```

```

Cisco-1-IOS#show interface fastEthernet 2/3 trunk

```

```

Port Mode Encapsulation Status Native vlan
Fa2/3 on 802.1q trunk-inbndl 1
(Po1)

```

```

Port Vlans allowed on trunk
Fa2/3 1-4094

```

```

Port Vlans allowed and active in management domain
Fa2/3 1,5,10

```

```

Port Vlans in spanning tree forwarding state and not pruned
Fa2/3 none

```

```

Cisco-1-IOS#show etherchannel summary

```

Flags: D - down P - in port-channel
 I - stand-alone s - suspended
 H - Hot-standby (LACP only)
 R - Layer3 S - Layer2
 U - in use f - failed to allocate aggregator

u - unsuitable for bundling
 Number of channel-groups in use: 2
 Number of aggregators: 2

Group	Port-channel	Protocol	Ports
1	Po1(SU)	-	Fa2/3(P) Fa2/4(P)
2	Po2(SU)	-	Fa2/5(P) Fa2/6(P)

=====
Cisco-1-IOS#show spanning-tree inter fastEthernet 2/3

Vlan	Role	Sts	Cost	Prio.Nbr	Type
VLAN0001	Altn	BLK	100	128.449	P2p
VLAN0005	Altn	BLK	100	128.449	P2p
VLAN0010	Altn	BLK	100	128.449	P2p

=====
Cisco-1-IOS#show spanning-tree inter fastEthernet 2/4

Vlan	Role	Sts	Cost	Prio.Nbr	Type
VLAN0001	Altn	BLK	100	128.449	P2p
VLAN0005	Altn	BLK	100	128.449	P2p
VLAN0010	Altn	BLK	100	128.449	P2p

=====
Cisco-1-IOS#show spanning-tree inter fastEthernet 2/5

Vlan	Role	Sts	Cost	Prio.Nbr	Type
VLAN0001	Altn	BLK	100	128.450	P2p
VLAN0005	Altn	BLK	100	128.450	P2p
VLAN0010	Altn	BLK	100	128.450	P2p

=====
Cisco-1-IOS#show spanning-tree inter fastEthernet 2/6

Vlan	Role	Sts	Cost	Prio.Nbr	Type
VLAN0001	Altn	BLK	100	128.450	P2p
VLAN0005	Altn	BLK	100	128.450	P2p
VLAN0010	Altn	BLK	100	128.450	P2p

=====
Cisco-1-IOS#show spanning-tree interface port-channel 1

Vlan	Role	Sts	Cost	Prio.Nbr	Type
VLAN0001	Altn	BLK	100	128.449	P2p
VLAN0005	Altn	BLK	100	128.449	P2p
VLAN0010	Altn	BLK	100	128.449	P2p

Example 5-34 includes some of the show commands taken from the Cisco-1-CatOS switch. In this example we can see that STP is blocking the links pointing to the GbESM.

Example 5-34 outputs of show commands from the Cisco-2-IOS switch

Cisco-2-IOS#**show interface fastEthernet 2/3 status**

Port	Name	Status	Vlan	Duplex	Speed	Type
Fa2/3		connected	trunk	a-full	a-100	10/100BaseTX

=====

Cisco-2-IOS#**show interface fastEthernet 2/3 trunk**

Port	Mode	Encapsulation	Status	Native vlan
Fa2/3	on	802.1q	trunk-inbndl (Po1)	1

Port	Vlans allowed on trunk
Fa2/3	1-4094

Port	Vlans allowed and active in management domain
Fa2/3	1,5,10

Port	Vlans in spanning tree forwarding state and not pruned
Fa2/3	1,5,10

=====

Cisco-2-IOS#**show etherchannel summary**

Flags: D - down P - in port-channel
 I - stand-alone s - suspended
 H - Hot-standby (LACP only)
 R - Layer3 S - Layer2
 U - in use f - failed to allocate aggregator

u - unsuitable for bundling

Number of channel-groups in use: 2

Number of aggregators: 2

Group	Port-channel	Protocol	Ports
1	Po1(SU)	-	Fa2/3(P) Fa2/4(P)
2	Po2(SU)	-	Fa2/5(P) Fa2/6(P)

=====

Cisco-2-IOS#**show spanning-tree inter fastEthernet 2/3**

Vlan	Role	Sts	Cost	Prio.Nbr	Type
VLAN0001	Desg	FWD	100	128.449	P2p
VLAN0005	Desg	FWD	100	128.449	P2p
VLAN0010	Desg	FWD	100	128.449	P2p

=====

Cisco-2-IOS#**show spanning-tree inter fastEthernet 2/4**

Vlan	Role	Sts	Cost	Prio.Nbr	Type
------	------	-----	------	----------	------

```
VLAN0001      Desg FWD 100      128.449 P2p
VLAN0005      Desg FWD 100      128.449 P2p
VLAN0010      Desg FWD 100      128.449 P2p
```

```
=====
```

```
Cisco-2-IOS#show spanning-tree inter fastEthernet 2/5
```

```
Vlan          Role Sts Cost      Prio.Nbr Type
-----
VLAN0001      Desg FWD 100      128.450 P2p
VLAN0005      Desg FWD 100      128.450 P2p
VLAN0010      Desg FWD 100      128.450 P2p
```

```
=====
```

```
Cisco-2-IOS#show spanning-tree inter fastEthernet 2/6
```

```
Vlan          Role Sts Cost      Prio.Nbr Type
-----
VLAN0001      Desg FWD 100      128.450 P2p
VLAN0005      Desg FWD 100      128.450 P2p
VLAN0010      Desg FWD 100      128.450 P2p
```

```
=====
```

```
Cisco-2-IOS#show spanning-tree interface port-channel 1
```

```
Vlan          Role Sts Cost      Prio.Nbr Type
-----
VLAN0001      Desg FWD 100      128.449 P2p
VLAN0005      Desg FWD 100      128.449 P2p
VLAN0010      Desg FWD 100      128.449 P2p
```

5.6 Troubleshooting GbESM connections to Cisco devices

A section on troubleshooting could realistically be an entire document in its own right. This section covers some basic items to look for when problems are experienced, but should not be considered a comprehensive resource for troubleshooting this environment.

Both IBM and Cisco have many excellent tools and documents to assist with troubleshooting, and it is recommended that these resources be utilized to their fullest.

There is no substitute for training and experience. It is strongly recommend that problems be directed to individuals with experience and training in the areas where problems occur.

5.6.1 Troubleshooting specifics

This section includes issues that are unique to deploying the GbESM in a Cisco environment.

Problem: Port will not come up

Link light not lighting on the GbESM or the Cisco switch for the port in question.

On the GbESM console interface, the command `/info/link` shows the status of the connection as Link Down.

Cause

The most likely cause is use of incorrect cable type (cross-over or straight-through).

Resolution

As discussed elsewhere in this chapter, the GbESM can support either a cross-over or a straight-through cable, as long as the port is configured for auto-negotiate for speed/duplex. When auto-negotiate is disabled (by hard coding the speed/duplex to something other than Auto), the auto-negotiation for cable type is also disabled, and the only cable that will work on a switch-to-switch link is a cross-over cable. Based on this, it is recommended always to use a cross-over cable to attach the GbESM to any switch, to ensure the link will always stay up, regardless of the speed/duplex setting.

Problem: Static link aggregation will not form

Aggregation link fails to pass traffic. Indications from both the GbESM and Cisco external switch are that the link has failed to come ready.

Cause

A common cause is that one of the links in the aggregation is configured differently (for example, EXT1 is configured for 100/Full while Ext2 is configured for Auto).

Resolution

When a link refuses to form, review the characteristics of all of the connections in the link (on both the Cisco external switch and the GbESM side). All ports in an aggregation must be configured identically.



Deploying the Layer 2-7 GbE Switch Module for IBM @server BladeCenter in a Nortel environment

The features and functionality of the Layer 2-7 GbE Switch Module (GbESM) component of the IBM @server BladeCenter allow the device to be connected to a network based on Nortel Networks equipment. The test configurations we used while preparing this document included the BayStack 380-24T, a gigabit aggregation layer switch, and the Passport 8600, a core L2/L3 switch. A brief description of the features and capabilities of the BayStack 5510 and the Passport 8300 is also included, though these switches were unavailable for use at the time the Redpaper was written. Depending on the organization's network design philosophy and business requirements, your designs may call for one type of switch over another.

In general, the BayStack 380-24T and BayStack 5510 are a low cost option for providing a large amount of throughput to the network from the GbESMs. Depending on the particular solution implemented, it can provide a good deal of resiliency as well as throughput. In general, the amount of throughput is tempered by the level of resiliency required, especially when Spanning Tree is implemented as the failover mechanism.

The Passport 8600 and Passport 8300 are higher cost options, but also higher performing ones. From one perspective, the GbESM already fulfills a typical aggregation layer switch function, thus moving the IBM @server BladeCenter one step closer to the core network from an architecture perspective. Thus it makes sense to consider how an IBM @server BladeCenter can connect to a core switch that provides L2/L3 capabilities and added levels of resiliency. The Passport 8600 can provide the greatest level of network reliability, including sub-second failover, while actively utilizing every single GbESM external port on a dual GbESM configuration for maximum throughput as well. A Passport 8600-based solution would most likely be used to connect to many IBM @server BladeCenter chassis units in the same configuration, such as in a server farm or datacenter application.

The BayStack 5510 is Nortel's newest product offering in the BayStack family. The 5510 is a 48 port layer 2 stackable 1u switch and can be stacked up to 8 units. Unfortunately, equipment for testing was unavailable at the time of the writing of this Redpaper. You can substitute the 5510 for the 380 in any of the test configurations listed below. However, the commands to configure the BayStack 5510 may be different from those listed for the 380.

The Passport 8300 is an enterprise version of the Passport 8600 used to write this Redpaper. As with the 5510, the equipment was not available in time for inclusion in this paper. The configurations listed below can be performed using the 8300, though the commands necessary to create the configurations may be different from those listed for the 8600.

Product test specifications

All configurations were tested with the following equipment and code revisions:

- ▶ IBM @server BladeCenter 8677
 - 3 Blade servers with Windows 2000 Advanced Server SP3
 - One Management module
 - 2 GbESM: L2-7 GbE Switch Modules
 - Software Version 20.0.1
 - PLD Firmware version: 3.6
- ▶ BayStack 380-24T
 - 24 10/100/1000Base-T ports
 - Software version: 2.0.1
 - Firmware version: 2.0.0.12
- ▶ Passport 8600 Switch 1
 - 8010 ten slot chassis
 - 8690SF (slot 5)
 - 8608GT (slot 1)
 - 8608 (slot 8)
 - Software version: 3.3.2.0
- ▶ Passport 8600 Switch 2
 - 8003 three slot chassis
 - 8691SF (slot 3)
 - 8608GT (slot 1)
 - 8608GT (slot 2)
 - Software version: 3.3.2.0

Note: Before you start using your Nortel Networks switches, you should ensure that the switches are set back to factory defaults.

6.1 Nortel Networks feature descriptions

Below is a partial list of the features supported by Nortel Networks equipment, some of which are proprietary to Nortel Networks. These features are listed here because they relate to the specific network designs being discussed in this document. Included in this section are caveats for the interworking of these features with IBM @server BladeCenter server. With the exception of routing, the following is primarily a list of Level 2 features of the switch. In addition, the example configurations demonstrate layer 2 configuration of the switch, again with the exception of routing. Configuring a higher layer function, such as load balancing, is beyond the scope of this paper.

Multi-Link Trunking

Multi-Link Trunking (MLT) is the Nortel Networks method of link aggregation that allows multiple Ethernet trunks to be aggregated together to provide a single logical trunk. An MLT provides the combined bandwidth of the multiple links and load sharing between these links, as well as the physical layer protection against the failure of any single link. Although MLT is a proprietary algorithm, it is completely interoperable with most link aggregation protocols, including 802.3ad static mode (no LACP) and Etherchannel.

Split MLT

Split-MLT (SMLT) is defined as an MLT which terminates on two different aggregation switches. This provides nodal redundancy with sub-second failure recovery. Additional terms in relation to the SMLT feature are defined as follows:

SMLT aggregation switch – One of two switches performing the role of SMLT aggregator, which connects to multiple wiring closet switches, edge switches or Customer Premise Equipment (CPE) devices. The Passport 8600 is an SMLT aggregation switch.

IST (Inter Switch Trunk) – One or more parallel point-to-point links that connect two aggregation switches together. The two aggregation switches utilize this channel to share information so that they can operate as a single logical switch. There can be only one IST per SMLT aggregation switch.

SMLT Client – A switch or server located at the edge of the network, such as in a wiring closet or CPE. An SMLT Client switch must be able to perform link aggregation (such as with MLT, 802.3ad static mode, or some other compatible method) but does not require any SMLT intelligence. The GbESM is an example of a client.

The IBM @server BladeCenter GbESM module does not need to support SMLT for it to serve as an SMLT Client. As a client, links will be aggregated using 802.3ad static mode, and the GbESM will be unaware that the individual links in the group terminate on different SMLT aggregation switches. This will be described further in the configuration that includes SMLT. For now it is only important to note that this feature interoperates with any standards compliant 802.3ad device acting in static mode.

Auto-MDI/MDI-X

Nortel Networks refers to this feature as Autopolarity. The terms can be considered interchangeable in the context of this document. The BayStack 380-24T switch supports Autopolarity, as does the GbESM.

Multiple Spanning Tree Groups

The GbESM and the Passport 8600 switch support multiple Spanning Tree Groups, but for purposes of simplicity with the GbESM, only one will be used in this Redpaper. The BayStack 380-24T does not support multiple Spanning Tree Groups.

Quality of Service

Both BayStack 380-24T and Passport 8600 have QoS features. The configuration of these features is beyond the scope of this document.

Routing

The GbESM by default has IP forwarding enabled. This feature means that the switch can function as router across multiple subnets (up to 127) if that many interfaces are configured. Routing requires that interfaces be properly configured and enabled, and that gateways be properly configured and enabled as well. Routing can be used to cross VLANs. The first test configuration we discuss demonstrates this. However, traffic cannot be routed onto or off of VLAN 4095, the management VLAN.

6.2 Limitations of configuration examples

The purpose of this paper is to show basic network designs and accompanying configurations. There are many advanced features and hardware options that could be incorporated into such network scenarios. For example, the GbESM supports load balancing traffic to and from an IBM @server BladeCenter server. However, such configurations are beyond the scope of this Redpaper.

6.3 Preliminary configuration for examples

This section includes high-level comments on the configurations used in the examples.

6.3.1 General recommendations

This section includes high-level comments on network design and recommendations for physical connection of Nortel Networks devices to an GbESM.

Terminology: VLAN trunking versus link aggregation (port trunking)

Chapter 5 defined *trunking* as a collection of VLANs on a single link. Grouping several ports together to form a single link was defined as *link aggregation*. On Nortel equipment, link aggregation is called a multi-link trunk. Since the command line and Web interfaces to the Nortel equipment use this terminology, we will as well. This is an effort to avoid confusion and make the example configurations easier to follow.

Use of Spanning Tree Protocol

Spanning Tree has been the traditional L2 method of providing resiliency, but in many cases, Spanning Tree becomes the limiting factor in network resiliency. This is why many Nortel Networks resiliency features, such as SMLT, do not depend on Spanning Tree, and in fact these features surpass the failover performance of Spanning Tree by orders of magnitude. It should be noted that Spanning Tree is not employed in most current network implementations.

Spanning Tree is often used to safeguard against improper cabling, or end-users that create loops by connecting two drops to a hub, and so forth. Since the BayStack 380-24T and

Passport 8600 switches are not performing a wiring closet switch role in any of these scenarios, there is no threat of an unsuspecting end user inadvertently creating a loop on these particular switches. However it is up to the network administrator to determine the exact nature of risks posed to the network. Knowing when to use Spanning Tree and when not to is of critical importance. In the configurations described here, Spanning Tree is shown as disabled on the GbESM. It is left enabled on the BayStack and Passport switches with one exception. In the split multi-link trunk design, Spanning Tree must be disabled. When the SMLT is correctly established on the Passport 8600s, the switch automatically disables Spanning Tree on those ports in the trunk. The other important issue with Spanning Tree is the path cost. It is crucial to consider path cost when designing a network that will be using Spanning Tree. If this is not done, the GbESM may be passing network traffic that does not have a destination inside the BladeCenter chassis. The only traffic that is sent into the switch should be traffic that is going to a device inside the chassis.

Important: In all of the following scenarios, all configuration is performed on all network hardware before any cables are connected. This ensures that even with Spanning Tree disabled, no loops appear in the network. In addition, prior to resetting the switches to factory defaults between scenarios, cables are removed from all switches.

Autonegotiation

Autonegotiation allows for a number of features other than speed and duplex settings to be enabled. In many cases, features specifically require autonegotiation. Because of this, it is recommended to use autonegotiation when connecting the GbESM to other Nortel Networks equipment. While it seems intuitive that manual configuration is the safer way to configure two devices that are connected to each other, this rule of thumb turns out to have the opposite effect in most cases. Human intervention is far more likely to introduce a misconfiguration, and this is compounded by the fact that many network administrators assume that only one of the devices need to be manually configured, while leaving the other to autonegotiate (resulting in a duplex mismatch). Furthermore, resiliency features such as RFI and FEFI (not applicable to 1000Base-T) require autonegotiation in order to be enabled, just like the autopolarity feature. Manual setting of ports thus sacrifices such features, resulting in less network resiliency rather than more. Manual setting should be reserved on a case-by-case basis for known interoperability problems.

Cabling

Since autonegotiation is being used to detect link speed and duplex settings, the auto-MDI/MDI-X ports on the GbESM will allow use of either a crossover or straight cable when connecting to other Nortel switches, even if the autonegotiation features are turned off. It is preferable to use a crossover cable since, when cabling other Nortel equipment such as two Alteon 180e switches together, a crossover cable is always required.

Tip: Since auto-MDI/MDI-X ports can use either type of cable, and non-autosensing ports must use crossover cables, use crossover cables whenever possible on all switch-to-switch connections to avoid any cabling problems.

6.3.2 IP addresses used in example configurations

Table 6-1 on page 164 lists the IP address settings for the switches and blades that were used in the following configurations. Some of the devices will have multiple IP addresses defined. These addresses are provided solely as a reference. The authors learned while writing this paper that having a record of the addresses assigned to each device made configuring the devices much easier and eliminated a host of potential problems. You are free

to use whatever address configuration you choose so long as all addresses are in the correct subnets.

The IP addresses for the Management Module and management interfaces on the GbESMs will not change. Also, the IP addresses for the Blades will not change. All subnet masks are 24-bit (255.255.255.0).

Table 6-1 Configuration example IP address listing

Device	Management IP address	Public IP address
BladeCenter Management Module	192.168.70.125	Not Applicable
GbESM 1	192.168.70.127	Interface 1: 192.168.47.127 Interface 2: 10.10.50.101
GbESM 2	192.168.70.128	Interface 1: 192.168.47.128 Interface 2: Not applicable
BayStack 1	Not applicable	192.168.47.230
BayStack 2	Not applicable	192.168.47.231
Passport 1	Not applicable	192.168.47.251
Passport 2	Not applicable	192.168.47.252
Blade 1	Not applicable	192.168.47.120
Blade 2	Not applicable	10.10.50.100
Blade 3	Not applicable	20.10.50.100
Client 1	Not applicable	192.168.47.150

6.3.3 Base configuration options common to all examples

The following are some configuration options established that are common to all of the examples. These are only for demonstration purposes in the examples, and more than likely will not be duplicated in your particular environment.

All configurations will have three VLANs configured: VLAN 1, VLAN 5, and VLAN 10. All VLANs should be tagged.

All configurations assume that all VLANs will be carried on all links. These scenarios all use link aggregation. As mentioned earlier, spanning tree does not provide any added benefit and will be disabled on the GbESM. If you prefer to leave Spanning Tree enabled, the BayStack 380-24T or Passport 8600 will be preferred over the GbESM to become the root bridge. Your existing network may already have a root bridge that is preferred over the BayStack 380-24T or Passport 8600. In any event, when using Spanning Tree it is important to plan the location of the root bridge to ensure optimal traffic flows. The GbESM should not be allowed to become the root bridge.

The following blade servers internal to the IBM @server BladeCenter will be placed in the specified VLANs during the configuration stage of each example:

- ▶ BladeServer 1: VLAN 1
- ▶ BladeServer 2: VLAN 5
- ▶ BladeServer 3: VLAN 10

6.3.4 Basic configuration procedures

Important: These test configurations were designed to be run on test equipment isolated from production networks. All of the configurations start with the requirement that the switches have been reset to factory defaults. Therefore, if any of these examples are tested on production equipment, network communications will be impacted and could possibly fail.

There are three ways to configure the BayStack 380-24T, each of which could be used to set up the scenarios described later in this chapter. In order to initially configure a BayStack 380-24T, an IP address must be assigned through the console port. The management interface on the console port is called the CLI. Once an IP address is assigned, the console port can be used to complete the configuration. Alternatively, telnet can be used, which presents the same CLI. Web-based Management can be used, or Device Manager. Both Device Manager and the Web-based interface present GUIs to the user. In this paper, the Web-based interface configuration screens are shown for configuration tasks beyond the initial switch setup of the BayStack 380-24T.

The Passport 8600 has two main user interfaces: the CLI and Device Manager. The CLI is accessible through the console interface and telnet. As with the BayStack 380-24T, Device Manager is a GUI-based utility that can be used to manage and configure the Passport 8600. In the configuration examples of this paper that include the Passport 8600, the CLI will be used to configure the switch. The initial configuration of IP address must be performed via the console interface. After this, telnet can be used for the remaining steps, though this is not required or explicitly noted in the configuration steps.

When configuring each of the scenarios that follow, it is recommended that the configuration be performed on both devices prior to connecting the cabling between the GbESM and Nortel Networks switch (BayStack 380-24T or Passport 8600). This will ensure that Spanning Tree loops are not formed inadvertently as the steps are followed. Alternatively, the ports can be manually disabled through any of the management interfaces mentioned previously. To do this on the BayStack 380-24T via the Web-based interface, go to **Configuration** → **Port Management**, select **Disabled** and click **Submit** for each port that connects to the GbESM, as shown in Figure 6-1.

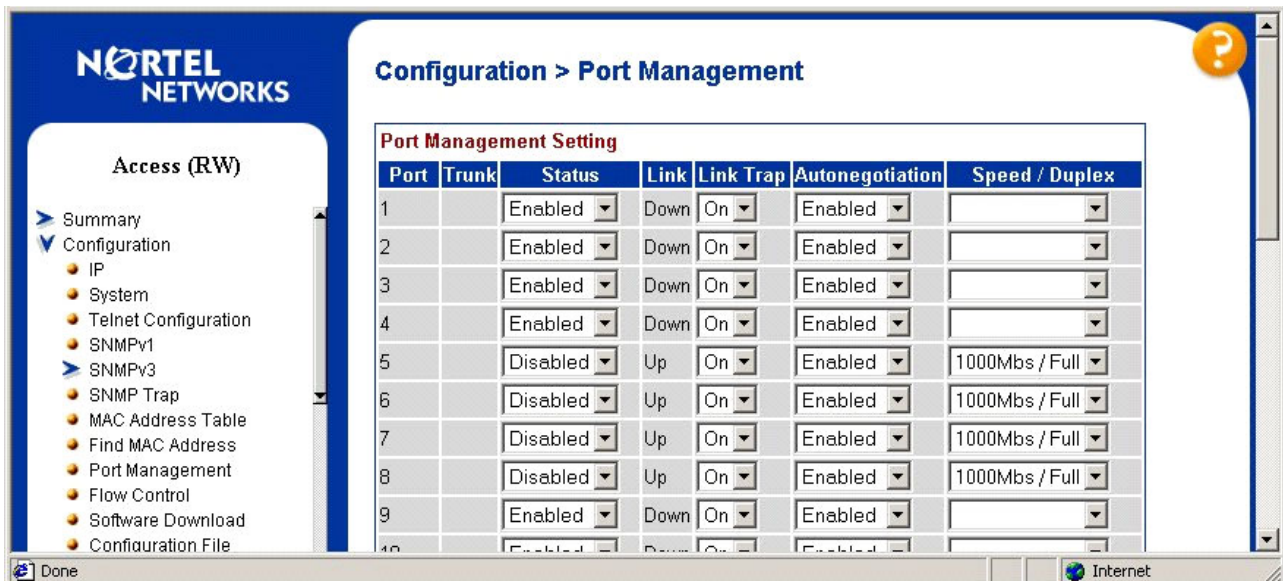


Figure 6-1 Port Management configuration window

The link status LEDs for each port should begin flashing. If any ports are currently members of an MLT group, the MLT group status will have to first be set to Disabled; select **Application** → **MultiLink Trunk** → **Group** to access the screen to disable the groups. Remember to re-enable the trunk after enabling the ports.

On the Passport 8600, ports can be disabled using the command shown in Example 6-1.

Example 6-1 Disabling ports on Passport 8600

```
config ethernet <portlist> state disable
```

Ports are re-enabled using the same command, but substituting the word “enable” as the state.

When making configuration changes this is the safest practice. For example, if the network grows to the point where it is necessary to change the configuration, then it is conceivable that a loop may be formed while performing configuration tasks. It is important that you understand the significance of each configuration change on a live network. While it is possible to migrate from one scenario to another without introducing problems, depending on the order of steps taken, this approach should only be used with extreme caution.

6.3.5 Base configuration tasks for the GbESM

Minor base configuration changes need to be applied to the GbESM.

Enabling external ports

After a factory reset, enable the external ports (Ext 1 through Ext 4). The external ports link states are set to Disable by default after a factory reset, and will need to be set to Enable for each external interface. You can check these settings at the switch by opening a telnet session and entering the command shown in Example 6-2.

Example 6-2 Listing link states

```
/info/link
```

This command will produce the listing shown in Figure 6-2 on page 167.

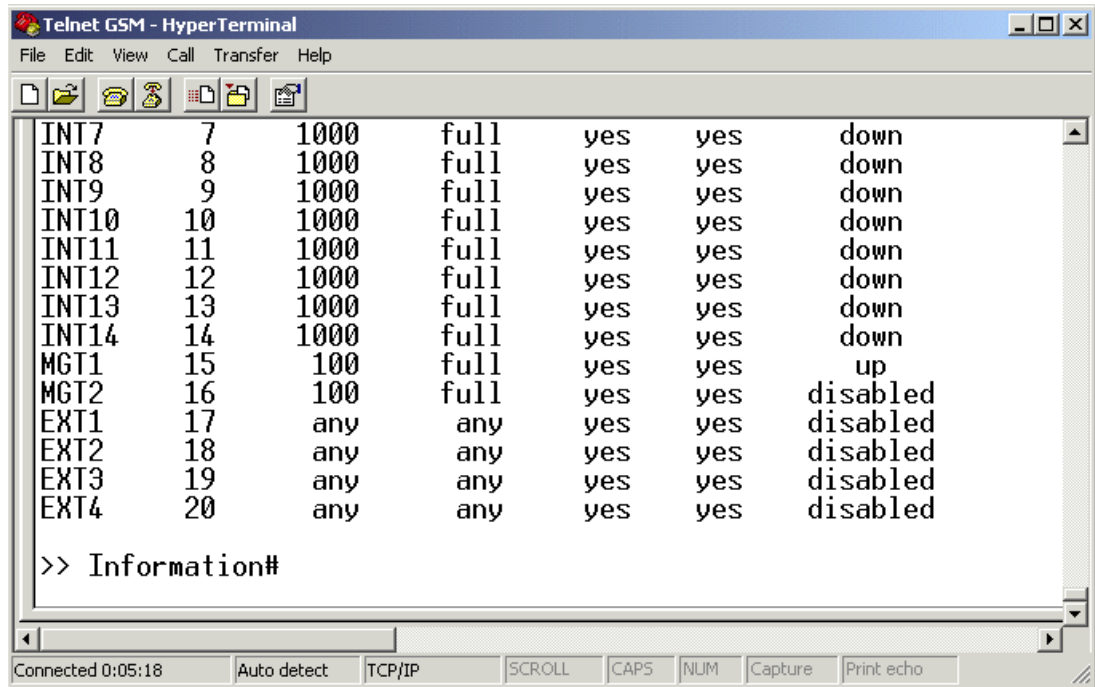


Figure 6-2 Factory reset link states

You can see from the listing that ports EXT1 through EXT4 are all disabled. The MGT2 port is also disabled. That is expected since that port is only enabled if a second management module is present and if that management module were to become active.

To enable these ports, log into the Management module via the Web interface. Click **I/O Module** → **Tasks** → **Management** and select the bay that the GbESM is in. Click **Advanced Management** and under **Advanced Setup** ensure that the drop-down box next to External Ports is set to Enabled. Click **Save**. Figure 6-3 on page 168 shows the port listing after the external ports are enabled. Ports EXT2 and EXT4 are listed as “down” because there are no devices attached to these ports. The status has changed from the previous listing of “disabled.”

```

Telnet GSM - HyperTerminal
File Edit View Call Transfer Help
[Icons]
INT7      7      1000    full    yes    yes    down
INT8      8      1000    full    yes    yes    down
INT9      9      1000    full    yes    yes    down
INT10     10     1000    full    yes    yes    down
INT11     11     1000    full    yes    yes    down
INT12     12     1000    full    yes    yes    down
INT13     13     1000    full    yes    yes    down
INT14     14     1000    full    yes    yes    down
MGT1      15     100     full    yes    yes    up
MGT2      16     100     full    yes    yes    disabled
EXT1      17     100     full    no     no     up
EXT2      18     any     any     yes    yes    down
EXT3      19     100     full    no     no     up
EXT4      20     any     any     yes    yes    down

>> Information#

Connected 0:01:24  ANSIW  TCP/IP  SCROLL  CAPS  NUM  Capture  Print echo

```

Figure 6-3 Enabling external ports

Default VLAN Settings

Verify the default VLAN settings. To check the default VLAN settings for each port on the GbESM, enter the command shown in Example 6-3.

Example 6-3 Listing port settings

```
/info/port
```

This command should produce the listing shown in Figure 6-4 on page 169. If it does not, the switch is not at factory settings and should be reset to factory settings before continuing with the scenarios. Use the command sequence shown in Example 6-4 to set the switch back to factory defaults. After you do this you will need to reassign the factory IP address as described previously.

Example 6-4 Reset to Factory Defaults

```
/boot/conf factory
/boot/reset
```

Alias	Port	Tag	RMON	PVID	NAME	VLAN(s)
INT1	1	n	d	1	INT1	1
INT2	2	n	d	1	INT2	1
INT3	3	n	d	1	INT3	1
INT4	4	n	d	1	INT4	1
INT5	5	n	d	1	INT5	1
INT6	6	n	d	1	INT6	1
INT7	7	n	d	1	INT7	1
INT8	8	n	d	1	INT8	1
INT9	9	n	d	1	INT9	1
INT10	10	n	d	1	INT10	1
INT11	11	n	d	1	INT11	1
INT12	12	n	d	1	INT12	1
INT13	13	n	d	1	INT13	1
INT14	14	n	d	1	INT14	1
MGT1	15	n	d	4095	MGT1	4095
MGT2	16	n	d	4095	MGT2	4095
EXT1	17	n	d	1	EXT1	1
EXT2	18	n	d	1	EXT2	1
EXT3	19	n	d	1	EXT3	1
EXT4	20	n	d	1	EXT4	1

>> Information# _

Figure 6-4 Default VLAN Listing

As described earlier in this Redpaper, you can see that the two management module ports are on VLAN 4095. It is not shown here, but IP Interface 128 is also on VLAN 4095. VLAN 4095 is the management VLAN and is immutable. You cannot add or remove ports or interfaces from this VLAN, nor can you delete the VLAN.

Multi-link trunk

To check the configuration of the multi-link trunk groups on the GbESM, open a telnet session to the switch and enter the command shown in Example 6-5.

Example 6-5 List multi-link trunk settings

```
/info/12/trunk
```

The switch should report back a message stating that all trunk groups are disabled. If it does not, the switch should be reset to factory defaults before continuing with the configuration examples.

6.3.6 Base configuration tasks for BayStack 380-24T

This configuration procedure assumes factory default settings. The console interface is the only interface that can be used to initially administer since the switch has no IP address.

The Main Menu is shown in Figure 6-5 on page 170.

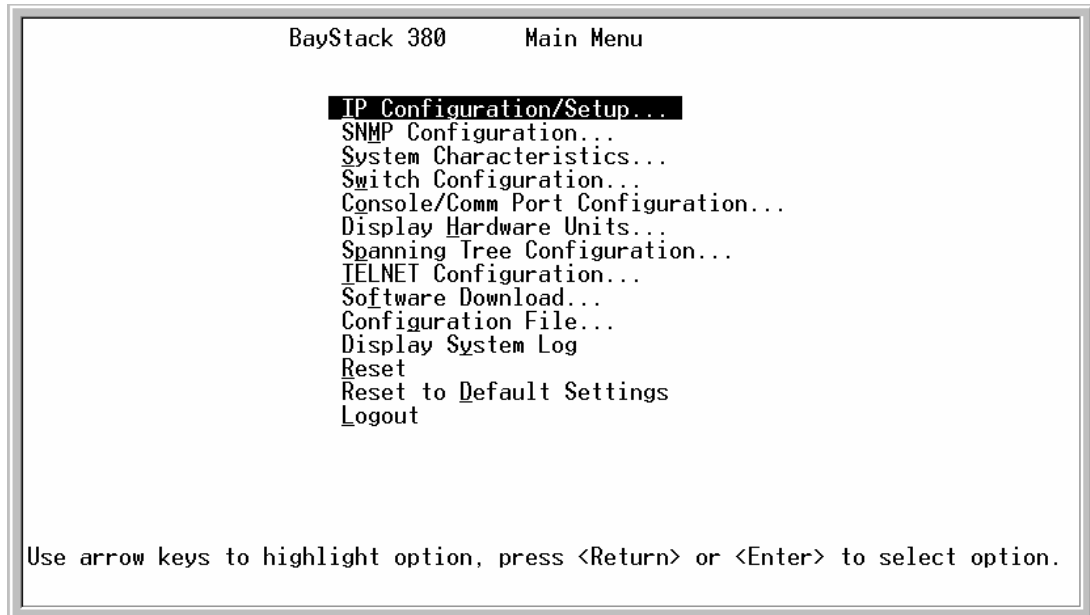


Figure 6-5 BayStack 380 Main Menu

Choose **IP Configuration/Setup** from the Main Menu. Assign an IP address to the switch as shown in Figure 6-6.

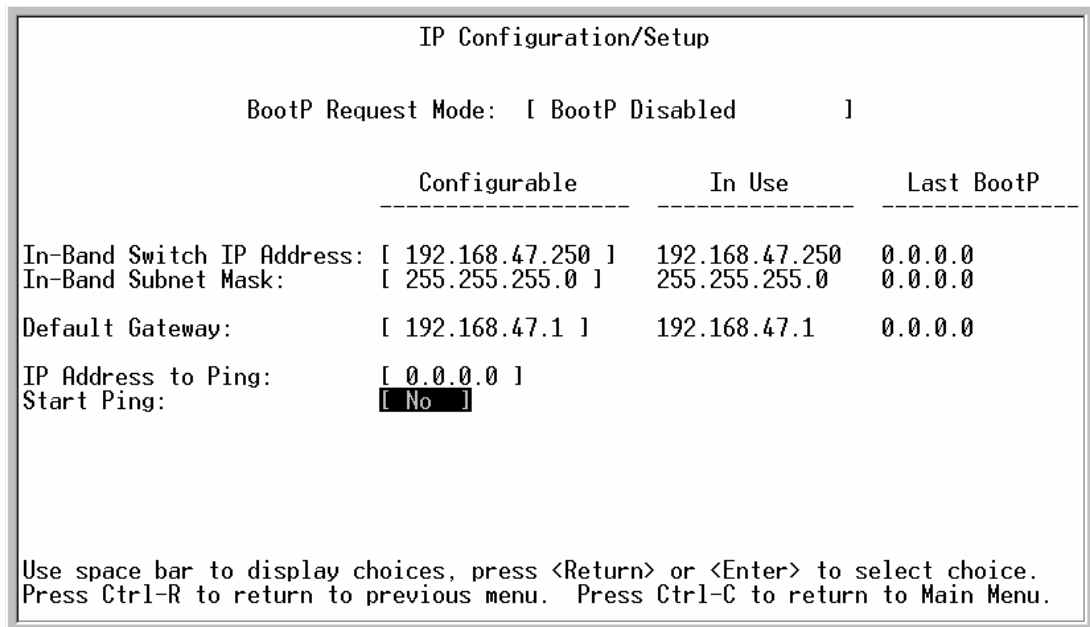


Figure 6-6 IP Configuration/Setup window

Set Switch Passwords. Select **Console/Comm Port Configuration** from the Main Menu. The resulting screen is shown in Figure 6-7 on page 171.

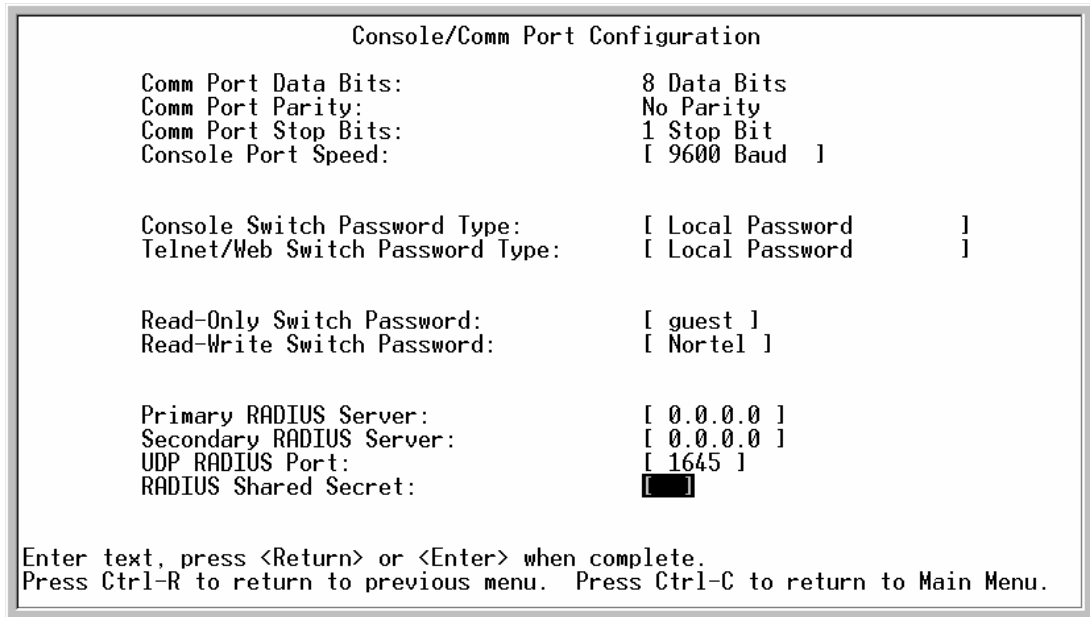


Figure 6-7 Console/Comm Port Configuration

Now that the switch has an IP interface configured, the Web interface can be used to configure the switch through a Web browser. To log in to the Web interface, use the Username “RW” and the correct password.

Go to **Configuration** → **System**, as shown in Figure 6-8. Define the system information and click **Submit**.

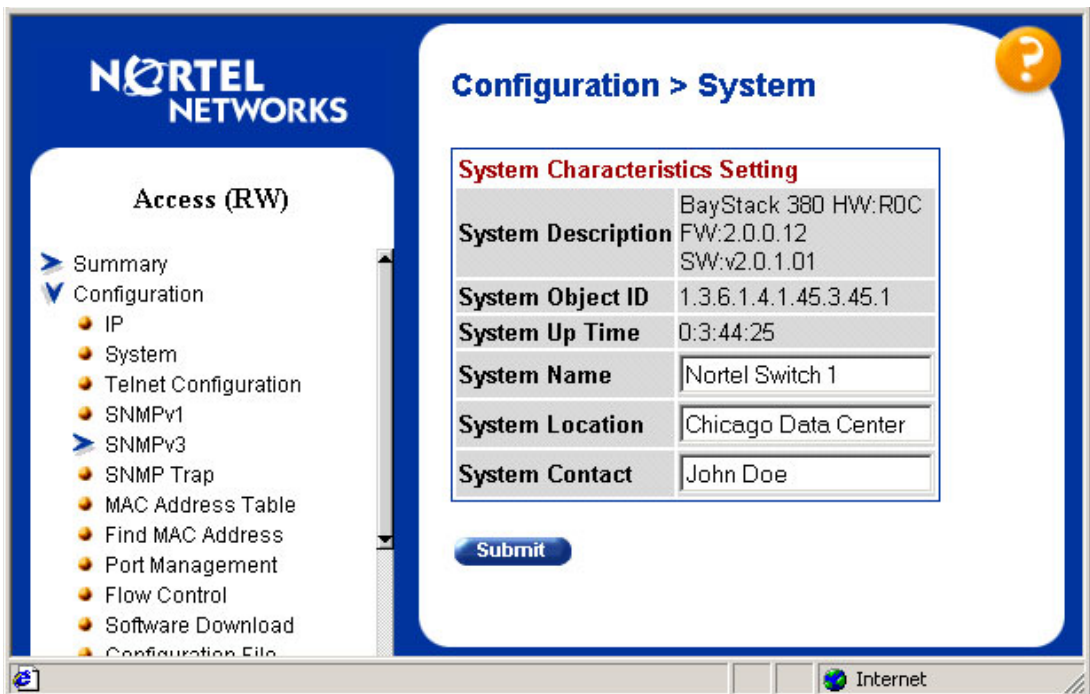


Figure 6-8 Configuration > System window

Go to **Application** → **VLAN** → **VLAN Configuration**. Create VLAN 5 as shown in Figure 6-9 and click **Create VLAN**.

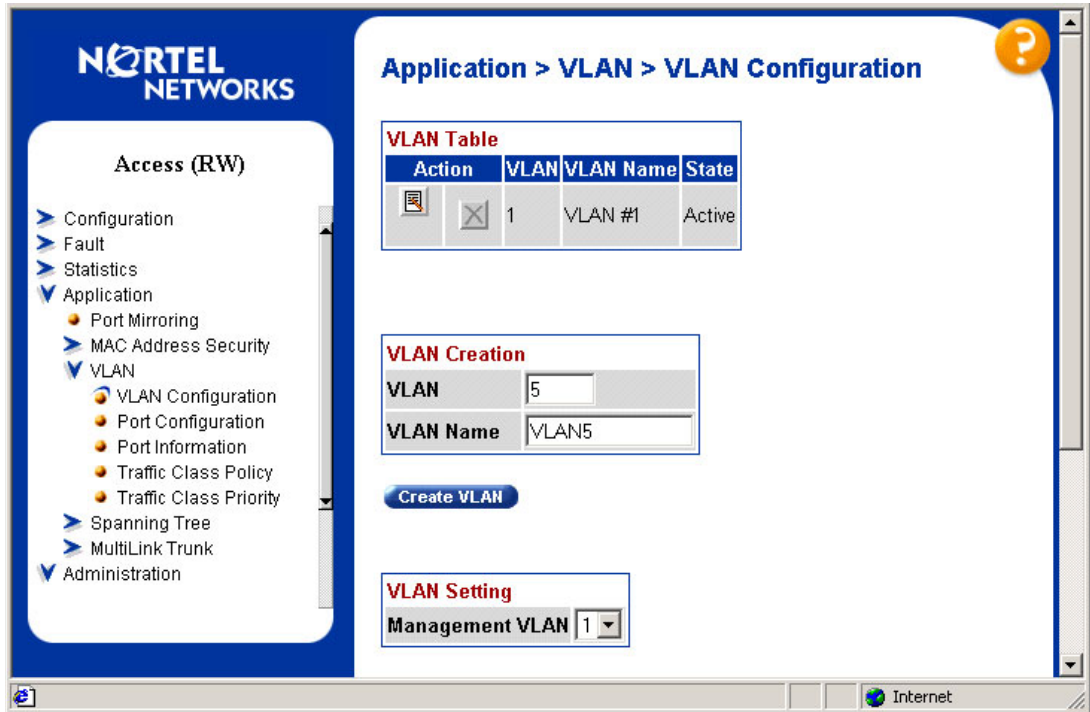


Figure 6-9 Application > VLAN > VLAN Configuration

The Spanning Tree bridge priority default for the BS380 (0x8000h) is the same as the default for the GbESM. In scenarios where Spanning Tree is to be used for redundancy, it is desirable that the GbESM not be elected as the root bridge. To ensure that this does not happen, the Bridge Priority of the BayStack 380 should be lowered. Even though some configurations recommend not running Spanning Tree between GbESM and BayStack 380 at all, it is good practice to go ahead and modify the bridge priority.

Go to **Application** → **Spanning Tree** → **Bridge Information** (Figure 6-10 on page 173). Set the STP bridge priority to a lower value (0x7000).

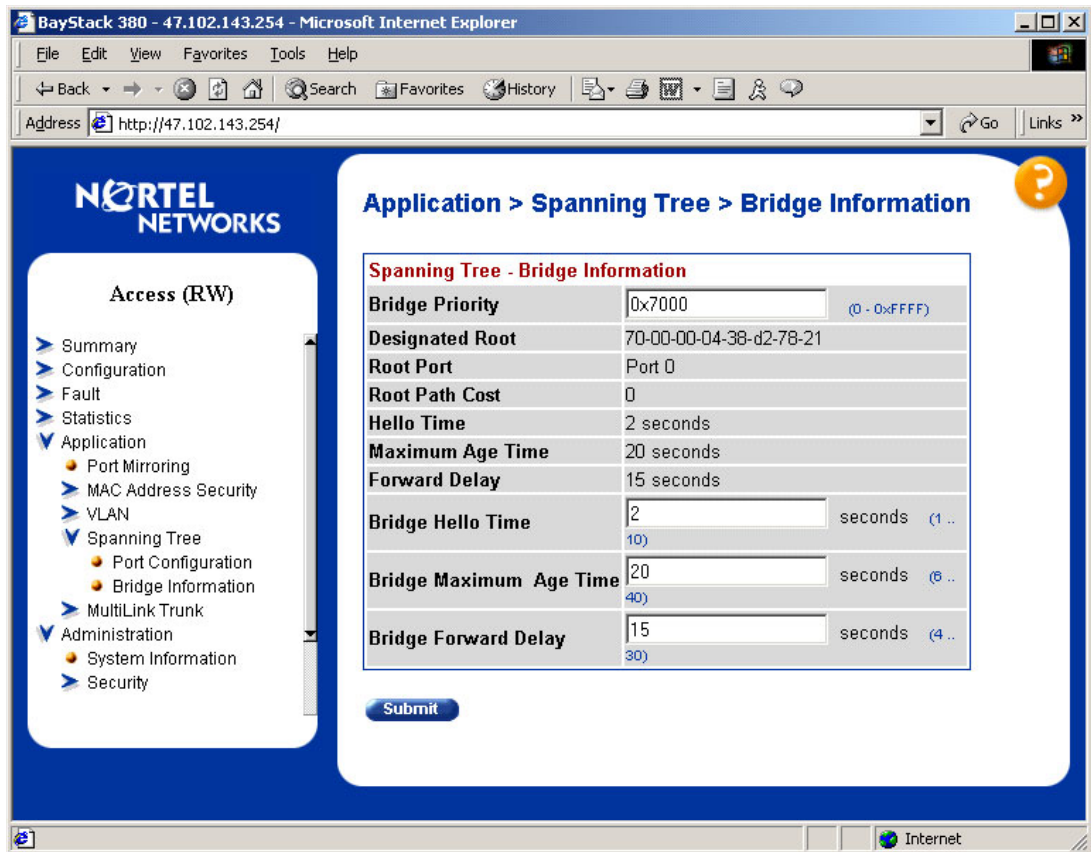


Figure 6-10 Application > Spanning Tree > Bridge Information

6.3.7 Base configuration tasks for Passport 8600

This configuration procedure assumes factory default settings. The console interface is the only interface that can be used to initially administer since the switch has no IP address. To log in to the switch use the default username "rwa" and password "rwa". The commands in Example 6-6 highlight the steps required to default an 8600. The out-of-band management IP should not be affected by defaulting the Passport 8600.

Example 6-6 Resetting the Passport 8600

```
config bootconfig flags factorydefault true
save boot
boot -y
```

Once the 8600 has been rebooted, the following command should be issued to save the default configuration:

```
save config
```

The commands in Figure 6-11 on page 174 can be used to assign an IP address to the management interface on the switch fabric. This is an out-of-band port.

```

Passport-8610:5# config bootconfig net mgmt
Passport-8610:5/config/bootconfig/net/mgmt# ip 192.168.47.252/24 cpu-slot 5
Passport-8610:5/config/bootconfig/net/mgmt#

```

Figure 6-11 Console interface: Assigning an IP address

The commands in Figure 6-12 will set the system level attributes, such as switch name, location, and contact. Also shown in this figure is a password change for the administration user account, "rwa". For security reasons, the other default passwords should be changed as well. The commands to change these additional passwords are not included in this example. The usernames that need password changes are: "rw", "ro", "l2", and "l3".

```

Passport-8610:5# config sys set
Passport-8610:5/config/sys/set# name PP8600-1
PP8600-1:5/config/sys/set# location SanFran Data Center
PP8600-1:5/config/sys/set# contact Jane Doe
PP8600-1:5/config/sys/set# back
PP8600-1:5/config/sys# top
PP8600-1:5#
PP8600-1:5# config cli
PP8600-1:5/config/cli# password rwa 123456

Enter the old password : ***
Enter the New password : *****
Re-enter the New password : *****

Password changed successfully
PP8600-1:5/config/cli# top
PP8600-1:5#

```

Figure 6-12 Console interface: Setting system attributes and changing passwords

VLANs 5 and 10 will be created next (as shown in Figure 6-13). VLAN 1 is the default VLAN so it is not necessary to explicitly create it.

```

PP8600-1:5# config vlan 5 create
PP8600-1:5/config/vlan/5/create# byport 1
PP8600-1:5/config/vlan/5/create# name VLAN5
PP8600-1:5/config/vlan/5/create# back
PP8600-1:5/config/vlan/5# back
PP8600-1:5/config# vlan 10 create
PP8600-1:5/config/vlan/10/create# byport 1
PP8600-1:5/config/vlan/10/create# name VLAN10
PP8600-1:5/config/vlan/10/create# top
PP8600-1:5#
PP8600-1:5# config stg 1
PP8600-1:5/config/stg/1# priority 24567
PP8600-1:5/config/stg/1# top
PP8600-1:5#

```

Figure 6-13 Console interface: Create VLANs 5 and 10

6.4 Configuration examples

This section contains the various design scenarios for connecting an IBM @server BladeCenter GbESM to a network based on Nortel Networks equipment. In all of these scenarios there are a few major items that must be configured on each switch to create a working configuration. These items are VLAN configuration, PVID assignments, and Multi-Link Trunking configuration. On the GbESM, the steps may be in a slightly different order than on the BayStack or the 8600 series switches. This is because while the steps could be performed in the same order on all the switches, the GbESM is more sensitive to the order in which commands are executed. The order of the steps as written here is the best we have found to avoid confusion or error when working through these configurations. Finally, the routing configuration commands are only used in the first two configuration sequences, though they can be applied to the rest of the configurations if you wish to do so.

6.4.1 Single GbESM with link aggregation and routing to single BayStack 380-24T

This connectivity scenario covers the most basic design. A single GbESM is connected to a single BayStack 380. Since the GbESM is aggregating as many as 14 servers, each with a Gigabit connection to the GbESM, uplink capacity to the rest of the network is important. That is why this basic scenario incorporates trunking four Gigabit links between GbESM and BayStack 380 (see Figure 6-14). This offers the maximum throughput between GbESM and BayStack 380. Fewer links can be used if desired, and the following configuration is easily modified accordingly.

This topology is resilient against single link or port failures, but not GbESM or switch failure. This design is useful for networks that need the maximum throughput of one GbESM but where redundancy of the switches is not a concern.

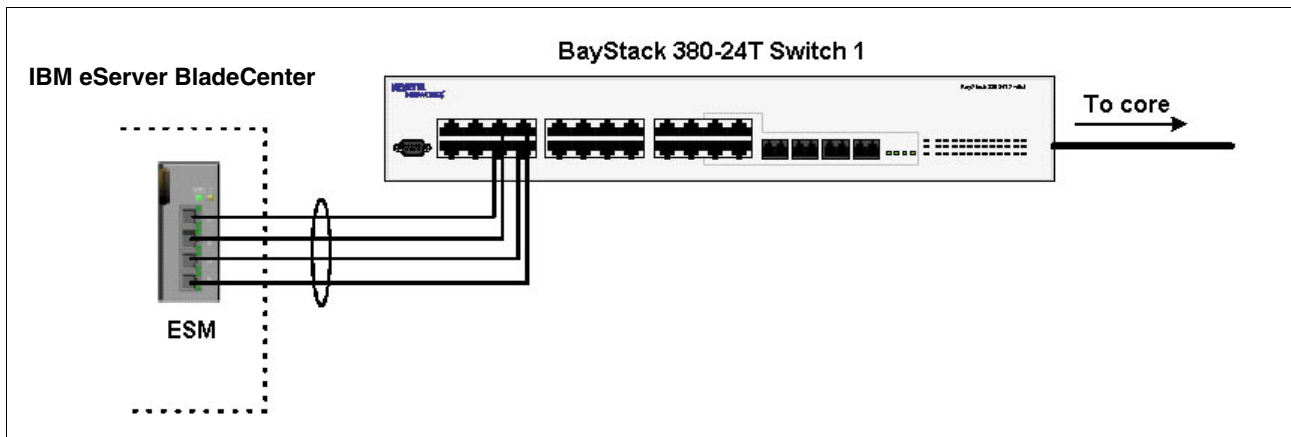


Figure 6-14 IBM eServer BladeCenter and BayStack 380-24T

Configuring the GbESM

Table 6-2 on page 176 walks through the actions required to configure the GbESM for this example. The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with admin level access.
- ▶ VLAN 1 on the GbESM is already enabled and all ports except the management ports are members. If any changes need to be made to VLAN 1 the instructions will note this.
- ▶ Ports EXT1 through EXT4 on the GbESM in Switch Module Bay 1 are used between the GbESM and Nortel.

- ▶ Commands are performed in the sequence shown.
- ▶ Nortel Networks switches are BayStack 380-24T gigabit switches. Autonegotiate will result in ports operating at 1000Base-T.
- ▶ All ports remain in Spanning Tree Group 1, and STP has been turned off.
- ▶ No cables are connected between any switches.
- ▶ The GbESM has been reset to factory defaults as shown in Example 6-4 on page 168.

Table 6-2 Configuring the GbESM

Description and comments	Instructions
<p>Step 1 - Configure the IP interfaces and routing</p> <p>This creates an interface on the public network. Since the management interface is on the management VLAN, we must configure an interface on the public network.</p>	<p>Enter the command /cfg/13/if 1/addr 192.168.47.127 to configure interface 1 with IP Address 192.168.47.127.</p> <p>Enter the command: /cfg/13/if 1/mask 255.255.255.0 to configure subnet mask 255.255.255.0.</p> <p>Enter the command: /c/13/if 1/ena to enable interface 1.</p> <p>Repeat these three steps with IP address 10.10.50.101 and subnet mask 255.255.255.0 on interface 2. Interface 2 will be used to route traffic to VLAN 5.</p> <p>Repeat these three steps with IP address 20.10.50.101 and subnet mask 255.255.255.0 on interface 3. Interface 3 will be used to route traffic to VLAN 10.</p> <p>Enter the command /cfg/13/frwd/on to enable IP forwarding. It should be enabled by default.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 2 - Configure internal ports</p> <p>This places the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command /cfg/12/stg 1/off to turn off Spanning Tree.</p> <p>Enter the command: /cfg/12/vlan 5/ena to enable VLAN 5.</p> <p>Enter the command: /cfg/12/vlan 5/add INT2 to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: /cfg/12/vlan 10/ena to enable VLAN 10.</p> <p>Enter the command: /cfg/12/vlan 10/add INT3 to add internal port 3 to VLAN 10. Repeat this command for INT4.</p> <p>All other PVIDs should be set for 1, with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: /info/port. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 3 - Configure tagging on external ports</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: /cfg/port EXT1/tag e to enable tagging on external port 1. Repeat this command for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>

<p>Step 4 - Configuring the VLANS for the external ports</p> <p>This puts the ports into VLANs 1, 5, and 10. This will allow external ports 1 through 4 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 because they are in it by default.</p>	<p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2 through 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 5 - Configure external ports and multi-link trunking</p> <p>This will allow external ports 1-4 to be aggregated together and treated as a single link.</p>	<p>Enter the command <code>/cfg/12/trunk 1/add EXT1</code> to add port EXT1 to Trunk Group 1.</p> <p>Repeat this step for external ports 2 through 4.</p> <p>Enter the command <code>/cfg/12/trunk 1/ena</code> to enable Trunk Group 1.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>

Configuring the BayStack 380-24T

This section describes the steps to configure a BayStack 380-24T.

Step 1: Configure ports

Click **Application** → **VLAN** → **Port Configuration**. The resulting screen is shown in Figure 6-15. Configure the links to be tagged. Port names can also be assigned. Click **Submit**.

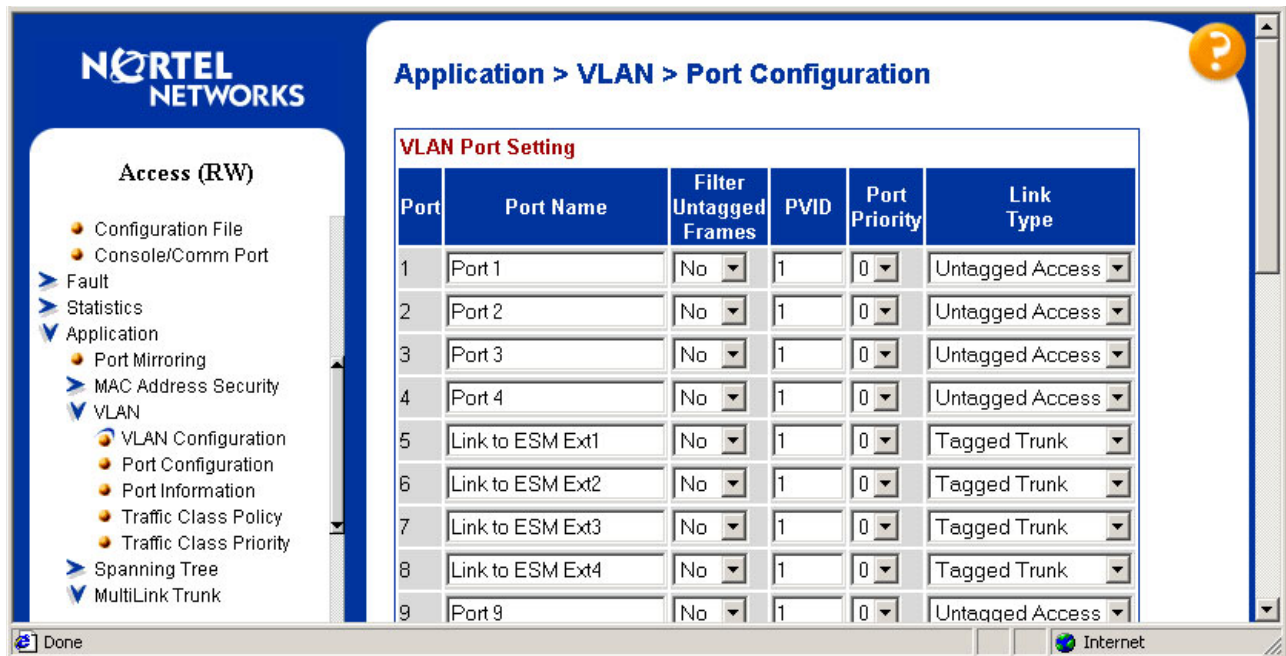


Figure 6-15 Application > VLAN > Port Configuration window

Step 2: Add VLANs to each link member

Click **Application** → **VLAN** → **VLAN Configuration**. Click the action button next to VLAN 5. Add the ports to VLAN 5 as shown in Figure 6-16 and click **Submit**. Repeat this for VLAN 10 and click **Submit**.

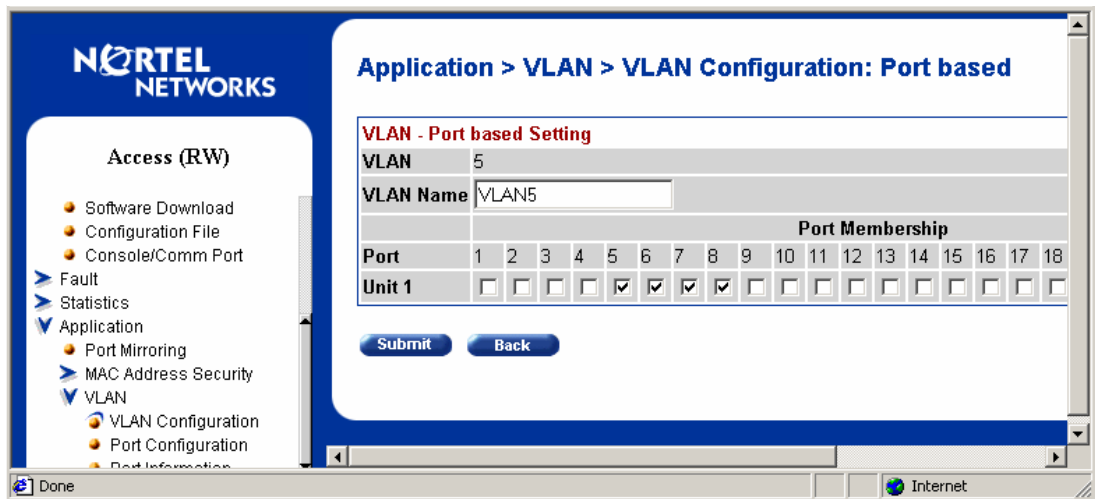


Figure 6-16 Application > VLAN > VLAN Configuration: Port based window

Step 3: Configure MLT

Click **Application** → **MultiLink Trunk** → **Group**. Specify ports 5, 6, 7, and 8, and set “STP Learning” to Fast (Figure 6-17 on page 178). Click **Submit**. The trunk is not enabled at this point. It must explicitly be enabled by changing the setting in the table to Enabled and clicking **Submit**.

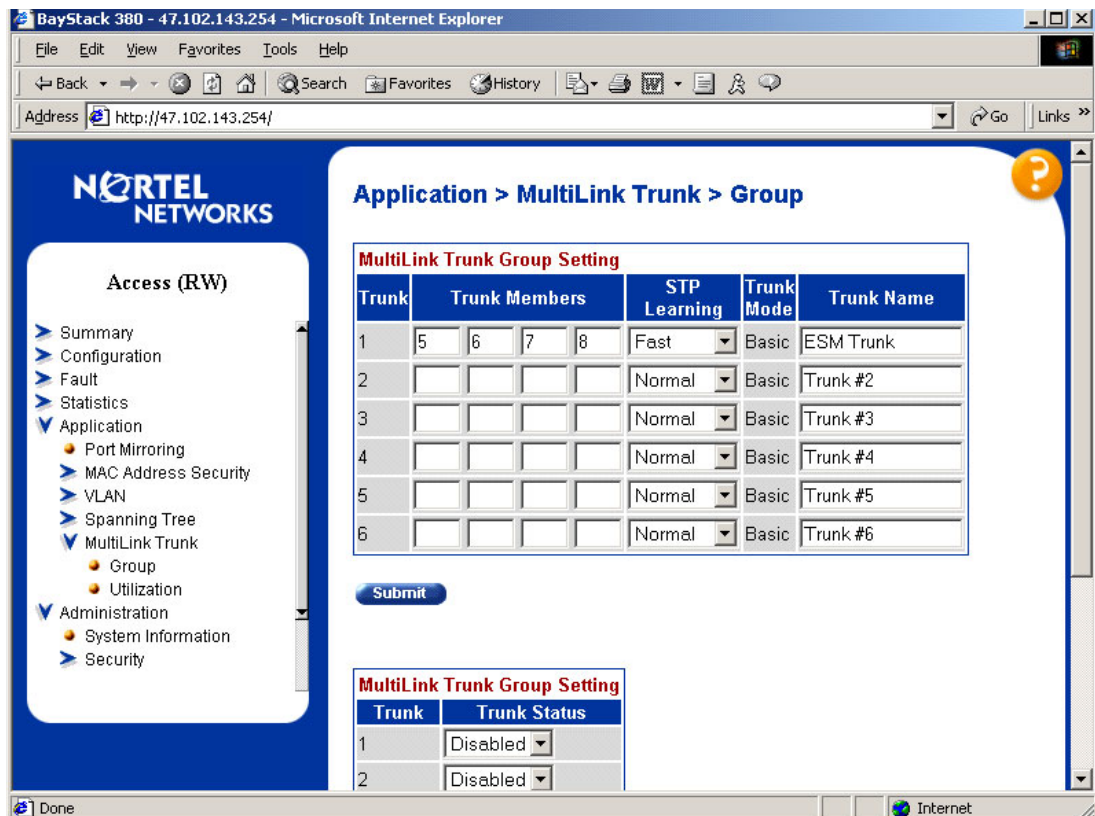


Figure 6-17 Application > MultiLink Trunk > Group window

6.4.2 Validation of GbESM configuration

To validate the setup of the GbESM, refer to “Step 5: Verify the configuration” on page 107.

6.4.3 Validation of BayStack 380-24T configuration

This section describes some steps that can be taken to quickly verify the configuration of the BayStack 380-24T.

Check port status, speed, and duplex settings from the Configuration > Port Management screen. Note that if autonegotiate is enabled, the Speed/Duplex box displays the negotiated speed/duplex setting. If autonegotiate is disabled, the Speed/Duplex box reflects the manual setting (Figure 6-18).

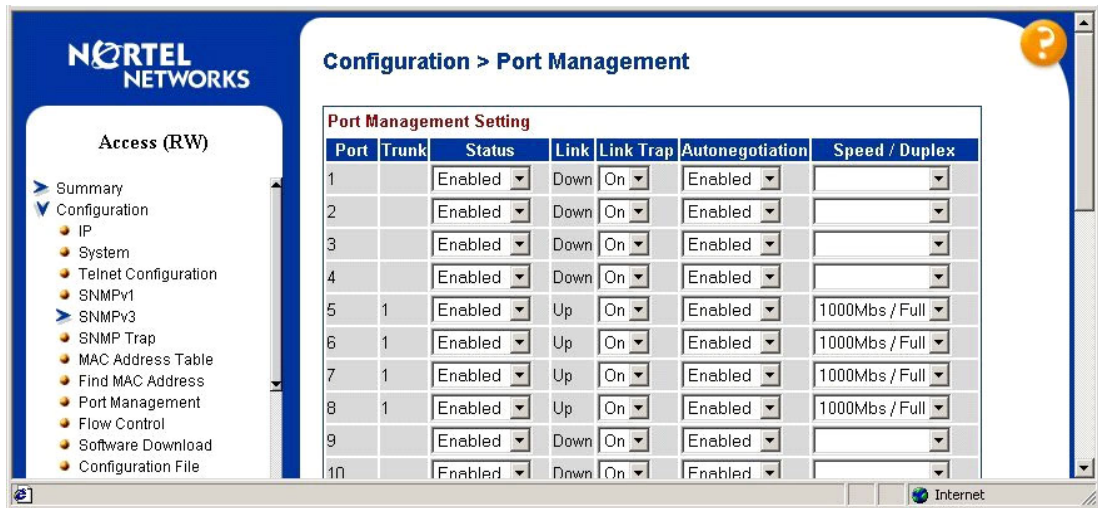


Figure 6-18 Port Management configuration window

Verify Spanning Tree forwarding state for the ports that connect to the GbESM. Go to Application > Spanning Tree > Port Configuration and confirm that the trunk ports are in a forwarding state (Figure 6-19).

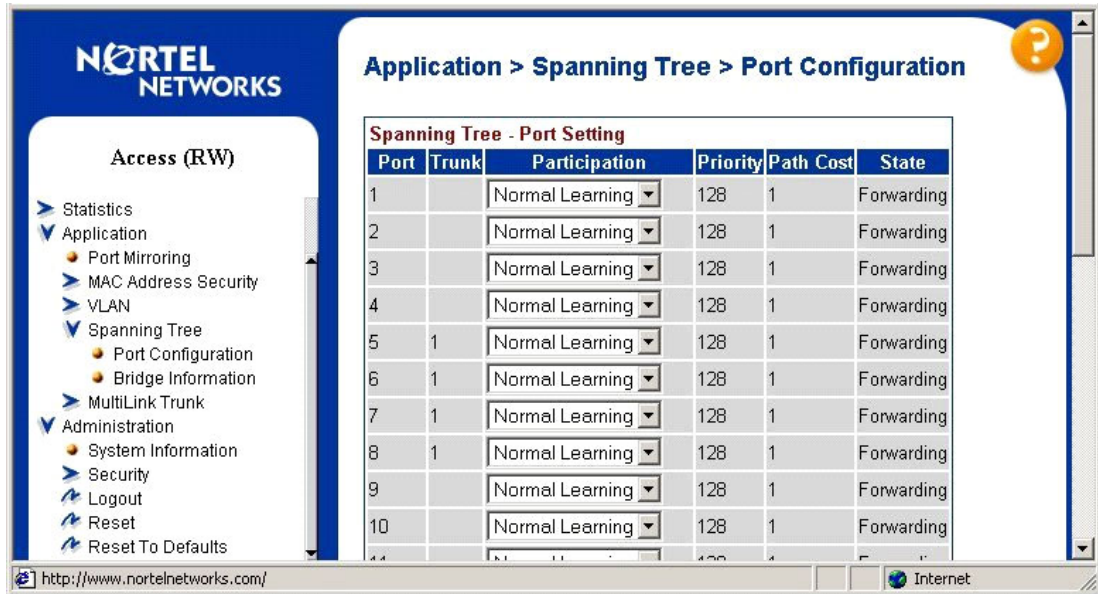


Figure 6-19 Port Configuration window

There are other ways to verify the proper operation. As long as the BayStack 380-24Ts IP address is on the same subnet as the GbESM's IP address, you can ping from the BayStack to the GbESM. The BayStack 380-24T ping utility is only available from the CI Menu interface (console or Telnet). From the main menu, choose **IP Configuration/Setup**. Set the address to ping and then start the ping by choosing the corresponding option. This is shown in Figure 6-20. To test the routing function of the GbESM, ping from either the client machine to Blade 2, or from the BayStack switch to Blade 2. If routing is configured properly on the GbESM, and the gateways are configured properly on the client machine and Blades 2 and 3, the GbESM will forward the traffic to the blades via their respective interfaces. Note that Blades 2 and 3 and the client machine are connected to ports on the network that are in different VLANs.

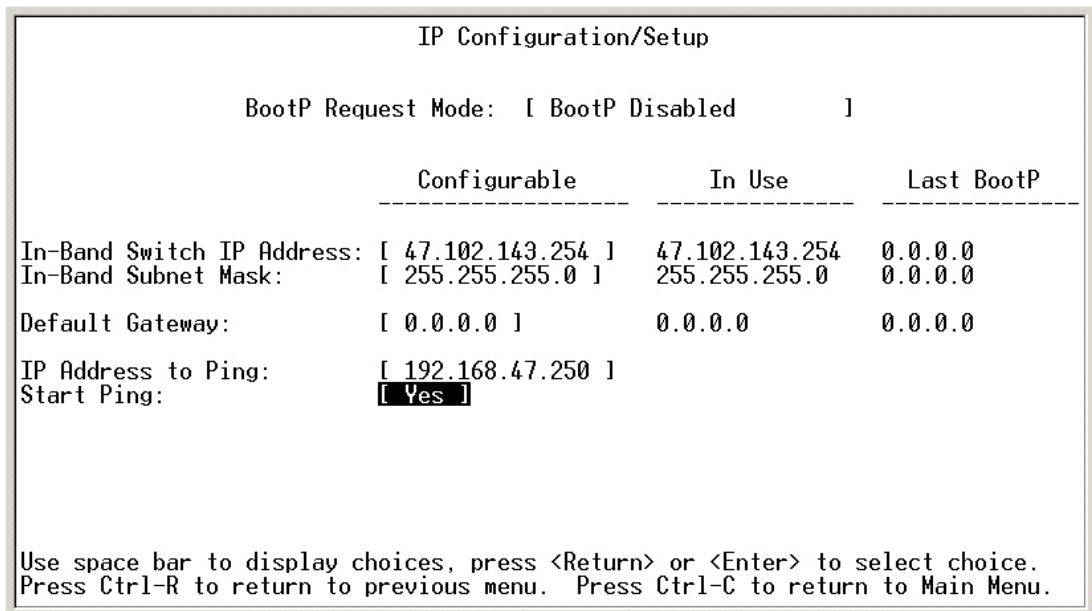


Figure 6-20 IP Configuration/Setup

Also, a BayStack 380-24T port can be assigned to a particular VLAN so that a workstation can be connected to the port in order to ping the corresponding server.

Not all steps must be performed to verify proper configuration; these are just possible tests that can be used.

6.4.4 Single GbESM with link aggregation to dual BayStack 380-24Ts

This network design offers greater protection against equipment failure, but at the cost of available throughput. Specifically, it incorporates a redundant BayStack 380, but only two Gigabit links are carrying traffic because the other two are blocked by Spanning Tree. This configuration (Figure 6-21) does not provide protection against GbESM failure. It can be used when the extra throughput is not as big a concern as the possibility of a BayStack 380 failure.

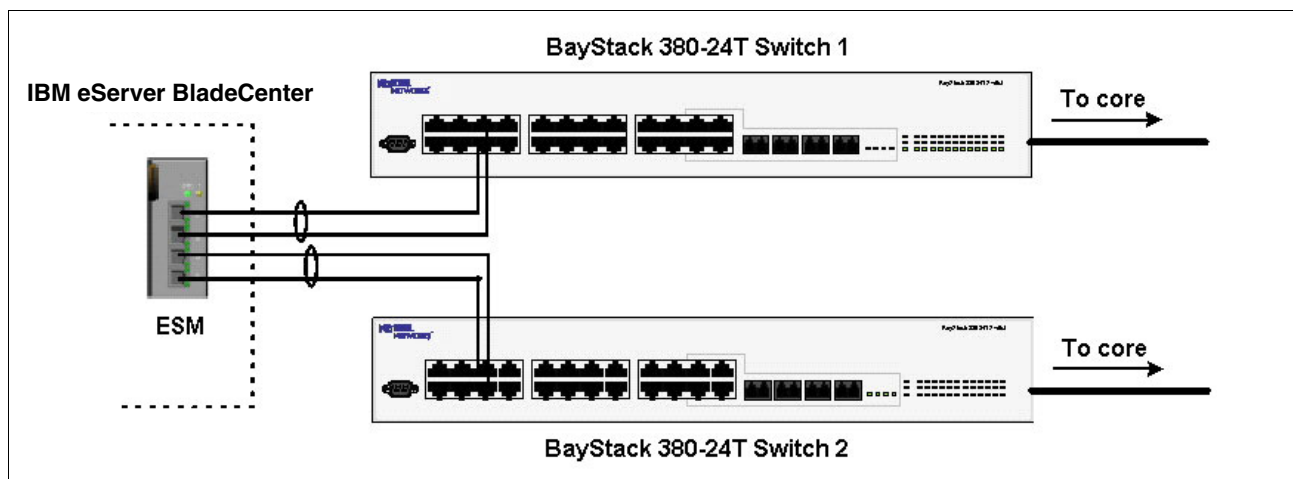


Figure 6-21 IBM eServer BladeCenter and dual BayStack 380-24T switches

Configuring the GbESM

Table 6-3 on page 182 walks through the actions required to configure the GbESM for this example. The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with root level access.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 1 are being used as the link between the @server BladeCenter and Nortel Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are being used as the link between the @serverBladeCenter and Nortel Switch 2.
- ▶ Commands are performed in the sequence shown.
- ▶ VLAN 1 on the GbESM is already enabled and all ports except the management ports are members. If any changes need to be made to VLAN 1 the instructions note this.
- ▶ Nortel Networks switches are BayStack 380-24T gigabit switches. Autonegotiate will result in ports operating at 1000Base-T.
- ▶ All ports remain in Spanning Tree Group 1, and STG 1 has been turned off.
- ▶ No cables are connected to either switch.
- ▶ The GbESM has been reset to factory defaults as shown in Example 6-4 on page 168.

Table 6-3 Configuring the GbESM

Description and comments	Instructions
<p>Step 1 - Configure the IP interfaces and routing.</p> <p>This creates an interface on the public network. Since the management interface is on the management VLAN, we must configure an interface on the public network.</p>	<p>Enter the command /cfg/13/if 1/addr 192.168.47.127 to configure interface 1 with IP Address 192.168.47.127.</p> <p>Enter the command: /cfg/13/if 1/mask 255.255.255.0 to configure subnet mask 255.255.255.0.</p> <p>Enter the command: /c/13/if 1/ena to enable interface 1.</p> <p>Repeat these three steps with IP address 10.10.50.101 and subnet mask 255.255.255.0 on interface 2. Interface 2 will be used to route traffic to VLAN 5.</p> <p>Repeat these three steps with IP address 20.10.50.101 and subnet mask 255.255.255.0 on interface 3. Interface 3 will be used to route traffic to VLAN 10.</p> <p>Enter the command /cfg/13/frwd/on to enable IP forwarding. It should be enabled by default.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 2 - Configure internal ports.</p> <p>This places the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command /cfg/12/stg 1/off to turn off Spanning Tree.</p> <p>Enter the command: /cfg/12/vlan 5/ena to enable VLAN 5.</p> <p>Enter the command: /cfg/12/vlan 5/add INT2 to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: /cfg/12/vlan 10/ena to enable VLAN 10.</p> <p>Enter the command: /cfg/12/vlan 10/add INT3 to add internal port 3 to VLAN 10.</p> <p>All other PVIDs should be set for 1, with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: /info/port. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 3 - Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANS 1, 5, and 10 simultaneously.</p>	<p>Enter the command: /cfg/port EXT1/tag e to enable tagging on external port 1. Repeat this command for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 4- Configure the VLANS for the external ports.</p> <p>This puts the ports into VLANs 1, 5, and 10. This will allow external ports 1 through 4 to carry traffic for VLANS 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p>	<p>Enter the command: /cfg/12/vlan 5/add EXT1 to add external port 1 to VLAN 5. Repeat for external ports 2 through 4.</p> <p>Enter the command: /cfg/12/vlan 10/add EXT1 to add external port 1 to VLAN 10. Repeat for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>

<p>Step 5 - Configure external ports and multi-link trunking.</p> <p>This scenario changes from the previous one and we need to create two trunks on the GbESM. Each trunk will have two ports.</p>	<p>Enter the command <code>/cfg/12/trunk 1/add EXT1</code> to add port EXT1 to Trunk Group 1.</p> <p>Repeat for external port 2.</p> <p>Enter the command <code>/cfg/12/trunk 1/ena</code> to enable Trunk Group 1.</p> <p>Enter the command <code>/cfg/12/trunk 2/add EXT3</code> to add port EXT3 to Trunk Group 2.</p> <p>Repeat for external port 4.</p> <p>Enter the command <code>/cfg/12/trunk 2/ena</code> to enable Trunk Group 2.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM</p>
---	--

Configuring the BayStack 380-24T

The configuration of Switch 2 will be similar to Switch 1. Only the configuration of Switch 1 is displayed here.

Step 1: Configure ports

Click **Application** → **VLAN** → **Port Configuration**. The resulting screen is shown in Figure 6-22. Configure the ports to be tagged. Port names can also be assigned. Click **Submit**.

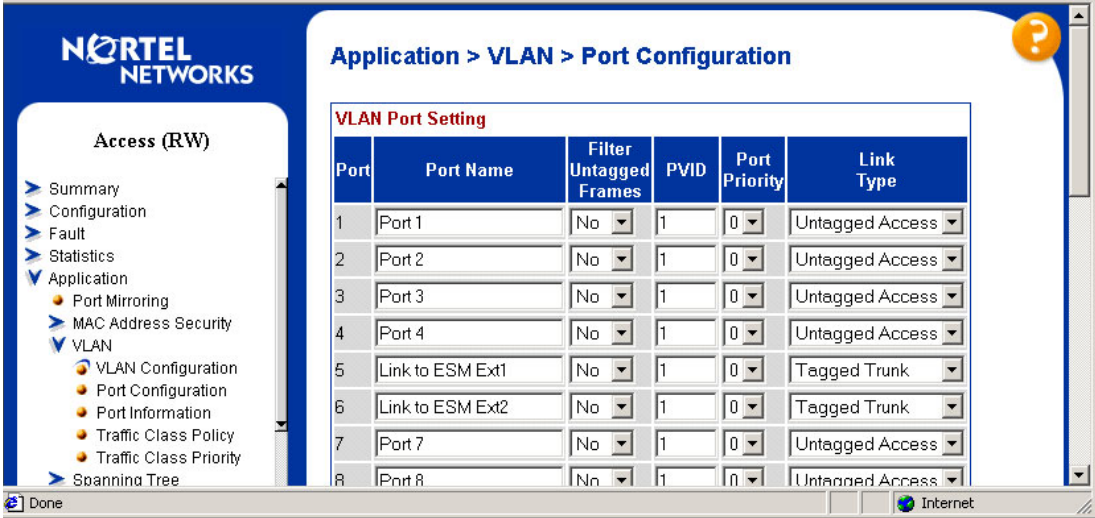


Figure 6-22 Application > VLAN > Port Configuration window

Step 2: Add VLANs to each link member

Click **Application** → **VLAN** → **VLAN Configuration**. Click the action button next to VLAN 5. Add the ports to VLAN 5 as shown in Figure 6-23 and click **Submit**. Repeat this for VLAN 10 and click **Submit**.

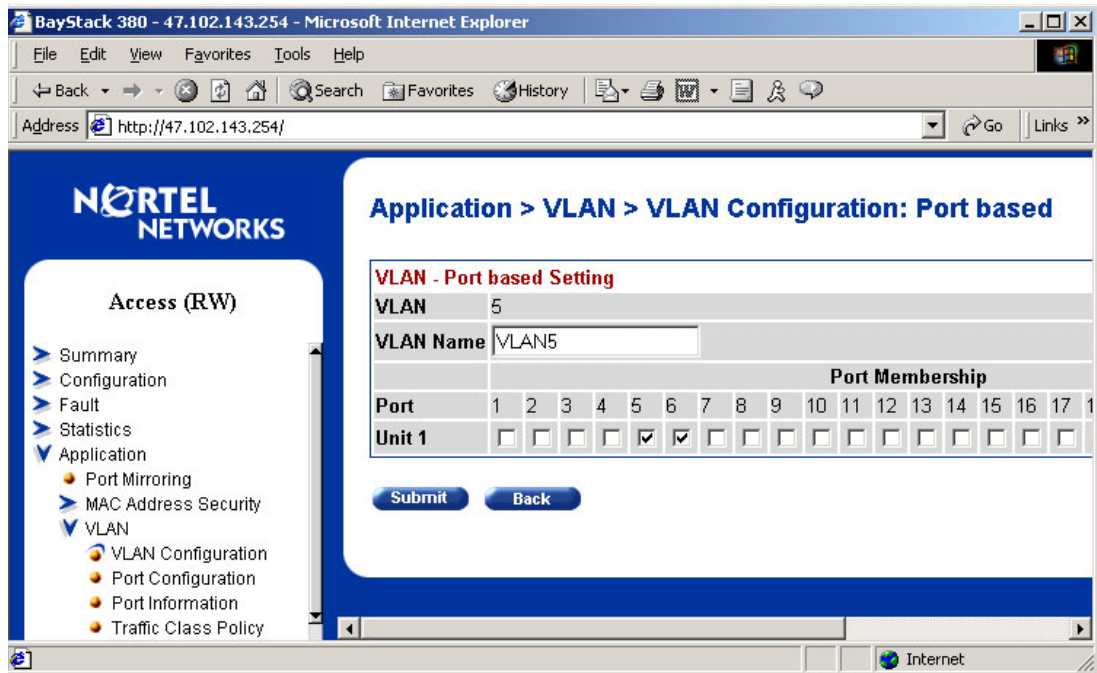


Figure 6-23 Application > VLAN > VLAN Configuration: Port based

Step 3: Configure MLT

Click **Application** → **MultiLink Trunk** → **Group**. Specify ports 5 and 6, and leave “STP Learning” as Normal (Figure 6-24). Click **Submit**. The trunk is not enabled at this point. It must explicitly be enabled by changing the setting in the table to Enabled and clicking **Submit**.

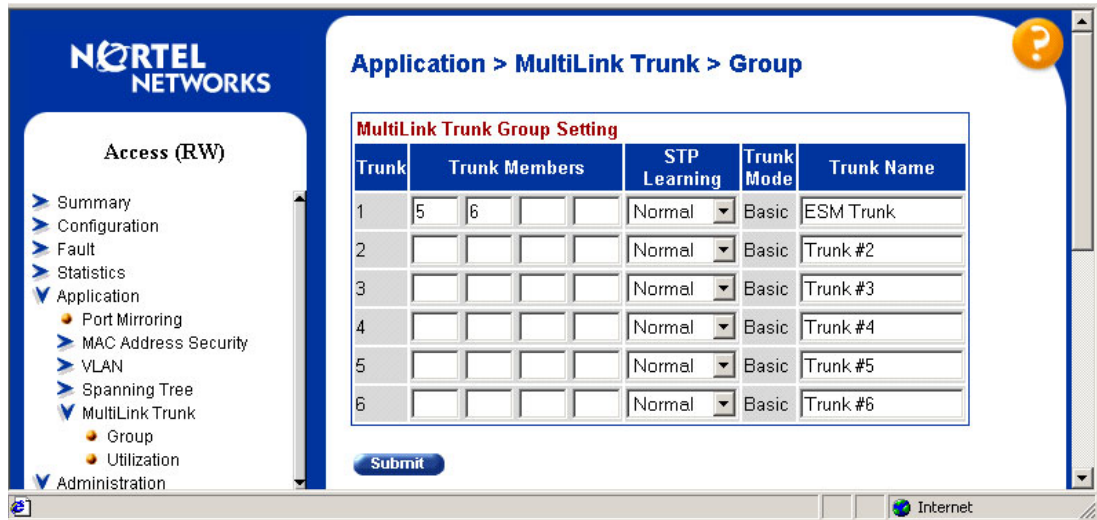


Figure 6-24 Application > MultiLink Trunk > Group window

6.4.5 Validation of GbESM configuration

To validate the setup of the GbESM, refer to “Step 5: Verify the configuration” on page 112.

6.4.6 Validation of BayStack 380-24T configuration

This section describes some steps that can be taken to quickly verify the configuration of the BayStack 380-24Ts. Only Switch 1 is shown below, though both BayStack switches should show the same essential information.

Check port status, speed, and duplex settings from the Configuration > Port Management screen. Note that if autonegotiate is enabled, the Speed/Duplex box displays the negotiated speed/duplex setting. If autonegotiate is disabled, the Speed/Duplex box reflects the manual setting (Figure 6-25).

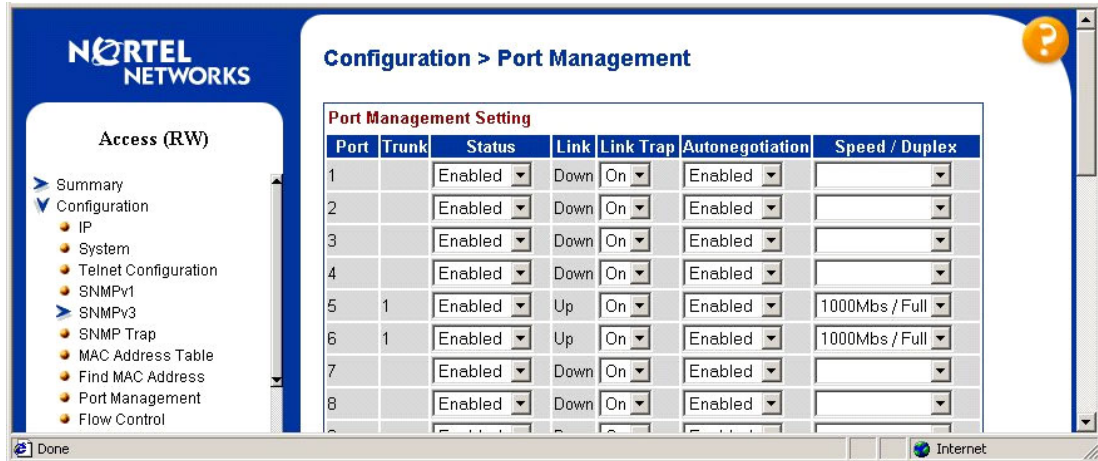


Figure 6-25 Port Management configuration window

Verify Spanning Tree forwarding state for the ports that connect to the GbESM. Go to Application > Spanning Tree > Port Configuration and confirm that the trunk ports are in a forwarding state. Both BayStack switches will be in the forwarding state (Figure 6-26). This is the proper result for Spanning Tree Protocol.

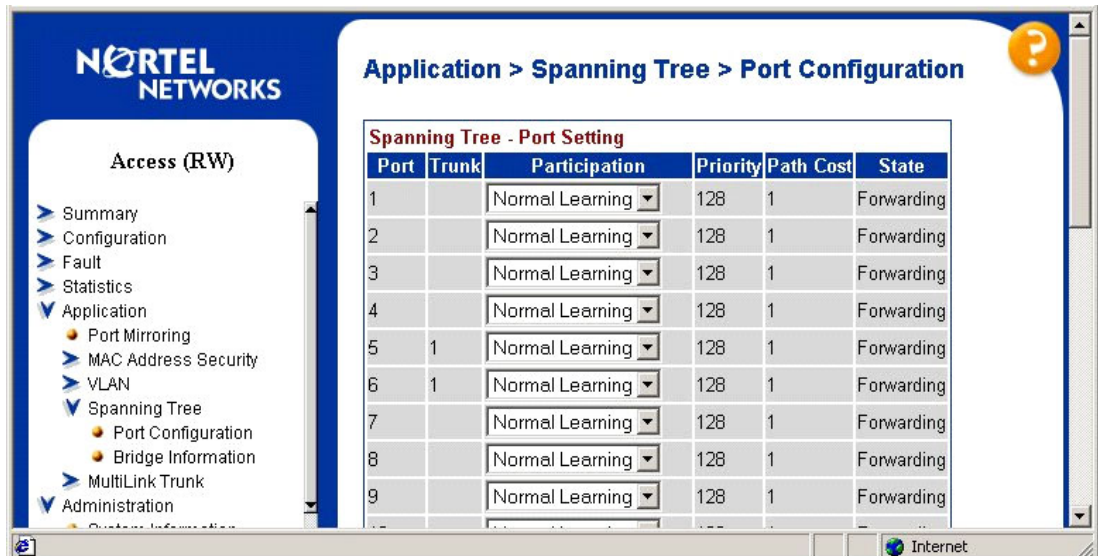


Figure 6-26 Port Configuration window

There are other ways to verify the proper operation. If GbESM's Interface 1 IP address is in the same subnets as the BayStack 380-24T's, you can ping from the BayStack 380-24T to the GbESM. The BayStack 380-24T ping utility is only available from the CI Menu interface

(console or Telnet). From the main menu, choose **IP Configuration/Setup**. Set the address as shown in Figure 6-27.

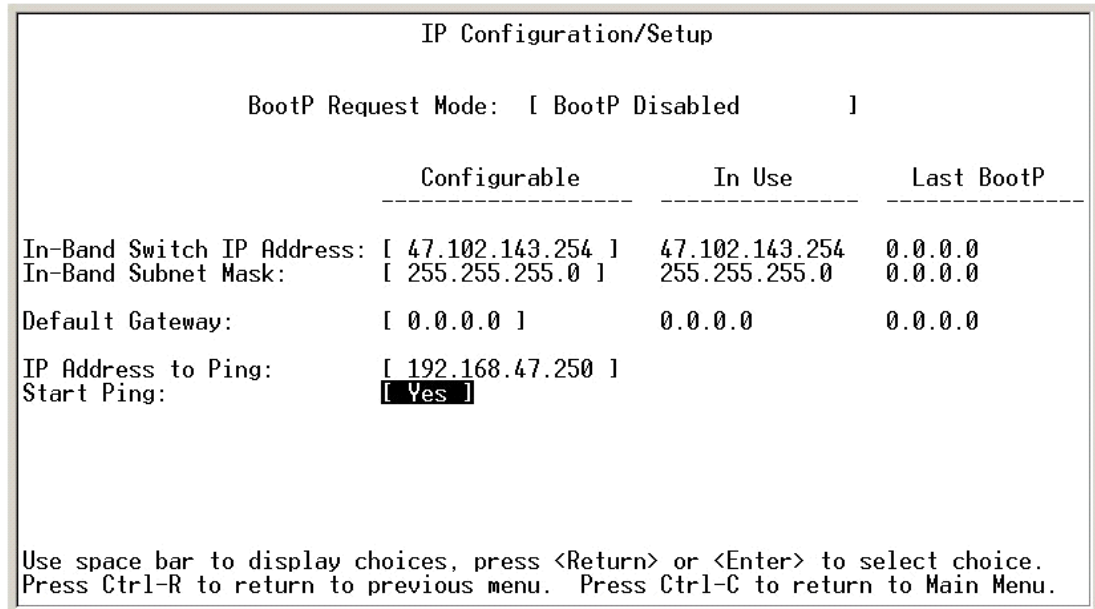


Figure 6-27 IP Configuration/Setup window

Not all steps must be performed to verify proper configuration; these are just possible tests that can be used. You can also test the routing by cabling one client to each of the BayStack switches. If routing is configured properly, the client will be able to ping any of the Blades in the BladeCenter.

6.4.7 Dual GbESMs with two port aggregation to Dual BayStack 380-24Ts

In this scenario (Figure 6-28 on page 187), two GbESMs and two BayStack 380s are connected by a series of two link trunks. Two of the four trunks will be blocked by Spanning Tree. Any one switch or GbESM can fail, as well as several ports or links, and it will still survive. This reliability is at the cost of throughput. This configuration can be used when a high degree of resiliency is required, but the maximum possible throughput is not required.

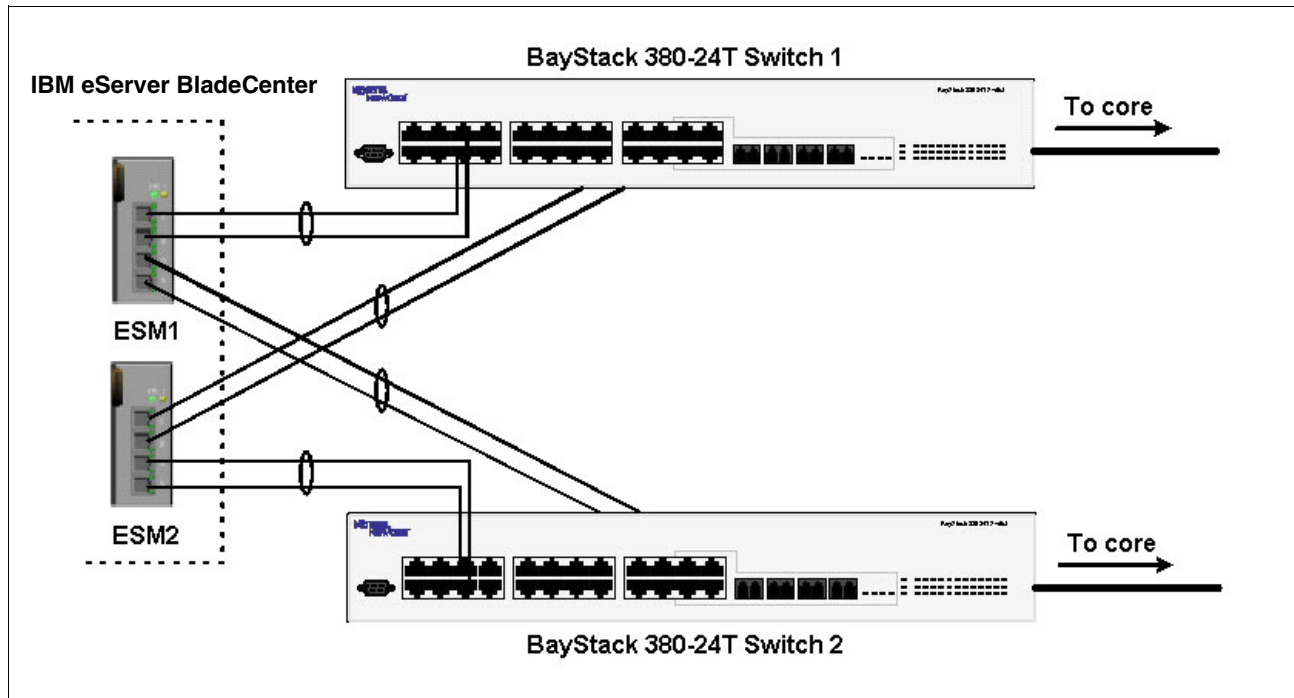


Figure 6-28 Dual GbESMs and dual BayStack 380 switches

Configuring the GbESM

Table 6-4 on page 188 walks through the actions required to configure the GbESM for this example. The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with admin level access.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 1 are being used as the link between the IBM @server BladeCenter and Nortel Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are being used as the link between the IBM @server BladeCenter and Nortel Switch 2.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 2 are being used as the link between the IBM @server BladeCenter and Nortel Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are being used as the link between the IBM @server BladeCenter and Nortel Switch 2.
- ▶ Commands are performed in the sequence shown.
- ▶ VLAN 1 on the GbESM is already enabled and all ports except the management ports are members. If any changes need to be made to VLAN 1 the instructions note this.
- ▶ Nortel Networks switches are BayStack 380-24T gigabit switches. Autonegotiate will result in ports operating at 1000Base-T.
- ▶ No cables are connected between any switches.
- ▶ The GbESM has been reset to factory defaults as shown in Example 6-4 on page 168.

The following steps apply to configuring the first GbESM. The same steps are used to configure the second GbESM as well. Repeat all steps in Table 6-4 for the second GbESM.

Table 6-4 Configuring the GbESMs

Description and comments	Instructions
<p>Step 1 - Configure internal ports.</p> <p>This places the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged. In this example, it is essential that Spanning Tree remains enabled.</p>	<p>Enter the command <code>/cfg/12/stg 1/on</code> to turn on Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer <code>y</code> to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. Repeat this command for INT4.</p> <p>All other PVIDs should be set for 1, with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 2- Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1. Repeat this command for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 3- Configuring the VLANS for the external ports.</p> <p>This puts the ports into VLANs 1, 5, and 10. This will allow external ports 1 through 4 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 as they are in it by default.</p>	<p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2 through 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 4- Configure external ports and multi-link trunking.</p> <p>This scenario changes from the previous one with the addition of a second GbESM, but we still need to create two trunks on each GbESM. Each trunk will have two ports.</p>	<p>Enter the command <code>/cfg/12/trunk 1/add EXT1</code> to add port EXT1 to Trunk Group 1.</p> <p>Repeat for external port 2.</p> <p>Enter the command <code>/cfg/12/trunk 1/ena</code> to enable Trunk Group 1.</p> <p>Enter the command <code>/cfg/12/trunk 2/add EXT3</code> to add port EXT3 to Trunk Group 2.</p> <p>Repeat for external port 4.</p> <p>Enter the command <code>/cfg/12/trunk 2/ena</code> to enable Trunk Group 2.</p> <p>Enter apply to apply the new configuration Enter save to save the configuration to NVRAM</p>

Configuring the BayStack 380-24T

The configuration of Switch 2 will be similar to Switch 1. Only the configuration of Switch 1 is described here.

Step 1: Configure ports

Click **Application** → **VLAN** → **Port Configuration**. The resulting screen is shown in Figure 6-29. Configure the ports to be tagged. Port names can also be assigned. Click **Submit**.

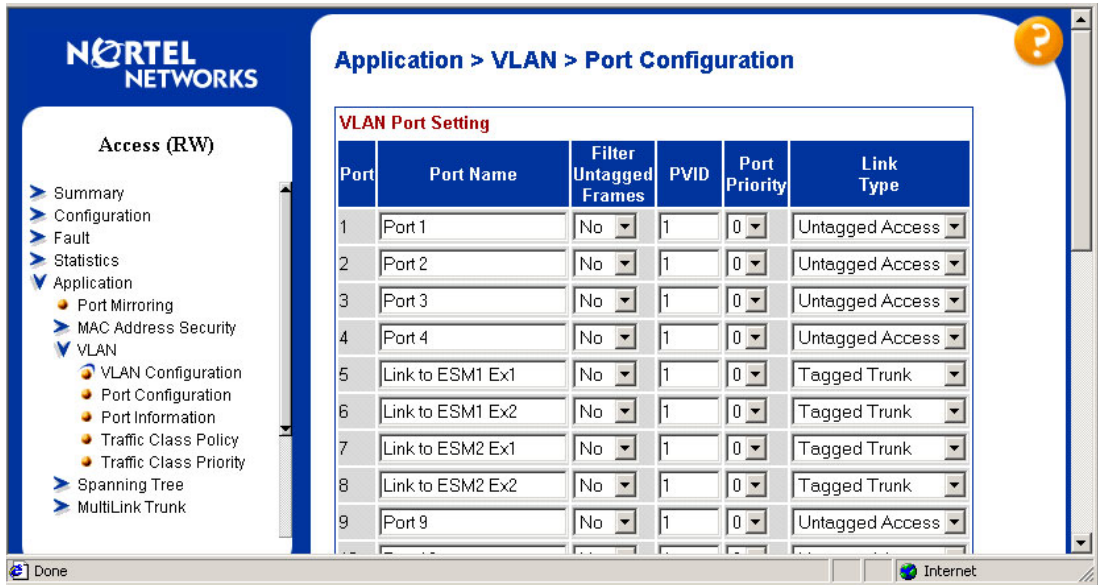


Figure 6-29 Application > VLAN > Port Configuration window

Step 2: Add VLANs to each link member

Click **Application** → **VLAN** → **VLAN Configuration**. Click the action button next to VLAN 5. Add the ports to VLAN 5 as shown in Figure 6-30 and click **Submit**. Repeat this for VLAN 10 and click **Submit**.

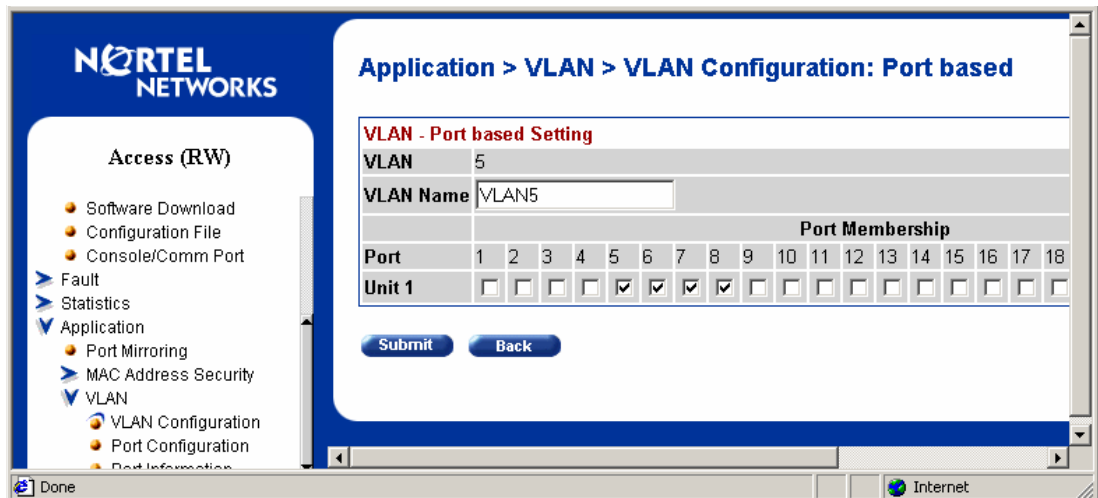


Figure 6-30 Application > VLAN > VLAN Configuration: Port based window

Step 3: Configure MLT

Click **Application** → **MultiLink Trunk** → **Group**. On the resulting screen, shown in Figure 6-31, specify ports 5 and 6, and leave “STP Learning” as Normal. Click **Submit**. Specify ports 7 and 8 as a second trunk with “STP Learning” as Normal. Click **Submit**. The trunks are not enabled at this point. Both trunks must explicitly be enabled by changing the setting in the table to Enabled and clicking **Submit**.

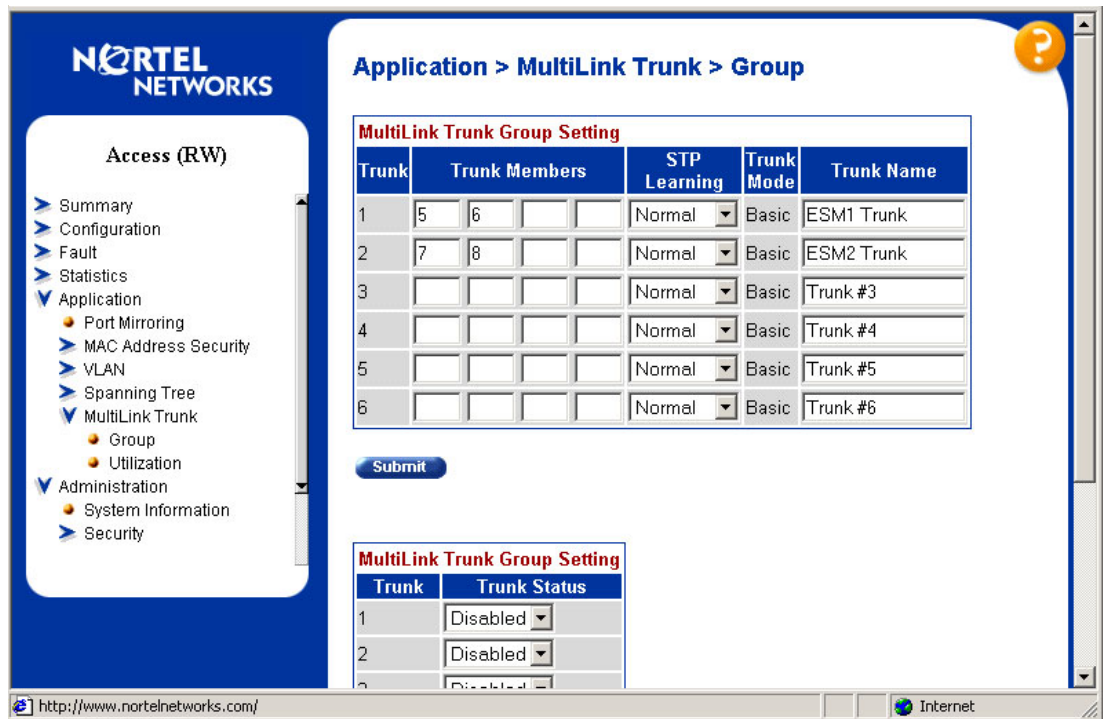


Figure 6-31 Application > MultiLink Trunk > Group window

6.4.8 Validation of GbESM configuration

To validate the setup of the GbESM, refer to “Step 5: Verify the configuration” on page 112.

6.4.9 Validation of BayStack 380-24T configuration

This section describes how to quickly verify the configuration of the BayStack 380-24Ts. Only Switch 1 is shown in Figure 6-32, although both BayStack switches should show the same essential information.

Check port status, speed, and duplex settings from the Configuration > Port Management screen. Note that if autonegotiate is enabled, the Speed/Duplex box displays the negotiated speed/duplex setting. If autonegotiate is disabled, the Speed/Duplex box reflects the manual setting.

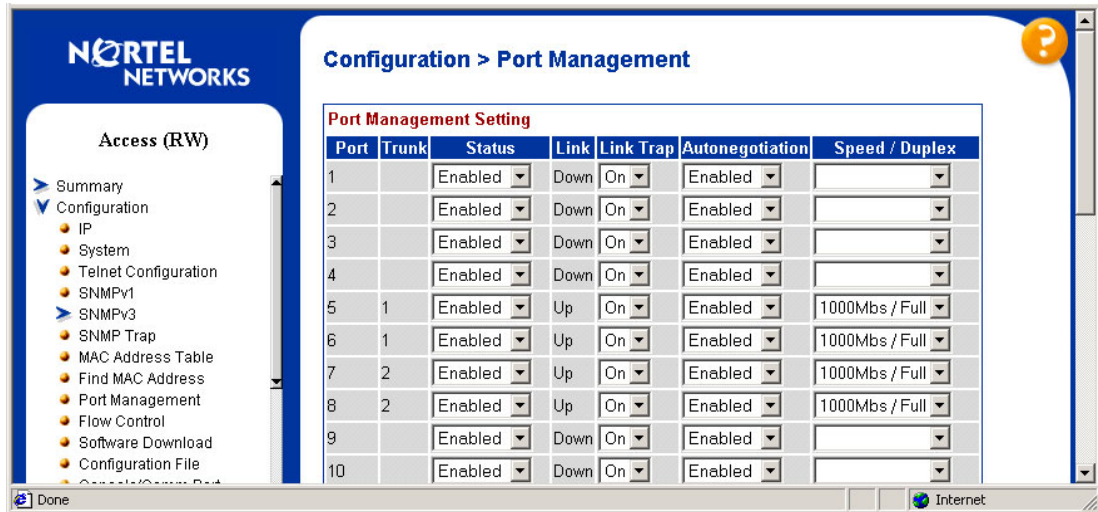


Figure 6-32 Port Management window

Verify Spanning Tree forwarding state for the ports that connect to the GbESM. Go to **Application** → **Spanning Tree** → **Port Configuration** and confirm that the trunk ports are in a forwarding state. Even though some of the GbESMs' link aggregation groups should show a blocking state, both BayStack switches will be in the forwarding state (Figure 6-33 on page 191). This is the proper result for Spanning Tree Protocol.

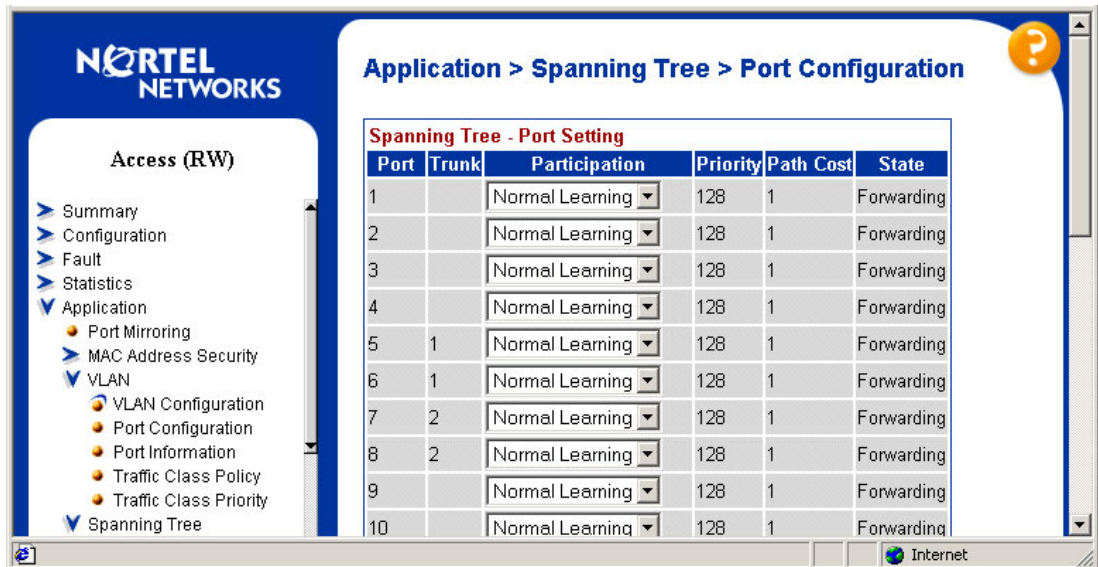


Figure 6-33 Port Configuration window

There are other steps that can also be used to verify the proper operation. If the GbESM management IP address is the same subnet as the BayStack 380-24T IP address, you can ping from BayStack 380-24T to GbESM. The BayStack 380-24T ping utility is only available from the CI Menu interface (console or Telnet). From the main menu, choose **IP Configuration/Setup**. Set the address to ping and then start the ping by choosing the corresponding option. This is shown in Figure 6-34 on page 192.

```

IP Configuration/Setup

BootP Request Mode: [ BootP Disabled ]

Configurable      In Use      Last BootP
-----
In-Band Switch IP Address: [ 47.102.143.254 ] 47.102.143.254 0.0.0.0
In-Band Subnet Mask:      [ 255.255.255.0 ] 255.255.255.0 0.0.0.0
Default Gateway:          [ 0.0.0.0 ] 0.0.0.0 0.0.0.0
IP Address to Ping:       [ 192.168.47.250 ]
Start Ping:               [ Yes ]

Use space bar to display choices, press <Return> or <Enter> to select choice.
Press Ctrl-R to return to previous menu. Press Ctrl-C to return to Main Menu.

```

Figure 6-34 IP Configuration/Setup window

Also, a BayStack 380-24T port can be assigned to a particular VLAN so that a workstation can be connected to the port in order to ping the corresponding server.

Not all steps must be performed to verify proper configuration; these are just possible tests that can be used.

6.4.10 Dual GbESMs with four port aggregation to dual BayStack 380-24Ts

This topology (Figure 6-35 on page 193) shows the maximum possible throughput for an IBM @server BladeCenter server equipped with two GbESMs. It does not sacrifice very much resiliency in gaining the throughput, but rather it moves the resiliency mechanisms to the servers themselves. This configuration can be used in networks which need the highest level of performance. This configuration is the same as the first scenario which had a single GbESM connected to a single BayStack. Each GbESM will have a single four-port multi-link trunk to a BayStack.

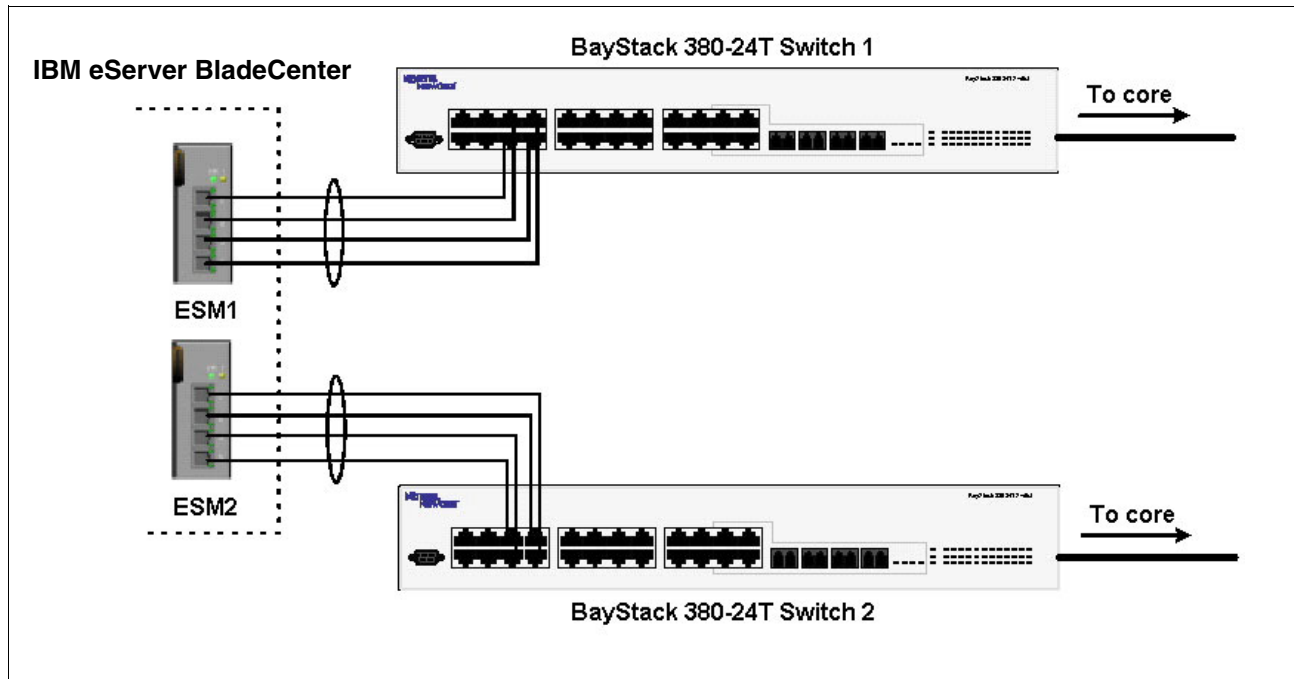


Figure 6-35 Dual GbESMs and dual BayStack 380 switches

Configuring the GbESM

Table 6-5 on page 194 walks through the actions required to configure the GbESM for this example. The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with root level access.
- ▶ Ports EXT1 through EXT4 on the GbESM in Switch Module Bay 1 are being used between IBM and Nortel Switch 1.
- ▶ Ports EXT1 through EXT4 on the GbESM in Switch Module Bay 2 are being used between IBM and Nortel Switch 2.
- ▶ The same VLAN setup is being used on both GbESMs. (This is not a requirement, the second GbESM could be configured to use different VLANs.)
- ▶ Commands are performed in the sequence shown.
- ▶ VLAN 1 on the GbESM is already enabled and all ports except the management ports are members. If any changes need to be made to VLAN 1 the instructions note this.
- ▶ Nortel Networks switches are BayStack 380-24T gigabit switches. Autonegotiate will result in ports operating at 1000Base-T.
- ▶ All ports remain in Spanning Tree Group 1, and STG 1 has been turned off.
- ▶ No cables are connected between any switches.
- ▶ The GbESM has been reset to factory defaults as shown in Example 6-4 on page 168.

The steps in Table 6-5 apply to configuring the first GbESM. Repeat the steps when configuring the second GbESM.

Table 6-5 Configuring the GbESM

Description and comments	Instructions
<p>Step 1 - Configure internal ports.</p> <p>This places the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. Repeat this command for INT4.</p> <p>All other PVIDs should be set for 1, with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 2 - Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1. Repeat this command for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 3 - Configuring the VLANS for the external ports.</p> <p>This puts the ports into VLANs 1, 5, and 10. This will allow external ports 1 through 4 to carry traffic for VLANs 1, 5, and 10. The ports do not need to be added to VLAN 1 because they are in it by default.</p>	<p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2 through 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 4 - Configure external ports and multi-link trunking.</p> <p>This scenario changes from the previous one and we no longer need two trunks on a single switch. We can move back to a single four port trunk.</p>	<p>Enter the command <code>/cfg/12/trunk 1/add EXT1</code> to add port EXT1 to Trunk Group 1.</p> <p>Repeat this step for external ports 2 through 4.</p> <p>Enter the command <code>/cfg/12/trunk 1/ena</code> to enable Trunk Group 1.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM</p>

Configuring the BayStack 380-24T

The configuration of Switch 2 will be similar to Switch 1. Only the configuration of Switch 1 is displayed here.

Step 1: Configure ports

Click **Application** → **VLAN** → **Port Configuration**. The resulting screen is shown in Figure 6-36. Configure the links to be tagged. Port names can also be assigned. Click **Submit**.

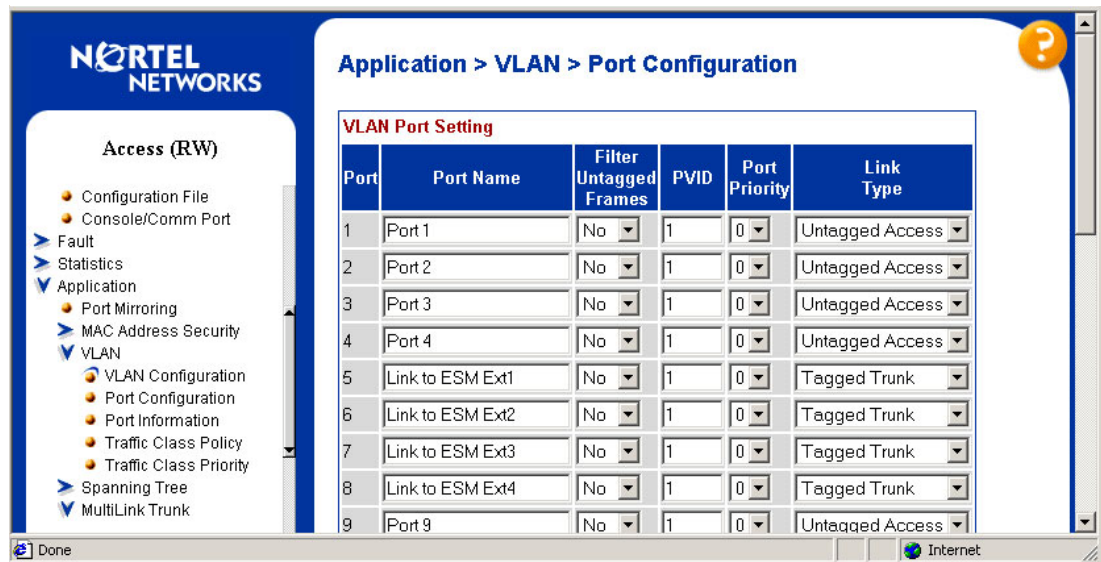


Figure 6-36 Application > VLAN > Port Configuration window

Step 2: Add VLANs to each link member

Click **Application** → **VLAN** → **VLAN Configuration**. On the resulting screen, shown in Figure 6-37, click the action button next to VLAN 5. Add the ports to VLAN 5 as shown in the figure and click **Submit**. Repeat this for VLAN 10 and click **Submit**.

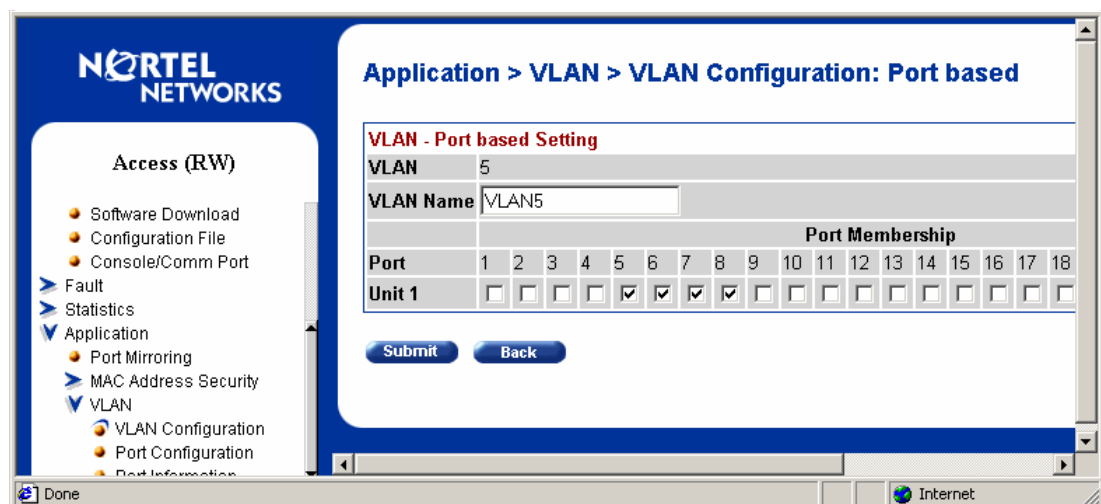


Figure 6-37 Application > VLAN > VLAN Configuration: Port based window

Step 3: Configure MLT

Click **Application** → **MultiLink Trunk** → **Group**. On the resulting screen, shown in Figure 6-38, specify ports 5, 6, 7, and 8, and set “STP Learning” to Fast. Click **Submit**. The trunk is not enabled at this point. It must explicitly be enabled by changing the setting in the table to Enabled and clicking **Submit**.

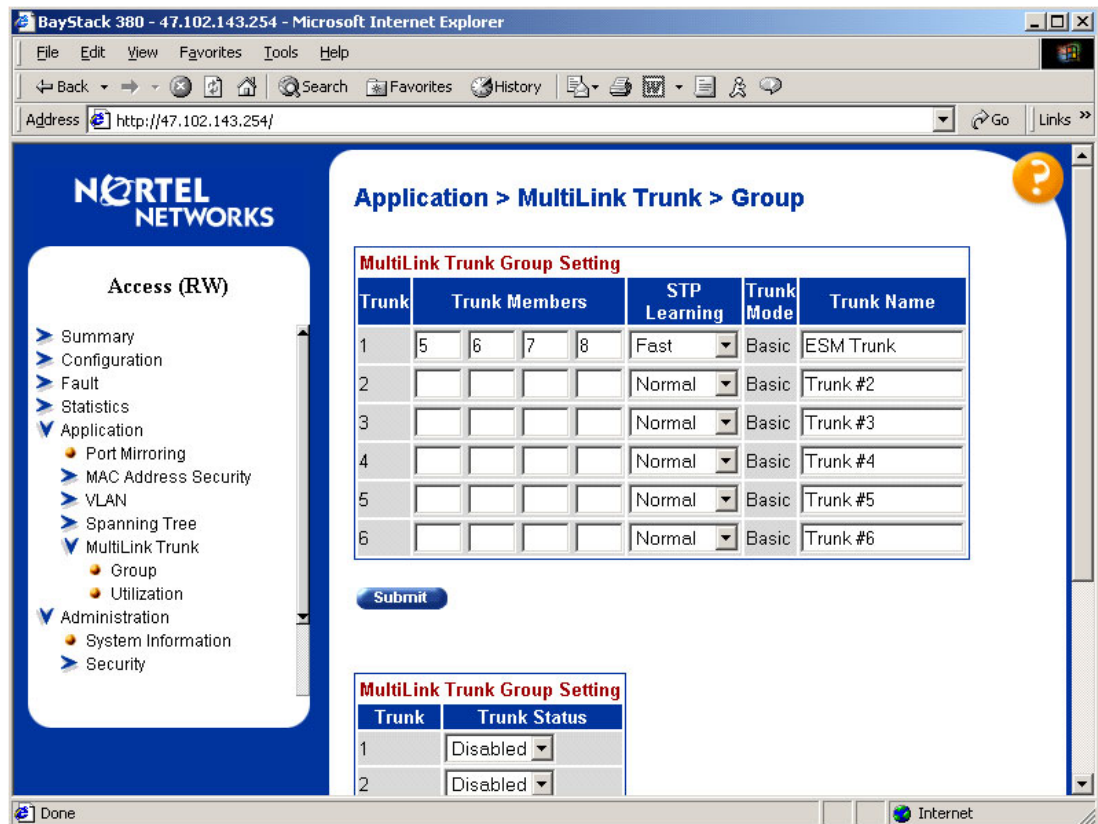


Figure 6-38 Application > MultiLink Trunk > Group window

6.4.11 Validation of GbESM configuration

To validate the setup of the GbESM, refer to “Step 5: Verify the configuration” on page 118.

6.4.12 Validation of BayStack 380-24T configuration

This section describes some steps that can be taken to quickly verify the configuration of the BayStack 380-24T. Validation should be performed on both switches, and the methods listed here are applicable to both switches.

Check port status, speed, and duplex settings from the Configuration > Port Management screen. Note that if autonegotiate is enabled, the Speed/Duplex box displays the negotiated speed/duplex setting. If autonegotiate is disabled, the Speed/Duplex box reflects the manual setting (Figure 6-39 on page 197).

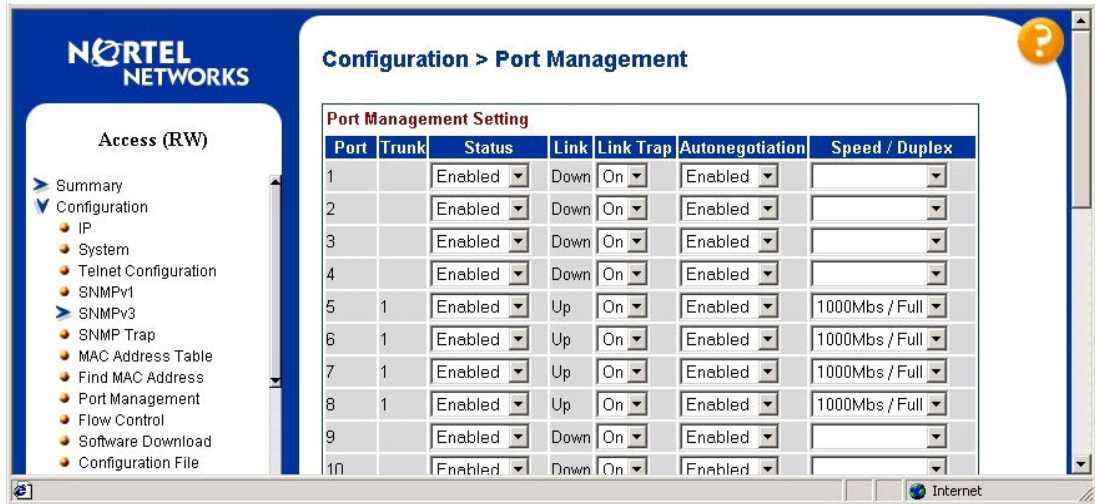


Figure 6-39 Port Management configuration window

Verify Spanning Tree forwarding state for the ports that connect to the GbESM. Go to Application > Spanning Tree > Port Configuration and confirm that the trunk ports are in a forwarding state (Figure 6-40).

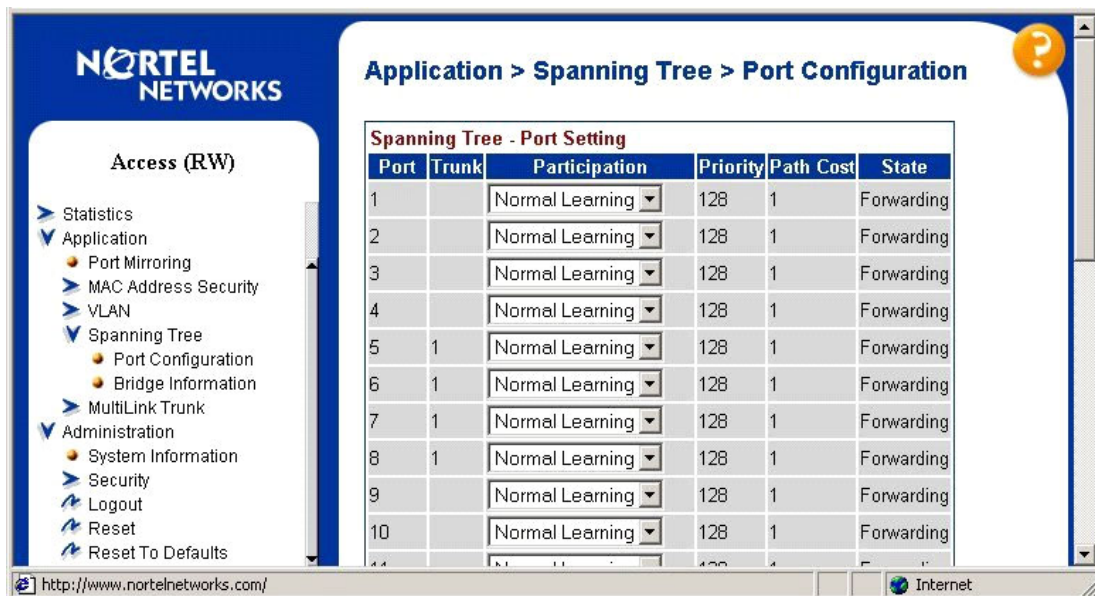


Figure 6-40 Port configuration window

Other steps can be taken to verify the proper operation. If the GbESM management IP address is on the same subnet as the BayStack 380-24Ts IP address, you can ping from BayStack 380-24T to GbESM. The BayStack 380-24T ping utility is only available from the CI Menu interface (console or Telnet). From the main menu, choose **IP Configuration/Setup**. Set the address to ping and then start the ping by choosing the corresponding option. This is shown in Figure 6-41 on page 198.

```

IP Configuration/Setup

BootP Request Mode: [ BootP Disabled ]

          Configurable          In Use          Last BootP
          -----          -
In-Band Switch IP Address: [ 47.102.143.254 ]  47.102.143.254  0.0.0.0
In-Band Subnet Mask:      [ 255.255.255.0 ]    255.255.255.0  0.0.0.0
Default Gateway:         [ 0.0.0.0 ]          0.0.0.0        0.0.0.0
IP Address to Ping:      [ 192.168.47.250 ]
Start Ping:              [ Yes ]

Use space bar to display choices, press <Return> or <Enter> to select choice.
Press Ctrl-R to return to previous menu. Press Ctrl-C to return to Main Menu.

```

Figure 6-41 IP Configuration/Setup window

Also, a BayStack 380-24T port can be assigned to a particular VLAN so that a workstation can be connected to the port in order to ping the corresponding server.

Not all steps must be performed to verify proper configuration; these are just possible tests that can be used.

6.4.13 Dual GbESMs with four port aggregation, each to a single Passport 8600

Many network architectures include an aggregation layer. The BayStack 380-24T serves as a gigabit aggregation layer switch in many networks. By the same measure, the GbESM is performing an aggregation function since it has 14 internal gigabit links and only 4 external gigabit links. Consider a BladeCenter server scenario where the 14 servers are capable of utilizing most of the four port (gigabit speed) aggregate trunk to an uplink device on a sustained basis. Now consider the expanded view of this sample network diagram if BayStack 380-24Ts are used to connect the GbESMs to the network. See Figure 6-42 on page 199.

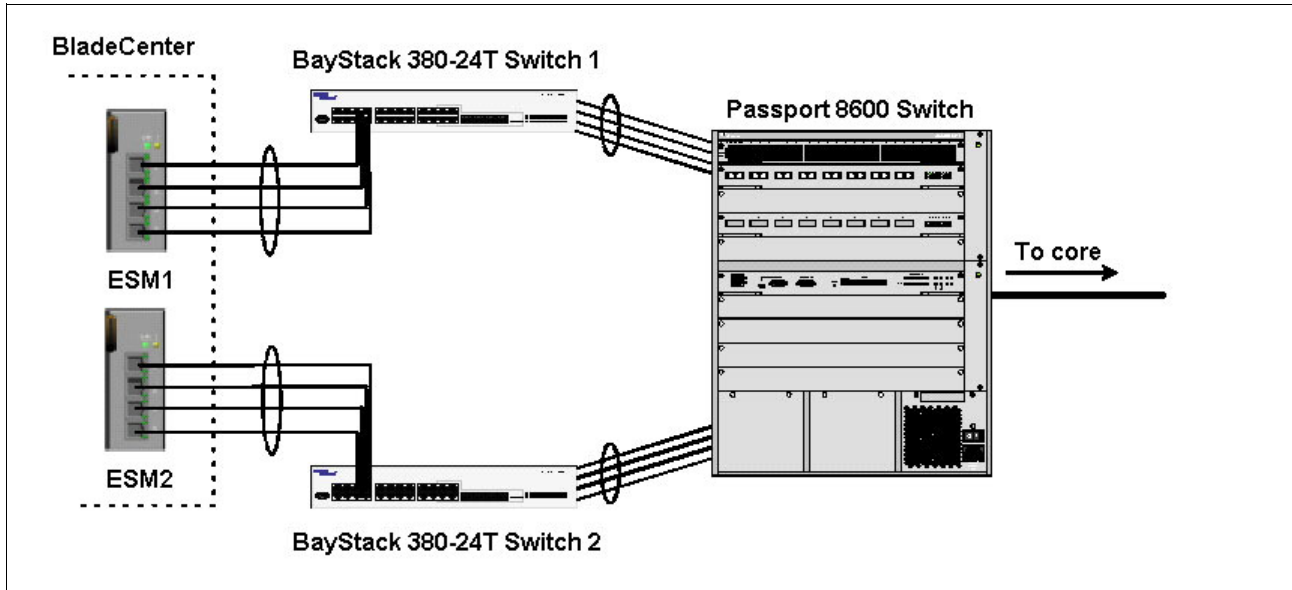


Figure 6-42 Dual GbESMs with four port aggregation, each to a single Passport 8600

Since the aggregate links between GbESM and BayStack 380-24T are mostly utilized, the aggregate links between the BayStack 380-24T and the Passport 8600 will be mostly utilized. In this case, it does not make much sense to use additional ports on the BayStack 380-24T to connect to other servers or end stations, otherwise oversubscription on the aggregate links between BayStack 380-24T and Passport 8600 will be likely. In this scenario, it makes much more sense to directly connect the GbESM to the core switch and bypass the BayStack 380-24T altogether. Hence, Figure 6-43 illustrates the next configuration example.

In short, this design is best implemented when all links are utilized highly enough to necessitate being connected to a core switch located in the data center. No ports will be in a Spanning Tree blocking state. This scenario does not provide resiliency in the case of a Passport 8600 failure, though GbESM and multiple link/port failures can be survived.

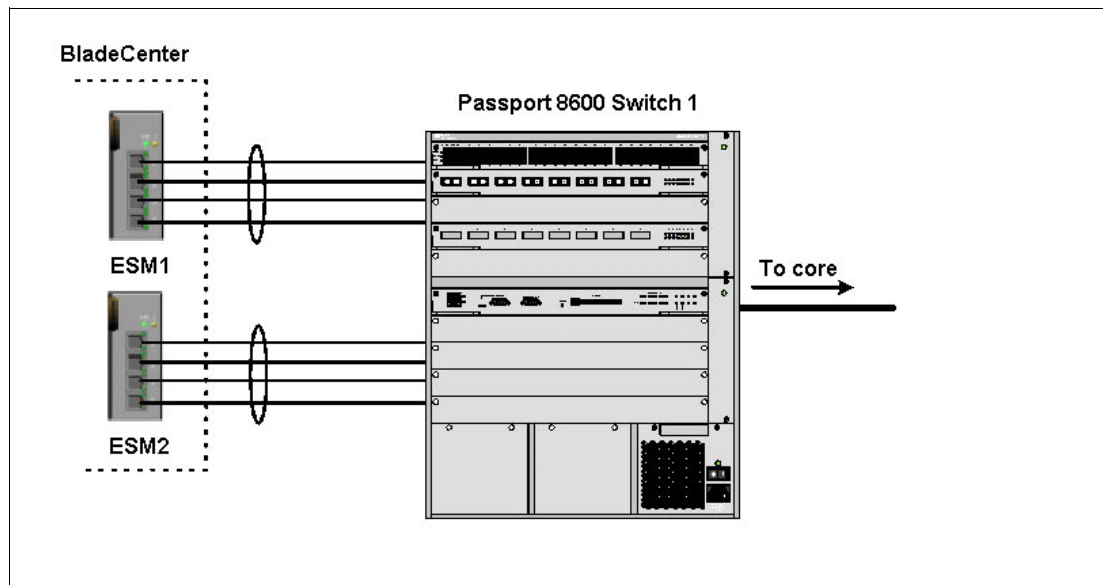


Figure 6-43 Dual GbESMs with four port aggregation, each to single Passport 8600

Configuring the GbESM

Table 6-6 walks through the actions required to configure the GbESM for this example. The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with root level access.
- ▶ Ports EXT1 through EXT4 on the GbESM in Switch Module Bay 1 are being used between IBM and Nortel Passport 8600 Switch 1.
- ▶ Ports EXT1 through EXT4 on the GbESM in Switch Module Bay 2 are being used between IBM and Nortel Passport 8600 Switch 1.
- ▶ The same VLAN setup is being used on both GbESMs. (This is not a requirement; the second GbESM could be configured to use different VLANs.)
- ▶ Commands are performed in the sequence shown.
- ▶ The Nortel switch is a Passport 8600. Autonegotiate will result in ports operating at 1000Base-T.
- ▶ All ports remain in Spanning Tree group 1 and STG 1 has been turned off.
- ▶ No cables are connected between any switches.
- ▶ The GbESM has been reset to factory defaults as shown in Example 6-4 on page 168.

The steps in Table 6-6 apply to configuring the first GbESM. Essentially the same steps are used to configure the second GbESM as well.

Table 6-6 Configuring the GbESM

Description and comments	Instructions
<p>Step 1 - Configure internal ports.</p> <p>This places the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command /cfg/12/stg 1/off to turn off Spanning Tree.</p> <p>Enter the command: /cfg/12/vlan 5/ena to enable VLAN 5.</p> <p>Enter the command: /cfg/12/vlan 5/add INT2 to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: /cfg/12/vlan 10/ena to enable VLAN 10.</p> <p>Enter the command: /cfg/12/vlan 10/add INT3 to add internal port 3 to VLAN 10. Repeat this command for INT4.</p> <p>All other PVIDs should be set for 1, with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: /info/port. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 2 - Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANs 1, 5, and 10 simultaneously.</p>	<p>Enter the command: /cfg/port EXT1/tag e to enable tagging on external port 1. Repeat this command for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 3 - Configure the VLANS for the external ports.</p> <p>This puts the ports into VLANs 1, 5, and 10. This will allow external ports1 through 4 to carry traffic for VLANS 1, 5, and 10. The ports do not need to be added to VLAN 1 because they are in it by default.</p>	<p>Enter the command: /cfg/12/vlan 5/add EXT1 to add external port 1 to VLAN 5. Repeat for external ports 2 through 4.</p> <p>Enter the command: /cfg/12/vlan 10/add EXT1 to add external port 1 to VLAN 10. Repeat for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>

<p>Step 4 - Configure external ports and multi-link trunking.</p> <p>We no longer need two trunks on a single switch. We can move back to a single four-port trunk.</p>	<p>Enter the command /cfg/12/trunk 1/add EXT1 to add port EXT1 to Trunk Group 1.</p> <p>Repeat this step for external ports 2 through 4.</p> <p>Enter the command /cfg/12/trunk 1/ena to enable Trunk Group 1.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM</p>
---	---

Configuring the Passport 8600

This section describes the actions required to configure the Passport 8600 for this example.

Step 1: Configure MLT

To do this you will need to create both MLT groups, enable tagging, and add specific ports. In this example, the 8608GT blade, which has 8 1000Base-T ports, is installed in slot 8. The first four ports are part of the first MLT group, and the last four ports are part of the second MLT group. After ports are added, VLANs are added to the MLT groups. The configuration is performed as shown in Figure 6-44.

```

PP8600-1:5#
PP8600-1:5# config mlt 1
PP8600-1:5/config/mlt/1# create
PP8600-1:5/config/mlt/1# name "ESM1 Trunk"
PP8600-1:5/config/mlt/1# perform-tagging enable
PP8600-1:5/config/mlt/1# add ports 8/1-8/4
PP8600-1:5/config/mlt/1# add vlan 1
PP8600-1:5/config/mlt/1# add vlan 5
PP8600-1:5/config/mlt/1# add vlan 10
PP8600-1:5/config/mlt/1# top
PP8600-1:5#
PP8600-1:5# config mlt 2
PP8600-1:5/config/mlt/2# create
PP8600-1:5/config/mlt/2# name "ESM2 Trunk"
PP8600-1:5/config/mlt/2# perform-tagging enable
PP8600-1:5/config/mlt/2# add ports 8/5-8/8
PP8600-1:5/config/mlt/2# add vlan 1
PP8600-1:5/config/mlt/2# add vlan 5
PP8600-1:5/config/mlt/2# add vlan 10
PP8600-1:5/config/mlt/2# top
PP8600-1:5#

```

Figure 6-44 Console interface: Configuring the MLT

In this configuration Spanning Tree is running on the trunks to the two GbESMs. Neither will end up in the blocking state since there is no loop. Therefore, you may additionally enable Faststart for the interfaces without consequence. To do this, use:

```
config ethernet <portlist> stg 1 faststart enable
```

6.4.14 Validation of GbESM configuration

To validate the setup of the GbESM, refer to “Step 5: Verify the configuration” on page 127.

6.4.15 Validation of Passport 8600 configuration

This section contains some steps that can be taken to quickly verify the configuration of the Passport 8600.

Check port status, speed, and duplex settings using the following command:

```
show ports info name <portlist>
```

If Autonegotiate is enabled, the information shown will be the negotiated speed/duplex settings. In Figure 6-45 is an example output of the command given the configuration provided in this scenario.

```
PP8600-1:5#
PP8600-1:5# show ports info name 8/1-8/8
```

Port Name						
PORT NUM	NAME	DESCRIPTION	OPERATE STATUS	OPERATE DUPLX	OPERATE SPEED	VLAN
8/1		1000BaseT	up	full	1000	Tagged
8/2		1000BaseT	up	full	1000	Tagged
8/3		1000BaseT	up	full	1000	Tagged
8/4		1000BaseT	up	full	1000	Tagged
8/5		1000BaseT	up	full	1000	Tagged
8/6		1000BaseT	up	full	1000	Tagged
8/7		1000BaseT	up	full	1000	Tagged
8/8		1000BaseT	up	full	1000	Tagged

```
PP8600-1:5#
```

Figure 6-45 Command output: Port configuration information

Verify Spanning Tree forwarding state for the ports that connect to the GbESMs. To do this, use the following command:

```
show port info stg main <portlist>
```

Confirm that the trunk ports are all in the forwarding state. The results of this command should look like Figure 6-46.

```
PP8600-1:5#
PP8600-1:5# show port info stg main 8/1-8/8
```

Port Stg								
SID	PORT_NUM	PRIO	STATE	ENABLE STP	FASTSTART	PATHCOST	FORWARD TRANSITION	CHANGE DETECTION
1	8/1	128	forwarding	true	true	1	2	true
1	8/2	128	forwarding	true	true	1	2	true
1	8/3	128	forwarding	true	true	1	2	true
1	8/4	128	forwarding	true	true	1	2	true
1	8/5	128	forwarding	true	true	1	2	true
1	8/6	128	forwarding	true	true	1	2	true
1	8/7	128	forwarding	true	true	1	2	true
1	8/8	128	forwarding	true	true	1	2	true

```
PP8600-1:5#
```

Figure 6-46 Command output: Spanning Tree forwarding state

Other steps can be taken to verify the proper operation. For example the Passport 8600's out-of-band management interface can be connected to a port that is a member of VLAN 1 (assuming the GbESM's management VLAN is 1). If this is done, it is possible to ping from Passport 8600 to GbESM. To ping from the Passport 8600 management interface to GbESM use the following command:

```
ping <ip_addr>
```

An example is shown in Figure 6-47.

```
PP8600-1:5#  
PP8600-1:5# ping 192.168.47.250  
192.168.47.250 is alive  
PP8600-1:5#
```

Figure 6-47 Command output: ping

Also, a Passport 8600 port can be assigned to a particular VLAN so that a workstation can be connected to the port in order to ping the corresponding server.

Not all steps must be performed to verify proper configuration; these are just possible tests that can be used.

6.4.16 Dual GbESMs with four port SMLT to dual passport 8600s

This last configuration best meets the needs of a mission-critical data center where 24/7 uptime is a requirement. Several advantages exist in this scenario. The first advantage is that by using SMLT there is no need to use Spanning Tree to block loops since SMLT inherently removes any loops. Thus all ports are forwarding all the time for maximum throughput. The second advantage is subsecond failover; since the two Passports share a common forwarding database, either Passport 8600 could fail or any link pairs could fail, and failover would occur in less than a second. If a GbESM fails, resiliency would be provided by the redundancy mechanism of the BladeCenter servers.

This configuration is cabled much like the earlier dual BayStack configuration where we had dual GbESMs connected to dual BayStack switches with a series of trunks. However, this configuration has one critical difference: the SMLT on the Passport 8600s is transparent to the GbESM switches, so we do not need to configure multiple trunks on each of the GbESMs. We only need to configure a single 4-port trunk on each GbESM, as shown in Figure 6-48.

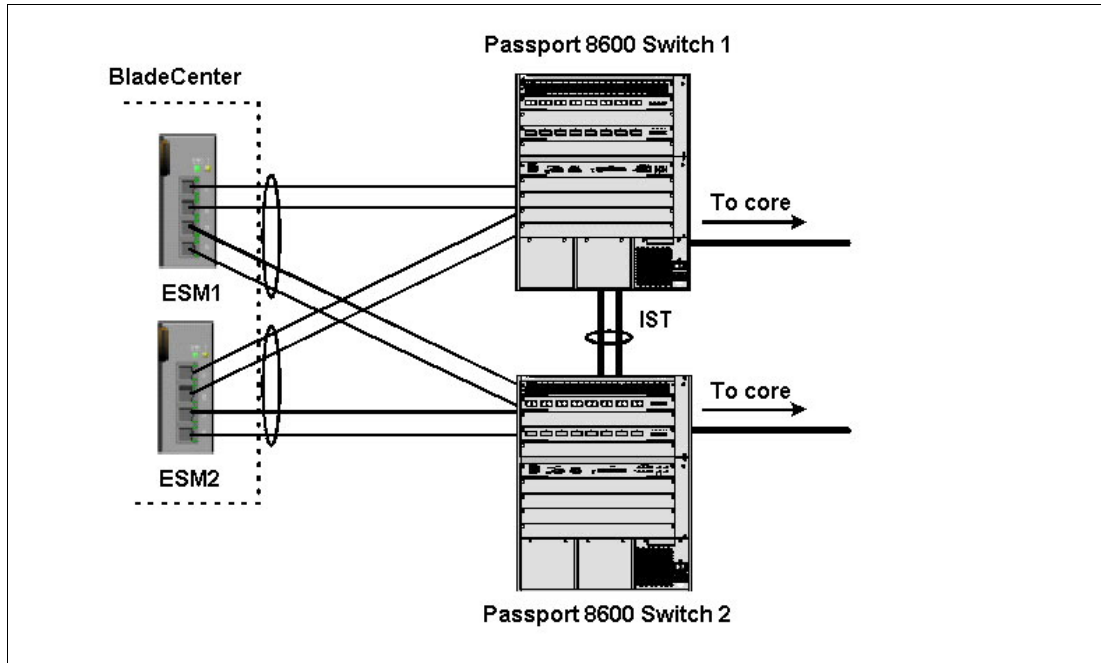


Figure 6-48 Dual GbESMs with four port SMLT to dual Passport 8600s

Configuring the GbESM

Table 6-7 on page 205 walks through the actions required to configure the GbESM for this example. The following assumptions have been made for this example:

- ▶ The user is already logged in to the GbESM with admin level access.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 1 are connected to ports 8/1 and 8/2 on Passport 8600 Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are connected to ports 2/1 and 2/2 on Passport 8600 Switch 2.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 2 are connected to ports 8/3 and 8/4 on Passport 8600 Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are connected to ports 2/3 and 2/4 on Passport 8600 Switch 2.
- ▶ The same VLAN setup is being used on both GbESMs. (This is not a requirement; the second GbESM could be configured to use different VLANs.)
- ▶ Commands are performed in the sequence shown.
- ▶ The Nortel switches are Passport 8600s with 1000Base-T ports. Autonegotiate will result in ports operating at 1000Base-T.
- ▶ All ports remain in Spanning Tree Group 1 and STG 1 has been turned off.
- ▶ No cables are connected between any switches.
- ▶ The GbESM has been reset to factory defaults as shown in Example 6-4 on page 168.

The steps in Table 6-7 apply to configuring the first GbESM. Essentially the same steps are used to configure the second GbESM as well.

Table 6-7 Configuring the GbESM

Description and comments	Instructions
<p>Step 1 - Configure internal ports.</p> <p>This places the desired blade server ports into the desired VLANs. The VLANs must be enabled before the ports can be added to them. Since the blade servers do not support tagged frames, the internal ports must not be tagged.</p>	<p>Enter the command <code>/cfg/12/stg 1/off</code> to turn off Spanning Tree.</p> <p>Enter the command: <code>/cfg/12/vlan 5/ena</code> to enable VLAN 5.</p> <p>Enter the command: <code>/cfg/12/vlan 5/add INT2</code> to add internal port 2 to VLAN 5. This corresponds to BladeServer bay 2. Answer y to the prompt to confirm changing the PVID on port INT2.</p> <p>Enter the command: <code>/cfg/12/vlan 10/ena</code> to enable VLAN 10.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add INT3</code> to add internal port 3 to VLAN 10. Repeat this command for INT4.</p> <p>All other PVIDs should be set for 1, with the exception of the management ports. The internal ports should all be untagged ports. To check this enter the command: <code>/info/port</code>. This lists the ports and tagging status.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 2 - Configure tagging on external ports.</p> <p>This enables tagging on the external ports so that they can be members of VLANS 1, 5, and 10 simultaneously.</p>	<p>Enter the command: <code>/cfg/port EXT1/tag e</code> to enable tagging on external port 1. Repeat this command for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 3 - Configuring the VLANS for the external ports.</p> <p>This puts the ports into VLANs 1, 5, and 10. This will allow external ports 1 through 4 to carry traffic for VLANS 1, 5, and 10. The ports do not need to be added to VLAN 1 because they are in it by default.</p>	<p>Enter the command: <code>/cfg/12/vlan 5/add EXT1</code> to add external port 1 to VLAN 5. Repeat for external ports 2 through 4.</p> <p>Enter the command: <code>/cfg/12/vlan 10/add EXT1</code> to add external port 1 to VLAN 10. Repeat for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>
<p>Step 4 - Configure external ports and multi-link trunking.</p> <p>We only need a single trunk on each switch. The SMLT configuration on the Passport 8600s is transparent to the GbESMs.</p>	<p>Enter the command <code>/cfg/12/trunk 1/add EXT1</code> to add port EXT1 to Trunk Group 1.</p> <p>Repeat this step for external ports 2 through 4.</p> <p>Enter apply to apply the new configuration. Enter save to save the configuration to NVRAM.</p>

Configuring the Passport 8600

This section walks through the sequence of actions required to configure the two Passport 8600 switches for this example. The following assumptions have been made for this example:

- ▶ Ports 1/1, 1/2, 1/3, 1/4 on Passport 8600 Switch 1 are connected to ports 1/1, 1/2, 1/3, 1/4 on Passport 8600 Switch 2. These links will be the IST trunk.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 1 are connected to ports 8/1 and 8/2 on Passport 8600 Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are connected to ports 2/1 and 2/2 on Passport 8600 Switch 2.
- ▶ Ports EXT1 and EXT2 on the GbESM in Switch Module Bay 2 are connected to ports 8/3 and 8/4 on Passport 8600 Switch 1.
- ▶ Ports EXT3 and EXT4 on the GbESM in Switch Module Bay 1 are connected to ports 2/3 and 2/4 on Passport 8600 Switch 2.

- Commands are performed in the sequence shown. Be sure to note which switch the commands are being applied to.

Step 1: Configure the IST on Passport 8600 Switch 1

To do this you will need to first configure a separate VLAN for IST communications. This VLAN and its subnet should remain local to the IST link.

An MLT group needs to be created, tagging enabled, and specific ports added. In this example, an 8608GT blade, which has eight 1000Base-T ports, is installed in slot 1 on Passport 8600 Switch 1 (the same blade is also installed in slot 1 of Passport 8600 Switch 2). Note: The first four ports of slot 1 on both Passport 8600s are connected to form the IST. This is strictly for purposes of keeping this example easy to follow. In a real network design, it is more robust to place IST links on different blades in the chassis to be able to survive an entire blade failure.

After this, VLANs need to be added and the MLT designated as an IST. Finally, the CP-limit feature is disabled on the IST links. These commands are shown in Figure 6-49.

```
PP8600-1:5#
PP8600-1:5# config vlan 1000 create
PP8600-1:5/config/vlan/1000/create# byport 1
PP8600-1:5/config/vlan/1000/create# name IST
PP8600-1:5/config/vlan/1000# ip create 10.1.1.1/24
PP8600-1:5/config/vlan/1000/create# top
PP8600-1:5#
PP8600-1:5# config mlt 1
PP8600-1:5/config/mlt/1# create
PP8600-1:5/config/mlt/1# name "IST"
PP8600-1:5/config/mlt/1# perform-tagging enable
PP8600-1:5/config/mlt/1# add ports 1/1-1/4
PP8600-1:5/config/mlt/1# add vlan 1000
PP8600-1:5/config/mlt/1# add vlan 1
PP8600-1:5/config/mlt/1# add vlan 5
PP8600-1:5/config/mlt/1# add vlan 10
PP8600-1:5/config/mlt/1# ist
PP8600-1:5/config/mlt/1/ist# create ip 10.1.1.2 vlan-id 1000

INFO : IST is created and enabled.
       The spanning tree protocol is disabled on the port(s) with IST enabled!

PP8600-1:5/config/mlt/1/ist# top
PP8600-1:5# config ethernet 1/1-1/4 cp-limit disable
PP8600-1:5#
```

Figure 6-49 Console interface: Configuring IST on switch 1

Step 2: Configure the IST on Passport 8600 Switch 2

This step is much like the first step except you will now be configuring the IST on the second switch. The same notes and caveats from step 1 apply to this step as well. The commands to configure Passport 8600 Switch 2 are shown in Figure 6-50.

```

PP8600-2:3#
PP8600-2:3# conf vlan 1000 create
PP8600-2:3/config/vlan/1000/create# byport 1
PP8600-2:3/config/vlan/1000/create# name IST
PP8600-2:3/config/vlan/1000/create# ip create 10.1.1.2/24
PP8600-2:3/config/vlan/1000/create# top
PP8600-2:3#
PP8600-2:3# conf mlt 1
PP8600-2:3/config/mlt/1# create
PP8600-2:3/config/mlt/1# name "IST"
PP8600-2:3/config/mlt/1# perform-tagging enable
PP8600-2:3/config/mlt/1# add ports 1/1-1/4
PP8600-2:3/config/mlt/1# add vlan 1000
PP8600-2:3/config/mlt/1# add vlan 1
PP8600-2:3/config/mlt/1# add vlan 5
PP8600-2:3/config/mlt/1# add vlan 10
PP8600-2:3/config/mlt/1# ist
PP8600-2:3/config/mlt/1/ist# create ip 10.1.1.1 vlan-id 1000

INFO : IST is created and enabled.
       The spanning tree protocol is disabled on the port(s) with IST enabled!

PP8600-2:3/config/mlt/1/ist# top
PP8600-2:3# config ethernet 1/1-1/4 cp-limit disable
PP8600-2:3#

```

Figure 6-50 Console interface: Configuring IST on switch 2

Step 3: Configure the SMLT trunks on Passport 8600 Switch 1

To do this you will need to create two MLT groups, enable tagging, and add specific ports. In this example, the 8608GT blade, which has eight 1000Base-T ports, is installed in slot 8 of Passport 8600 Switch 1. The first two ports, 8/1 and 8/2 are part of the first SMLT group, and the next two ports, 8/3 and 8/4 are part of the second SMLT group. After ports are added, VLANs are added to the SMLT groups. Finally, the trunk is designated as an SMLT group. The configuration is performed as shown in Figure 6-51.

```

PP8600-1:5#
PP8600-1:5# config mlt 2
PP8600-1:5/config/mlt/2# create
PP8600-1:5/config/mlt/2# name "ESM1 Trunk"
PP8600-1:5/config/mlt/2# perform-tagging enable
PP8600-1:5/config/mlt/2# add ports 8/1-8/2
PP8600-1:5/config/mlt/2# add vlan 1
PP8600-1:5/config/mlt/2# add vlan 5
PP8600-1:5/config/mlt/2# add vlan 10
PP8600-1:5/config/mlt/2# smlt create smlt-id 1

INFO : The spanning tree protocol is disabled on the port(s)
       while configure the port(s) with SMLT!

PP8600-1:5/config/mlt/2# top
PP8600-1:5#
PP8600-1:5# config mlt 3
PP8600-1:5/config/mlt/3# create
PP8600-1:5/config/mlt/3# name "ESM2 Trunk"
PP8600-1:5/config/mlt/3# perform-tagging enable
PP8600-1:5/config/mlt/3# add ports 8/3-8/4
PP8600-1:5/config/mlt/3# add vlan 1
PP8600-1:5/config/mlt/3# add vlan 5
PP8600-1:5/config/mlt/3# add vlan 10
PP8600-1:5/config/mlt/3# smlt create smlt-id 2

INFO : The spanning tree protocol is disabled on the port(s)
       while configure the port(s) with SMLT!

PP8600-1:5/config/mlt/3# top
PP8600-1:5#

```

Figure 6-51 Console interface: Configuring SMLT on switch 1

Step 4: Configure the SMLT trunks on Passport 8600 Switch 2

To do this you will need to create two MLT groups, enable tagging, and add specific ports. In this example, the 8608GT blade, which has eight 1000Base-T ports, is installed in slot 2 of Passport 8600 Switch 2. The first two ports, 2/1 and 2/2 are part of the first SMLT group, and the next two ports, 2/3 and 2/4 are part of the second SMLT group. After ports are added, VLANs are added to the SMLT groups. Finally, the trunk is designated as an SMLT group. The configuration is performed as shown in Figure 6-52.

```
PP8600-2:3#
PP8600-2:3# config mlt 2
PP8600-2:3/config/mlt/2# create
PP8600-2:3/config/mlt/2# name "ESM1 Trunk"
PP8600-2:3/config/mlt/2# perform-tagging enable
PP8600-2:3/config/mlt/2# add ports 2/1-2/2
PP8600-2:3/config/mlt/2# add vlan 1
PP8600-2:3/config/mlt/2# add vlan 5
PP8600-2:3/config/mlt/2# add vlan 10
PP8600-2:3/config/mlt/2# smlt create smlt-id 1

INFO : The spanning tree protocol is disabled on the port(s)
       while configure the port(s) with SMLT!

PP8600-2:3/config/mlt/2# top
PP8600-2:3#
PP8600-2:3# config mlt 3
PP8600-2:3/config/mlt/3# create
PP8600-2:3/config/mlt/3# name "ESM2 Trunk"
PP8600-2:3/config/mlt/3# perform-tagging enable
PP8600-2:3/config/mlt/3# add ports 2/3-2/4
PP8600-2:3/config/mlt/3# add vlan 1
PP8600-2:3/config/mlt/3# add vlan 5
PP8600-2:3/config/mlt/3# add vlan 10
PP8600-2:3/config/mlt/3# smlt create smlt-id 2

INFO : The spanning tree protocol is disabled on the port(s)
       while configure the port(s) with SMLT!

PP8600-2:3/config/mlt/3# top
PP8600-2:3#
```

Figure 6-52 Console interface: Configuring SMLT on switch 2

6.4.17 Validation of GbESM configuration

To validate the setup of the GbESM, refer to “Step 5: Verify the configuration” on page 137.

6.4.18 Validation of Passport 8600 configuration

This section contains some steps that can be taken to quickly verify the configuration of the SMLT setup of the two Passport 8600s. The same basic verification steps discussed in the previous scenario apply here as well. Note that Spanning Tree was disabled on all SMLT and IST trunk groups, so it may not be necessary to check for forwarding state on the Passport 8600.

Check port status, speed, and duplex settings:

```
show ports info name <portlist>
```

If autonegotiate is enabled, the information shown will be the negotiated speed/duplex settings. For sample output refer back to the previous scenario.

Check the status of the IST trunk using the following command:

```
show mlt ist info
```

This will confirm that the IST trunk is up and receiving heartbeat messages. Sample output of what you should see is shown in Figure 6-53.

```

PP8600-1:5#
PP8600-1:5# show mlt ist info
=====
                                Mlt IST Info
=====
MLT   IP          VLAN   ENABLE   IST
ID    ADDRESS      ID     IST      STATUS
-----
 1    10.1.1.2      1000   true     up
PP8600-1:5#

```

Figure 6-53 Command output

More detailed information on the IST will be shown with the following command:

```
show mlt ist stat
```

The status of the SMLT can be verified using this command:

```
show mlt smlt info
```

The “current type” field should match the “admin type” field. In the case of SMLT, both should have a value of “smlt”. A working configuration is shown in Figure 6-54.

```

PP8600-1:5#
PP8600-1:5# show mlt smlt info
=====
                                Mlt SMLT Info
=====
MLT   SMLT   ADMIN   CURRENT
ID    ID      TYPE    TYPE
-----
 2    1       smlt    smlt
 3    2       smlt    smlt
PP8600-1:5#

```

Figure 6-54 Command output: Working configuration

Other steps can be taken to verify the proper operation. For example, the Passport 8600's out-of-band management interface can be connected to a port that is a member of VLAN 1 (assuming the GbESM's management VLAN is 1). If this is done, it is possible to ping from Passport 8600 to GbESM. To ping from the Passport 8600 management interface to GbESM use the following command:

```
ping <ip_addr>
```

Also, a Passport 8600 port can be assigned to a particular VLAN so that a workstation can be connected to the port in order to ping the corresponding server. To fully ensure that the SMLT is working under failure conditions, start a continuous ping from workstation to BladeCenter server, and begin disconnecting cables between Passport 8600 switches and an GbESM. As long as at least one IST link is up, and as long as at least one link between GbESM and either of the two Passports is up, pings should continue to get through. At most one ping should drop, but in many cases not even a single ping will drop.

Not all steps must be performed to verify proper configuration; these are just possible tests that can be used.

6.5 Troubleshooting GbESM connections to Nortel Networks devices

Troubleshooting a network can range from a simple problem resolution to a lengthy process. Covering all possible steps to isolate and fix problems is beyond the scope of this document. The purpose of this section is to give some helpful tips to solving some of the more obvious problems. There are numerous tools and documents available to aid in troubleshooting both IBM and Nortel Networks equipment. This document should not be treated as a substitute for training and experience. It is recommended that problems that are not easily resolved by the suggestions contained here be directed to individuals with experience and training in the areas that are experiencing issues.

Issue: Link is down

Link lights are off, or administration interface to either device is showing the link as down.

Resolution:

Verify that autonegotiate is enabled on both ends of the link. If it is not, or if manual setting of speed/duplex is required for some reason, verify that both sides are configured the same way and that a crossover cable is being used. Some Nortel switches such as the Alteon 180 series are required to use a crossover cable for switch-to-switch connections other than to the MIDI/MID-X auto-sensing GbESM. If you are not using a crossover cable with this type of switch, you must have autonegotiation enabled on the GbESM. Also check to make sure that ports are not administratively disabled on either device.

Issue: Traffic for a specific VLAN or all VLANs will not pass between GbESM and Nortel device

Links are up, but pings between a workstation and BladeCenter server fail.

Resolution:

Ensure proper VLAN configurations. Make sure that VLANs are added to the correct trunk groups. Make sure that tagging is enabled for all VLANs on the GbESM and on the Nortel device. Make sure that routing is configured properly.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this Redpaper.

IBM Redbooks

For information on ordering these publications, see “How to get IBM Redbooks” on page 212. Note that some of the documents referenced here may be available in softcopy only.

- ▶ *The Cutting Edge: IBM eServer BladeCenter*, REDP3581
- ▶ *IBM eServer BladeCenter Systems Management*, REDP3582
- ▶ *Deploying Citrix MetaFrame on IBM eServer BladeCenter*, REDP3583
- ▶ *Deploying Lotus Domino on IBM eServer BladeCenter*, REDP3584
- ▶ *Deploying Microsoft Exchange on IBM eServer BladeCenter*, REDP3585
- ▶ *Deploying Samba on IBM eServer BladeCenter*, REDP3595
- ▶ *Deploying Apache on IBM eServer BladeCenter*, REDP3588

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM @server BladeCenter Type 8677 Installation and User's Guide*
- ▶ *IBM 4-Port GB Ethernet Switch Module for BladeCenter Installation Guide*
- ▶ *IBM @server BladeCenter Management Module User's Guide*
- ▶ *Alteon OS 20.0 Application Guide*, Part 215654-A (Nortel)
- ▶ *Alteon OS 20.0 Command Reference*, Part 215655-A (Nortel)

Online resources

These Web sites and URLs are also relevant as further information sources:

- ▶ IBM eServer xSeries® products
<http://www.ibm.com/eserver/xseries/>
- ▶ Documentation for Cisco switches
<http://www.cisco.com/en/US/products/hw/switches/>
- ▶ Documentation for Cisco routers
<http://www.cisco.com/en/US/products/hw/routers/>
- ▶ IEEE LAN specifications online
<http://standards.ieee.org/reading/ieee/std/lanman/>
- ▶ Documentation for troubleshooting Spanning Tree issues
http://www.cisco.com/en/US/tech/tk389/tk621/technologies_tech_note09186a00800951ac.shtml

- ▶ Layer 2 technology documents
<http://www.cisco.com/cgi-bin/Support/browse/index.pl?i=Technologies&f=1324>
- ▶ BayStack 380-24T Product Literature
http://www.nortelnetworks.com/products/02/bstk/switches/baystack_380/doclib.html
- ▶ Technical documentation for BayStack 380-24T
<http://www130.nortelnetworks.com/cgi-bin/eserv/cs/main.jsp?cscat=documentation&tranProduct=11501>
- ▶ Passport 8600 Product Literature
<http://www.nortelnetworks.com/products/01/passport/lan/doclib.html#8600>
- ▶ Technical documentation for Passport 8600
<http://www130nortelnetworks.com/cgi-bin/eserv/cs/main.jsp?cscat=documentation&tranProduct=9015>
- ▶ IEEE LAN specifications online
<http://standards.ieee.org/reading/ieee/std/lanman/>
- ▶ Cisco Internetworking Technology Handbook
http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/index.htm
- ▶ Cisco datacenter architectures
http://www.cisco.com/en/US/netso1/ns110/ns53/net_solution_home.html
- ▶ Cisco Connection Online
http://www.cisco.com/en/US/partner/netso1/ns110/ns53/ns224/networking_solutions_packages_list.html
- ▶ xSeries products
<http://www.ibm.com/eserver/xseries/>

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Index

Numerics

- 10/100/1000 Mbps connections 8
- 1000Base-T 9, 181, 193
- 100BASE-TX 9
- 100-ohm STP 9
- 10BASE-T 9
- 4 Gigabit 72
- 802.1D Spanning Tree support 8
- 802.1P/Q MIB 9
- 802.1Q 77, 93, 111
 - tagging 78
- 802.1Q Tagged 8
- 802.1Q trunking 137
 - Cisco configuration 112
 - configure 118, 145
 - configuring 127
- 802.3ad 93, 161
- 8677 2

A

- access layer 87–88
- Advanced Switch Management 20
- aggregated link
 - port cost 127
 - status 123, 130, 150
- aggregated links 94, 141
- aggregation
 - verifying 120
- aggregation element 89
- aggregation layer 87–88
- aggregation level 88
 - recommended Cisco hardware 89
- aging time 34–35
- ANSI/IEEE 802.3 NWay auto-negotiation 9
- application health checking 32
- application-specific integrated circuits (ASICs) 34
- Apply vs Save 13
- Architectural Reference 86
- authentication method 28
- auto-MDI/MDI-X 161, 163
- autonegotiate 92, 181, 187, 193
- auto-negotiation 92
- autonegotiation 163
- autopolarity 161
- autosensing 8

B

- backbone 8
- BackboneFast 86
- bandwidth 5
- base configuration 100, 166
 - Cisco switches 100
 - for Cisco IOS based switch 102

- basic configuration example with Cisco 104
- BayStack 380-24T 159–160, 164, 175, 181
- BayStack 5510 160
- BBI 20–21, 97
 - usage tips 21
- blade server 8, 61
- BladeCenter 159
- blocking specific ports 94
- blocking state 37
- boot PROM 90, 160
- bootstrap protocol (BOOTP) 9
- BPDU 36, 95
- bridge aging time 83
- bridge forward delay 84
- bridge hello time 83–84
- bridge maximum age 83–84
- bridge priority 83–84
- Bridge Protocol Data Units 36
- broadcast domains 40
- broadcast frames 39
- Browser Based Interface 20–21

C

- cache redirection 32
- Catalyst 3550 89
- Catalyst 3750 89
- Catalyst 4000 89
- CatOS 90
 - resetting 90
- CDP 86
- chassis 2, 7, 12
- Cisco
 - configuration 106
 - configuration verification 108, 148
 - disable a range of ports 99
 - enable ports 99
 - features not in Nortel switch 86
 - IOS
 - disable port 99
 - show speed and duplex 123, 130
 - show trunking status 130
 - traditional architecture 87
 - verification 130
 - verify configuration and operation 139
 - verifying configuration 114
 - verifying operation 122
- Cisco Catalyst 4006 89–90
- Cisco Catalyst 6500 89
- Cisco Discovery Protocol 86
- Cisco DTP 94
- Cisco EtherChannel 86
- Cisco ISL
 - support 93
- CLI

- configuring the switch 22
- collision 31
- collision domains 34
- collisions 34
- command
 - /boot/conf 19
 - /boot/conf active 25
 - /boot/gtimg 26
 - /boot/reset 19, 26
 - /cfg 25
 - /cfg/l2/stg 105, 111, 135, 143
 - /cfg/l2/stg 1/off 125
 - /cfg/l2/stg/brg
 - syntax 83
 - /cfg/l2/stg/port
 - syntax 84
 - /cfg/l2/trunk 74, 117, 144
 - /cfg/l2/trunk/ena 73
 - /cfg/l2/vlan 77, 105, 111, 117, 125, 135, 143–144
 - syntax 82
 - /cfg/l3/ip/if 24
 - /cfg/port 72, 79, 111, 136, 143
 - /cfg/port /tag enable 79
 - /cfg/port EXT1/tag e 125
 - /cfg/ptcfg 26
 - /cfg/sys/access/user 29
 - /info/l2/port 77
 - /info/L2/trunk 82
 - /info/l2/trunk 75
 - /info/l2/vlan 76
 - /info/l3/ip 23
 - /info/port 105, 125, 135
 - /oper/port 23
 - add 25
 - addr 24
 - apply 24, 99
 - boot 22, 26
 - boot/conf factory 91
 - cfg 22
 - cfg/port 99, 105, 125
 - channel-group 1 mode on 145
 - clear config all 90
 - Configure t 118, 137, 145
 - diff 23
 - ena 24
 - enable 117
 - erase cat4000_flash
 - 91
 - help 23
 - info/l2/stg 108, 114, 120, 129, 139, 147
 - info/l2/trunk 120, 129, 148
 - info/link 107, 113, 119, 128, 137, 146
 - info/port 107, 113, 119, 128, 138, 147
 - int port-channel1 146
 - int range FastEthernet 2/3 - 4 145
 - int range FastEthernet 2/3 - 6 118
 - l2 25
 - maint 22
 - mask 24
 - oper 22
 - reload 91
 - reset 91
 - save 24
 - set port channel 2/3-4 1 145
 - set port channel 2/3-4 mode on 145
 - set port disable 2/3 99
 - set spantree portcost 2/3 100 137
 - set spantree portcost 2/3-4 100 127
 - set spantree portcost 2/3-6 100 145
 - set trunk 106
 - set trunk 2/3 nonegotiate 118, 127, 137, 145
 - sh span root 96
 - show channel traffic 123, 130, 150–151, 153
 - show config 115, 123, 130, 139, 149
 - show etherchannel summary 123, 130, 150, 153
 - show int FastEthernet 2/3 status 109, 115, 149
 - show int FastEthernet 2/3 trunk 109, 115, 149
 - show int Port-channel 1 trunk 123
 - show interface fastethernet 2/3 status 153
 - show interface fastEthernet 2/3 trunk 153
 - show port channel 123, 130, 150–152
 - show port status 2/1 109, 115
 - show port status 2/3 149, 151–152
 - show port trunk 2/1 109, 115
 - show port trunk 2/3 123, 149, 151–152
 - show run 115, 123, 130, 149
 - show spanning int FastEthernet 2/3 150
 - show spanning int po1 150
 - show spanning-tree inter fastEthernet 2/3 154
 - show spanning-tree interface port-channel 1 154
 - show spantree 2/3-6 150, 152–153
 - spanning-tree cost 100 127, 137, 145–146
 - stats 22
 - stats/port ext1/if 120, 122
 - trunk 25
 - vlan 24
 - write erase 91
 - write mem 118, 137, 146
- command line interface 20, 74
- configuration
 - best practice 13
 - capture 26
 - verficiation 118
 - verification 112
 - verifying 127, 137, 146
- Configuration changes
 - applying 13
- configuration sequence 97
- Configure 802.1Q trunking 177, 183, 188, 194, 201, 205
- configure MLT 178
- Configure PVIDs 105, 111, 117, 125, 135, 143, 176, 182, 188, 194, 200, 205
- configuring the BayStack 380-24T 177, 188, 195
- configuring the ESM 175, 181, 187, 193
- copper ports 8
- core layer 87
- CRC error 31, 35
- cross-over 157
- cross-over cable 13, 74
 - recommended 92

- crossover cable 163
- current configuration
 - capture 26
- Customer Premise Equipment (CPE) 161
- cut-through 35

D

- data traffic 6
- datacenter 87
 - recomendations 93
- default addresses 3
- default gateway 6
- default password
 - Management Module 100
- design scenarios 175
- destination machine 33
- DHCP server 12
- disconnect procedure 98
- distribution layer 87
- drivers 6
- DTP 86
- Dual GbESM configuration 131
- duplex
 - setting 125
- dynamic host configuration protocol (DHCP) 9
- Dynamic Trunking Protocol 86, 94

E

- egress 7
- egress packets 78
- EIA/TIA-568 100-ohm STP 9
- EIA/TIA-568B 100-ohm STP 9
- enable external Ethernet ports 16
- ESM 175
- ESM IP address 17
- ESM management session 17
- ESM telnet interface 19
- ESM Web interface 19
- EtherChannel 72, 93
- Ethernet activity 5
- Ethernet bridge 33
- Ethernet connectivity 2
- Ethernet daughter card 2, 4
- Ethernet link 5
- Ethernet port 12
- Ethernet protocol 31
- Ethernet switch error 5
- Ethernet Switch Module (ESM) 90
- Ethernet switches 34
- Ethernet switching 2
- example
 - connected to 2 Cisco switches 110
- excessive flooding 35
- Ext1 2
- external copper ports 8
- external Ethernet interface 5
- External Port
 - configuration 136
 - configure tagging 105

- External Ports
 - configure tagging 125
 - configuring 117, 125, 144
 - setting VLANs 111
 - tagging 111
- external ports
 - configure tagging 136
 - enable for management 3

F

- factory default setting 75
- factory defaults
 - resetting 18
- Factory Reset 100
- factory reset 166
 - preserve IP address 17
- Fast Learning 163
- filtering data frames 34
- firmware 6, 11
 - files 26
 - management module 11
 - upgrade 25
- Firmware version 90
- firmware version 90, 160
- flash memory 8
- floods 35
- flow control 61
- forwarding database aging time 84
- forwarding state 37
- forwarding table 33
 - static entries 35
- forwarding table age time 8
- four port static aggregation 116
- frame 31
 - header 79
- frames
 - tagged 79
- front-end layer 87

G

- GbESM
 - resetting 91
- general switch information 6
- getting started 12
- gigabit aggregation layer switch 159
- GVRP 86

H

- hardware configuration in this book 89
- Hello BPDU 37
- high availability 133
- high redundancy 133
- http
 - [//www.cisco.com/en/US/partner/net-sol/ns110/ns53/ns224/networking_solutions_packages_list.html](http://www.cisco.com/en/US/partner/net-sol/ns110/ns53/ns224/networking_solutions_packages_list.html) 212
- HTTP Web interface 13
- HyperTerminal

text capture 27

I

- IBM BladeCenter 4-Port Ethernet Switch Module 1
- IBM BladeCenter 8677 160
- IBM Director 13
- identification label 2
- IEEE 802.1D 35–36
- IEEE 802.1Q 7, 35, 79
- IEEE 802.1Q Tagged VLAN 9
- IEEE 802.3 10BASE-T Ethernet 9
- IEEE 802.3ab 1000BASE-T 9
- IEEE 802.3u 100BASE-TX Fast Ethernet 9
- IEEE 802.3x 61
- IEEE 802.3x Full-duplex Flow Control 9
- IEEE 802.3z Gigabit Ethernet 9
- IEEE LACP 74
- in-band 8
- ingress 7
- initial recommended switch settings 14
- INT15 80
- INT16 80
- Interface 128 2
- interface MIB 9
- internal full-duplex 10/100 Mbps ports 8
- internal full-duplex gigabit ports 8
- Inter-Switch Link 93
- IP address
 - change only with management module 17
 - how to switch 17
- IP Interface
 - configuring with CLI 24
- ISL 86
- IST (Inter Switch Trunk) 161

J

- Java 13
- JavaScript 1.2 13
- jumbo frames 76, 82

L

- L2 method 162
- latency 35
- Layer 2
 - guidelines 71
 - layer 2 104
 - Layer 2 network 87, 97
 - Layer 3 88
 - layer 3 network 97
- LED 4–5
- link
 - access 94
 - access vs trunk 94
- Link Aggregation 23, 175
 - Spanning Tree considerations 36
- link aggregation 5, 8, 72, 93, 97–98
 - Cisco hint/tip 106, 112, 118, 126, 136, 144
 - configure 145

- configuring 24
- Link Aggregation Group 93
- link aggregation protocols 161
- link trunks 186
- link utilization 150
- logical network segment 7
- logical network topology 7

M

- MAC address 3, 31, 33
- MAC addresses 7
- management information base (MIB) 9
- Management Module 2, 8, 12, 100
 - default IP 14
 - default password 100
 - NVRAM 18
- management module Web interface 6, 19
- management network 13
- management ports 80
- management station 12
- maximum throughput 116
- MDI/MDI-X 92
- MDI-X port 13
- media access control (MAC) 8
- media dependent interface (MDI) 13
- mini-RMON MIB 9
- misconfiguration 163
- MM 6
- monitor port
 - trunk member 73
- Mono Spanning Tree 95
- multi-link trunking (MLT) 161
- multiple VLANs
 - single connection 93

N

- Native IOS 90
 - resetting 91
- Native VLAN 94
 - definition 93
- network architecture 97
- network design 97, 181
- network loop 94
- network management 8
- network monitoring 6
- network packets 7
- nonegotiate 94
- Nortel Networks 159
- NVRAM 18, 25, 106, 111, 117, 127, 136–137, 144, 146
 - Cisco 112

O

- OSI 88
- OSI reference model 31–32
 - data link layer 33
 - layers 32
 - physical layer 33
- out of band port 173

out-of-band 8

P

- packet forwarding 7
- packet header 7
- PAgP 74, 86
- Passport 8300 159–160
- Passport 8600 159–160, 164, 173
- password
 - change 29
 - default 20
- path cost 37
- path costs 94, 139
- performance
 - maximum configuration 141
- performance and availability compromise 141
- port
 - configurable default VLAN number 77
 - default configuration settings 77
 - Ext1 2
 - path cost 36
 - setting speed 125
 - showing current state 23
 - trunk vs access 93
- Port Aggregation 192
- Port Aggregation Protocol 86
- port aliases 75
- port cost 95–96, 110, 124, 134, 141, 145
 - changing 96
 - configure 127, 137
 - setting 146
- port mirroring 80
- port numbers 75
- port path cost 84
- port priority 84
- port statistics 6
- Port VLAN Identifier 79
- Port VLAN identifier 79
- port-mirroring 73
- power-on self-test (POST) 5
- Preserve IP address 17
- primary link 7
- priority queues 8
- processor blades 2
- production network 13
- production switch
 - considerations 96
- protocols 8
- PVID 76–79, 105, 125, 143
 - changing 77
 - viewing configuration 77
- PVID 4095 80
- PVST+ 95

Q

Quality of Service (QoS) 162

R

- RADIUS 28
- random-access memory (RAM) 8
- Red Hat Linux 160
- Redbooks Web site 212
 - Contact us x
- redundancy 133
 - test 150
- redundant switch 181
- Remote Access Server 28
- Remote Authentication Dial-in User Service 28
- remote management 6
- remote monitoring (RMON) 9
- resetting systems 90
- Restore Factory Defaults 18
- revert flash 13
- root bridge 36, 84, 98
- root cost 95
- root ID 37
- root switch 95, 98, 110, 139

S

- Save vs Apply 13
- secondary link 7
- segment 31
- serial number 2
- server load balancing 32
- servers 8
- set intervals 18
- shared media 31, 34
- show speed and duplex 115
- show trunking status 115
 - Cisco 123
- simple network management protocol (SNMP) 8
- SMLT aggregation switch 161
- SMLT Client 161
- SNMP 6
- SNMP traps 17
- Source learning 34
- source learning 34
- source machine 33
- Spanning Tree 98, 109, 124, 127, 132–133, 139, 141, 162, 186
 - configure port cost 112
 - considerations 104
 - guidelines 94
 - recommendation 94
 - show status 108, 131
 - status 115, 140, 150
 - turning off 105, 125
- spanning tree 36, 72, 131
- spanning tree bridge priority 172
- Spanning Tree Group 80
- Spanning Tree Group index 82
- Spanning Tree Groups 162
- Spanning Tree loops 93–94
- Spanning Tree Protocol 7, 9, 162
- Split-MLT (SMLT) 161
- SSH 6

- SSL appliances 32
- SSL encryption/decryption 32
- standards 8
- Start Telnet/Web Session 20
- startup-config 91
- static link aggregation 118
- stations 8
- statistically distributed
 - network traffic 72
- status LEDs 4
- STG
 - individual port parameters 84
- STG root bridge 84
- store-and-forward 8, 35
- STP 88
- STP bridge protocol 35
- straight cable 163
- subnet
 - default 13
- switch forwarding table 34
- switch information panel 21
- switch maintenance 6
- switch management 6, 8
- switch module 12
- switch parameters 6
- switch TCP/IP address 6

T

- Tagged frame 79
- tagged frames 79
- tagging 7, 77
 - enabling for a port 79
- telnet interface 6
- telnet remote console 8
- terminal emulator 27
- Test redundancy 140
- TFTP 25
- TFTP server 26
- third-party device 74
- transmission method 8
- traps 6
- trivial file transfer protocol (TFTP) 8
- troubleshooting 156
 - port will not come up 156
- trunk
 - definition 93
 - spanning tree considerations 36
- trunk group 74, 148
 - before you configure 72
 - configuration commands 75
 - configuration rules 73
 - defining 117
 - fault tolerant 72
 - traffic distribution 72
 - viewing information 75
- trunk groups 72
 - best performance 72
 - defining 144
- trunk members

- spanning tree 73
- trunks
 - showing configuration 82

U

- UDLD 86
- UniDirectional Link Detection 86
- unshielded twisted pair (UTP) 13
- Untagged frame 79
- untagged frames 78
- untagged VLAN 93
- untagging 7
- UplinkFast 86
- user accounts 27
- User guides 6
- userID
 - default 97
- userid
 - default 20
- UTP Category 3 9
- UTP Category 4 9
- UTP Category 5 9
- UTP Category 5e 9

V

- verifying configuration 107
- VID 79
- virtual link 72
- virtual local area network (VLAN) 7–8, 39
- VLAN 40, 164, 174
 - 4095 2
 - 4095 reserved 76
 - automatic creation 76
 - belonging to more than one 78
 - carrying 94
 - configuration rules 80
 - default 76
 - jumbo frames 76
 - membership 82
 - number per switch 76
 - port in multiple 79
 - trunk group 72
 - verification 124
 - verify configurations 107
 - viewing configuration 76
- VLAN 1 98
- VLAN 10 98
- VLAN 5 98
- VLAN identifier 79
- VLAN Management Policy Server 86
- VLAN Tagging 79
- VLAN tags 7
- VLAN Trunk 81
- VLAN Trunking 86
- VLAN Trunking Protocol 86
- vlan.dat 91
- VMPS 86
- VTP 86

W

Web browser 6, 13, 171

Web-based management 8

WS-X4013 Catalyst 4000 Supervisor II 90

WS-X4515 Catalyst 4500 Supervisor IV 90



IBM @server BladeCenter Layer 2-7 Network Switching



Step-by-step configuration instructions for the Layer 2-7 Gigabit Ethernet Switch Module

Working network configurations to get your IBM @server BladeCenter solution up and running quickly

Helpful information for rapid deployment in a Cisco or Nortel network infrastructure

This IBM Redpaper will help you install, tailor, and configure the new IBM @server BladeCenter and the Layer 2-7 GbE Switch Module (GbESM) in various network environments.

In this Redpaper, we discuss the features and functions of the GbESM and how it is managed. We also introduce some networking terms and concepts to the users and administrators of the IBM @server BladeCenter product. This material is meant to give the non-networking professional an overall view of the switching and bridging environment, not to substitute for in-depth training on networking fundamentals.

We then provide step-by-step instructions for establishing your initial network configurations to ensure your Layer 2-7 GbE Switch Module is configured correctly and is operable so you can begin immediate communications across the network. We base this discussion on actual working configurations that we built and tested in our labs.

Also featured in this Redpaper are several configuration examples of deploying the Layer 2-7 GbE Switch Module in Cisco and Nortel environments.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks