

BLADEOS™

Application Guide

BNT Layer2/3 GbE Switch Module for IBM BladeCenter®

Version 5.1

Part Number: BMD00099, December 2009

BLADE
NETWORK TECHNOLOGIES

2350 Mission College Blvd.
Suite 600
Santa Clara, CA 95054
www.bladenetwork.net

Copyright © 2009 BLADE Network Technologies, Inc., 2350 Mission College Blvd., Suite 600, Santa Clara, California, 95054, USA. All rights reserved. Part Number: BMD00099.

This document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this document may be reproduced in any form by any means without prior written authorization of BLADE Network Technologies, Inc. Documentation is provided “as is” without warranty of any kind, either express or implied, including any kind of implied or express warranty of non-infringement or the implied warranties of merchantability or fitness for a particular purpose.

U.S. Government End Users: This document is provided with a “commercial item” as defined by FAR 2.101 (Oct. 1995) and contains “commercial technical data” and “commercial software documentation” as those terms are used in FAR 12.211-12.212 (Oct. 1995). Government End Users are authorized to use this documentation only in accordance with those rights and restrictions set forth herein, consistent with FAR 12.211- 12.212 (Oct. 1995), DFARS 227.7202 (JUN 1995) and DFARS 252.227-7015 (Nov. 1995).

BLADE Network Technologies, Inc. reserves the right to change any products described herein at any time, and without notice. BLADE Network Technologies, Inc. assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by BLADE Network Technologies, Inc. The use and purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of BLADE Network Technologies, Inc.

BLADE Network Technologies, the BLADE logo, BLADEHarmony, BNT, NMotion, RackSwitch, Rackonomics, RackSwitch Solution Partner, ServerMobility, SmartConnect and VMready are trademarks of BLADE Network Technologies, Inc. in the United States and certain other countries. Cisco® and EtherChannel® are registered trademarks of Cisco Systems, Inc. in the United States and certain other countries. Any other trademarks appearing in this manual are owned by their respective companies.

Originated in the USA.

Contents

Preface ■ 17

- Who Should Use This Guide ■ 17
- What You'll Find in This Guide ■ 17
- Additional References ■ 19
- Typographic Conventions ■ 20
- How to Get Help ■ 20

Part 1: Basic Switching ■ 21

Chapter 1: Accessing the Switch ■ 23

- Management Module Setup ■ 24
 - Factory-Default vs. MM-Assigned IP Addresses ■ 24
 - Default Gateway ■ 25
 - Configuring Management Module for Switch Access ■ 25
- Using Telnet ■ 27
 - Connect to the Switch via SSH ■ 27
 - BOOTP Relay Agent ■ 27
 - DHCP Relay Agent ■ 29
- Using the Browser-Based Interface ■ 31
 - Configuring BBI Access via HTTP ■ 31
 - Configuring BBI Access via HTTPS ■ 31
 - BBI Summary ■ 32

Using Simple Network Management Protocol	■	35
SNMP Version 1	■	35
SNMP Version 3	■	35
Configuring SNMP Trap Hosts	■	38
RADIUS Authentication and Authorization	■	42
How RADIUS Authentication Works	■	43
Configuring RADIUS on the Switch	■	43
RADIUS Authentication Features in BLADEOS	■	44
Switch User Accounts	■	45
RADIUS Attributes for BLADEOS User Privileges	■	45
TACACS+ Authentication	■	46
How TACACS+ Authentication Works	■	46
TACACS+ Authentication Features in BLADEOS	■	46
Authorization	■	47
Accounting	■	48
Command Authorization and Logging	■	48
TACACS+ Password Change	■	49
Configuring TACACS+ Authentication on the Switch	■	50
LDAP Authentication and Authorization	■	51
Secure Shell and Secure Copy	■	53
Configuring SSH/SCP Features on the Switch	■	54
Configuring the SCP Administrator Password	■	55
Using SSH and SCP Client Commands	■	55
SSH and SCP Encryption of Management Messages	■	56
Generating RSA Host and Server Keys for SSH Access	■	57
SSH/SCP Integration with Radius Authentication	■	57
SSH/SCP Integration with TACACS+ Authentication	■	58
End User Access Control	■	59
Considerations for Configuring End User Accounts	■	59
Strong Passwords	■	59
User Access Control Menu	■	60
Listing Current Users	■	61
Logging In to an End User Account	■	61
Protected Mode	■	62

Chapter 2: Port-Based Network Access Control ■ 63

Extensible Authentication Protocol over LAN ■ 64

EAPoL Authentication Process ■ 65

EAPoL Message Exchange ■ 66

EAPoL Port States ■ 67

Guest VLAN ■ 67

Supported RADIUS Attributes ■ 68

EAPoL Configuration Guidelines ■ 69

Chapter 3: VLANs ■ 71

VLANs Overview ■ 71

VLANs and Port VLAN ID Numbers ■ 72

VLAN Tagging ■ 74

VLAN Topologies and Design Considerations ■ 78

VLAN Bandwidth Management ■ 80

Calculating the Rate ■ 80

Calculating the Maximum Burst Size ■ 81

Bandwidth Management Configuration Guidelines ■ 81

Bandwidth Configuration Example ■ 82

Protocol-Based VLANs ■ 83

Port-Based vs. Protocol-Based VLANs ■ 84

PVLAN Priority Levels ■ 84

PVLAN Tagging ■ 84

PVLAN Configuration Guidelines ■ 85

Configuring PVLAN ■ 85

Private VLANs ■ 88

Private VLAN Ports ■ 89

Configuration Guidelines ■ 89

Configuration Example ■ 90

Chapter 4: Ports and Trunking ■ 91

Trunking Overview ■ 92

Statistical Load Distribution ■ 93

Built-In Fault Tolerance ■ 93

Before You Configure Static Trunks ■ 93

Inter-Switch Link ■ 94

Trunk Group Configuration Rules ■ 95

Port Trunking Example ■ 96

Configurable Trunk Hash Algorithm ■ 98

Link Aggregation Control Protocol ■ 99

Chapter 5: Spanning Tree Group	103
STG Overview	103
Bridge Protocol Data Units (BPDUs)	104
BDPU Overview	104
Determining the Path for Forwarding BPDUs	105
Spanning Tree Group Guidelines	106
Adding a VLAN to a Spanning Tree Group	106
Creating a VLAN	106
Rules for VLAN Tagged Ports	106
Adding and Removing Ports from STGs	107
Multiple Spanning Trees	108
Default Spanning Tree Configuration	108
Why Do We Need Multiple Spanning Trees?	108
Switch-Centric Spanning Tree Group	109
VLAN Participation in Spanning Tree Groups	110
Configuring Multiple Spanning Tree Groups	110
Port Fast Forwarding	112
Fast Uplink Convergence	112
Hot Links	113
Forward Delay	113
Pre-emption	113
FDB Update	113
Configuration Guidelines	114
Configuring Hot Links	114
Chapter 6: RSTP and MSTP	115
Rapid Spanning Tree Protocol	116
Port State Changes	116
Port Type and Link Type	117
RSTP Configuration Guidelines	117
RSTP Configuration Example	118
Multiple Spanning Tree Protocol	119
MSTP Region	119
Common Internal Spanning Tree	119
MSTP Configuration Guidelines	120
MSTP Configuration Example	120

Chapter 7: Link Layer Discovery Protocol ■ 121

- LLDP Overview ■ 121
- Enabling or Disabling LLDP ■ 122
 - Global LLDP Setting ■ 122
 - Transmit and Receive Control ■ 122
- LLDP Transmit Features ■ 123
 - Scheduled Interval ■ 123
 - Minimum Interval ■ 123
 - Time-to-Live for Transmitted Information ■ 124
 - Trap Notifications ■ 124
 - Changing the LLDP Transmit State ■ 125
 - Types of Information Transmitted ■ 125
- LLDP Receive Features ■ 127
 - Types of Information Received ■ 127
 - Viewing Remote Device Information ■ 127
 - Time-to-Live for Received Information ■ 128
- LLDP Example Configuration ■ 129

Chapter 8: Quality of Service ■ 131

- QoS Overview ■ 131
- Using ACL Filters ■ 133
 - Summary of Packet Classifiers ■ 133
 - Summary of ACL Actions ■ 135
 - ACL Order of Precedence ■ 135
 - ACL Blocks and ACL Groups ■ 135
 - Assigning ACLs to a Port ■ 137
 - ACL Metering and Re-Marking ■ 138
 - Viewing ACL Statistics ■ 139
 - ACL Configuration Examples ■ 139
- Using DSCP Values to Provide QoS ■ 141
 - Differentiated Services Concepts ■ 141
 - Per-Hop Behavior ■ 142
 - QoS Levels ■ 143
 - DSCP Re-Marking and Mapping ■ 144
- Using 802.1p Priorities to Provide QoS ■ 146
- Queuing and Scheduling ■ 147

Part 2: IP Routing ■ 149

Chapter 9: Basic IP Routing ■ 151

IP Routing Benefits ■ 151

Routing Between IP Subnets ■ 152

 Subnet Routing Example ■ 154

 Using VLANs to Segregate Broadcast Domains ■ 156

Dynamic Host Configuration Protocol ■ 158

 DHCP Relay Agent ■ 159

 DHCP Relay Agent Configuration ■ 160

Chapter 10: Routing Information Protocol ■ 161

Distance Vector Protocol ■ 161

Stability ■ 161

Routing Updates ■ 162

RIPv1 ■ 162

RIPv2 ■ 162

RIPv2 in RIPv1 Compatibility Mode ■ 163

RIP Features ■ 163

RIP Configuration Example ■ 165

Chapter 11: IGMP ■ 167

IGMP Snooping ■ 168

 IGMPv3 ■ 169

 IGMP Snooping Configuration Example ■ 169

 Static Multicast Router ■ 171

IGMP Relay ■ 172

 Configuration Guidelines ■ 172

 Configure IGMP Relay ■ 173

Additional IGMP Features ■ 174

 FastLeave ■ 174

 IGMP Filtering ■ 175

Chapter 12: Border Gateway Protocol ■ 177

Internal Routing Versus External Routing ■ 177

Forming BGP Peer Routers ■ 179

What is a Route Map? ■ 179

Incoming and Outgoing Route Maps ■ 180

Precedence ■ 181

Configuration Overview ■ 181

Aggregating Routes ■ 183

Redistributing Routes ■ 183

BGP Attributes ■ 184

Selecting Route Paths in BGP ■ 185

BGP Failover Configuration ■ 186

Default Redistribution and Route Aggregation Example ■ 188

Chapter 13: OSPF ■ 191

OSPF Overview ■ 191

Types of OSPF Areas ■ 192

Types of OSPF Routing Devices ■ 193

Neighbors and Adjacencies ■ 194

The Link-State Database ■ 194

The Shortest Path First Tree ■ 195

Internal Versus External Routing ■ 195

OSPF Implementation in BLADEOS ■ 196

Configurable Parameters ■ 196

Defining Areas ■ 197

Assigning the Area Index ■ 197

Using the Area ID to Assign the OSPF Area Number ■ 198

Attaching an Area to a Network ■ 198

Interface Cost ■ 199

Electing the Designated Router and Backup ■ 199

Summarizing Routes ■ 199

Default Routes ■ 200

Virtual Links ■ 201

Router ID ■ 201

- Authentication ■ 202
 - Configuring Plain Text OSPF Passwords ■ 202
 - Configuring MD5 Authentication ■ 203
- Host Routes for Load Balancing ■ 204
- Loopback Interfaces in OSPF ■ 204
- OSPF Features Not Supported in This Release ■ 205
- OSPF Configuration Examples ■ 205
 - Example 1: Simple OSPF Domain ■ 206
 - Example 2: Virtual Links ■ 208
 - Example 3: Summarizing Routes ■ 212
 - Verifying OSPF Configuration ■ 214
- Chapter 14: Remote Monitoring ■ 215**
 - RMON Overview ■ 215
 - RMON Group 1—Statistics ■ 216
 - RMON Group 2—History ■ 217
 - History MIB Objects ■ 217
 - Configuring RMON History ■ 218
 - RMON Group 3—Alarms ■ 219
 - Alarm MIB Objects ■ 219
 - Configuring RMON Alarms ■ 220
 - RMON Group 9—Events ■ 221

Part 3: High Availability Fundamentals ■ 223

Chapter 15: High Availability ■ 225

- Layer 2 Failover ■ 226
 - Auto Monitoring Trunk Links ■ 226
 - VLAN Monitor ■ 226
 - Auto Monitor Configurations ■ 227
 - Setting the Failover Limit ■ 229
 - Manually Monitoring Port Links ■ 229
 - L2 Failover with Other Features ■ 230
 - LACP ■ 230
 - Spanning Tree Protocol ■ 230
 - Configuration Guidelines ■ 230
 - Auto Monitor Guidelines ■ 231
 - Manual Monitor Guidelines ■ 231
 - Configuring Layer 2 Failover ■ 232
 - Auto Monitor Example ■ 232
 - Manual Monitor Example ■ 233
- VRRP Overview ■ 234
 - VRRP Components ■ 234
 - VRRP Operation ■ 235
 - Selecting the Master VRRP Router ■ 236
- Failover Methods ■ 237
 - Active-Active Redundancy ■ 238
 - Hot-Standby Redundancy ■ 238
 - Virtual Router Group ■ 239
- BLADEOS Extensions to VRRP ■ 240
- Virtual Router Deployment Considerations ■ 241
- High Availability Configurations ■ 242
 - Active-Active Configuration ■ 242
 - Hot-Standby Configuration ■ 247

Part 4: Appendices ■ 251

Appendix A: Monitoring Ports for Troubleshooting ■ 253

sFlow ■ 253

sFlow Statistical Counters ■ 253

sFlow Example Configuration ■ 254

Port Mirroring ■ 255

Port Mirroring Behavior ■ 256

Configuring Port Mirroring ■ 256

Appendix B: RADIUS Server Configuration Notes ■ 259

Glossary ■ 261

Index ■ 263

Figures

Chapter 1: Accessing the Switch

Figure 1: Switch Management on the BladeCenter Management Module ■ 25

Figure 2: BOOTP Relay Agent Configuration ■ 28

Figure 3: DHCP Relay Agent Configuration ■ 30

Chapter 2: Port-Based Network Access Control

Figure 4: Authenticating a Port Using EAPoL ■ 65

Chapter 3: VLANs

Figure 5: Default VLAN settings ■ 75

Figure 6: Port-based VLAN assignment ■ 76

Figure 7: 802.1Q tagging (after port-based VLAN assignment) ■ 76

Figure 8: 802.1Q tag assignment ■ 77

Figure 9: 802.1Q tagging (after 802.1Q tag assignment) ■ 77

Figure 10: Multiple VLANs with VLAN-Tagged Gigabit Adapters ■ 78

Chapter 4: Ports and Trunking

Figure 11: Port Trunk Group ■ 92

Figure 12: Port Trunk Group Configuration Example ■ 96

Chapter 5: Spanning Tree Group

Figure 13: Using Multiple Instances of Spanning Tree Group ■ 109

Figure 14: Implementing Multiple Spanning Tree Groups ■ 109

Chapter 8: Quality of Service

Figure 15: QoS Model ■ 132

Figure 16: Layer 3 IPv4 Packet ■ 141

Figure 17: Layer 2 802.1q/802.1p VLAN Tagged Packet ■ 146

Chapter 9: Basic IP Routing

Figure 18: The Router Legacy Network ■ 152

Figure 19: Switch-Based Routing Topology ■ 153

Figure 20: DHCP Relay Agent Configuration ■ 160

Chapter 12: Border Gateway Protocol

Figure 21: iBGP and eBGP ■ 178

Figure 22: Distributing Network Filters in Access Lists and Route Maps ■ 180

Figure 23: BGP Failover Configuration Example ■ 186

Figure 24: Route Aggregation and Default Route Redistribution ■ 188

Chapter 13: OSPF

Figure 25: OSPF Area Types ■ 192

Figure 26: OSPF Domain and an Autonomous System ■ 193

Figure 27: Injecting Default Routes ■ 200

Figure 28: OSPF Authentication ■ 202

Figure 29: A Simple OSPF Domain ■ 206

Figure 30: Configuring a Virtual Link ■ 208

Figure 31: Summarizing Routes ■ 212

Chapter 15: High Availability

Figure 32: Basic Layer 2 Failover ■ 227

Figure 33: Two trunks, each in a different Failover Trigger ■ 228

Figure 34: Two trunks, one Failover Trigger ■ 228

Figure 35: A Non-VRRP, Hot-Standby Configuration ■ 237

Figure 36: Active-Active Redundancy ■ 238

Figure 37: Hot-Standby Redundancy ■ 238

Figure 38: Active-Active High-Availability Configuration ■ 242

Figure 39: Hot-Standby Configuration ■ 247

Appendix A: Monitoring Ports for Troubleshooting

Figure 40: Mirroring Ports ■ 255

Tables

Preface

Table 1: Typographic Conventions ■ 20

Chapter 1: Accessing the Switch

Table 2: GbESM IP addresses, based on switch-module bay numbers ■ 24

Table 3: User Access Levels ■ 45

Table 4: BLADEOS-proprietary Attributes for RADIUS ■ 45

Table 5: Default TACACS+ Authorization Levels ■ 47

Table 6: Alternate TACACS+ Authorization Levels ■ 47

Chapter 2: Port-Based Network Access Control

Table 7: Support for RADIUS Attributes ■ 68

Chapter 4: Ports and Trunking

Table 8: Actor vs. Partner LACP configuration ■ 99

Chapter 5: Spanning Tree Group

Table 9: Ports, Trunk Groups, and VLANs ■ 104

Chapter 6: RSTP and MSTP

Table 10: RSTP vs. STP Port states ■ 116

Chapter 7: Link Layer Discovery Protocol

Table 11: LLDP Optional Information Types ■ 126

Chapter 8: Quality of Service

Table 12: Well-Known Protocol Types ■ 133

Table 13: Well-Known Application Ports ■ 134

Table 14: Well-Known TCP flag values ■ 134

Table 15: Default QoS Service Levels ■ 143

Chapter 9: Basic IP Routing

Table 16: Subnet Routing Example: IP Address Assignments ■ 154

Table 17: Subnet Routing Example: IP Interface Assignments ■ 154

Table 18: Subnet Routing Example: Optional VLAN Ports ■ 156

Chapter 15: High Availability

Table 19: VRRP Tracking Parameters ■ 240

Preface

The *BLADEOS 5.1 Application Guide* describes how to configure and use the BLADEOS 5.1 software on the BNT Layer2/3 GbE Switch Module for IBM BladeCenter. For documentation on installing the switch physically, see the *Installation Guide* for your GbE Switch Module (GbESM).

Who Should Use This Guide

This guide is intended for network installers and system administrators engaged in configuring and maintaining a network. The administrator should be familiar with Ethernet concepts, IP addressing, Spanning Tree Protocol, and SNMP configuration parameters.

What You'll Find in This Guide

This guide will help you plan, implement, and administer BLADEOS software. Where possible, each section provides feature overviews, usage examples, and configuration instructions.

Part 1: Basic Switching

- **Chapter 1, “Accessing the Switch,”** describes how to access the GbESM to configure the switch and view switch information and statistics. This chapter also discusses different methods to manage the switch for remote administrators, such as setting specific IP addresses and using Remote Authentication Dial-in User Service (RADIUS), Secure Shell (SSH), Secure Copy (SCP), and other methods.
- **Chapter 2, “Port-Based Network Access Control,”** describes how to authenticate devices attached to a LAN port that has point-to-point connection characteristics. This feature prevents access to ports that fail authentication and authorization and provides security to ports of the GbESM that connect to blade servers.
- **Chapter 3, “VLANs,”** describes how to configure Virtual Local Area Networks (VLANs) for creating separate network segments, including how to use VLAN tagging for devices that use multiple VLANs. This chapter also describes Protocol-based VLANs, and Private VLANs.

- [Chapter 4, “Ports and Trunking,”](#) describes how to group multiple physical ports together to aggregate the bandwidth between large-scale network devices.
- [Chapter 5, “Spanning Tree Group,”](#) discusses how Spanning Trees configure the network so that the switch uses the most efficient path when multiple paths exist.
- [Chapter 6, “RSTP and MSTP,”](#) describes the Rapid Spanning Tree Protocol (RSTP) and Multiple Spanning Tree Protocol (MSTP) extensions to Spanning Tree Protocol (STP), providing rapid convergence of spanning trees for quicker network reconfiguration.
- [Chapter 7, “Link Layer Discovery Protocol,”](#) describes how Link Layer Discovery Protocol helps neighboring network devices learn about each others’ ports and capabilities.
- [Chapter 8, “Quality of Service,”](#) discusses Quality of Service (QoS) features, including IP filtering using Access Control Lists (ACLs), Differentiated Services, and IEEE 802.1p priority values.

Part 2: IP Routing

- [Chapter 9, “Basic IP Routing,”](#) describes how to configure the GbESM for IP routing using IP subnets and DHCP Relay.
- [Chapter 10, “Routing Information Protocol,”](#) describes how the BLADEOS software implements standard Routing Information Protocol (RIP) for exchanging TCP/IP route information with other routers.
- [Chapter 11, “IGMP,”](#) describes how the BLADEOS software implements IGMP Snooping or IGMP Relay to conserve bandwidth in a multicast-switching environment.
- [Chapter 12, “Border Gateway Protocol,”](#) describes Border Gateway Protocol (BGP) concepts and features supported in BLADEOS.
- [Chapter 13, “OSPF,”](#) describes key Open Shortest Path First (OSPF) concepts and their implemented in BLADEOS, and provides examples of how to configure your switch for OSPF support.
- [Chapter 14, “Remote Monitoring,”](#) describes how to configure the RMON agent on the switch, so that the switch can exchange network monitoring data.

Part 3: High Availability Fundamentals

- [Chapter 15, “High Availability,”](#) describes how the GbESM supports high-availability network topologies using features such as Layer 2 Failover or Virtual Router Redundancy Protocol (VRRP).

Part 4: Appendices

- [Appendix A, “Monitoring Ports for Troubleshooting,”](#) discusses tools for troubleshooting your switch, such as port mirroring and sFlow.
- [Appendix B, “RADIUS Server Configuration Notes,”](#) discusses how to modify RADIUS configuration files for the Nortel Networks BaySecure Access Control RADIUS server, to provide authentication for users of the GbESM.

Additional References

Additional information about installing and configuring the GbESM is available in the following guides:

- *BNT Layer2/3 GbE Switch Module Installation Guide*
- *BLADEOS 5.1 Command Reference*
- *BLADEOS 5.1 ISCLI Reference Guide*
- *BLADEOS 5.1 BBI Quick Guide*

Typographic Conventions

The following table describes the typographic styles used in this book.

Table 1 Typographic Conventions

Typface or Symbol	Meaning	Example
AaBbCc123	This type is used for names of commands, files, and directories used within the text. It also depicts on-screen computer output and prompts.	View the <code>readme.txt</code> file. Main#
AaBbCc123	This bold type appears in command examples. It shows text that must be typed in exactly as shown.	Main# sys
<AaBbCc123>	This italicized type appears in command examples as a parameter placeholder. Replace the indicated text with the appropriate real name or value when using the command. Do not type the brackets. This also shows book titles, special terms, or words to be emphasized.	To establish a Telnet session, enter: host# telnet <IP address> Read your <i>User's Guide</i> thoroughly.
[]	Command items shown inside brackets are optional and can be used or excluded as the situation demands. Do not type the brackets.	host# ls [-a]
	The vertical bar () is used in command examples to separate choices where multiple options exist. Select only one of the listed options. Do not type the vertical bar.	host# set left right center
AaBbCc123	This block type depicts menus, buttons, and other controls that appear in Web browsers and other graphical interfaces.	Click the Save button.

How to Get Help

If you need help, service, or technical assistance, see the “Getting help and technical assistance” appendix in the *BNT Layer2/3 GbE Switch Module Installation Guide*.

Part 1: Basic Switching

This section discusses basic switching functions. This includes how to access and manage the switch:

- Accessing the switch
- Port-Based Network Access Control
- VLANs
- Port Trunking
- Spanning Tree Protocol
- Rapid Spanning Tree and Protocol and Multiple Spanning Tree Protocol
- Link Layer Detection Protocol
- Quality of Service Converged Enhanced Ethernet and Fibre Channel over Ethernet

CHAPTER 1

Accessing the Switch

The BLADEOS software provides means for accessing, configuring, and viewing information and statistics about the GbE Switch Module (GbESM). This chapter discusses different methods of accessing the switch and ways to secure the switch for remote administrators:

- “Management Module Setup” on page 24
- “Using the Browser-Based Interface” on page 31
- “Using Simple Network Management Protocol” on page 35
- “Securing Access to the Switch” on page 42
 - “RADIUS Authentication and Authorization” on page 42
 - “TACACS+ Authentication” on page 46
 - “LDAP Authentication and Authorization” on page 51
 - “Secure Shell and Secure Copy” on page 53
 - “End User Access Control” on page 59
 - “Protected Mode” on page 62

Management Module Setup

The BladeCenter GbESM is an integral subsystem within the overall BladeCenter system. The BladeCenter chassis includes a management module as the central element for overall chassis management and control.

You can use the management module to configure and manage the GbESM. The GbESM communicates with the management module(s) through its internal port 15 (MGT1) and port 16 (MGT2), which you can access through the 100 Mbps Ethernet port on each management module. The factory default settings permit management and control access to the switch module *only* through the management module, or the built-in serial port.

You can use the external Ethernet ports (EXT x) on the switch module for management and control of the switch, by selecting this mode as an option through the management module configuration utility program (see the applicable *BladeCenter Installation and User's Guide* publications for more information).

Note – Support for each management module is provided by a separate management port (MGT1 and MGT2). One port is active, and the other port is used as a backup.

Factory-Default vs. MM-Assigned IP Addresses

Each GbESM must be assigned its own Internet Protocol address, which is used for communication with an SNMP network manager or other transmission control protocol/Internet Protocol (TCP/IP) applications (for example, BootP or TFTP). The factory-default IP address is 10.90.90. x , where x is based on the number of the bay into which the GbESM is installed. For additional information, see the *Installation Guide*. The management module assigns an IP address of 192.168.70.1 xx , where xx is also based on the number of the bay into which each GbESM is installed, as shown in the following table:

Table 2 GbESM IP addresses, based on switch-module bay numbers

Bay Number	Factory-Default IP Address	IP Address Assigned by MM
Bay 1	10.90.90.91	192.168.70.127
Bay 2	10.90.90.92	192.168.70.128
Bay 3	10.90.90.94	192.168.70.129
Bay 4	10.90.90.97	192.168.70.130

Note – GbESMs installed in Bay 1 and Bay 2 connect to server NICs 1 and 2, respectively. However, Windows operating systems using older I/O expansion adapters show that GbESMs installed in Bay 3 and Bay 4 connect to server NICs 4 and 3, respectively.

Default Gateway

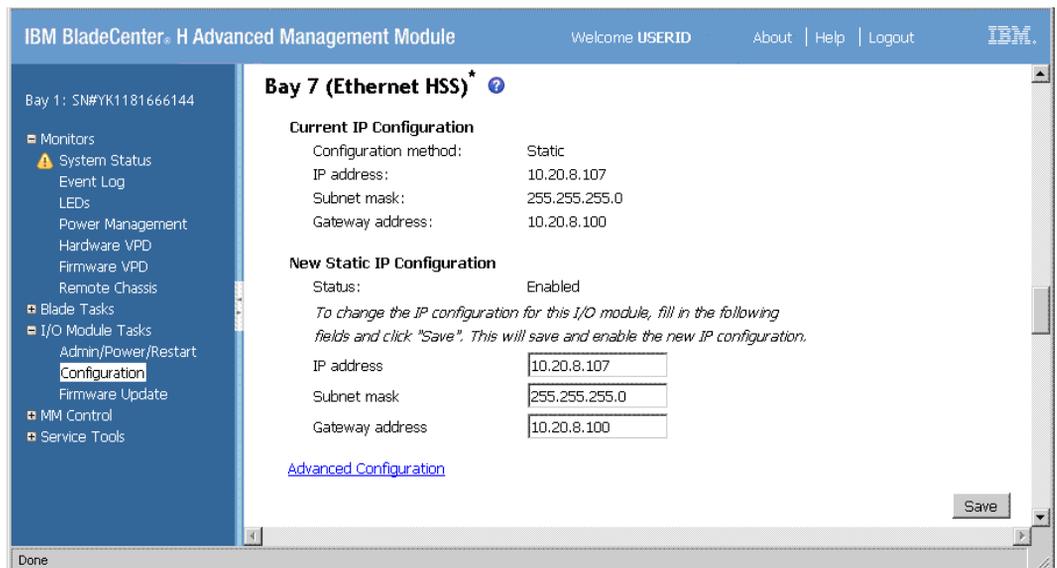
The default Gateway IP address determines where packets with a destination address outside the current subnet should be sent. Usually, the default Gateway is a router or host acting as an IP gateway to handle connections to other subnets of other TCP/IP networks. If you want to access the GbESM from outside your local network, use the management module to assign a default Gateway address to the GbESM. Choose **I/O Module Tasks > Configuration** from the navigation pane on the left, and enter the default Gateway IP address (for example, 192.168.70.125). Click **Save**.

Configuring Management Module for Switch Access

Complete the following initial configuration steps:

1. Connect the Ethernet port of the management module to a 10/100 Mbps network (with access to a management station) or directly to a management station.
2. Access and log on to the management module, as described in the *BladeCenter Management Module User's Guide*. The management module provides the appropriate IP addresses for network access (see the applicable *BladeCenter Installation and User's Guide* publications for more information).
3. Select **Configuration** on the **I/O Module Tasks** menu on the left side of the BladeCenter Management Module window. See [Figure 1](#).

Figure 1 Switch Management on the BladeCenter Management Module



Note – The screen shown is an example only. Actual screen output may differ, depending on the type, version, and configuration of hardware and software installed in your systems.

4. You can use the default IP addresses provided by the management module, or you can assign a new IP address to the switch module through the management module. You can assign this IP address through one of the following methods:
 - Manually through the BladeCenter management module
 - Automatically through the IBM Director Configuration Wizard

Note – If you change the IP address of the GbESM, make sure that the switch module and the management module both reside on the same subnet.

5. Enable the following features in the management module:
 - External Ports (I/O Module Tasks > Admin/Power/Restart > Advanced Setup)
 - External management over all ports (Configuration > Advanced Configuration)

This setting is required if you want to access the management network through the external data ports (EXT x) on the GbESM.

The default value is `Disabled` for both features. If these features are not already enabled, change the value to `Enabled`, then click **Save**.

Note – In `Advanced Configuration > Advanced Setup`, enable “Preserve new IP configuration on all switch resets,” to retain the switch’s IP interface when you restore factory defaults. This setting preserves the management port’s IP address in the management module’s memory, so you maintain connectivity to the management module after a reset.

You can now start a Telnet session, Browser-Based Interface (Web) session, a Secure Shell session, or a secure HTTPS session to the GbESM.

Using Telnet

Use the management module to access the GbESM through Telnet. Choose **I/O Module Tasks > Configuration** from the navigation pane on the left. Select a bay number and click **Advanced Configuration > Start Telnet/Web Session > Start Telnet Session**. A Telnet window opens a connection to the Switch Module (requires Java 1.4 Plug-in).

Once you have configured the GbESM with an IP address and gateway, you can access the switch from any workstation connected to the management network. Telnet access provides the same options for user and administrator access as those available through the management module.

To establish a Telnet connection with the switch, you can run the Telnet program on your workstation and issue the Telnet command, followed by the switch IP address:

```
telnet <switch IP address>
```

Connect to the Switch via SSH

The SSH (Secure Shell) protocol enables you to securely log into another computer over a network to execute commands remotely. As a secure alternative to using Telnet to manage switch configuration, SSH ensures that all data sent over the network is encrypted and secure.

By default, the SSH feature is disabled. For information on enabling and using SSH for switch access, see [“Secure Shell and Secure Copy” on page 53](#). For more information on using the CLI for switch configuration, see the *BLADEOS 5.1 Command Reference*.

BOOTP Relay Agent

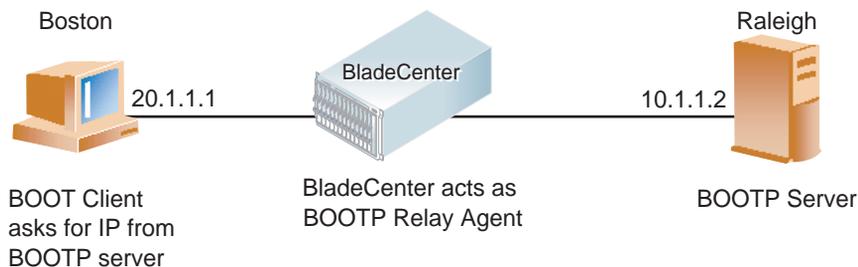
BOOTP Relay Agent Overview

The GbESM can function as a Bootstrap Protocol relay agent, enabling the switch to forward a client request for an IP address up to two BOOTP servers with IP addresses that have been configured on the switch.

When a switch receives a BOOTP request from a BOOTP client requesting an IP address, the switch acts as a proxy for the client. The request is then forwarded as a UDP Unicast MAC layer message to two BOOTP servers whose IP addresses are configured on the switch. The servers respond to the switch with a Unicast reply that contains the default gateway and IP address for the client. The switch then forwards this reply back to the client.

Figure 2 shows a basic BOOTP network example.

Figure 2 BOOTP Relay Agent Configuration



The use of two servers provide failover redundancy. The client request is forwarded to both BOOTP servers configured on the switch. However, no health checking is supported.

BOOTP Relay Agent Configuration

To enable the GbESM to be the BOOTP forwarder, you need to configure the BOOTP server IP addresses on the switch, and enable BOOTP relay on the interface(s) on which the BOOTP requests are received.

Generally, you should configure the command on the switch IP interface that is closest to the client, so that the BOOTP server knows from which IP subnet the newly allocated IP address should come.

Use the following commands to configure the switch as a BOOTP relay agent:

```
>> # /cfg/13/bootp
>> Bootstrap Protocol Relay# addr <IP address> (IP address of BOOTP server)
>> Bootstrap Protocol Relay# addr2 <IP address> (IP address of 2nd BOOTP server)
>> Bootstrap Protocol Relay# on (Globally turn BOOTP relay on)
>> Bootstrap Protocol Relay# off (Globally turn BOOTP relay off)
>> Bootstrap Protocol Relay# cur (Display current configuration)
```

Use the following command to enable the Relay functionality on an IP interface:

```
>> # /cfg/13/if <interface number> /relay ena
```

DHCP Relay Agent

DHCP Relay Agent Overview

DHCP is described in RFC 2131, and the DHCP relay agent supported on the GbESM is described in RFC 1542. DHCP uses UDP as its transport protocol. The client sends messages to the server on port 67 and the server sends messages to the client on port 68.

DHCP defines the methods through which clients can be assigned an IP address for a finite lease period and allowing reassignment of the IP address to another client later. Additionally, DHCP provides the mechanism for a client to gather other IP configuration parameters it needs to operate in the TCP/IP network.

In the DHCP environment, the switch acts as a relay agent. The DHCP relay feature (`/cfg/13/bootp`) enables the switch to forward a client request for an IP address to two BOOTP servers with IP addresses that have been configured on the switch.

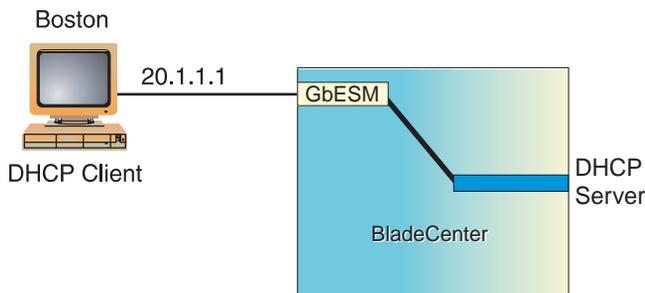
When a switch receives a UDP broadcast on port 67 from a DHCP client requesting an IP address, the switch acts as a proxy for the client, replacing the client source IP (SIP) and destination IP (DIP) addresses. The request is then forwarded as a UDP Unicast MAC layer message to two BOOTP servers whose IP addresses are configured on the switch. The servers respond as a UDP Unicast message back to the switch, with the default gateway and IP address for the client. The destination IP address in the server response represents the interface address on the switch that received the client request. This interface address tells the switch on which VLAN to send the server response to the client.

DHCP Relay Agent Configuration

To enable the GbESM to be the BOOTP forwarder, you need to configure the DHCP/BOOTP server IP addresses on the switch. Generally, you should configure the switch IP interface on the client side to match the client's subnet, and configure VLANs to separate clients and server subnets. The DHCP server knows from which IP subnet the newly allocated IP address should come.

The following figure shows a basic DHCP network example:

Figure 3 DHCP Relay Agent Configuration



In GbESM implementation, there is no need for primary or secondary servers. The client request is forwarded to the BOOTP servers configured on the switch. The use of two servers provide failover redundancy. However, no health checking is supported.

Use the following commands to configure the switch as a DHCP relay agent:

```
>> # /cfg/13/bootp
>> Bootstrap Protocol Relay# addr           (Set IP address of BOOTP server)
>> Bootstrap Protocol Relay# addr2        (Set IP address of 2nd BOOTP server)
>> Bootstrap Protocol Relay# on           (Globally turn BOOTP relay on)
>> Bootstrap Protocol Relay# off          (Globally turn BOOTP relay off)
>> Bootstrap Protocol Relay# cur          (Display current configuration)
```

Additionally, DHCP Relay functionality can be assigned on a per interface basis. Use the following command to enable the Relay functionality:

```
>> # /cfg/13/if <interface number>/relay ena
```

Using the Browser-Based Interface

Use the management module to access the GbESM through a Web session. Choose I/O Module Tasks > Configuration from the navigation pane on the left. Select a bay number and click Advanced Configuration > Start Telnet/Web Session > Start Web Session. A browser window opens a connection to the Switch Module.

The Browser-based Interface (BBI) provides access to the common configuration, management and operation features of the GbESM through your Web browser. For more information, refer to the *BBI Quick Guide*.

By default, BBI access is enabled on the switch (`/cfg/sys/access/http ena`).

Configuring BBI Access via HTTP

To enable or disable BBI access on the switch via HTTP, use the following commands:

```
>> # /cfg/sys/access/http ena           (Enable HTTP access)
      -or-
>> # /cfg/sys/access/http dis          (Disable HTTP access)
```

The management module requires the default HTTP web server port (port 80) to access the BBI. However, you can change the default Web server port with the following command:

```
>> # /cfg/sys/access/wport <x>
```

For workstation access to your switch via the Browser-Based Interface, open a Web browser window and type in the URL using the IP interface address of the switch, such as `http://10.10.10.1`.

Configuring BBI Access via HTTPS

The BBI can also be accessed via a secure HTTPS connection over management and data ports.

1. Enable HTTPS

By default, BBI access via HTTPS is disabled on the switch. To enable BBI Access via HTTPS, use the following command:

```
>> # /cfg/sys/access/https/access ena
```

2. Set the HTTPS server port number (optional)

To change the HTTPS Web server port number from the default port 443, use the following command:

```
>> # /cfg/sys/access/https/port <x>
```

3. Use the `apply` and `save` commands to activate and store the configuration changes.

4. Generate the HTTPS certificate.

Accessing the BBI via HTTPS requires that you generate a certificate to be used during the key exchange. A default certificate is created the first time HTTPS is enabled, but you can create a new certificate defining the information you want to be used in the various fields.

```
>> /cfg/sys/access/https/generate
Country Name (2 letter code) []: <country code>
State or Province Name (full name) []: <state>
Locality Name (eg, city) []: <city>
Organization Name (eg, company) []: <company>
Organizational Unit Name (eg, section) []: <org. unit>
Common Name (eg, YOUR name) []: <name>
Email (eg, email address) []: <email address>
Confirm generating certificate? [y/n]: y
Generating certificate. Please wait (approx 30 seconds)
restarting SSL agent
```

5. Save the HTTPS certificate.

The certificate is valid only until the switch is rebooted. In order to save the certificate so that it is retained beyond reboot or power cycles, use the following command:

```
>> # /cfg/sys/access/https/certsave
```

When a client (e.g. web browser) connects to the switch, they will be asked if they accept the certificate and can verify that the fields are what expected. Once BBI access is granted to the client, the BBI can be used as described in the *BLADEOS 5.1 BBI Quick Guide*.

BBI Summary

The BBI is organized at a high level as follows:

Context buttons – allow you to select the type of action you wish to perform. The *Configuration* button provides access to the configuration elements for the entire switch. The *Statistics* button provides access to the switch statistics and state information. The *Dashboard* button allows you to display settings and operating status of a variety of switch features.

Navigation Window – provides a menu list of switch features and functions, as follows:

- **System** – this folder provides access to the configuration elements for the entire switch.
 - General
 - User Table
 - Radius
 - TACACS+
 - LDAP
 - NTP
 - sFlow
 - Syslog/Trap Features
 - Config/Image Control
 - Management Network
 - Transceiver
 - Protected Mode
 - Chassis
- **Switch Ports** – configure each of the physical ports on the switch.
- **Port-Based Port Mirroring** – configure port mirroring and mirror port.
- **Layer 2** – Configure Quality of Service (QoS) features for the switch.
 - 802.1X
 - FDB
 - Virtual LANs
 - Spanning Tree Groups
 - MSTP/RSTP
 - LLDP
 - Failover
 - Hot Links
 - Trunk Groups
 - Trunk Hash
 - LACP
 - Uplink Fast
 - BPDU Guard
 - PVST+ compatibility
 - MAC Address Notification
- **RMON** – Configure Remote Monitoring (RMON) features for the switch.

- **Layer 3** – Configure Layer 3 features for the switch.
 - IP Interfaces
 - IP Loopback Interfaces
 - Network Routes
 - Static IPMC Routes
 - ARP
 - Network Filters
 - Route Maps
 - Border Gateway Protocol
 - Default Gateways
 - IPv6 Default Gateways
 - IGMP
 - OSPF Routing Protocol
 - Routing Information Protocol
 - Virtual Router Redundancy Protocol
 - Domain Name System
 - Bootstrap Protocol Relay
 - General
- **QoS** – Configure Quality of Service (QoS) features for the switch.
 - 802.1p
 - DSCP
- **Access Control** – Configure Access Control Lists to filter IP packets.
 - Access Control Lists
 - Access Control List Blocks
 - Access Control List Groups

Using Simple Network Management Protocol

BLADEOS provides Simple Network Management Protocol (SNMP) version 1, version 2, and version 3 support for access through any network management software, such as IBM Director or HP-OpenView.

SNMP Version 1

To access the SNMP agent on the GbESM, the read and write community strings on the SNMP manager should be configured to match those on the switch. The default read community string on the switch is `public` and the default write community string is `private`.

The read and write community strings on the switch can be changed using the following commands on the CLI:

```
>> # /cfg/sys/ssnmp/rcomm
      -and-
>> # /cfg/sys/ssnmp/wcomm
```

The SNMP manager should be able to reach the management interface or any one of the IP interfaces on the switch.

For the SNMP manager to receive the SNMPv1 traps sent out by the SNMP agent on the switch, configure the trap host on the switch with the following command:

```
>> # /cfg/sys/ssnmp/trsrc <1-128>
```

SNMP Version 3

SNMP version 3 (SNMPv3) is an enhanced version of the Simple Network Management Protocol, approved by the Internet Engineering Steering Group in March, 2002. SNMPv3 contains additional security and authentication features that provide data origin authentication, data integrity checks, timeliness indicators and encryption to protect against threats such as masquerade, modification of information, message stream modification and disclosure.

SNMPv3 allows clients to query the MIBs securely.

To access the SNMPv3 menu, enter the following command in the CLI:

```
>> # /cfg/sys/ssnmp/snmpv3
```

For more information on SNMP MIBs and the commands used to configure SNMP on the switch, see the *BLADEOS 5.1 Command Reference*.

Default Configuration

BLADEOS has two SNMPv3 users by default. Both of the following users have access to all the MIBs supported by the switch:

- User 1 name/password: adminmd5/adminmd5. Authentication used is MD5.
- User 2 name/password: adminsha/adminsha. Authentication used is SHA.

To configure an SNMP user name, enter the following command from the CLI:

```
>> # /cfg/sys/ssnmp/snmpv3/usm 1
```

User Configuration

Users can be configured to use the authentication/privacy options. The GbESM support two authentication algorithms: MD5 and SHA, as specified in the following command:

```
>> # /cfg/sys/ssnmp/snmpv3/usm <user number>/auth {md5 | sha | none }
```

1. To configure a user with name “admin,” authentication type MD5, and authentication password of “admin,” privacy option DES with privacy password of “admin,” use the following CLI commands.

```
>> # /cfg/sys/ssnmp/snmpv3/usm 5
>> SNMPv3 usmUser 5# name "admin"                (Configure 'admin' user type)
>> SNMPv3 usmUser 5# auth md5
>> SNMPv3 usmUser 5# authpw admin
>> SNMPv3 usmUser 5# priv des
>> SNMPv3 usmUser 5# privpw admin
```

2. Configure a user access group, along with the views the group may access. Use the access table to configure the group’s access level.

```
>> # /cfg/sys/ssnmp/snmpv3/access 5
>> SNMPv3 vacmAccess 5# name "admingrp"          (Configure an access group)
>> SNMPv3 vacmAccess 5# level authPriv
>> SNMPv3 vacmAccess 5# rview "iso"
>> SNMPv3 vacmAccess 5# wview "iso"
>> SNMPv3 vacmAccess 5# nview "iso"
```

Because the read view (rview), write view (wview), and notify view (nview) are all set to “iso,” the user type has access to all private and public MIBs.

3. Assign the user to the user group. Use the group table to link the user to a particular access group.

```
>> # /cfg/sys/ssnmp/snmpv3/group 5
>> SNMPv3 vacmSecurityToGroup 5# uname admin
>> SNMPv3 vacmSecurityToGroup 5# gname admingrp
```

If you want to allow user access only to certain MIBs, see “View-Based Configuration,” next.

View-Based Configurations

■ CLI User Equivalent

To configure an SNMP user equivalent to the CLI “user” login, use the following configuration:

```

/c/sys/ssnmp/snmpv3/usm 4                               (Configure the user)
  name "usr"
/c/sys/ssnmp/snmpv3/access 3                           (Configure access group 3)
  name "usrgrp"
  rview "usr"
  wview "usr"
  nview "usr"
/c/sys/ssnmp/snmpv3/group 4                           (Assign user to access group 3)
  uname usr
  gname usrgrp
/c/sys/ssnmp/snmpv3/view 6                             (Create views for user)
  name "usr"
  tree "1.3.6.1.4.1.1872.2.5.1.2"                     (Agent statistics)
/c/sys/ssnmp/snmpv3/view 7
  name "usr"
  tree "1.3.6.1.4.1.1872.2.5.1.3"                     (Agent information)
/c/sys/ssnmp/snmpv3/view 8
  name "usr"
  tree "1.3.6.1.4.1.1872.2.5.2.2"                     (L2 statistics)
/c/sys/ssnmp/snmpv3/view 9
  name "usr"
  tree "1.3.6.1.4.1.1872.2.5.2.3"                     (L2 information)
/c/sys/ssnmp/snmpv3/view 10
  name "usr"
  tree "1.3.6.1.4.1.1872.2.5.3.2"                     (L3 statistics)
/c/sys/ssnmp/snmpv3/view 11
  name "usr"
  tree "1.3.6.1.4.1.1872.2.5.3.3"                     (L3 information)

```

■ CLI Oper Equivalent

<code>/c/sys/ssnmp/snmpv3/usm 5</code>	<i>(Configure the oper)</i>
<code>name "oper"</code>	
<code>/c/sys/ssnmp/snmpv3/access 4</code>	<i>(Configure access group 4)</i>
<code>name "opergrp"</code>	
<code>rview "oper"</code>	
<code>wview "oper"</code>	
<code>nview "oper"</code>	
<code>/c/sys/ssnmp/snmpv3/group 4</code>	<i>(Assign oper to access group 4)</i>
<code>uname oper</code>	
<code>gname opergrp</code>	
<code>/c/sys/ssnmp/snmpv3/view 20</code>	<i>(Create views for oper)</i>
<code>name "usr"</code>	
<code>tree "1.3.6.1.4.1.1872.2.5.1.2"</code>	<i>(Agent statistics)</i>
<code>/c/sys/ssnmp/snmpv3/view 21</code>	
<code>name "usr"</code>	
<code>tree "1.3.6.1.4.1.1872.2.5.1.3"</code>	<i>(Agent information)</i>
<code>/c/sys/ssnmp/snmpv3/view 22</code>	
<code>name "usr"</code>	
<code>tree "1.3.6.1.4.1.1872.2.5.2.2"</code>	<i>(L2 statistics)</i>
<code>/c/sys/ssnmp/snmpv3/view 23</code>	
<code>name "usr"</code>	
<code>tree "1.3.6.1.4.1.1872.2.5.2.3"</code>	<i>(L2 information)</i>
<code>/c/sys/ssnmp/snmpv3/view 24</code>	
<code>name "usr"</code>	
<code>tree "1.3.6.1.4.1.1872.2.5.3.2"</code>	<i>(L3 statistics)</i>
<code>/c/sys/ssnmp/snmpv3/view 25</code>	
<code>name "usr"</code>	
<code>tree "1.3.6.1.4.1.1872.2.5.3.3"</code>	<i>(L3 information)</i>

Configuring SNMP Trap Hosts

SNMPv1 Trap Host

1. Configure a user with no authentication and password.

```
/cfg/sys/ssnmp/snmpv3/usm 10/name "v1trap"
```

2. Configure an access group and group table entries for the user. Use the following menu to specify which traps can be received by the user:

```
/cfg/sys/ssnmp/snmpv3/access <user number>
```

In the example below the user will receive the traps sent by the switch.

```

/c/sys/ssnmp/snmpv3/access 10          (Access group to view SNMPv1 traps)
    name "vtrap"
    model snmpv1
    nview "iso"
/c/sys/ssnmp/snmpv3/group 10         (Assign user to the access group)
    model snmpv1
    uname vtrap
    gname vtrap

```

3. Configure an entry in the notify table.

```

/c/sys/ssnmp/snmpv3/notify 10       (Assign user to the notify table)
    name vtrap
    tag vtrap

```

4. Specify the IP address and other trap parameters in the targetAddr and targetParam tables. Use the following menus:

```

/cfg/sys/ssnmp/snmpv3/taddr <target address index number>
    -and-
/cfg/sys/ssnmp/snmpv3/tparam <target parameter index number>

```

For example

```

/c/sys/ssnmp/snmpv3/taddr 10        (Define an IP address to send traps)
    name vtrap
    addr 47.80.23.245
    taglist vtrap
    pname vlparam
/c/sys/ssnmp/snmpv3/tparam 10      (Specify SNMPv1 traps to send)
    name vlparam
    mpmodel snmpv1
    uname vtrap
    model snmpv1

```

5. Use the community table to specify which community string is used in the trap.

```

/c/sys/ssnmp/snmpv3/comm 10        (Define the community string)
    index vtrap
    name public
    uname vtrap

```

SNMPv2 Trap Host Configuration

The SNMPv2 trap host configuration is similar to the SNMPv1 trap host configuration. Wherever you specify the model, use `snmpv2` instead of `snmpv1`.

```

c/sys/ssnmp/snmpv3/usm 10                (Configure user named "v2trap")
name "v2trap"
/c/sys/ssnmp/snmpv3/access 10           (Access group to view SNMPv2 traps)
  name "v2trap"
  model snmpv2
  nview "iso"
/c/sys/ssnmp/snmpv3/group 10           (Assign user to the access group)
  model snmpv2
  uname v2trap
  gname v2trap
/c/sys/ssnmp/snmpv3/notify 10         (Assign user to the notify table)
  name v2trap
  tag v2trap
/c/sys/ssnmp/snmpv3/taddr 10          (Define an IP address to send traps)
  name v2trap
  addr 47.81.25.66
  taglist v2trap
  pname v2param
/c/sys/ssnmp/snmpv3/tparam 10         (Specify SNMPv2 traps to send)
  name v2param
  mpmodel snmpv2c
  uname v2trap
  model snmpv2
/c/sys/ssnmp/snmpv3/comm 10          (Define the community string)
  index v2trap
  name public
  uname v2trap

```

SNMPv3 Trap Host Configuration

To configure a user for SNMPv3 traps, you can choose to send the traps with both privacy and authentication, with authentication only, or without privacy or authentication.

This is configured in the access table using the following commands:

```
/cfg/sys/ssnmp/snmpv3/access <x>/level
/cfg/sys/ssnmp/snmpv3/tparam <x>
```

Configure the user in the user table accordingly.

It is not necessary to configure the community table for SNMPv3 traps because the community string is not used by SNMPv3.

The following example shows how to configure a SNMPv3 user `v3trap` with authentication only:

<code>/c/sys/ssnmp/snmpv3/usm 11</code>	<i>(Configure user named "v3trap")</i>
<code>name "v3trap"</code>	
<code>auth md5</code>	
<code>authpw v3trap</code>	
<code>/c/sys/ssnmp/snmpv3/access 11</code>	<i>(Access group to view SNMPv3 traps)</i>
<code>name "v3trap"</code>	
<code>level authNoPriv</code>	
<code>nview "iso"</code>	
<code>/c/sys/ssnmp/snmpv3/group 11</code>	<i>(Assign user to the access group)</i>
<code>uname v3trap</code>	
<code>gname v3trap</code>	
<code>/c/sys/ssnmp/snmpv3/notify 11</code>	<i>(Assign user to the notify table)</i>
<code>name v3trap</code>	
<code>tag v3trap</code>	
<code>/c/sys/ssnmp/snmpv3/taddr 11</code>	<i>(Define an IP address to send traps)</i>
<code>name v3trap</code>	
<code>addr 47.81.25.66</code>	
<code>taglist v3trap</code>	
<code>pname v3param</code>	
<code>/c/sys/ssnmp/snmpv3/tparam 11</code>	<i>(Specify SNMPv3 traps to send)</i>
<code>name v3param</code>	
<code>uname v3trap</code>	
<code>level authNoPriv</code>	<i>(Set the authentication level)</i>

Securing Access to the Switch

Secure switch management is needed for environments that perform significant management functions across the Internet. The following are some of the functions for secured management:

- Authentication and authorization of remote administrators: see [“RADIUS Authentication and Authorization”](#) on page 42
- Encryption of management information exchanged between the remote administrator and the switch: see [“Secure Shell and Secure Copy”](#) on page 53

The following sections are addressed in this section:

- [“RADIUS Authentication and Authorization”](#) on page 42
- [“TACACS+ Authentication”](#) on page 46
- [“LDAP Authentication and Authorization”](#) on page 51
- [“Secure Shell and Secure Copy”](#) on page 53
- [“End User Access Control”](#) on page 59
- [“Protected Mode”](#) on page 62

RADIUS Authentication and Authorization

BLADEOS supports the RADIUS (Remote Authentication Dial-in User Service) method to authenticate and authorize remote administrators for managing the switch. This method is based on a client/server model. The Remote Access Server (RAS)—the switch—is a client to the back-end database server. A remote user (the remote administrator) interacts only with the RAS, not the back-end server and database.

RADIUS authentication consists of the following components:

- A protocol with a frame format that utilizes UDP over IP (based on RFC 2138 and 2866)
- A centralized server that stores all the user authorization information
- A client, in this case, the switch

The GbESM—acting as the RADIUS client—communicates to the RADIUS server to authenticate and authorize a remote administrator using the protocol definitions specified in RFC 2138 and 2866. Transactions between the client and the RADIUS server are authenticated using a shared key that is not sent over the network. In addition, the remote administrator passwords are sent encrypted between the RADIUS client (the switch) and the back-end RADIUS server.

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 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.

Configuring RADIUS on the Switch

Use the following procedure to configure Radius authentication on your GbESM. For more information, see “[RADIUS Server Configuration Notes](#)” on page 259.

1. Turn RADIUS authentication on, then configure the Primary and Secondary RADIUS servers.

```
>> Main# /cfg/sys/radius                (Select the RADIUS Server menu)
>> RADIUS Server# on                    (Turn RADIUS on)
Current status: OFF
New status: ON
>> RADIUS Server# prisrv 10.10.1.1      (Enter primary server IP)
Current primary RADIUS server: 0.0.0.0
New pending primary RADIUS server: 10.10.1.1
>> RADIUS Server# secsrv 10.10.1.2      (Enter secondary server IP)
Current secondary RADIUS server: 0.0.0.0
New pending secondary RADIUS server: 10.10.1.2
```

2. Configure the RADIUS secret.

```
>> RADIUS Server# secret
Enter new RADIUS secret: <1-32 character secret>
>> RADIUS Server# secret2
Enter new secondary RADIUS server secret: <1-32 character secret>
```



Caution—If you configure the RADIUS secret using any method other than through the console port or management module, the secret may be transmitted over the network as clear text.

3. If desired, you may change the default UDP port number used to listen to RADIUS.

The well-known port for RADIUS is 1645.

```
>> RADIUS Server# port
Current RADIUS port: 1645
Enter new RADIUS port [1500-3000]: <UDP port number>
```

4. Configure the number retry attempts for contacting the RADIUS server, and the timeout period.

```

>> RADIUS Server# retries
Current RADIUS server retries: 3
Enter new RADIUS server retries [1-3]: <server retries>
>> RADIUS Server# timeout
Current RADIUS server timeout: 3
Enter new RADIUS server timeout [1-10]: <the timeout period in minutes>

```

RADIUS Authentication Features in BLADEOS

BLADEOS supports the following RADIUS authentication features:

- Supports RADIUS client on the switch, based on the protocol definitions in RFC 2138 and RFC 2866.
- Allows a RADIUS secret password of up to 32 characters.
- Supports *secondary authentication server* so that when the primary authentication server is unreachable, the switch can send client authentication requests to the secondary authentication server. Use the `/cfg/sys/radius/cur` command to show the currently active RADIUS authentication server.
- Supports user-configurable RADIUS server retry and time-out values:
 - Time-out value = 1-10 seconds
 - Retries = 1-3

The switch will time out if it does not receive a response from the RADIUS server within 1-10 seconds. The switch automatically retries connecting to the RADIUS server 1-3 times before it declares the server down.

- Supports user-configurable RADIUS application port. The default is 1645/UDP-based on RFC 2138. Port 1812 is also supported.
- Allows network administrator to define privileges for one or more specific users to access the switch at the RADIUS user database.
- SecurID is supported if the RADIUS server can do an ACE/Server client proxy. The password is the PIN number, plus the token code of the SecurID card.

Switch User Accounts

The user accounts listed in [Table 3](#) can be defined in the RADIUS server dictionary file.

Table 3 User Access Levels

User Account	Description and Tasks Performed	Password
User	The User has no direct responsibility for switch management. He/she can view all switch status information and statistics but cannot make any configuration changes to the switch.	user
Operator	In addition to User capabilities, the Operator has limited switch management access, including the ability to make temporary, operational configuration changes to some switch features, and to reset switch ports (other than management ports).	oper
Administrator	The super-user Administrator has complete access to all menus, information, and configuration commands on the switch, including the ability to change both the user and administrator passwords.	admin

RADIUS Attributes for BLADEOS User Privileges

When the user logs in, the switch authenticates his/her level of access by sending the RADIUS access request, that is, the client authentication request, to the RADIUS authentication server.

If the remote user is successfully authenticated by the authentication server, the switch will verify the *privileges* of the remote user and authorize the appropriate access. The administrator has an option to allow *backdoor* access via Telnet, SSH, HTTP, and HTTPS. The default GbESM setting for backdoor access is `disabled`. Backdoor access is always enabled on the console port.

Note – To obtain the RADIUS backdoor password for your GbESM, contact your IBM Service and Support line.

All user privileges, other than those assigned to the Administrator, have to be defined in the RADIUS dictionary. RADIUS attribute 6 which is built into all RADIUS servers defines the administrator. The file name of the dictionary is RADIUS vendor-dependent. The following RADIUS attributes are defined for BLADEOS user privileges levels:

Table 4 BLADEOS-proprietary Attributes for RADIUS

User Name/Access	User-Service-Type	Value
User	<i>Vendor-supplied</i>	255
Operator	<i>Vendor-supplied</i>	252
Admin	<i>Vendor-supplied</i>	6

TACACS+ Authentication

BLADEOS supports authentication, authorization, and accounting with networks using the Cisco Systems TACACS+ protocol. The GbESM functions as the Network Access Server (NAS) by interacting with the remote client and initiating authentication and authorization sessions with the TACACS+ access server. The remote user is defined as someone requiring management access to the GbESM either through a data or management port.

TACACS+ offers the following advantages over RADIUS:

- TACACS+ uses TCP-based connection-oriented transport; whereas RADIUS is UDP-based. TCP offers a connection-oriented transport, while UDP offers best-effort delivery. RADIUS requires additional programmable variables such as re-transmit attempts and time-outs to compensate for best-effort transport, but it lacks the level of built-in support that a TCP transport offers.
- TACACS+ offers full packet encryption whereas RADIUS offers password-only encryption in authentication requests.
- TACACS+ separates authentication, authorization and accounting.

How TACACS+ Authentication Works

TACACS+ works much in the same way as RADIUS authentication as described on [page 42](#).

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

During a session, if additional authorization checking is needed, the switch checks with a TACACS+ server to determine if the user is granted permission to use a particular command.

TACACS+ Authentication Features in BLADEOS

Authentication is the action of determining the identity of a user, and is generally done when the user first attempts to log in to a device or gain access to its services. BLADEOS supports ASCII inbound login to the device. PAP, CHAP and ARAP login methods, TACACS+ change password requests, and one-time password authentication are not supported.

Authorization

Authorization is the action of determining a user's privileges on the device, and usually takes place after authentication.

The default mapping between TACACS+ authorization levels and BLADEOS management access levels is shown in [Table 5](#). The authorization levels listed in this table must be defined on the TACACS+ server.

Table 5 Default TACACS+ Authorization Levels

BLADEOS User Access Level	TACACS+ Level
user	0
oper	3
admin	6

Alternate mapping between TACACS+ authorization levels and BLADEOS management access levels is shown in [Table 6](#). Use the `/cfg/sys/tacacs/cmap ena` command to use the alternate TACACS+ authorization levels.

Table 6 Alternate TACACS+ Authorization Levels

BLADEOS User Access Level	TACACS+ Level
user	0–1
oper	6–8
admin	14–15

You can customize the mapping between TACACS+ privilege levels and GbESM management access levels. Use the `/cfg/sys/tacacs/usermap` command to manually map each TACACS+ privilege level (0-15) to a corresponding GbESM management access level.

If the remote user is successfully authenticated by the authentication server, the switch verifies the *privileges* of the remote user and authorizes the appropriate access. The administrator has an option to allow *backdoor* access via Telnet (`/cfg/sys/tacacs/bckdoor`). The default value for Telnet access is `disabled`. The administrator also can enable *secure backdoor* (`/cfg/sys/tacacs/secbd`), to allow access if both the primary and the secondary TACACS+ servers fail to respond.

Note – To obtain the TACACS+ backdoor password for your switch, contact your IBM Service and Support line.

Accounting

Accounting is the action of recording a user's activities on the device for the purposes of billing and/or security. It follows the authentication and authorization actions. If the authentication and authorization is not performed via TACACS+, there are no TACACS+ accounting messages sent out.

You can use TACACS+ to record and track software logins, configuration changes, and interactive commands.

The GbESM supports the following TACACS+ accounting attributes:

- protocol (console/telnet/ssh/http)
- start_time
- stop_time
- elapsed_time
- disc-cause

Note – When using the Browser-Based Interface, the TACACS+ Accounting Stop records are sent only if the Quit button on the browser is clicked.

Command Authorization and Logging

When TACACS+ Command Authorization is enabled (`/cfg/sys/tacacs/cauth ena`), BLADEOS configuration commands are sent to the TACACS+ server for authorization. When TACACS+ Command Logging is enabled (`/cfg/sys/tacacs/clog ena`), BLADEOS configuration commands are logged on the TACACS+ server.

The following examples illustrate the format of BLADEOS commands sent to the TACACS+ server:

```
authorization request, cmd=cfgtree, cmd-arg=/cfg/l3/if
accounting request, cmd=/cfg/l3/if, cmd-arg=1
authorization request, cmd=cfgtree, cmd-arg=/cfg/l3/if/ena
accounting request, cmd=/cfg/l3/if/ena
authorization request, cmd=cfgtree, cmd-arg=/cfg/l3/if/addr
accounting request, cmd=/cfg/l3/if/addr, cmd-arg=10.90.90.91

authorization request, cmd=apply
accounting request, cmd=apply
```

The following rules apply to TACACS+ command authorization and logging:

- Only commands from a Console, Telnet, or SSH connection are sent for authorization and logging. SNMP, BBI, or file-copy commands (for example, TFTP or sync) are not sent.
- Only leaf-level commands are sent for authorization and logging. For example, `/cfg` is not sent, but `/cfg/sys/tacacs/cauth` is sent.
- The full path of each command is sent for authorization and logging. For example:
`/cfg/sys/tacacs/cauth`
- Command arguments are not sent for authorization. For `/cauth ena`, only `/cauth` is authorized. The command and its first argument are logged, if issued on the same line.
- Only executed commands are logged.
- Invalid commands are checked by BLADEOS, and are not sent for authorization or logging.
- Authorization is performed on each leaf-level command separately. If the user issues multiple commands at once, each command is sent separately as a full path.
- Only the following global commands are sent for authorization and logging:

```

apply
diff
ping
revert
save
telnet
traceroute

```

TACACS+ Password Change

BLADEOS supports TACACS+ password change. When enabled, users can change their passwords after successful TACACS+ authorization. Use the command `/cfg/sys/tacacs/passch` to enable or disable this feature.

Use the following commands to change the password for the primary and secondary TACACS+ servers:

<code>>> # /cfg/sys/tacacs/chpass_p</code>	<i>(Change primary TACACS+ password)</i>
<code>>> # /cfg/sys/tacacs/chpass_s</code>	<i>(Change secondary TACACS+ password)</i>

Configuring TACACS+ Authentication on the Switch

1. Turn TACACS+ authentication on, then configure the Primary and Secondary TACACS+ servers.

```
>> Main# /cfg/sys/tacacs+                (Select the TACACS+ Server menu)
>> TACACS+ Server# on                    (Turn TACACS+ on)
Current status: OFF
New status:      ON
>> TACACS+ Server# prisrv 10.10.1.1      (Enter primary server IP)
Current primary TACACS+ server:      0.0.0.0
New pending primary TACACS+ server: 10.10.1.1
>> TACACS+ Server# secsrv 10.10.1.2      (Enter secondary server IP)
Current secondary TACACS+ server:    0.0.0.0
New pending secondary TACACS+ server: 10.10.1.2
```

2. Configure the TACACS+ secret and second secret.

```
>> TACACS+ Server# secret
Enter new TACACS+ secret: <1-32 character secret>
>> TACACS+ Server# secret2
Enter new TACACS+ second secret: <1-32 character secret>
```



Caution—If you configure the TACACS+ secret using any method other than a direct console connection or through a secure management module connection, the secret may be transmitted over the network as clear text.

3. If desired, you may change the default TCP port number used to listen to TACACS+. The well-known port for TACACS+ is 49.

```
>> TACACS+ Server# port
Current TACACS+ port: 49
Enter new TACACS+ port [1-65000]: <port number>
```

4. Configure the number of retry attempts, and the timeout period.

```
>> TACACS+ Server# retries
Current TACACS+ server retries: 3
Enter new TACACS+ server retries [1-3]: <server retries>
>> TACACS+ Server# time
Current TACACS+ server timeout: 5
Enter new TACACS+ server timeout [4-15]: <timeout period in minutes>
```

5. Configure custom privilege-level mapping (optional).

```

>> TACACS+ Server# usermap 2
Current privilege mapping for remote privilege 2: not set
Enter new local privilege mapping: user
>> TACACS+ Server# usermap 3 user
>> TACACS+ Server# usermap 4 user
>> TACACS+ Server# usermap 5 oper

```

6. Apply and save the configuration.

LDAP Authentication and Authorization

BLADEOS supports the LDAP (Lightweight Directory Access Protocol) method to authenticate and authorize remote administrators to manage the switch. LDAP is based on a client/server model. The switch acts as a client to the LDAP server. A remote user (the remote administrator) interacts only with the switch, not the back-end server and database.

LDAP authentication consists of the following components:

- A protocol with a frame format that utilizes TCP over IP
- A centralized server that stores all the user authorization information
- A client, in this case, the switch

Each entry in the LDAP server is referenced by its Distinguished Name (DN). The DN consists of the user-account name concatenated with the LDAP domain name. If the user-account name is John, the following is an example DN:

```
uid=John,ou=people,dc=domain,dc=com
```

Configuring the LDAP Server

GbESM user groups and user accounts must reside within the same domain. On the LDAP server, configure the domain to include GbESM user groups and user accounts, as follows:

- User Accounts:
 - Use the *uid* attribute to define each individual user account.
- User Groups:
 - Use the *members* attribute in the *groupOfNames* object class to create the user groups. The first word of the common name for each user group must be equal to the user group names defined in the GbESM, as follows:
 - admin
 - oper
 - user

Configuring LDAP Authentication on the Switch

1. Turn LDAP authentication on, then configure the Primary and Secondary LDAP servers.

```
>> Main# /cfg/sys/ldap                               (Select the LDAP Server menu)
>> LDAP Server# on                                   (Turn LDAP on)
Current status: OFF
New status:     ON
>> LDAP Server# prisrv 10.10.1.1                   (Enter primary server IP)
Current primary LDAP server: 0.0.0.0
New pending primary LDAP server: 10.10.1.1
>> LDAP Server# secsrv 10.10.1.2                   (Enter secondary server IP)
Current secondary LDAP server: 0.0.0.0
New pending secondary LDAP server: 10.10.1.2
```

2. Configure the domain name.

```
>> LDAP Server# domain
Current LDAP domain name: ou=people,dc=domain,dc=com
Enter new LDAP domain name: ou=people,dc=mydomain,dc=com
```

3. If desired, you may change the default TCP port number used to listen to LDAP.

The well-known port for LDAP is 389.

```
>> LDAP Server# port
Current LDAP port: 389
Enter new LDAP port [1-65000]: <port number>
```

4. Configure the number of retry attempts for contacting the LDAP server, and the timeout period.

```
>> LDAP Server# retries
Current LDAP server retries: 3
Enter new LDAP server retries [1-3]: < server retries>
>> LDAP Server# timeout
Current LDAP server timeout: 5
Enter new LDAP server timeout [4-15]: 10 (Enter the timeout period in minutes)
```

5. Apply and save the configuration.

Secure Shell and Secure Copy

Because using Telnet does not provide a secure connection for managing a GbESM, Secure Shell (SSH) and Secure Copy (SCP) features have been included for GbESM management. SSH and SCP use secure tunnels to encrypt and secure messages between a remote administrator and the switch.

SSH is a protocol that enables remote administrators to log securely into the GbESM over a network to execute management commands.

SCP is typically used to copy files securely from one machine to another. SCP uses SSH for encryption of data on the network. On a GbESM, SCP is used to download and upload the switch configuration via secure channels.

Although SSH and SCP are disabled by default, enabling and using these features provides the following benefits:

- Identifying the administrator using Name/Password
- Authentication of remote administrators
- Authorization of remote administrators
- Determining the permitted actions and customizing service for individual administrators
- Encryption of management messages
- Encrypting messages between the remote administrator and switch
- Secure copy support

The BLADEOS implementation of SSH supports both versions 1.5 and 2.0 and supports SSH clients version 1.5 - 2.x. The following SSH clients have been tested:

- SSH 1.2.23 and SSH 1.2.27 for Linux (freeware)
- SecureCRT 3.0.2 and SecureCRT 3.0.3 for Windows NT (Van Dyke Technologies, Inc.)
- F-Secure SSH 1.1 for Windows (Data Fellows)
- Putty SSH
- Cygwin OpenSSH
- Mac X OpenSSH
- Solaris 8 OpenSSH
- AxeSSH SSHPro
- SSH Communications Vandyke SSH A
- F-Secure

Configuring SSH/SCP Features on the Switch

SSH and SCP features are disabled by default. To change the SSH/SCP settings, using the following procedures.

To Enable or Disable the SSH Feature

Begin a Telnet session from the console port and enter the following commands:

```
>> # /cfg/sys/sshd/on                               (Turn SSH on)
Current status: OFF
New status: ON

>> # /cfg/sys/sshd/off                               (Turn SSH off)
Current status: ON
New status: OFF
```

To Enable or Disable SCP Apply and Save

Enter the following commands from the switch CLI to enable the SCP `putcfg_apply` and `putcfg_apply_save` commands:

```
>> # /cfg/sys/sshd/ena                               (Enable SCP apply and save)
SSHHD# apply                                         (Apply changes to start generating
                                                         RSA host and server keys)

RSA host key generation starts.....
.....
RSA host key generation completes (lasts 212549 ms)
RSA host key is being saved to Flash ROM, please don't reboot
the box immediately.
RSA server key generation starts.....
RSA server key generation completes (lasts 75503 ms)
RSA server key is being saved to Flash ROM, please don't reboot
the box immediately.
-----
Apply complete; don't forget to "save" updated configuration.

>> # /cfg/sys/sshd/dis                               (Disable SSH/SCP apply/save)
```

Configuring the SCP Administrator Password

To configure the `scpadm` (SCP Administrator) password, enter the following command (the default password is `admin`):

```
>> /cfg/sys/sshd/scpadmin
Changing SCP-only Administrator password; validation required...
Enter current administrator password: <password>
Enter new SCP-only administrator password: <new password>
Re-enter new SCP-only administrator password: <new password>
New SCP-only administrator password accepted.
```

Using SSH and SCP Client Commands

This section shows the format for using some client commands. The examples below use 205.178.15.157 as the IP address of a sample switch.

To Log In to the Switch

Syntax:

```
ssh <switch IP address>
-or-
ssh -l <login name> <switch IP address>
```

Example:

```
>> # ssh 205.178.15.157
>> # ssh -l <login-name> 205.178.15.157      (Login to the switch)
```

To Download the Switch Configuration Using SCP

Syntax:

```
scp <username>@<switch IP address>:getcfg <local filename>
```

Example:

```
>> # scp scpadmin@205.178.15.157:getcfg ad4.cfg
```

To Upload the Configuration to the Switch

Syntax:

```
scp <local filename> <username>@<switch IP address>:putcfg
```

Example:

```
>> # scp ad4.cfg scpadmin@205.178.15.157:putcfg
```

To Apply and Save the Configuration

The `apply` and `save` commands are still needed after the last command, or use the following commands:

```
>> # scp ad4.cfg scpadmin@205.178.15.157:putcfg_apply
>> # scp ad4.cfg scpadmin@205.178.15.157:putcfg_apply_save
```

- The `diff` command is automatically executed at the end of `putcfg` to notify the remote client of the difference between the new and the current configurations.
- `putcfg_apply` runs the `apply` command after the `putcfg` is done.
- `putcfg_apply_save` saves the new configuration to the flash after `putcfg_apply` is done.
- The `putcfg_apply` and `putcfg_apply_save` commands are provided because extra `apply` and `save` commands are usually required after a `putcfg`; however, an SCP session is not in an interactive mode at all.

SSH and SCP Encryption of Management Messages

The following encryption and authentication methods are supported for SSH and SCP:

- Server Host Authentication: Client RSA authenticates the switch at the beginning of every connection
- Key Exchange: RSA
- Encryption: 3DES-CBC, DES
- User Authentication: Local password authentication, RADIUS, SecurID (via RADIUS, TACACS+, for SSH only—does not apply to SCP)

Generating RSA Host and Server Keys for SSH Access

To support the SSH server feature, two sets of RSA keys (host and server keys) are required. The host key is 1024 bits and is used to identify the GbESM. The server key is 768 bits and is used to make it impossible to decipher a captured session by breaking into the GbESM at a later time.

When the SSH server is first enabled and applied, the switch automatically generates the RSA host and server keys and is stored in the FLASH memory.

Note – To configure RSA host and server keys, first connect to the GbESM through the console port (commands are not available via external Telnet connection), and enter the following commands to generate them manually.

```
>> # /cfg/sys/sshd/hkeygen           (Generates the host key)
>> # /cfg/sys/sshd/skeygen          (Generates the server key)
```

These two commands take effect immediately without the need of an apply command.

When the switch reboots, it will retrieve the host and server keys from the FLASH memory. If these two keys are not available in the flash and if the SSH server feature is enabled, the switch automatically generates them during the system reboot. This process may take several minutes to complete.

The switch can also automatically regenerate the RSA server key. To set the interval of RSA server key autogeneration, use this command:

```
>> # /cfg/sys/sshd/intrval <number of hours (0-24)>
```

A value of 0 (zero) denotes that RSA server key autogeneration is disabled. When greater than 0, the switch will autogenerate the RSA server key every specified interval; however, RSA server key generation is skipped if the switch is busy doing other key or cipher generation when the timer expires.

Note – The switch will perform only one session of key/cipher generation at a time. Thus, an SSH/SCP client will not be able to log in if the switch is performing key generation at that time. Also, key generation will fail if an SSH/SCP client is logging in at that time.

SSH/SCP Integration with Radius Authentication

SSH/SCP is integrated with RADIUS authentication. After the RADIUS server is enabled on the switch, all subsequent SSH authentication requests will be redirected to the specified RADIUS servers for authentication. The redirection is transparent to the SSH clients.

SSH/SCP Integration with TACACS+ Authentication

SSH/SCP is integrated with TACACS+ authentication. After the TACACS+ server is enabled on the switch, all subsequent SSH authentication requests will be redirected to the specified TACACS+ servers for authentication. The redirection is transparent to the SSH clients.

SecurID Support

SSH/SCP can also work with SecurID, a token card-based authentication method. The use of SecurID requires the interactive mode during login, which is not provided by the SSH connection.

Note – There is no SNMP or Browser-Based Interface (BBI) support for SecurID because the SecurID server, ACE, is a one-time password authentication and requires an interactive session.

Using SecurID with SSH

Using SecurID with SSH involves the following tasks.

- To log in using SSH, use a special username, “ace,” to bypass the SSH authentication.
- After an SSH connection is established, you are prompted to enter the username and password (the SecurID authentication is being performed now).
- Provide your username and the token in your SecurID card as a regular Telnet user.

Using SecurID with SCP

Using SecurID with SCP can be accomplished in two ways:

- Using a RADIUS server to store an administrator password.
You can configure a regular administrator with a fixed password in the RADIUS server if it can be supported. A regular administrator with a fixed password in the RADIUS server can perform both SSH and SCP with no additional authentication required.
- Using an SCP-only administrator password.
Use the command, `/cfg/sys/sshd/scpadm` to bypass the checking of SecurID.
An SCP-only administrator’s password is typically used when SecurID is used. For example, it can be used in an automation program (in which the tokens of SecurID are not available) to back up (download) the switch configurations each day.

Note – The SCP-only administrator’s password must be different from the regular administrator’s password. If the two passwords are the same, the administrator using that password will not be allowed to log in as an SSH user because the switch will recognize him as the SCP-only administrator. The switch will only allow the administrator access to SCP commands.

End User Access Control

BLADEOS allows an administrator to define end user accounts that permit end users to perform operation tasks via the switch CLI commands. Once end user accounts are configured and enabled, the switch requires username/password authentication.

For example, an administrator can assign a user, who can then log into the switch and perform operational commands (effective only until the next switch reboot).

Considerations for Configuring End User Accounts

- A maximum of 10 user IDs are supported on the switch.
- BLADEOS supports end user support for Console, Telnet, BBI, and SSHv1/v2 access to the switch.
- If RADIUS authentication is used, the user password on the Radius server will override the user password on the GbESM. Also note that the password change command on the switch only modifies the use switch password and has no effect on the user password on the Radius server. Radius authentication and user password cannot be used concurrently to access the switch.
- Passwords can be up to 128 characters in length for TACACS, RADIUS, Telnet, SSH, Console, and Web access.

Strong Passwords

The administrator can require use of Strong Passwords for users to access the GbESM. Strong Passwords enhance security because they make password guessing more difficult.

The following rules apply when Strong Passwords are enabled:

- Each passwords must be 8 to 14 characters
- Within the first 8 characters, the password:
 - must have at least one number or one symbol
 - must have both upper and lower case letters
 - cannot be the same as any four previously used passwords

The following are examples of strong passwords:

- 1234AbcXyz
- Super+User
- Exo1cet2

The administrator can choose the number of days allowed before each password expires. When a strong password expires, the user is allowed to log in one last time (last time) to change the password. A warning provides advance notice for users to change the password.

Use the Strong Password menu to configure Strong Passwords.

```
>> # /cfg/sys/access/user/strongpw
```

User Access Control Menu

The end user access control menu is located in the System access menu.

```
>> # /cfg/sys/access/user
```

Setting Up User IDs

Up to 10 user IDs can be configured in the User ID menu.

```
>> # /cfg/sys/access/user/uid 1
```

Defining User Names and Passwords

Use the User ID menu to define user names and passwords.

```
>> User ID 1 # name user1                (Assign name to user ID 1)
Current user name:
New user name:      user1
>> User ID 1 # pswd                    (Assign password to user ID 1)
Changing user1 password; validation required:
Enter current admin password: <current administrator password>
Enter new user1 password: <new user password>
Re-enter new user1 password: <new user password>
New user1 password accepted.
```

Defining a User's Access Level

The end user is by default assigned to the user access level (also known as class of service, or CoS). CoS for all user accounts have global access to all resources except for User CoS, which has access to view only resources that the user owns. For more information, see [Table 3 on page 45](#).

To change the user's level, enter the class of service `cos` command, and select one of the following options:

```
>> User ID 1 # cos <user/oper/admin>
```

Validating a User's Configuration

```
>> User ID 2 # cur
      name jane      , dis, cos user      , password valid, offline
```

Enabling or Disabling a User

An end user account must be enabled before the switch recognizes and permits login under the account. Once enabled, the switch requires any user to enter both username and password.

```
>> # /cfg/sys/access/user/uid <user ID>/ena
>> # /cfg/sys/access/user/uid <user ID>/dis
```

Listing Current Users

The `cur` command displays defined user accounts and whether or not each user is currently logged into the switch.

```
>> # /cfg/sys/access/user/cur

Usernames:
  user      - Enabled - offline
  oper      - Disabled - offline
  admin     - Always Enabled - online 1 session

Current User ID table:
  1: name jane      , ena, cos user      , password valid, online
  2: name john     , ena, cos user      , password valid, online
```

Logging In to an End User Account

Once an end user account is configured and enabled, the user can login to the switch, using the username/password combination. The level of switch access is determined by the CoS established for the end user account.

Protected Mode

Protected Mode settings allow the switch administrator to block the management module from making configuration changes that affect switch operation. The switch retains control over those functions.

The following management module functions are disabled when Protected Mode is turned **on**:

- External Ports: Enabled/Disabled
- External management over all ports: Enabled/Disabled
- Restore Factory Defaults
- New Static IP Configuration

In this release, configuration of the functions listed above are restricted to the local switch when you turn Protected Mode **on**. In future releases, individual control over each function may be added.

Note – Before you turn Protected Mode on, make sure that external management (Telnet) access to one of the switch's IP interfaces is enabled.

Use the following command to turn Protected Mode **on**: `/oper/prm/on`

If you lose access to the switch through the external ports, use the console port to connect directly to the switch, and configure an IP interface with Telnet access.

CHAPTER 2

Port-Based Network Access Control

Port-Based Network Access control provides a means of authenticating and authorizing devices attached to a LAN port that has point-to-point connection characteristics. It prevents access to ports that fail authentication and authorization. This feature provides security to ports of the GbE Switch Module (GbESM) that connect to blade servers.

The following topics are discussed in this section:

- [“Extensible Authentication Protocol over LAN” on page 64](#)
- [“EAPoL Authentication Process” on page 65](#)
- [“EAPoL Port States” on page 67](#)
- [“Guest VLAN” on page 67](#)
- [“Supported RADIUS Attributes” on page 68](#)
- [“EAPoL Configuration Guidelines” on page 69](#)

Extensible Authentication Protocol over LAN

BLADEOS can provide user-level security for its ports using the IEEE 802.1X protocol, which is a more secure alternative to other methods of port-based network access control. Any device attached to an 802.1X-enabled port that fails authentication is prevented access to the network and denied services offered through that port.

The 802.1X standard describes port-based network access control using Extensible Authentication Protocol over LAN (EAPoL). EAPoL provides a means of authenticating and authorizing devices attached to a LAN port that has point-to-point connection characteristics and of preventing access to that port in cases of authentication and authorization failures.

EAPoL is a client-server protocol that has the following components:

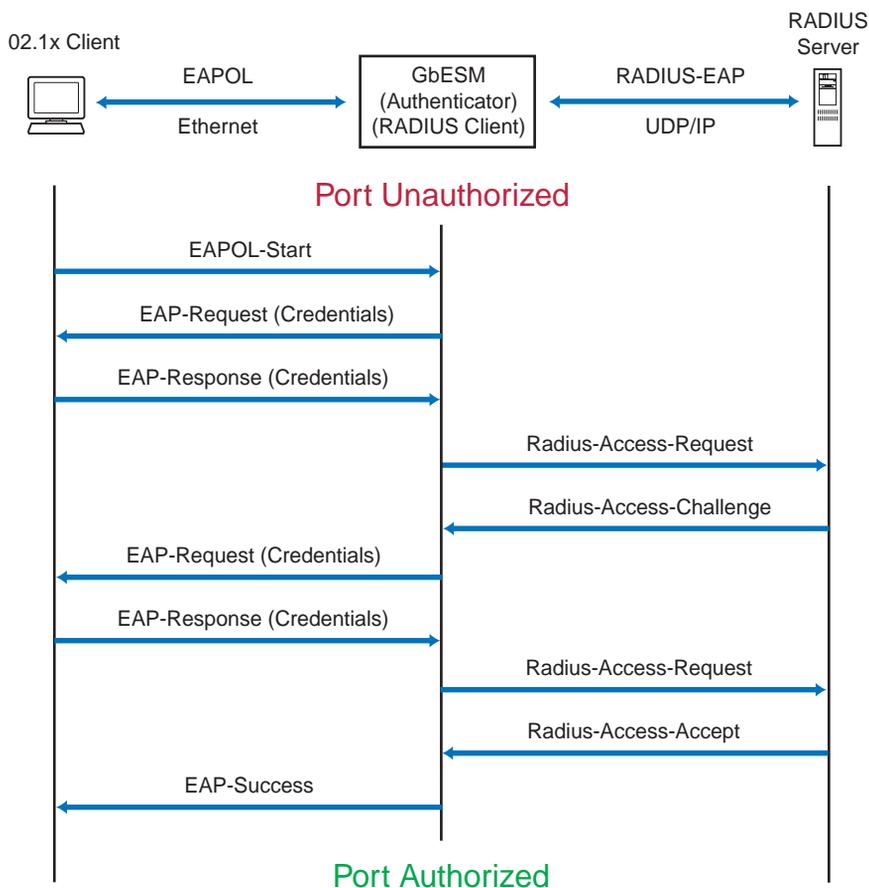
- **Supplicant or Client**
The Supplicant is a device that requests network access and provides the required credentials (user name and password) to the Authenticator and the Authenticator Server.
- **Authenticator**
The Authenticator enforces authentication and controls access to the network. The Authenticator grants network access based on the information provided by the Supplicant and the response from the Authentication Server. The Authenticator acts as an intermediary between the Supplicant and the Authentication Server: requesting identity information from the client, forwarding that information to the Authentication Server for validation, relaying the server's responses to the client, and authorizing network access based on the results of the authentication exchange. The GbESM acts as an Authenticator.
- **Authentication Server**
The Authentication Server validates the credentials provided by the Supplicant to determine if the Authenticator should grant access to the network. The Authentication Server may be co-located with the Authenticator. The GbESM relies on external RADIUS servers for authentication.

Upon a successful authentication of the client by the server, the 802.1X-controlled port transitions from unauthorized to authorized state, and the client is allowed full access to services through the port. When the client sends an EAP-Logoff message to the authenticator, the port will transition from authorized to unauthorized state.

EAPoL Authentication Process

The clients and authenticators communicate using Extensible Authentication Protocol (EAP), which was originally designed to run over PPP, and for which the IEEE 802.1X Standard has defined an encapsulation method over Ethernet frames, called EAP over LAN (EAPoL). [Figure 4](#) shows a typical message exchange initiated by the client.

Figure 4 Authenticating a Port Using EAPoL



EAPoL Message Exchange

During authentication, EAPoL messages are exchanged between the client and the GbESM authenticator, while RADIUS-EAP messages are exchanged between the GbESM authenticator and the RADIUS server.

Authentication is initiated by one of the following methods:

- GbESM authenticator sends an EAP-Request/Identity packet to the client
- Client sends an EAPoL-Start frame to the GbESM authenticator, which responds with an EAP-Request/Identity frame.

The client confirms its identity by sending an EAP-Response/Identity frame to the GbESM authenticator, which forwards the frame encapsulated in a RADIUS packet to the server.

The RADIUS authentication server chooses an EAP-supported authentication algorithm to verify the client's identity, and sends an EAP-Request packet to the client via the GbESM authenticator. The client then replies to the RADIUS server with an EAP-Response containing its credentials.

Upon a successful authentication of the client by the server, the 802.1X-controlled port transitions from unauthorized to authorized state, and the client is allowed full access to services through the controlled port. When the client later sends an EAPoL-Logoff message to the GbESM authenticator, the port transitions from authorized to unauthorized state.

If a client that does not support 802.1X connects to an 802.1X-controlled port, the GbESM authenticator requests the client's identity when it detects a change in the operational state of the port. The client does not respond to the request, and the port remains in the unauthorized state.

Note – When an 802.1X-enabled client connects to a port that is not 802.1X-controlled, the client initiates the authentication process by sending an EAPoL-Start frame. When no response is received, the client retransmits the request for a fixed number of times. If no response is received, the client assumes the port is in authorized state, and begins sending frames, even if the port is unauthorized.

EAPoL Port States

The state of the port determines whether the client is granted access to the network, as follows:

- **Unauthorized**
While in this state the port discards all ingress and egress traffic except EAP packets.
- **Authorized**
When the client is successfully authenticated, the port transitions to the authorized state allowing all traffic to and from the client to flow normally.
- **Force Unauthorized**
You can configure this state that denies all access to the port.
- **Force Authorized**
You can configure this state that allows full access to the port.

Use the 802.1X Global Configuration Menu (`/cfg/12/8021x/global`) to configure 802.1X authentication for all ports in the switch. Use the 802.1X Port Menu (`/cfg/12/8021x/port <x>`) to configure a single port.

Guest VLAN

The guest VLAN provides limited access to unauthenticated ports. Use the following command to configure a guest VLAN: `/cfg/12/8021x/global/gvlan`

Client ports that have not received an EAPoL response are placed into the Guest VLAN, if one is configured on the switch. Once the port is authenticated, it is moved from the Guest VLAN to its configured VLAN.

When Guest VLAN enabled, the following considerations apply while a port is in the unauthenticated state:

- The port is placed in the guest VLAN.
- The Port VLAN ID (PVID) is changed to the Guest VLAN ID.
- Port tagging is disabled on the port.

Supported RADIUS Attributes

The 802.1X Authenticator relies on external RADIUS servers for authentication with EAP. [Table 7](#) lists the RADIUS attributes that are supported as part of RADIUS-EAP authentication based on the guidelines specified in Annex D of the 802.1X standard and RFC 3580.

Table 7 Support for RADIUS Attributes

#	Attribute	Attribute Value	A-R	A-A	A-C	A-R
1	User-Name	The value of the Type-Data field from the supplicant's EAP-Response/Identity message. If the Identity is unknown (i.e. Type-Data field is zero bytes in length), this attribute will have the same value as the Calling-Station-Id.	1	0-1	0	0
4	NAS-IP-Address	IP address of the authenticator used for Radius communication.	1	0	0	0
5	NAS-Port	Port number of the authenticator port to which the supplicant is attached.	1	0	0	0
24	State	Server-specific value. This is sent unmodified back to the server in an Access-Request that is in response to an Access-Challenge.	0-1	0-1	0-1	0
30	Called-Station-ID	The MAC address of the authenticator encoded as an ASCII string in canonical format, e.g. 000D5622E3 9F.	1	0	0	0
31	Calling-Station-ID	The MAC address of the supplicant encoded as an ASCII string in canonical format, e.g. 00034B436206.	1	0	0	0
79	EAP-Message	Encapsulated EAP packets from the supplicant to the authentication server (Radius) and vice-versa. The authenticator relays the decoded packet to both devices.	1+	1+	1+	1+
80	Message-Authenticator	Always present whenever an EAP-Message attribute is also included. Used to integrity-protect a packet.	1	1	1	1
87	NAS-Port-ID	Name assigned to the authenticator port, e.g. Server1_Port3	1	0	0	0

Legend:

RADIUS Packet Types: A-R (Access-Request), A-A (Access-Accept), A-C (Access-Challenge), A-R (Access-Reject)

RADIUS Attribute Support:

- 0 This attribute MUST NOT be present in a packet.
- 0+ Zero or more instances of this attribute MAY be present in a packet.
- 0-1 Zero or one instance of this attribute MAY be present in a packet.
- 1 Exactly one instance of this attribute MUST be present in a packet.
- 1+ One or more of these attributes MUST be present.

EAPoL Configuration Guidelines

When configuring EAPoL, consider the following guidelines:

- The 802.1X port-based authentication is currently supported only in point-to-point configurations, that is, with a single supplicant connected to an 802.1X-enabled switch port.
- 802.1X authentication is not supported on ISL ports or on any port that is part of a trunk.
- When 802.1X is enabled, a port has to be in the authorized state before any other Layer 2 feature can be operationally enabled. For example, the STG state of a port is operationally disabled while the port is in the unauthorized state.
- The 802.1X supplicant capability is not supported. Therefore, none of its ports can successfully connect to an 802.1X-enabled port of another device, such as another switch, that acts as an authenticator, unless access control on the remote port is disabled or is configured in forced-authorized mode. For example, if a GbESM is connected to another GbESM, and if 802.1X is enabled on both switches, the two connected ports must be configured in force-authorized mode.
- The 802.1X standard has optional provisions for supporting dynamic virtual LAN assignment via RADIUS tunnelling attributes, for example, Tunnel-Type (=VLAN), Tunnel-Medium-Type (=802), and Tunnel-Private-Group-ID (=VLAN id). These attributes are not supported and might affect 802.1X operations. Other unsupported attributes include Service-Type, Session-Timeout, and Termination-Action.
- RADIUS accounting service for 802.1X-authenticated devices or users is not supported.
- Configuration changes performed using SNMP and the standard 802.1X MIB will take effect immediately.

CHAPTER 3

VLANs

This chapter describes network design and topology considerations for using Virtual Local Area Networks (VLANs). VLANs are commonly used to split up groups of network users into manageable broadcast domains, to create logical segmentation of workgroups, and to enforce security policies among logical segments. The following topics are discussed in this chapter:

- [“VLANs and Port VLAN ID Numbers” on page 72](#)
- [“VLAN Tagging” on page 74](#)
- [“VLAN Topologies and Design Considerations” on page 78](#)

This section discusses how you can logically connect users and segments to a host that supports many logical segments or subnets by using the flexibility of the multiple VLAN system.

- [“VLAN Bandwith Management” on page 80](#)
- [“Protocol-Based VLANs” on page 83](#)
- [“Private VLANs” on page 88](#)

Note – Basic VLANs can be configured during initial switch configuration (see “Using the Setup Utility” in the *BLADEOS 5.1 Command Reference*). More comprehensive VLAN configuration can be done from the Command Line Interface (see “VLAN Configuration” as well as “Port Configuration” in the *BLADEOS 5.1 Command Reference*).

VLANs Overview

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.

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.

The GbE Switch Module (GbESM) supports jumbo frames with a Maximum Transmission Unit (MTU) of 9,216 bytes. Within the frame, 18 bytes are reserved for the Ethernet header and CRC trailer. The remaining 9,198 bytes comprise the packet, which includes the payload of up to 9,000 bytes, and any additional overhead such as 802.1q or VLAN tags.

Note – Jumbo frames are not supported for traffic sent to switch management interfaces.

VLANs and Port VLAN ID Numbers

VLAN Numbers

BLADEOS supports up to 1024 VLANs per switch. Even though the maximum number of VLANs supported at any given time is 1024, 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 for use by the management network, which includes internal management ports (MGT1 and MGT2) and (by default) internal ports. This configuration allows Serial over LAN (SoL) management—a feature available on certain server blades. Management functions can also be assigned to other VLANs (using the `/cfg/l2/vlan <x>/mgmt ena` command).

Use the following command to view VLAN information:

```
>> /info/l2/vlan
```

VLAN	Name	Status	Bandwidth	MGT	Ports
1	Default VLAN	ena	ena	dis	INT1-INT14 EXT1-EXT6
4095	Mgmt VLAN	ena	ena	ena	INT1-INT14 MGT1 MGT2

PVLAN	Protocol	FrameType	EtherType	Priority	Status	Ports
1	2	empty	0000	0	dis	empty

PVLAN	PVLAN-Tagged Ports
none	none

Note – The sample screens that appear in this document might differ slightly from the screens displayed by your system. Screen content varies based on the type of BladeCenter unit that you are using and the firmware versions and options that are installed.

PVID Numbers

Each port in the switch has a configurable default VLAN number, known as its *PVID*. By default, the PVID for all non-management ports is set to 1, which correlates to the default VLAN ID. The PVID for each port can be configured to any VLAN number between 1 and 4094.

Use the following CLI commands to view PVIDs:

- Port information:

```
>> /info/port
```

Alias	Port	Tag	Fast	RMON	Lrn	Fld	PVID	NAME	VLAN(s)
INT1	1	y	n	d	e	e	1	INT1	1 4095
INT2	2	y	n	d	e	e	1	INT2	1 4095
INT3	3	y	n	d	e	e	1	INT3	1 4095
INT4	4	y	n	d	e	e	1	INT4	1 4095
INT5	5	y	n	d	e	e	1	INT5	1 4095
INT6	6	y	n	d	e	e	1	INT6	1 4095
INT7	7	y	n	d	e	e	1	INT7	1 4095
INT8	8	y	n	d	e	e	1	INT8	1 4095
INT9	9	y	n	d	e	e	1	INT9	1 4095
INT10	10	y	n	d	e	e	1	INT10	1 4095
INT11	11	y	n	d	e	e	1	INT11	1 4095
INT12	12	y	n	d	e	e	1	INT12	1 4095
ISL1	13	n	n	d	e	e	1	ISL1	1
ISL2	14	n	n	d	e	e	1	ISL2	1
MGT1	15	y	n	d	e	e	4095*	MGT1	4095
MGT2	16	y	n	d	e	e	4095*	MGT2	4095
EXT1	17	n	n	d	e	e	1	EXT1	1
EXT2	18	n	n	d	e	e	1	EXT2	1
EXT3	19	n	n	d	e	e	1	EXT3	1
...									

* = PVID is tagged.

Note – The sample output that appears in this document might differ slightly from that displayed by your system. Output varies based on the type of BladeCenter unit that you are using and the firmware versions and options that are installed.

- Port Configuration:

```
>> /cfg/port INT7/pvid 7
Current port VLAN ID: 1
New pending port VLAN ID: 7
```

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 (see “VLAN Tagging” on page 74).

VLAN Tagging

BLADEOS software supports 802.1Q VLAN *tagging*, providing standards-based VLAN support for Ethernet systems.

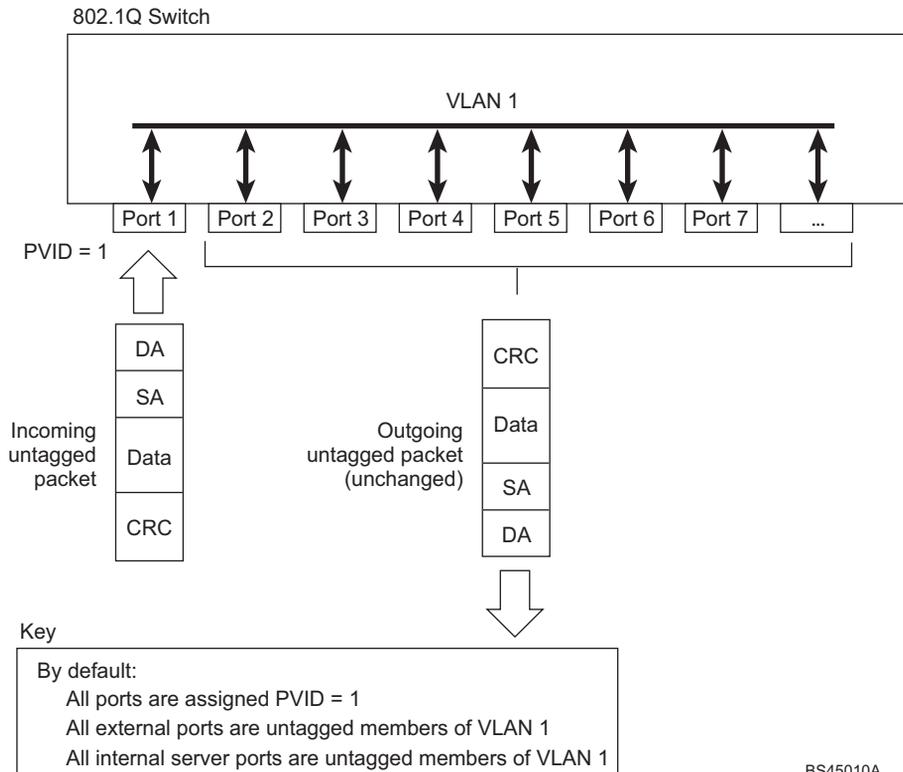
Tagging places the VLAN identifier in the frame header of a packet, allowing each port to belong to multiple VLANs. When you add a port to multiple VLANs, you also must enable tagging on that port.

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.

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. Any untagged frames received by the switch are classified with the PVID of the receiving port.
- 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.
- Untagged member—a port that has been configured as an untagged member of a specific VLAN. When an untagged frame exits the switch through an untagged member port, the frame header remains unchanged. When a tagged frame exits the switch through an untagged member port, the tag is stripped and the tagged frame is changed to an untagged frame.
- Tagged member—a port that has been configured as a tagged member of a specific VLAN. When an untagged frame exits the switch through a tagged member port, the frame header is modified to include the 32-bit tag associated with the PVID. When a tagged frame exits the switch through a tagged member port, the frame header remains unchanged (original VID remains).

Note – If a 802.1Q tagged frame is received by a port that has VLAN-tagging disabled, then the frame is dropped at the ingress port.

Figure 5 Default VLAN settings

Note – The port numbers specified in these illustrations may not directly correspond to the physical port configuration of your switch model.

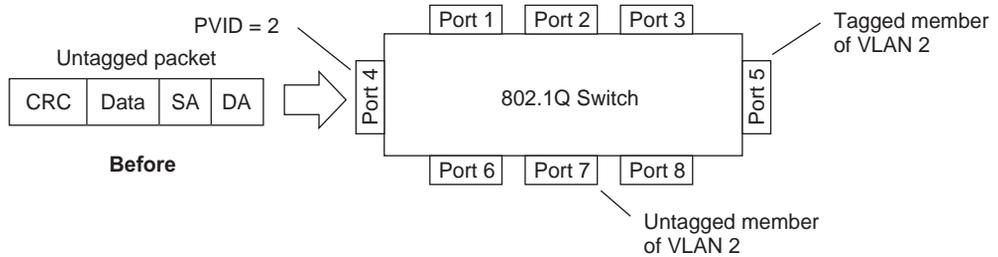
When a VLAN is configured, ports are added as members of the VLAN, and the ports are defined as either *tagged* or *untagged* (see [Figure 6](#) through [Figure 9](#)).

The default configuration settings for GbESMs 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 5](#), all incoming packets are assigned to VLAN 1 by the default port VLAN identifier (PVID = 1).

Figure 6 through Figure 9 illustrate generic examples of VLAN tagging. In Figure 6, untagged incoming packets are assigned directly to VLAN 2 (PVID = 2). Port 5 is configured as a *tagged* member of VLAN 2, and port 7 is configured as an *untagged* member of VLAN 2.

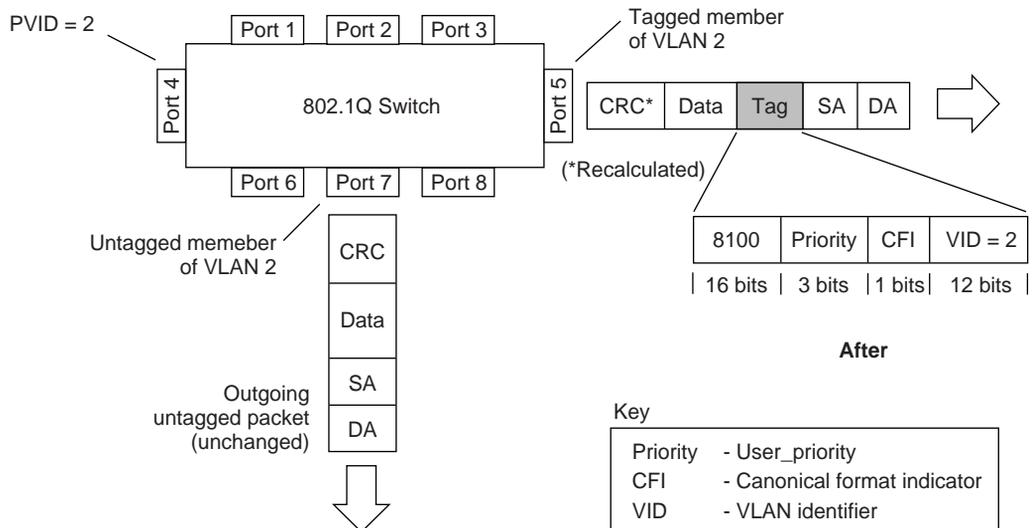
Note – The port assignments in the following figures are general examples and are not meant to match any specific GbESM.

Figure 6 Port-based VLAN assignment



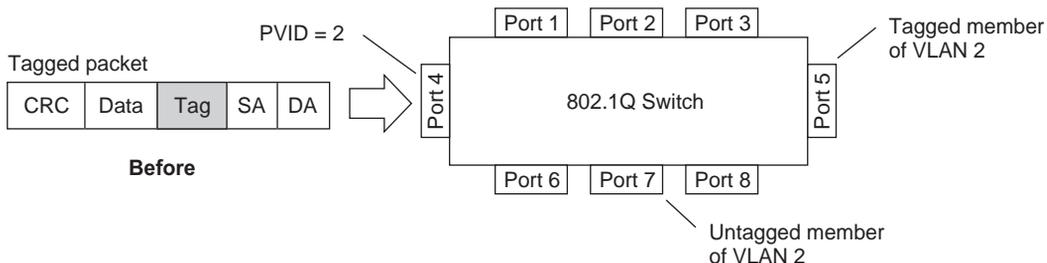
As shown in Figure 7, the untagged packet is marked (tagged) as it leaves the switch through port 5, which is configured as a tagged member of VLAN 2. The untagged packet remains unchanged as it leaves the switch through port 7, which is configured as an untagged member of VLAN 2.

Figure 7 802.1Q tagging (after port-based VLAN assignment)



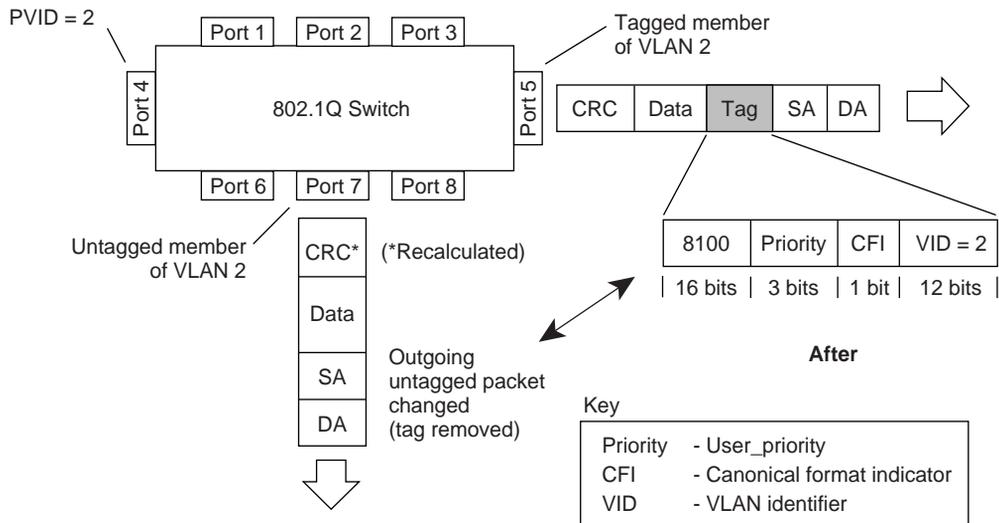
In **Figure 8**, tagged incoming packets are assigned directly to VLAN 2 because of the tag assignment in the packet. Port 5 is configured as a *tagged* member of VLAN 2, and port 7 is configured as an *untagged* member of VLAN 2.

Figure 8 802.1Q tag assignment



As shown in **Figure 9**, the tagged packet remains unchanged as it leaves the switch through port 5, which is configured as a tagged member of VLAN 2. However, the tagged packet is stripped (untagged) as it leaves the switch through port 7, which is configured as an untagged member of VLAN 2.

Figure 9 802.1Q tagging (after 802.1Q tag assignment)



Note – Set the configuration to factory default (`/boot/conf factory`) to reset all non-management ports to VLAN 1.

VLAN Topologies and Design Considerations

- By default, the BLADEOS software is configured so that tagging is disabled on all external ports, and enabled on all internal ports.
- By default, the BLADEOS software is configured so that all internal ports are members of VLAN 1. Internal ports are also members of VLAN 4095 (the default management VLAN) to allow Serial over LAN (SoL) management, a feature of certain server blades.
- By default, the BLADEOS software is configured so that the management ports (MGT1 and MGT2) are members of the default management VLAN 4095.
- Multiple management VLANs can be configured on the switch, in addition to the default VLAN 4095, using the `/cfg/l2/vlan <x>/mgmt ena` command.
- If configuring Spanning Tree Groups (STG), note that Spanning Tree Groups 2-128 may contain only one VLAN.

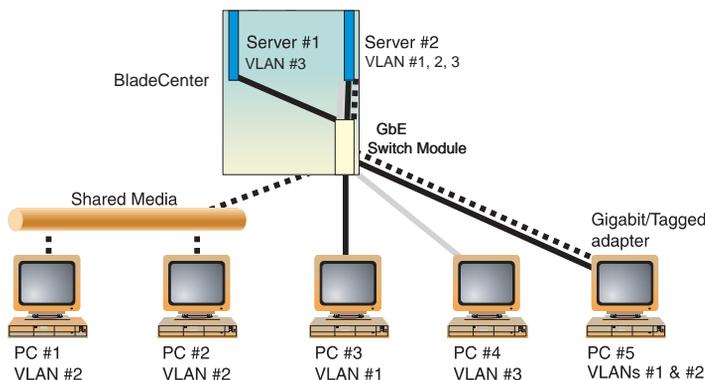
VLAN Configuration Rules

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. For more information trunk groups, see [“Port Trunking Example” on page 96](#).
- 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. For more information on configuring port mirroring, see [“Monitoring Ports for Troubleshooting” on page 253](#).
- Management VLANs must contain the management ports (MGT1 and MGT2), and can include one or more internal ports (INT:x). External ports (EXT:x) cannot be members of any management VLAN.

Example: Multiple VLANs with Tagging Adapters

Figure 10 Multiple VLANs with VLAN-Tagged Gigabit Adapters



The features of this VLAN are described below:

Component	Description
GbE Switch Module	This switch is configured for three VLANs that represent three different IP subnets. Two servers and five clients are attached to the switch.
Server #1	This server is a member of VLAN 3 and has presence in only one IP subnet. The associated internal switch port is only a member of VLAN 3, so tagging is disabled.
Server #2	This high-use server needs to be accessed from all VLANs and IP subnets. The server has a VLAN-tagging adapter installed with VLAN tagging turned on. The adapter is attached to one of the internal switch ports, that is a member of VLANs 1, 2, and 3, and has tagging enabled. Because of the VLAN tagging capabilities of both the adapter 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 PC 5. The associated external switch port has tagging disabled.
PC #3	A member of VLAN 1, this PC can only communicate with Server 2 and PC 5. The associated external switch port has tagging disabled.
PC #4	A member of VLAN 3, this PC can only communicate with Server 1 and Server 2. The associated external switch port has tagging disabled.
PC #5	A member of both VLAN 1 and VLAN 2, this PC has a VLAN-tagging Gigabit Ethernet adapter installed. It can communicate with Server 2 and PC 3 via VLAN 1, and to Server 2, PC 1 and PC 2 via VLAN 2. The associated external switch port is a member of VLAN 1 and VLAN 2, and has tagging enabled.

Note – VLAN tagging is required only on ports that are connected to other GbESMs or on ports that connect to tag-capable end-stations, such as servers with VLAN-tagging adapters.

VLAN Bandwith Management

Bandwidth management allows you to allocate uplink bandwidth to a group of servers by placing them into a particular VLAN. Configure bandwidth management on the VLAN, so the group of server ports in the VLAN cannot exceed the bandwidth limits.

Access Control Lists (ACLs) are used to provide bandwidth management on the VLAN. ACLs used for bandwidth management can be configured only in the VLAN Bandwidth Management menu (`/cfg/12/vlan <x>/bandwidth`). When you view ACL information, ACLs used for bandwidth management are visible.

Calculating the Rate

Calculate the traffic flow rate based on the number of ports in the VLAN. Each time you change the number of ports in the VLAN, it is recommended that you re-calculate and re-enter the `rate` value.

To calculate the traffic flow rate, multiply 64 kbps by the total number of ports in the VLAN. Acceptable rate values will be a multiple of the result.

Example 1: If the VLAN contains two member ports ($2 \times 64 = 128$), then the following values will be accepted:

128 (1×128)

256 ($2 \times 2 \times 128$)

...

1280 (10×128)

...

12800 (100×128)

...

2000000 (15625×128)

Example 2: If the VLAN contains fourteen member ports ($14 \times 64 = 896$), then the following values will be accepted:

896 (1×896)

1792 (2×896)

...

8960 (10×896)

...

89600 (100×896)

...

14000000 (15625×896)

Calculating the Maximum Burst Size

Calculate the maximum burst size based on the number of ports in the VLAN. Each time you change the number of ports in the VLAN, it is recommended that you re-calculate and re-enter the `mbsize` value.

To calculate the maximum burst size, multiply one of the following discreet values by the total number of ports in the VLAN: 32, 64, 128, 256, 512, 1024, 2048, 4096.

Example 1: If the VLAN contains two member ports, then the following values will be accepted:

64 (2 × 32)
128 (2 × 64)
256 (2 × 128)
512 (2 × 256)
1024 (2 × 512)
2048 (2 × 1024)
4096 (2 × 2048)
8192 (2 × 4096)

Example 2: If the VLAN contains fourteen member ports, then the following values will be accepted:

448 (14 × 32)
896 (14 × 64)
1792 (14 × 128)
3584 (14 × 256)
7168 (14 × 512)
14336 (14 × 1024)
28672 (14 × 2048)
57344 (14 × 4096)

Bandwidth Management Configuration Guidelines

Consider the following guidelines when you configure bandwidth management on a VLAN:

- For each port in the VLAN, one Access Control List (ACL) is used to provide bandwidth management. The total number of ACLs available are reduced by the total number of ports in the VLAN.
- If you add or remove ports, the VLAN bandwidth settings may be retained if the new configuration complies with the rules described in [“Calculating the Rate” on page 80](#) and [“Calculating the Maximum Burst Size” on page 81](#). However, it is recommended that you re-calculate the settings, to ensure proper operation.

Bandwidth Configuration Example

Follow this procedure to configure a bandwidth management on a VLAN.

1. Create a new VLAN and assign one or more internal ports.

```
>> /cfg/12/vlan 2                               (Select VLAN 2)
>> VLAN 2# ena                                  (Enable the VLAN)
>> VLAN 2# add int1                             (Add port INT1 to VLAN 2)
>> VLAN 2# add int2                             (Add port INT2 to VLAN 2)
```

2. Configure bandwidth management settings for the VLAN.

```
>> VLAN 2# bwidth ena                           (Enable bandwidth management)
>> Bandwidth Management# rate 12800
>> Bandwidth Management# mbsize 2048
>> Bandwidth Management# cur
    Bandwidth Management Support: enabled
    Number of internal ports: 2
    Bandwidth rate (KiloBits/sec): 12800kbps
    Bandwidth maximum burst size (KiloBits): 2048kb
```

3. Apply and save the configuration.

```
>> Bandwidth Management# apply                  (Apply the configuration)
>> Bandwidth Management# save                   (Save your changes)
```

Protocol-Based VLANs

Protocol-based VLANs (PVLANS) allow you to segment network traffic according to the network protocols in use. Traffic for supported network protocols can be confined to a particular port-based VLAN. You can give different priority levels to traffic generated by different network protocols.

With PVLAN, the switch classifies incoming packets by Ethernet protocol of the packets, not by the configuration of the ingress port. When an untagged or priority-tagged frame arrives at an ingress port, the protocol information carried in the frame is used to determine a VLAN to which the frame belongs. If a frame's protocol is not recognized as a pre-defined PVLAN type, the ingress port's PVID is assigned to the frame. When a tagged frame arrives, the VLAN ID in the frame's tag is used.

Each VLAN can contain up to eight different PVLANS. You can configure separate PVLANS on different VLANs, with each PVLAN segmenting traffic for the same protocol type. For example, you can configure PVLAN 1 on VLAN 2 to segment IPv4 traffic, and PVLAN 8 on VLAN 100 to segment IPv4 traffic.

To define a PVLAN on a VLAN, configure a PVLAN number (1-8) and specify the frame type and the Ethernet type of the PVLAN protocol. You must assign at least one port to the PVLAN before it can function. Define the PVLAN frame type and Ethernet type as follows:

- Frame type—consists of one of the following values:
 - Ether2 (Ethernet II)
 - SNAP (Subnetwork Access Protocol)
 - LLC (Logical Link Control)
- Ethernet type—consists of a 4-digit (16 bit) hex value that defines the Ethernet type. You can use common Ethernet protocol values, or define your own values. Following are examples of common Ethernet protocol values:
 - IPv4 = 0800
 - IPv6 = 86dd
 - ARP = 0806

Port-Based vs. Protocol-Based VLANs

Each VLAN supports both port-based and protocol-based association, as follows:

- The default VLAN configuration is port-based. All data ports are members of VLAN 1, with no PVLAN association.
- When you add ports to a PVLAN, the ports become members of both the port-based VLAN and the PVLAN. For example, if you add port EXT1 to PVLAN 1 on VLAN 2, the port also becomes a member of VLAN 2.
- When you delete a PVLAN, its member ports remain members of the port-based VLAN. For example, if you delete PVLAN 1 from VLAN 2, port EXT1 remains a member of VLAN 2.
- When you delete a port from a VLAN, the port is deleted from all corresponding PVLANS.

PVLAN Priority Levels

You can assign each PVLAN a priority value of 0-7, used for Quality of Service (QoS). PVLAN priority takes precedence over a port's configured priority level. If no priority level is configured for the PVLAN (priority = 0), each port's priority is used (if configured).

All member ports of a PVLAN have the same PVLAN priority level.

PVLAN Tagging

When PVLAN tagging is enabled, the switch tags frames that match the PVLAN protocol. For more information about tagging, see [“VLAN Tagging” on page 74](#).

Untagged ports must have PVLAN tagging disabled. Tagged ports can have PVLAN tagging either enabled or disabled.

PVLAN tagging has higher precedence than port-based tagging. If a port is tag enabled (`/cfg/port <x>/tag`), and the port is a member of a PVLAN, the PVLAN tags egress frames that match the PVLAN protocol.

Use the tag list command (`/cfg/l2/vlan <x>/pvlan <x>/taglist`) to define the complete list of tag-enabled ports in the PVLAN. Note that all ports not included in the PVLAN tag list will have PVLAN tagging disabled.

PVLAN Configuration Guidelines

Consider the following guidelines when you configure protocol-based VLANs:

- Each port can support up to 16 VLAN protocols.
- The GbESM can support up to 16 protocols simultaneously.
- Each PVLAN must have at least one port assigned before it can be activated.
- The same port within a port-based VLAN can belong to multiple PVLANS.
- An untagged port can be a member of multiple PVLANS.
- A port cannot be a member of different VLANs with the same protocol association.

Configuring PVLAN

Follow this procedure to configure a Protocol-based VLAN (PVLAN).

1. Create a VLAN and define the protocol type(s) supported by the VLAN.

```

>> /cfg/12/vlan 2                               (Select VLAN 2)
>> VLAN 2# ena                                  (enable VLAN 2)
Current status: disabled
New status:      enabled
>> VLAN 2# pvlan
  Enter protocol number [1-8]: 1                 (Select a protocol number)
>> VLAN 2 Protocol 1# pty
Current FrameType: empty; EtherType: empty
Enter new frame type(Ether2/SNAP/LLC): ether2   (Define the frame type)
Enter new Ether type:      0800                 (Define the Ethernet type)
New pending FrameType: Ether2; EtherType: 0800

```

2. Configure the priority value for the protocol.

```

>> VLAN 2 Protocol 1# prio 1                     (Configure the priority value)

```

3. Add member ports for this PVLAN.

```

>> VLAN 2 Protocol 1# add int1
Port INT1 is an UNTAGGED port and its current PVID is 1.
Confirm changing PVID from 1 to 2 [y/n]: y
Current ports for VLAN 2:      empty
Current ports for VLAN 1, Protocol 3:      empty
Pending new ports for VLAN 2:      INT1
Pending new ports for VLAN 2, Protocol 1:  INT1

>> VLAN 2 Protocol 1# add ext1
Port EXT1 is an UNTAGGED port and its current PVID is 1.
Confirm changing PVID from 1 to 2 [y/n]: y
Current ports for VLAN 2:      empty
Current ports for VLAN 1, Protocol 2:      empty
Pending new ports for VLAN 2:      INT1 EXT1
Pending new ports for VLAN 2, Protocol 1:  INT1 EXT1

```

Note – If VLAN tagging is turned on and the port being added to the VLAN has a different default VLAN (PVID), you will be asked to confirm changing the PVID to the current VLAN, as shown in the example.

4. Configure VLAN tagging for ports.

```

>> VLAN 2 Protocol 1# /cfg/port int1/tag ena      (Enable tagging on port)
Current VLAN tag support: disabled
New VLAN tag support:      enabled
Port INT1 changed to tagged.

>> Port INT1# /cfg/l2/vlan 2/pvlan 1/tagpvl      (Enable PVLAN tagging)
Enter port to be tagged:      int1
Ena/Dis pvlan tag:      ena
Current status: disabled
New status:      enabled
WARN: Tagging status of Port 1 in VLAN 2 will be changed for
      all protocols.
Confirm changing port's pvlan tagging status [y/n]: y

```

5. Enable the PVLAN.

```

>> VLAN 2 Protocol 1# ena      (Enable protocol-based VLAN)
Current status: disabled
New status:      enabled
>> VLAN 2 Protocol 1# apply      (Apply the configuration)
>> VLAN 2 Protocol 1# save      (Save your changes)

```

6. Verify PVLAN operation.

```

>> /info/12/vlan

VLAN          Name                Status Bandwidth MGT          Ports
-----
1             Default VLAN        ena     ena dis  INT1-INT14 EXT1-EXT6
4095          Mgmt VLAN          ena     ena ena  INT1-INT14 MGT1 MGT2

PVLAN Protocol  FrameType  EtherType  Priority  Status  Ports
-----
1       2          empty      0000       0         dis     empty

PVLAN          PVLAN-Tagged Ports
-----
none           none

```

Private VLANs

Private VLANs provide Layer 2 isolation between the ports within the same broadcast domain. Private VLANs can control traffic within a VLAN domain, and provide port-based security for host servers.

Use Private VLANs to partition a VLAN domain into sub-domains. Each sub-domain is comprised of one primary VLAN and one secondary VLAN, as follows:

- **Primary VLAN**—carries unidirectional traffic downstream from promiscuous ports. Each Private VLAN has only one primary VLAN. All ports in the Private VLAN are members of the primary VLAN.
- **Secondary VLAN**—Secondary VLANs are internal to a private VLAN domain, and are defined as follows:
 - **Isolated VLAN**—carries unidirectional traffic upstream from the host servers toward ports in the primary VLAN and the gateway. Each Private VLAN can contain only one Isolated VLAN.
 - **Community VLAN**—carries upstream traffic from ports in the community VLAN to other ports in the same community, and to ports in the primary VLAN and the gateway. Each Private VLAN can contain multiple community VLANs.

After you define the primary VLAN and one or more secondary VLANs, you map the secondary VLAN(s) to the primary VLAN.

Private VLAN Ports

Private VLAN ports are defined as follows:

- **Promiscuous**—A promiscuous port is an external port that belongs to the primary VLAN. The promiscuous port can communicate with all the interfaces, including ports in the secondary VLANs (Isolated VLAN and Community VLANs). Each promiscuous port can belong to only one Private VLAN.
- **Isolated**—An isolated port is a host port that belongs to an isolated VLAN. Each isolated port has complete layer 2 separation from other ports within the same private VLAN (including other isolated ports), except for the promiscuous ports.
 - Traffic sent to an isolated port is blocked by the Private VLAN, except the traffic from promiscuous ports.
 - Traffic received from an isolated port is forwarded only to promiscuous ports.
- **Community**—A community port is a host port that belongs to a community VLAN. Community ports can communicate with other ports in the same community VLAN, and with promiscuous ports. These interfaces are isolated at layer 2 from all other interfaces in other communities and from isolated ports within the Private VLAN.

Only external ports are promiscuous ports. Only internal ports may be isolated or community ports.

Configuration Guidelines

The following guidelines apply when configuring Private VLANs:

- Management VLANs cannot be Private VLANs. Management ports (MGT1 and MGT2) cannot be members of a Private VLAN.
- The default VLAN 1 cannot be a Private VLAN.
- Protocol-based VLANs must be disabled when you use Private VLANs.
- IGMP Snooping must be disabled on isolated VLANs.
- Each secondary port's (isolated port and community ports) PVID must match its corresponding secondary VLAN ID.
- Private VLAN ports cannot be members of a trunk group. Link Aggregation Control Protocol (LACP) must be turned off on ports within a Private VLAN.
- Ports within a secondary VLAN cannot be members of other VLANs.
- All VLANs that comprise the Private VLAN must belong to the same Spanning Tree Group.
- Blade servers connected to internal ports (secondary VLAN ports) must be configured to tag packets with the primary VLAN number.

Configuration Example

Follow this procedure to configure a Private VLAN.

1. Select a VLAN and define the Private VLAN type as primary.

```
>> /cfg/12/vlan 100                                (Select VLAN 100)
>> VLAN 100# privlan/type primary                  (Define the Private VLAN type)
Current Private-VLAN type:
Pending Private-VLAN type: primary
>> privlan# ena
```

2. Configure a secondary VLAN and map it to the primary VLAN.

```
>> /cfg/12/vlan 110                                (Select VLAN 110)
>> VLAN 110# privlan/type isolated                (Define the Private VLAN type)
Current Private-VLAN type:
Pending Private-VLAN type: isolated
>> privlan# map 100                                (Map to the primary VLAN)
Vlan 110 is mapped to the primary vlan 100.
Vlan 110 port(s) will be added to vlan 100.
>> privlan# ena
>> privlan# apply                                (Apply the configuration)
>> privlan# save                                (Save your changes)
```

CHAPTER 4

Ports and Trunking

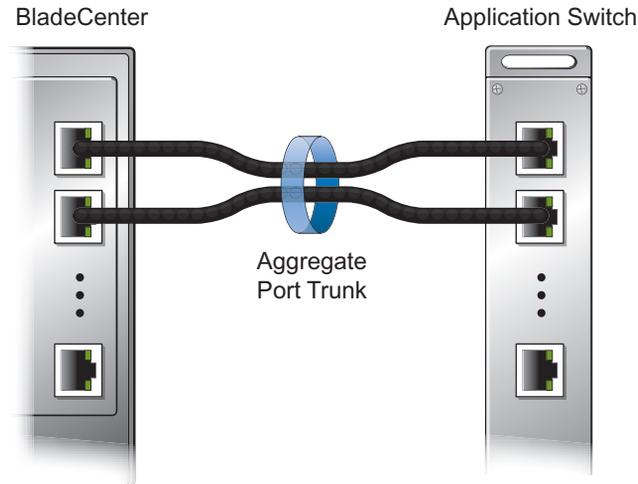
Trunk groups can provide super-bandwidth, multi-link connections between GbE Switch Modules (GbESMs) or other trunk-capable devices. A trunk group is a group of ports that act together, combining their bandwidth to create a single, larger virtual link. This chapter provides configuration background and examples for trunking multiple ports together:

- [“Trunking Overview” on page 92](#)
- [“Port Trunking Example” on page 96](#)
- [“Configurable Trunk Hash Algorithm” on page 98](#)
- [“Link Aggregation Control Protocol” on page 99](#)

Trunking Overview

When using port trunk groups between two switches, as shown in [Figure 11](#), you can create a virtual link between the switches, operating with combined throughput levels that depends on how many physical ports are included. Each GbESM supports up to 13 trunk groups, and each trunk group can contain up to 8 member ports, depending on the port type and availability.

Figure 11 Port Trunk Group



Trunk groups are also useful for connecting a GbESM to third-party devices that support link aggregation, such as Cisco routers and switches with EtherChannel technology (*not* ISL trunking technology) and Sun's Quad Fast Ethernet Adapter. Trunk Group technology is compatible with these devices when they are configured manually.

Statistical Load Distribution

Network traffic is statistically distributed between the ports in a trunk group. The BLADEOS-powered switch uses the Layer 2 MAC address information present in each transmitted frame for determining load distribution.

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.

Note – Broadcast, multicast, and DLF packets cannot be distributed across the links in a trunk based on MAC addresses. These packets will use one particular link unless another hashing method is specified (see [“Configurable Trunk Hash Algorithm” on page 98](#)).

Built-In Fault Tolerance

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 Static Trunks

When you create and enable a static 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 the configuration rules provided in the section, [“Trunk Group Configuration Rules” on page 95.](#)
2. Determine which switch ports 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. See [“Spanning Tree Group” on page 103](#) for Spanning Tree Group configuration guidelines.
4. Consider how existing VLANs will be affected by the addition of a trunk.

Inter-Switch Link

When the GbESM resides in a BladeCenter HT chassis (BCHT), internal port 13 (ISL1) and internal port 14 (ISL2) are statically-configured as members of Trunk Group 13. These ports can provide an Inter-Switch Link (ISL) between two GbESMs in the chassis. The ISL provides fixed links between switch modules.

The ISL trunk configuration is as follows:

ISL1 on Bay 1 GbESM connects to ISL1 on Bay 2 GbESM

ISL2 on Bay 1 GbESM connects to ISL2 on Bay 2 GbESM

ISL1 on Bay 3 GbESM connects to ISL1 on Bay 4 GbESM

ISL2 on Bay 3 GbESM connects to ISL2 on Bay 4 GbESM

The default configuration for the ISL ports is **disabled**. The ISL Trunk Group is **enabled**, and cannot be disabled or deleted. You can add up to 6 external ports to each ISL trunk (for a total of up to 8 ports), but no internal ports can be added.

When the ISL option is detected, the GbESM includes the ISL ports in its configuration and status menus. For example:

```
>> /info/port
```

Alias	Port	Tag	Fast	Ln	Fld	PVID	NAME	VLAN(s)
INT1	1	y	n	e	e	1	INT1	1 4095
INT2	2	y	n	e	e	1	INT2	1 4095
...								
INT11	11	y	n	e	e	1	INT11	1 4095
INT12	12	y	n	e	e	1	INT12	1 4095
ISL1	13	y	n	e	e	1	ISL1	1
ISL2	14	y	n	e	e	1	ISL2	1
MGT1	15	y	n	e	e	4095*	MGT	4095
MGT2	16	y	n	e	e	4095*	MGT	4095
EXT1	17	n	n	e	e	1	EXT1	1
...								

The trunk information command displays information about the ISL trunk, as follows:

```
>> # /info/l2/trunk
Trunk group 11: Enabled
Protocol - Static
port state:
  ISL1: STG 1 DOWN
Reminder: Port 13 needs to be enabled.
  ISL2: STG 1 DOWN
Reminder: Port 14 needs to be enabled.
```

Trunk Group Configuration Rules

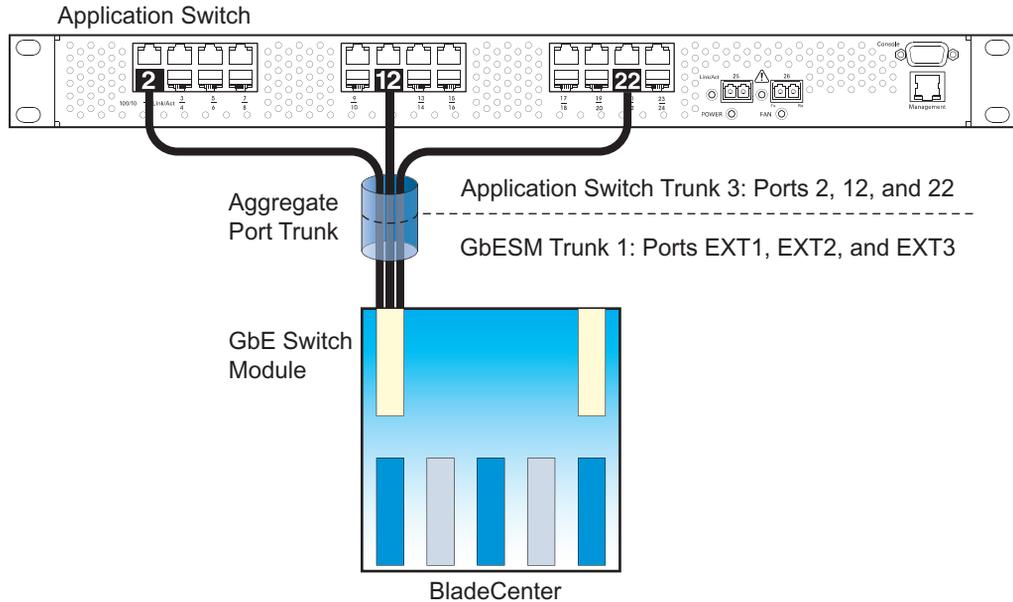
The trunking feature operates according to specific configuration rules. When creating trunks, consider the following rules that determine how a trunk group reacts in any network topology:

- All trunks must originate from one device, and lead to one destination device. For example, you cannot combine a link from Server 1 and a link from Server 2, into one trunk group.
- Any physical switch port can belong to only one trunk group.
- Depending on port availability, the switch supports up to 8 ports in each trunk group.
- Internal (INT x) and external ports (EXT x) cannot become members of the same trunk group.
- Trunking from third-party devices must comply with Cisco[®] EtherChannel[®] technology.
- All trunk member ports 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 following command (`/cfg/12/trunk <x>/ena`). 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 (STG). 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.
- 802.1X authentication is not supported on ISL ports or on any port that is part of a trunk.
- 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.
- All ports in static trunks must have the same link configuration (speed, duplex, flow control).

Port Trunking Example

In the example below, three ports are trunked between two switches.

Figure 12 Port Trunk Group Configuration Example



Prior to configuring each switch in the above 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 *BLADEOS Command Reference*.

1. Connect the switch ports that will be members in the trunk group.
2. Configure the trunk using these steps on the GbESM:
 - a. Define a trunk group.

```

>> # /cfg/12/trunk 1                (Select trunk group 1)
>> Trunk group 1# add EXT1          (Add port EXT1 to 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# ena                (Enable trunk group 1)
```

- b. Apply and verify the configuration.

```

>> Trunk group 1# apply              (Make your changes active)
>> Trunk group 1# cur                (View current trunking configuration)
```

Examine the resulting information. If any settings are incorrect, make appropriate changes.

- c. Save your new configuration changes.

```

>> Trunk group 1# save                (Save for restore after reboot)
```

3. Repeat the process on the other switch.

```

>> # /cfg/12/trunk 3                (Select trunk group 3)
>> Trunk group 3# add 2              (Add port 2 to trunk group 3)
>> Trunk group 3# add 12             (Add port 12 to trunk group 3)
>> Trunk group 3# add 22             (Add port 22 to trunk group 3)
>> Trunk group 3# ena                (Enable trunk group 3)
>> Trunk group 3# apply              (Make your changes active)
>> Trunk group 3# cur                (View current trunking configuration)
>> Trunk group 3# save                (Save for restore after reboot)
```

Trunk group 1 (on the GbESM) is now connected to trunk group 3 on the Application Switch.

Note – In this example, a GbESM and an application switch are used. If a third-party device supporting link aggregation is used (such as Cisco routers and switches with EtherChannel technology or Sun's Quad Fast Ethernet Adapter), trunk groups on the third-party device should be configured manually. Connection problems could arise when using automatic trunk group negotiation on the third-party device.

4. Examine the trunking information on each switch.

```

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

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

Configurable Trunk Hash Algorithm

This feature allows you to configure parameters for the GbESM trunk hash algorithm, instead of using the default values.

Use the Trunk Hash commands(`/cfg/12/thash`) to configure new default behavior for Layer 2 traffic and Layer 3 traffic. The trunk hash settings affect both static trunks and LACP trunks.

You can select a minimum of one parameter or a maximum of two parameters to create one of the following configurations:

- Source IP (SIP)
- Destination IP (DIP)
- Source MAC (SMAC)
- Destination MAC (DMAC)
- Source IP (SIP) + Destination IP (DIP)
- Source MAC (SMAC) + Destination MAC (DMAC)

Note – Hashing based on SMAC or DMAC does not apply to broadcast, multicast, or DLF packets. These packets will always use one particular link in the trunk unless different hashing criteria is used.

Link Aggregation Control Protocol

LACP Overview

Link Aggregation Control Protocol (LACP) is an IEEE 802.3ad standard for grouping several physical ports into one logical port (known as a dynamic trunk group or Link Aggregation group) with any device that supports the standard. Please refer to IEEE 802.3ad-2002 for a full description of the standard.

IEEE 802.3ad allows standard Ethernet links to form a single Layer 2 link using the Link Aggregation Control Protocol (LACP). Link aggregation is a method of grouping physical link segments of the same media type and speed in full duplex, and treating them as if they were part of a single, logical link segment. If a link in a LACP trunk group fails, traffic is reassigned dynamically to the remaining link or links of the dynamic trunk group.

The GbESM supports up to 13 LACP trunks, each with up to 8 ports.

Note – LACP implementation in BLADEOS does not support the Churn machine, an option used to detect if the port is operable within a bounded time period between the actor and the partner. Only the Marker Responder is implemented, and there is no marker protocol generator.

A port's Link Aggregation Identifier (LAG ID) determines how the port can be aggregated. The Link Aggregation ID (LAG ID) is constructed mainly from the *system ID* and the port's *admin key*, as follows:

- **System ID** is an integer value based on the switch's MAC address and the system priority assigned in the CLI.
- **Admin key.** A port's Admin key is an integer value (1 - 65535) that you can configure in the CLI. Each GbESM port that participates in the same LACP trunk group must have the same *admin key* value. The Admin key is *local significant*, which means the partner switch does not need to use the same Admin key value.

For example, consider two switches, an Actor (the GbESM) and a Partner (another switch), as shown in [Table 8](#).

Table 8 Actor vs. Partner LACP configuration

Actor Switch	Partner Switch 1	Partner Switch 2
Port EXT1 (admin key = 100)	Port 1 (admin key = 50)	
Port EXT2 (admin key = 100)	Port 2 (admin key = 50)	

In the configuration shown in [Table 8](#), Actor switch ports EXT1 and EXT2 aggregate to form an LACP trunk group with Partner switch ports 1 and 2.

LACP automatically determines which member links can be aggregated and then aggregates them. It provides for the controlled addition and removal of physical links for the link aggregation.

Each port in the GbESM can have one of the following LACP modes.

- `off` (default)

The user can configure this port in to a regular static trunk group.

- `active`

The port is capable of forming an LACP trunk. This port sends LACPDU packets to partner system ports.

- `passive`

The port is capable of forming an LACP trunk. This port only responds to the LACPDU packets sent from an LACP *active* port.

Each active LACP port transmits LACP data units (LACPDUs), while each passive LACP port listens for LACPDUs. During LACP negotiation, the admin key is exchanged. The LACP trunk group is enabled as long as the information matches at both ends of the link. If the admin key value changes for a port at either end of the link, that port's association with the LACP trunk group is lost.

When the system is initialized, all ports by default are in LACP *off* mode and are assigned unique *admin keys*. To make a group of ports aggregatable, you assign them all the same *admin key*. You must set the port's LACP mode to *active* to activate LACP negotiation. You can set other port's LACP mode to passive, to reduce the amount of LACPDU traffic at the initial trunk-forming stage.

Use the `/info/12/trunk` command or the `/info/12/lacp/dump` command to check whether the ports are trunked. Static trunks are listed as trunks 1 through 8. Dynamic trunks are listed as 14 through 26.

Note – If you configure LACP on ports with 802.1X network access control, make sure the ports on both sides of the connection are properly configured for both LACP and 802.1X.

Configuring LACP

Use the following procedure to configure LACP for port EXT1 and port EXT2 to participate in link aggregation.

1. Set the LACP mode on port EXT1.

```
>> # /cfg/l2/lacp/port EXT1          (Select port EXT1)
>> LACP port EXT1# mode active      (Set port EXT1 to LACP active mode)
```

2. Define the admin key on port EXT1. Only ports with the same admin key can form a LACP trunk group.

```
>> LACP port EXT1# adminkey 100      (Set port EXT1 adminkey to 100)
Current LACP port adminkey: 17
New pending LACP port adminkey: 100
```

3. Set the LACP mode on port EXT2.

```
>> # /cfg/l2/lacp/port EXT2          (Select port EXT2)
>> LACP port EXT2# mode active      (Set port EXT2 to LACP active mode)
```

4. Define the admin key on port EXT2.

```
>> LACP port EXT2# adminkey 100      (Set port EXT2 adminkey to 100)
Current LACP port adminkey: 18
New pending LACP port adminkey: 100
```

5. Apply and verify the configuration.

```
>> LACP port EXT2# apply             (Make your changes active)
>> LACP port EXT2# cur              (View current trunking configuration)
```

6. Save your new configuration changes.

```
>> LACP port EXT2# save             (Save for restore after reboot)
```


CHAPTER 5

Spanning Tree Group

When multiple paths exist on a network, Spanning Tree Group (STG) configures the network so that a switch uses only the most efficient path. The following topics are discussed in this chapter:

- “STG Overview” on page 103
- “Bridge Protocol Data Units (BPDUs)” on page 104
- “Multiple Spanning Trees” on page 108
- “Port Fast Forwarding” on page 112
- “Fast Uplink Convergence” on page 112
- “Hot Links” on page 113

STG Overview

Spanning Tree Group (STG) detects and eliminates logical loops in a bridged or switched network. When multiple paths exist, Spanning Tree configures the network so that a switch uses only the most efficient path. If that path fails, Spanning Tree automatically sets up another active path on the network to sustain network operations.

The GbE Switch Module (GbESM) supports IEEE 802.1D Spanning Tree Protocol. It is compatible with PVST+ by configuring each STP Group in different STP instances.

Note – The GbESM also supports IEEE 802.1w Rapid Spanning Tree Protocol, and IEEE 802.1s Multiple Spanning Tree Protocol. For more information, see Chapter 6, “Rapid Spanning Tree Protocol/Multiple Spanning Tree Protocol.”

The relationship between port, trunk groups, VLANs, and Spanning Trees is shown in [Table 9](#).

Table 9 Ports, Trunk Groups, and VLANs

Switch Element	Belongs To
Port	Trunk group or One or more VLANs
Trunk group	One or more VLANs
VLAN (non-default)	One Spanning Tree Group

Note – Due to Spanning Tree’s sequence of listening, learning, and forwarding or blocking, lengthy delays may occur. You can use Port Fast Forwarding (`/cfg/port <x>/fastfwd/ena`) to permit a port that participates in Spanning Tree to bypass the Listening and Learning states and enter directly into the Forwarding state. While in the Forwarding state, the port listens to the BPDUs to learn if there is a loop and, if dictated by normal STG behavior (following priorities, and so on), the port transitions into the Blocking state. This feature permits the GbESM to interoperate well within Rapid Spanning Tree networks.

Bridge Protocol Data Units (BPDUs)

BDPU Overview

To create a Spanning Tree, the switch generates a configuration Bridge Protocol Data Unit (BPDU), which it then forwards out of its ports. All switches in the Layer 2 network participating in the Spanning Tree gather information about other switches in the network through an exchange of BPDUs.

A BPDU is a 64-byte packet that is sent out at a configurable interval, which is typically set for two seconds. The BPDU is used to establish a path, much like a “hello” packet in IP routing. BPDUs contain information about the transmitting bridge and its ports, including bridge and MAC addresses, bridge priority, port priority, and path cost. If the ports are tagged, each port sends out a special BPDU containing the tagged information.

The generic action of a switch on receiving a BPDU is to compare the received BPDU to its own BPDU that it will transmit. If the received BPDU is better than its own BPDU, it will replace its BPDU with the received BPDU. Then, the switch adds its own bridge ID number and increments the path cost of the BPDU. The switch uses this information to block any necessary ports.

Determining the Path for Forwarding BPDUs

When determining which port to use for forwarding and which port to block, the GbE Switch Module uses information in the BPDU, including each bridge priority ID. A technique based on the “lowest root cost” is then computed to determine the most efficient path for forwarding.

Bridge Priority

The bridge priority parameter controls which bridge on the network is the STG root bridge. To make one switch become 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. Use the following command to configure the bridge priority:

```
/cfg/l2/stg <x>/brg/prio
```

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. Use the following command to configure the port priority:

```
/cfg/l2/stg <x>/port <x>/prio
```

Port Path Cost

The port path cost assigns lower values to high-bandwidth ports, such as Gigabit Ethernet, to encourage their use. The cost of a port also depends on whether the port operates at full-duplex (lower cost) or half-duplex (higher cost). For example, if a 100-Mbps (Fast Ethernet) link has a “cost” of 10 in half-duplex mode, it will have a cost of 5 in full-duplex mode. The objective is to use the fastest links so that the route with the lowest cost is chosen. A value of 0 (the default) indicates that the default cost will be computed for an auto-negotiated link or trunk speed.

Spanning Tree Group Guidelines

This section provides important information on configuring Spanning Tree Groups (STGs):

Adding a VLAN to a Spanning Tree Group

- If no VLANs exist beyond the default VLAN 1 see [“Creating a VLAN” on page 106](#) for information on adding ports to VLANs.
- Add the VLAN to the STG using the following command:

```
/cfg/12/stg <STG number>/add <VLAN number>
```

Note – To ensure proper operation with switches that use Cisco Per VLAN Spanning Tree (PVST+), you must either create a separate STG for each VLAN, or manually add all associated VLANs into a single STG.

Creating a VLAN

When you create a VLAN, that VLAN automatically belongs to STG 1, the default STG. If you want the VLAN in another STG, you must move the VLAN by assigning it to another STG.

- Move a newly created VLAN to an existing STG by following this order:
 - Create the VLAN
 - Add the VLAN to an existing STG
- VLANs must be contained *within* a single STG; a VLAN cannot span multiple STGs. By confining VLANs within a single STG, you avoid problems with spanning tree blocking ports and causing a loss of connectivity within the VLAN. When a VLAN spans multiple switches, it is recommended that the VLAN remain within the same Spanning Tree Group (have the same STG ID) across all the switches.
- If ports are tagged, all trunked ports can belong to multiple STGs.
- A port that is not a member of any VLAN cannot be added to any STG. The port must be added to a VLAN, and that VLAN added to the desired STG.

Rules for VLAN Tagged Ports

- Tagged ports can belong to more than one STG, but untagged ports can belong to only one STG.
- When a tagged port belongs to more than one STG, the egress BPDUs are tagged to distinguish the BPDUs of one STG from those of another STG.
- An untagged port cannot span multiple STGs.

Adding and Removing Ports from STGs

- When you add a port to a VLAN that belongs to an STG, the port is also added to the STG. However, if the port you are adding is an untagged port and is already a member of an STG, that port will not be added to an additional STG because an untagged port cannot belong to more than one STG.

For example, assume that VLAN 1 belongs to STG 1. You add an untagged port, port 1, that does not belong to any STG to VLAN 1, and port 1 will become part of STG 1.

If you add *untagged* port 5 (which is a member to STG 2) to STG 1, the switch will prompt you to change the PVID from 2 to 1:

```
"Port 5 is an UNTAGGED port and its current PVID is 2.
Confirm changing PVID from 2 to 1 [y/n]:" y
```

- When you remove a port from VLAN that belongs to an STG, that port will also be removed from the STG. However, if that port belongs to another VLAN in the same STG, the port remains in the STG.

As an example, assume that port 1 belongs to VLAN 1, and VLAN 1 belongs to STG 1. When you remove port 1 from VLAN 1, port 1 is also removed from STG 1.

However, if port 1 belongs to both VLAN 1 and VLAN 2 and both VLANs belong to STG 1, removing port 1 from VLAN 1 does not remove port 1 from STG 1 because VLAN 2 is still a member of STG 1.

- An STG cannot be deleted, only disabled. If you disable the STG while it still contains VLAN members, Spanning Tree will be off on all ports belonging to that VLAN.

The relationship between port, trunk groups, VLANs, and Spanning Trees is shown in [Table 9](#).

Multiple Spanning Trees

Each GbE Switch Module supports a maximum of 128 Spanning Tree Groups (STGs). Multiple STGs provide multiple data paths, which can be used for load-balancing and redundancy.

You enable load balancing between two GbESMs using multiple STGs by configuring each path with a different VLAN and then assigning each VLAN to a separate STG. Each STG is independent. Each STG sends its own Bridge Protocol Data Units (BPDUs), and each STG must be independently configured.

The STG, or bridge group, forms a loop-free topology that includes one or more virtual LANs (VLANs). The switch supports 128 STGs running simultaneously. The default STG 1 may contain an unlimited number of VLANs. All other STGs 2-128 may contain only one VLAN each.

Default Spanning Tree Configuration

In the default configuration, a single STG with the ID of 1 includes all non-management ports on the switch. It is called the default STG. Although ports can be added to or deleted from the default STG, the default STG (STG 1) itself cannot be deleted from the system.

All other STGs, except the default STG 1 and management STG 128, are empty and VLANs must be added by the user. However, you cannot assign ports directly to an STG. Instead, add ports to a VLAN first and then add the VLAN to the STG.

The default configuration of management STG 128 contains management VLAN 4095.

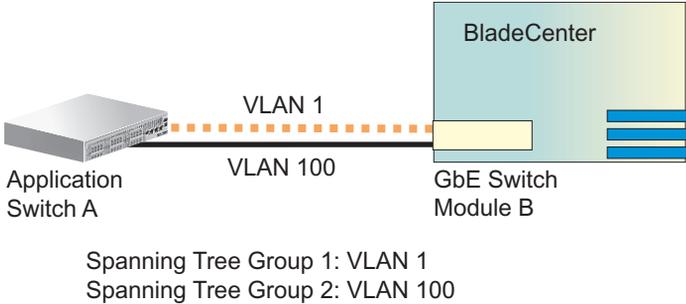
Each STG is enabled by default, and assigned an ID number from 2 to 127. By default, the spanning tree on the management ports is turned off in both STP/PVST+ mode and in MSTP/RSTP mode.

Why Do We Need Multiple Spanning Trees?

Figure 13 shows a simple example of why we need multiple Spanning Trees. Two VLANs, VLAN 1 and VLAN 100 exist between application switch A and GbESM B. If you have a single Spanning Tree Group, the switches see an apparent loop, and one VLAN may become blocked, affecting connectivity, even though no actual loop exists.

If VLAN 1 and VLAN 100 belong to different Spanning Tree Groups, then the two instances of Spanning Tree separate the topology without forming a loop. Both VLANs can forward packets between the switches without losing connectivity.

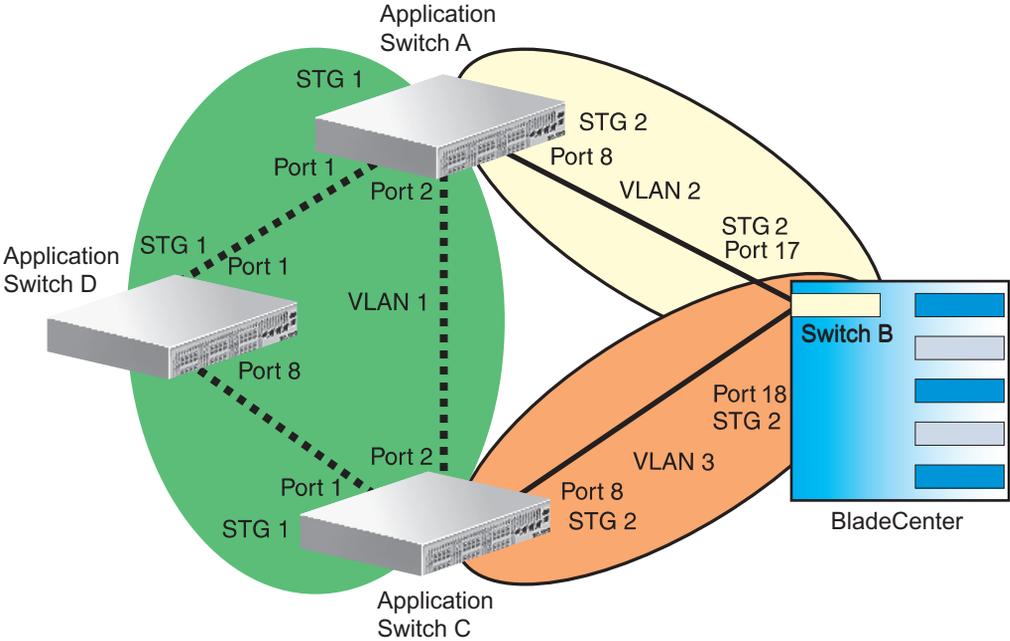
Figure 13 Using Multiple Instances of Spanning Tree Group



Switch-Centric Spanning Tree Group

In Figure 14 on page 109, VLAN 2 is shared by application switch A and GbESM B on ports 8 and 17 respectively. Application Switch A identifies VLAN 2 in Spanning Tree Group 2 and GbESM B identifies VLAN 2 in Spanning Tree Group 2. Spanning Tree Group is switch-centric—it is used to identify the VLANs participating in the Spanning Tree Groups. The Spanning Tree Group ID is not transmitted in the BPDU. Each Spanning Tree decision is based on the configuration of that switch.

Figure 14 Implementing Multiple Spanning Tree Groups



VLAN Participation in Spanning Tree Groups

The VLAN participation for each Spanning Tree Group in [Figure 14 on page 109](#) is discussed in the following sections:

■ VLAN 1 Participation

If application switch A is the root bridge, then application switch A will transmit the BPDU for VLAN 1 on ports 1 and 2. Application switch C receives the BPDU on its port 2 and application switch D receives the BPDU on its port 1. Application switch D will block port 8 or application switch C will block port 1 depending on the information provided in the BPDU.

■ VLAN 2 Participation

Application switch A, the root bridge generates another BPDU for Spanning Tree Group 2 and forwards it out from port 8. GbESM B receives this BPDU on its port 17. Port 17 on GbESM B is on VLAN 2, Spanning Tree Group 2. Because switch B has no additional ports participating in Spanning Tree Group 1, this BPDU is not forwarded to any additional ports and application switch A remains the designated root.

■ VLAN 3 Participation

For VLAN 3 you can have GbESM B or application switch C to be the root bridge. If switch B is the root bridge for VLAN 3, Spanning Tree Group 2, then switch B transmits the BPDU out from port 18. Application switch C receives this BPDU on port 8 and is identified as participating in VLAN 3, Spanning Tree Group 2. Since application switch C has no additional ports participating in Spanning Tree Group 2, this BPDU is not forwarded to any additional ports and GbESM B remains the designated root.

Configuring Multiple Spanning Tree Groups

This configuration shows how to configure the three instances of Spanning Tree Groups on the switches A, B, C, and D illustrated in [Figure 14 on page 109](#).

By default Spanning Trees 2 to 127 are empty, and Spanning Tree Group 1 contains all configured VLANs until individual VLANs are explicitly assigned to other Spanning Tree Groups. You can have only one VLAN per Spanning Tree Group except for Spanning Tree Group 1.

1. Configure the following on application switch A:

Add port 8 to VLAN 2 and define Spanning Tree Group 2 for VLAN 2.

```
>> # /cfg/12/vlan2                (Select VLAN 2 menu)
>> VLAN 2# add 8                  (Add port 8)
>> VLAN 2# ../stg 2              (Select Spanning Tree Group 2)
>> Spanning Tree Group 2# add 2   (Add VLAN 2)
```

VLAN 2 is automatically removed from Spanning Tree Group 1.

2. Configure the following on GbESM B:

Add port 17 to VLAN 2, port 18 to VLAN 3 and define Spanning Tree Group 2 for VLAN 3.

```
>> # /cfg/12/vlan2                (Select VLAN 2 menu)
>> VLAN 2# add 17                 (Add port 17)
>> VLAN 2# ../vlan3              (Select VLAN 3 menu)
>> VLAN 3# add 18                 (Add port 18)
>> VLAN 3# ../stg 2              (Select Spanning Tree Group 2)
>> Spanning Tree Group 2# add 3   (Add VLAN 3)
```

VLAN 3 is removed from Spanning Tree Group 1 and, by default, VLAN 2 remains in Spanning Tree Group 1.

Note – Each instance of Spanning Tree Group is enabled by default.

3. Configure the following on application switch C:

Add port 8 to VLAN 3 and define Spanning Tree Group 3 for VLAN 3.

```
>> # /cfg/12/vlan3                (Select VLAN 3 menu)
>> VLAN 3# add 8                  (Add port 8)
>> VLAN 3# ../stg 2              (Select Spanning Tree Group 2)
>> Spanning Tree Group 2# add 3   (Add VLAN 3)
```

VLAN 3 is automatically removed from Spanning Tree Group 1 and by default VLAN 2 remains in Spanning Tree Group 1.

Note – Application Switch D does not require any special configuration for multiple Spanning Trees, because it is configured for the default Spanning Tree Group (STG 1) only.

Port Fast Forwarding

Port Fast Forwarding permits a port that participates in Spanning Tree to bypass the Listening and Learning states and enter directly into the Forwarding state. While in the Forwarding state, the port listens to the BPDUs to learn if there is a loop and, if dictated by normal STG behavior (following priorities, etc.), the port transitions into the Blocking state. This feature permits the GbESM to interoperate well within Rapid Spanning Tree (RSTP) networks.

Use the following CLI commands to enable Port Fast Forwarding on an external port.

```
>> # /cfg/port ext1                (Select port EXT1)
>> Port EXT1# fastfwd ena          (Enable Port Fast Forwarding)
```

Remember to apply and save your configuration changes.

Fast Uplink Convergence

Fast Uplink Convergence enables the GbESM to quickly recover from the failure of the primary link or trunk group in a Layer 2 network using Spanning Tree Protocol. Normal recovery can take as long as 50 seconds, while the backup link transitions from Blocking to Listening to Learning and then Forwarding states. With Fast Uplink Convergence enabled, the GbESM immediately places the secondary path into Forwarding state, and multicasts addresses in the forwarding database (FDB) and ARP table over the secondary link so that upstream switches can learn the new path.

Configuration Guidelines

When you enable Fast Uplink Convergence, BLADEOS automatically makes the following configuration changes:

- Sets the bridge priority to 65535 so that it does not become the root switch.
- Increases the cost of all of the external ports by 3000, across all VLANs and Spanning Tree Groups. This ensures that traffic never flows through the GbESM to get to another switch unless there is no other path.

These changes are reversed if the feature is disabled.

Configuring Fast Uplink Convergence

Use the following CLI command to enable Fast Uplink Convergence on external ports.

```
>> # /cfg/12/upfast ena
```

Remember to apply and save your configuration changes.

Hot Links

For network topologies that require Spanning Tree to be turned off, Hot Links provides basic link redundancy with fast recovery.

Hot Links consists of up to 25 triggers. A trigger consists of a pair of layer 2 interfaces, each containing an individual port, trunk, or LACP adminkey. One interface is the Master, and the other is a Backup. While the Master interface is set to the active state and forwards traffic, the Backup interface is set to the standby state and blocks traffic until the Master interface fails. If the Master interface fails, the Backup interface is set to active and forwards traffic. Once the Master interface is restored, it transitions to the standby state and blocks traffic until the Backup interface fails.

You may select a physical port, static trunk, or an LACP adminkey as a Hot Link interface. Only external ports (EXT x) and Inter-Switch Link (ISL) ports can be members of a Hot Links trigger interface.

Forward Delay

The Forward Delay timer allows Hot Links to monitor the Master and Backup interfaces for link stability before selecting one interface to transition to the active state. Before the transition occurs, the interface must maintain a stable link for the duration of the Forward Delay interval.

For example, if you set the Forward delay timer to 10 seconds (`/cfg/l2/hotlink/trigger <x>/fdelay 10`), the switch will select an interface to become active only if a link remained stable for the duration of the Forward Delay period. If the link is unstable, the Forward Delay period starts again.

Pre-emption

You can configure the Master interface to resume the active state whenever it becomes available. With Hot Links preemption enabled (`/cfg/l2/hotlink/trigger <x>/preempt ena`), the Master interface transitions to the active state immediately upon recovery. The Backup interface immediately transitions to the standby state. If Forward Delay is enabled, the transition occurs when an interface has maintained link stability for the duration of the Forward Delay period.

FDB Update

Use the FDB update option to notify other devices on the network about updates to the Forwarding Database (FDB). When you enable FDB update (`/cfg/l2/hotlinks/sndfdb ena`), the switch sends multicasts of addresses in the forwarding database (FDB) over the active interface, so that other devices on the network can learn the new path. The Hot Links FDB update option uses the station update rate (`/cfg/l2/update`) to determine the rate at which to send FDB packets.

Configuration Guidelines

The following configuration guidelines apply to Hot links:

- Only external ports and inter-switch links can be configured as Hot Links.
- Ports that are configured as Hot Link interfaces must have STP disabled.
- When Hot Links is turned on, MSTP, RSTP, and PVRST must be turned off (`/cfg/l2/mrst/off`).
- When Hot Links is turned on, UplinkFast must be disabled (`/cfg/l2/upfast d`).
- A port that is a member of the Master interface cannot be a member of the Backup interface. A port that is a member of one Hot Links trigger cannot be a member of another Hot Links trigger.
- An individual port that is configured as a Hot Link interface cannot be a member of a trunk.

Configuring Hot Links

Use the following CLI commands to configure Hot Links.

```

>> # /cfg/l2/hotlink/trigger 1 ena           (Enable Hot Links Trigger 1)
>> Trigger 1# master/port ext1             (Add port to Master interface)
>> Master# ..
>> Trigger 1# backup/port ext2             (Add port to Backup interface)
>> Backup# ..
>> Trigger 1# ..
>> Hot Links# on                           (Turn on Hot Links)
>> Hot Links# apply                         (Make your changes active)
>> Hot Links# save                          (Save for restore after reboot)

```

CHAPTER 6

RSTP and MSTP

IEEE 802.1w Rapid Spanning Tree Protocol enhances the Spanning Tree Protocol to provide rapid convergence on Spanning Tree Group 1. IEEE 802.1s Multiple Spanning Tree Protocol extends the Rapid Spanning Tree Protocol, to provide both rapid convergence and load balancing in a VLAN environment.

The following topics are discussed in this chapter:

- “Rapid Spanning Tree Protocol” on page 116
 - “Port State Changes” on page 116
 - “Port Type and Link Type” on page 117
 - “RSTP Configuration Guidelines” on page 117
 - “RSTP Configuration Example” on page 118
- “Multiple Spanning Tree Protocol” on page 119
 - “MSTP Region” on page 119
 - “Common Internal Spanning Tree” on page 119
 - “MSTP Configuration Guidelines” on page 120
 - “MSTP Configuration Example” on page 120

Rapid Spanning Tree Protocol

Rapid Spanning Tree Protocol (RSTP) provides rapid convergence of the spanning tree and provides for fast re-configuration critical for networks carrying delay-sensitive traffic such as voice and video. RSTP significantly reduces the time to reconfigure the active topology of the network when changes occur to the physical topology or its configuration parameters. RSTP reduces the bridged-LAN topology to a single Spanning Tree.

For more information about Spanning Tree Protocol, see [“Spanning Tree Group” on page 103](#).

RSTP parameters are configured in Spanning Tree Group 1. STP Groups 2-128 do not apply to RSTP, and must be cleared. There are new STP parameters to support RSTP, and some values to existing parameters are different.

RSTP is compatible with devices that run 802.1D Spanning Tree Protocol. If the switch detects 802.1D BPDUs, it responds with 802.1D-compatible data units. RSTP is not compatible with Per VLAN Spanning Tree (PVST+) protocol.

Port State Changes

The port state controls the forwarding and learning processes of Spanning Tree. In RSTP, the port state has been consolidated to the following: discarding, learning, and forwarding. [Table 10](#) compares the port states between 802.1D Spanning Tree and 802.1w Rapid Spanning Trees.

Table 10 RSTP vs. STP Port states

Operational Status	STP Port State	RSTP Port State
Enabled	Blocking	Discarding
Enabled	Listening	Discarding
Enabled	Learning	Learning
Enabled	Forwarding	Forwarding
Disabled	Disabled	Discarding

Port Type and Link Type

Spanning Tree configuration includes the following parameters to support RSTP and MSTP: edge port and link type. Although these parameters are configured for Spanning Tree Groups 1-128 (`/cfg/l2/stg <x>/port <x>`), they only take effect when RSTP/MSTP is turned on.

Edge Port

A port that does not connect to a bridge is called an *edge port*. Edge ports generally connect to a server, therefore, ports INT1-INT14 should have **edge** enabled. Edge ports can start forwarding as soon as the link is up.

Edge ports do not take part in Spanning Tree, and should not receive BPDUs. If a port with **edge** enabled does receive a BPDU, it begins STP processing only if it is connected to a spanning tree bridge. If it is connected to a host, the edge port ignores BPDUs.

Link Type

The link type determines how the port behaves in regard to Rapid Spanning Tree. The link type corresponds to the duplex mode of the port, as follows:

- `p2p`: A full-duplex link to another device (point-to-point)
- `shared`: A half-duplex link is a shared segment and can contain more than one device.
- `auto`: The switch dynamically configures the link type.

RSTP Configuration Guidelines

This section provides important information about configuring Rapid Spanning Tree Groups. When RSTP is turned on, the following occurs:

- STP parameters apply only to STP Group 1.
- Only STG 1 is available. All other STGs are turned off.
- All VLANs, including management VLANs, are moved to STG 1.

RSTP Configuration Example

This section provides steps to configure Rapid Spanning Tree on the GbE Switch Module (GbESM), using the Command-Line Interface (CLI).

1. Configure port and VLAN membership on the switch.
2. Disable and clear STP groups (other than STG 1).

```
>> /cfg/12/stg 2                               (Select Spanning Tree Group 2)
>> Spanning Tree Group 2# clear                 (Clear STP Group 2 parameters)
>> Spanning Tree Group 2# off                   (Turn off STP Group 2)
```

3. Set the Spanning Tree mode to Rapid Spanning Tree.

```
>> /cfg/12/mrst                                 (Select Multiple Spanning Tree menu)
>> Multiple Spanning Tree# mode rstp           (Set mode to Rapid Spanning Tree)
>> Multiple Spanning Tree# on                  (Turn Rapid Spanning Tree on)
```

4. Configure STP Group 1 parameters.

```
>> /cfg/12/stg 1                               (Select Spanning Tree Protocol menu)
>> Spanning Tree Group 1# add 2                 (Add VLAN 2 STP Group 1)
>> Spanning Tree Group 1# apply                 (Apply the configurations)
```

Multiple Spanning Tree Protocol

IEEE 802.1s Multiple Spanning Tree extends the IEEE 802.1w Rapid Spanning Tree Protocol through multiple Spanning Tree Groups. MSTP maintains up to 32 spanning-tree instances, that correspond to STP Groups 1-32.

For more information about Spanning Tree Protocol, see [“Spanning Tree Group” on page 103](#).

In Multiple Spanning Tree Protocol (MSTP), several VLANs can be mapped to each Spanning-Tree instance. Each Spanning-Tree instance is independent of other instances. MSTP allows frames assigned to different VLANs to follow separate paths, each path based on an independent Spanning-Tree instance. This approach provides multiple forwarding paths for data traffic, enabling load-balancing, and reducing the number of Spanning-Tree instances required to support a large number of VLANs.

By default, the spanning tree on the management ports is turned off in both STP/PVST+ mode and in MSTP/RSTP mode.

MSTP Region

A group of interconnected bridges that share the same attributes is called an MST region. Each bridge within the region must share the following attributes:

- Alphanumeric name
- Revision number
- VLAN-to STG mapping scheme

MSTP provides rapid re-configuration, scalability and control due to the support of regions, and multiple Spanning-Tree instances support within each region.

Common Internal Spanning Tree

The Common Internal Spanning Tree (CIST) provides a common form of Spanning Tree Protocol, with one Spanning-Tree instance that can be used throughout the MSTP region. CIST allows the switch to interoperate with legacy equipment, including devices that run IEEE 802.1D (STP).

CIST allows the MSTP region to act as a virtual bridge to other bridges outside of the region, and provides a single Spanning-Tree instance to interact with them.

CIST port configuration includes Hello time, Edge port enable/disable, and Link Type. These parameters do not affect Spanning Tree Groups 1–128. They apply only when the CIST is used.

MSTP Configuration Guidelines

This section provides important information about configuring Multiple Spanning Tree Groups:

- When MSTP is turned on, the switch automatically moves management VLAN 4095 to the CIST. When MSTP is turned off, the switch moves VLAN 4095 from the CIST to Spanning Tree Group 128.
- When you enable MSTP, you must configure the Region Name, and a default version number of 1 is configured automatically.
- Each bridge in the region must have the same name, version number, and VLAN mapping.

MSTP Configuration Example

This section provides steps to configure Multiple Spanning Tree Protocol on the GbE Switch Module, using the Command-Line Interface (CLI).

1. Configure port and VLAN membership on the switch.
2. Set the mode to Multiple Spanning Tree, and configure MSTP region parameters.

```

>> /cfg/12/mrst                               (Select Multiple Spanning Tree menu)
>> Multiple Spanning Tree# mode mstp          (Set mode to Multiple Spanning Trees)
>> Multiple Spanning Tree# on                 (Turn Multiple Spanning Trees on)
>> Multiple Spanning Tree# name <name>       (Define the Region name)
```

3. Assign VLANs to Spanning Tree Groups.

```

>> /cfg/12/stg 2                               (Select Spanning Tree Group 2)
>> Spanning Tree Group 2# add 2                (Add VLAN 2)
```

CHAPTER 7

Link Layer Discovery Protocol

The BLADEOS software support Link Layer Discovery Protocol (LLDP). This chapter discusses the use and configuration of LLDP on the switch:

- [“LLDP Overview” on page 121](#)
- [“Enabling or Disabling LLDP” on page 122](#)
- [“LLDP Transmit Features” on page 123](#)
- [“LLDP Receive Features” on page 127](#)
- [“LLDP Example Configuration” on page 129](#)

LLDP Overview

Link Layer Discovery Protocol (LLDP) is an IEEE 802.1AB-2005 standard for discovering and managing network devices. LLDP uses Layer 2 (the data link layer), and allows network management applications to extend their awareness of the network by discovering devices that are direct neighbors of already known devices.

With LLDP, the GbESM can advertise the presence of its ports, their major capabilities, and their current status to other LLDP stations in the same LAN. LLDP transmissions occur on ports at regular intervals or whenever there is a relevant change to their status. The switch can also receive LLDP information advertised from adjacent LLDP-capable network devices.

In addition to discovery of network resources, and notification of network changes, LLDP can help administrators quickly recognize a variety of common network configuration problems, such as unintended VLAN exclusions or mis-matched port aggregation membership.

The LLDP transmit function and receive function can be independently configured on a per-port basis. The administrator can allow any given port to transmit only, receive only, or both transmit and receive LLDP information.

The LLDP information to be distributed by the GbESM ports, and that which has been collected from other LLDP stations, is stored in the switch's Management Information Base (MIB). Network Management Systems (NMS) can use Simple Network Management Protocol (SNMP) to access this MIB information. LLDP-related MIB information is read-only.

Changes, either to the local switch LLDP information or to the remotely received LLDP information, are flagged within the MIB for convenient tracking by SNMP-based management systems.

For LLDP to provide expected benefits, all network devices that support LLDP should be consistent in their LLDP configuration.

Enabling or Disabling LLDP

Global LLDP Setting

By default, LLDP is enabled on the GbESM. To turn LLDP off or on, use the following commands:

```
>> # /cfg/12/lldp/off (Turn LLDP off globally)
    or
>> # /cfg/12/lldp/on (Turn LLDP on globally)
```

Transmit and Receive Control

The GbESM can also be configured to transmit or receive LLDP information on a port-by-port basis. By default, when LLDP is globally enabled on the switch, GbESM ports transmit and receive LLDP information (see the `tx_rx` option below). To change the LLDP transmit and receive state, the following commands are available:

```
>> # /cfg/12/lldp/port <n> (Select a switch port)
>> LLDP Port# admstat tx_rx (Transmit and receive LLDP)
>> LLDP Port# admstat tx_only (Only transmit LLDP)
>> LLDP Port# admstat rx_only (Only receive LLDP)
>> LLDP Port# admstat disabled (Do not participate in LLDP)
```

To view the LLDP transmit and receive status, use the following commands:

```
>> # /cfg/12/lldp/cur (View LLDP status of all ports)
    or
>> # /cfg/12/lldp/port <n>/cur (View status of the selected port)
```

LLDP Transmit Features

Numerous LLDP transmit options are available, including scheduled and minimum transmit interval, expiration on remote systems, SNMP trap notification, and the types of information permitted to be shared.

Scheduled Interval

The GbESM can be configured to transmit LLDP information to neighboring devices once each 5 to 32768 seconds. The scheduled interval is global; the same interval value applies to all LLDP transmit-enabled ports. However, to help balance LLDP transmissions and keep them from being sent simultaneously on all ports, each port maintains its own interval clock, based on its own initialization or reset time. This allows switch-wide LLDP transmissions to be spread out over time, though individual ports comply with the configured interval.

The global transmit interval can be configured using the following command:

```
>> # /cfg/12/lldp/msgtxint <interval>
```

where *interval* is the number of seconds between LLDP transmissions. The range is 5 to 32768. The default is 30 seconds.

Minimum Interval

In addition to sending LLDP information at scheduled intervals, LLDP information is also sent when the GbESM detects relevant changes to its configuration or status (such as when ports are enabled or disabled). To prevent the GbESM from sending multiple LLDP packets in rapid succession when port status is in flux, a transmit delay timer can be configured.

The transmit delay timer represents the minimum time permitted between successive LLDP transmissions on a port. Any interval-driven or change-driven updates will be consolidated until the configured transmit delay expires.

The minimum transmit interval can be configured using the following command:

```
>> # /cfg/12/lldp/txdelay <interval>
```

where *interval* is the minimum number of seconds permitted between successive LLDP transmissions on any port. The range is 1 to one-quarter of the scheduled transmit interval (*msgtxint*), up to 8192. The default is 2 seconds.

Time-to-Live for Transmitted Information

The transmitted LLDP information is held by remote systems for a limited time. A time-to-live parameter allows the switch to determine how long the transmitted data should be held before it expires. The hold time is configured as a multiple of the configured transmission interval.

```
>> # /cfg/12/lldp/msgtxhld <multiplier>
```

where *multiplier* is a value between 2 and 10. The default value is 4, meaning that remote systems will hold the port's LLDP information for 4 x the 30-second `msgtxint` value, or 120 seconds, before removing it from their MIB.

Trap Notifications

If SNMP is enabled on the GbESM (see [“Using Simple Network Management Protocol” on page 35](#)), each port can be configured to send SNMP trap notifications whenever LLDP transmissions are sent. By default, trap notification is disabled for each port. The trap notification state can be changed using the following commands:

```
>> # /cfg/12/lldp/port <n>/snmptrap ena      (Send SNMP trap notifications)
      or
>> # /cfg/12/lldp/port <n>/snmptrap dis     (Do not send trap notifications)
```

In addition to sending LLDP information at scheduled intervals, LLDP information is also sent when the GbESM detects relevant changes to its configuration or status (such as when ports are enabled or disabled). To prevent the GbESM from sending multiple trap notifications in rapid succession when port status is in flux, a global trap delay timer can be configured.

The trap delay timer represents the minimum time permitted between successive trap notifications on any port. Any interval-driven or change-driven trap notices from the port will be consolidated until the configured trap delay expires.

The minimum trap notification interval can be configured using the following command:

```
>> # /cfg/12/lldp/notifint <interval>
```

where *interval* is the minimum number of seconds permitted between successive LLDP transmissions on any port. The range is 1 to 3600. The default is 5 seconds.

If SNMP trap notification is enabled, the notification messages can also appear in the system log. This is enabled by default. To change whether the SNMP trap notifications for LLDP events appear in the system log, use the following commands:

```
>> # /cfg/sys/syslog/log lldp ena           (Add LLDP notification to Syslog)
      or
>> # /cfg/sys/syslog/log lldp dis         (Do not log LLDP notifications)
```

Changing the LLDP Transmit State

When the port is disabled, or when LLDP transmit is turned off for the port using the `admstat` command's `rx_only` or `disabled` options (see [“Transmit and Receive Control” on page 122](#)), a final LLDP packet is transmitted with a time-to-live value of 0. Neighbors that receive this packet will remove the LLDP information associated with the GbESM port from their MIB.

In addition, if LLDP is fully disabled on a port (using `admstat disabled`) and later re-enabled, the GbESM will temporarily delay resuming LLDP transmissions on the port in order to allow the port LLDP information to stabilize. The reinitialization delay interval can be globally configured for all ports using the following command:

```
>> # /cfg/12/lldp/redelay <interval>
```

where *interval* is the number of seconds to wait before resuming LLDP transmissions. The range is between 1 and 10. The default is 2 seconds.

Types of Information Transmitted

When LLDP transmission is permitted on the port (see [“Enabling or Disabling LLDP” on page 122](#)), the port advertises the following required information in type/length/value (TLV) format:

- Chassis ID
- Port ID
- LLDP Time-to-Live

LLDP transmissions can also be configured to enable or disable inclusion of optional information, using the following command:

```
>> # /cfg/12/1ldp/port <n>/tlv/<type> {ena|dis}
```

where *type* is an LLDP information option from [Table 11](#):

Table 11 LLDP Optional Information Types

Type	Description
portdesc	Port Description
sysname	System Name
sysdescr	System Description
syscap	System Capabilities
mgmtaddr	Management Address
portvid	IEEE 802.1 Port VLAN ID
portprot	IEEE 802.1 Port and Protocol VLAN ID
vlanname	IEEE 802.1 VLAN Name
protid	IEEE 802.1 Protocol Identity
macphy	IEEE 802.3 MAC/PHY Configuration/Status, including the auto-negotiation, duplex, and speed status of the port.
powermdi	IEEE 802.3 Power via MDI, indicating the capabilities and status of devices that require or provide power over twisted-pair copper links.
linkaggr	IEEE 802.3 Link Aggregation status for the port.
framesz	IEEE 802.3 Maximum Frame Size for the port.
all	Select all optional LLDP information for inclusion or exclusion.

By default, all optional LLDP information is included in LLDP transmissions.

LLDP Receive Features

Types of Information Received

When the LLDP receive option is enabled on a port (see [“Enabling or Disabling LLDP” on page 122](#)), the port may receive the following information from LLDP-capable remote systems:

- Chassis Information
- Port Information
- LLDP Time-to-Live
- Port Description
- System Name
- System Description
- System Capabilities Supported/Enabled
- Remote Management Address

The GbESM stores the collected LLDP information in the MIB. Each remote LLDP-capable device is responsible for transmitting regular LLDP updates. If the received updates contain LLDP information changes (to port state, configuration, LLDP MIB structures, deletion), the switch will set a change flag within the MIB for convenient notification to SNMP-based management systems.

Viewing Remote Device Information

LLDP information collected from neighboring systems can be viewed in numerous ways:

- Using a centrally-connected LLDP analysis server
- Using an SNMP agent to examine the GbESM MIB
- Using the GbESM Browser-Based Interface (BBI)
- Using CLI or isCLI commands on the GbESM

Using the CLI the following command displays remote LLDP information:

```
>> # /info/12/lldp/remodev [ <index number> ]
```

To view a summary of remote information, omit the *Index number* parameter. For example:

```
>> # /info/12/lldp/remodev
LLDP Remote Devices Information
```

LocalPort	Index	Remote Chassis ID	Remote Port	Remote System Name
EXT3	1	00 18 b1 33 1d 00	23	

To view detailed information for a remote device, specify the *Index number* as found in the summary. For example, in keeping with the sample summary, to list details for the first remote device (with an Index value of 1), use the following command:

```
>> # /info/12/lldp/remodev 1
Local Port Alias: EXT3
    Remote Device Index      : 1
    Remote Device TTL        : 99
    Remote Device RxChanges  : false
    Chassis Type             : Mac Address
    Chassis Id               : 00-18-b1-33-1d-00
    Port Type                 : Locally Assigned
    Port Id                  : 23
    Port Description         : EXT7

    System Name              :
    System Description       : BNT 1/10Gb Uplink Ethernet Switch Module for IBM
                             BladeCenter, flash image: version 0.0.0,
                             boot image: version 5.1.0.12

    System Capabilities Supported : bridge, router
    System Capabilities Enabled   : bridge, router

    Remote Management Address:
        Subtype                : IPv4
        Address                 : 10.100.120.181
        Interface Subtype       : ifIndex
        Interface Number        : 128
        Object Identifier       :
```

Note – Received LLDP information can change very quickly. When using `/info/12/lldp/rx` or `/info/12/lldp/remodev` commands, it is possible that flags for some expected events may be too short-lived to be observed in the output.

Time-to-Live for Received Information

Each remote device LLDP packet includes an expiration time. If the switch port does not receive an LLDP update from the remote device before the time-to-live clock expires, the switch will consider the remote information to be invalid, and will remove all associated information from the MIB.

Remote devices can also intentionally set their LLDP time-to-live to 0, indicating to the switch that the LLDP information is invalid and should be immediately removed.

LLDP Example Configuration

1. Turn LLDP on globally.

```
>> # /cfg/12/lldp/on
```

2. Set the global LLDP timer features.

```
>> LLDP# msgtxint 30           (Schedule transmit every 30 seconds)
>> LLDP# txdelay 2           (Never more often than 2 seconds)
>> LLDP# msgtxhld 4         (Hold on remote side for 4 intervals)
>> LLDP# redelay 2         (Wait 2 seconds after reinitialization)
>> LLDP# notifint 5       (Minimum 5 seconds between traps)
```

3. Set LLDP options for each port.

```
>> LLDP# port <n>           (Select a switch port)
>> LLDP Port# admstat tx_rx (Transmit and receive LLDP)
>> LLDP Port# snmptrap ena (Enable SNMP trap notifications)
>> LLDP Port# tlv/all ena  (Transmit all optional information)
```

4. Enable syslog reporting.

```
>> # /cfg/sys/syslog/log lldp ena
```

5. Apply and Save the configuration.

6. Verify the configuration settings:

```
>> # /cfg/12/lldp/cur
```

7. View remote device information as needed.

```
>> # /info/12/lldp/remodev
    or
>> # /info/12/lldp/remodev <index number>
```


CHAPTER 8

Quality of Service

Quality of Service (QoS) features allow you to allocate network resources to mission-critical applications at the expense of applications that are less sensitive to such factors as time delays or network congestion. You can configure your network to prioritize specific types of traffic, ensuring that each type receives the appropriate QoS level.

The following topics are discussed in this section:

- [“QoS Overview” on page 131](#)
- [“Using ACL Filters” on page 133](#)
- [“Using DSCP Values to Provide QoS” on page 141](#)
- [“Using 802.Ip Priorities to Provide QoS” on page 146](#)
- [“Queuing and Scheduling” on page 147](#)

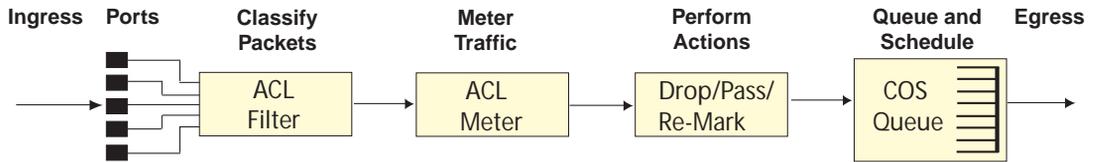
QoS Overview

QoS helps you allocate guaranteed bandwidth to critical applications, and limit bandwidth for less critical applications. Applications such as video and voice must have a certain amount of bandwidth to work correctly; using QoS, you can provide that bandwidth when necessary. Also, you can put a high priority on applications that are sensitive to timing out or those that cannot tolerate delay, assigning that traffic to a high-priority queue.

By assigning QoS levels to traffic flows on your network, you can ensure that network resources are allocated where they are needed most. QoS features allow you to prioritize network traffic, thereby providing better service for selected applications.

[Figure 15 on page 132](#) shows the basic QoS model used by the GbE Switch Module (GbESM).

Figure 15 QoS Model



The GbESM uses the Differentiated Services (DiffServ) architecture to provide QoS functions. DiffServ is described in IETF RFC 2474 and RFC 2475.

With DiffServ, you can establish policies for directing traffic. A policy is a traffic-controlling mechanism that monitors the characteristics of the traffic (for example, its source, destination, and protocol) and performs a controlling action on the traffic when certain characteristics are matched.

The GbESM can classify traffic by reading the DiffServ Code Point (DSCP) or IEEE 802.1p priority value, or by using filters to match specific criteria. When network traffic attributes match those specified in a traffic pattern, the policy instructs the GbESM to perform specified actions on each packet that passes through it. The packets are assigned to different Class of Service (COS) queues and scheduled for transmission.

The basic GbESM QoS model works as follows:

- Classify traffic:
 - Read DSCP
 - Read 802.1p Priority
 - Match ACL filter parameters
- Meter traffic:
 - Define bandwidth and burst parameters
 - Select actions to perform on in-profile and out-of-profile traffic
- Perform actions:
 - Drop packets
 - Pass packets
 - Mark DSCP or 802.1p Priority
 - Set COS queue (with or without re-marking)
- Queue and schedule traffic:
 - Place packets in one of the available COS queues
 - Schedule transmission based on the COS queue weight

Using ACL Filters

Access Control Lists (ACLs) are filters that allow you to classify and segment traffic, so you can provide different levels of service to different traffic types. Each filter defines the conditions that must match for inclusion in the filter, and also the actions that are performed when a match is made.

Summary of Packet Classifiers

The GbESM allows you to classify packets based on various parameters, such as:

- Ethernet
 - Source MAC address
 - Destination MAC address
 - VLAN number/mask
 - Ethernet type
 - Ethernet Priority, which is the IEEE 802.1p Priority
- IPv4
 - Source IP address/mask
 - Destination address/mask
 - Type of Service value
 - IP protocol number protocol number or name as shown in [Table 12](#):

Table 12 Well-Known Protocol Types

Number	Protocol Name
1	icmp
2	igmp
6	tcp
17	udp
89	ospf
112	vrrp

- TCP/UDP
 - TCP/UDP application source port as shown in [Table 13](#)
 - TCP/UDP application destination port as shown in [Table 13](#)
 - TCP/UDP flag value as shown in [Table 14](#)

Table 13 Well-Known Application Ports

TCP/UDP Number	TCP/UDP Application	TCP/UDP Number	TCP/UDP Application	TCP/UDP Number	TCP/UDP Application
20	ftp-data	79	finger	179	bgp
21	ftp	80	http	194	irc
22	ssh	109	pop2	220	imap3
23	telnet	110	pop3	389	ldap
25	smtp	111	sunrpc	443	https
37	time	119	nntp	520	rip
42	name	123	ntp	554	rtsp
43	whois	143	imap	1645, 1812	Radius
53	domain	144	news	1813	Radius Accounting
69	tftp	161	snmp	1985	hsrp
70	gopher	162	snmptrap		

Table 14 Well-Known TCP flag values

Flag	Value
URG	0x0020
ACK	0x0010
PSH	0x0008
RST	0x0004
SYN	0x0002
FIN	0x0001

- Packet format
 - Ethernet format (eth2, SNAP, LLC)
 - Ethernet tagging format
 - IP format (IPv4, IPv6)
- Egress port

Summary of ACL Actions

Actions determine how the traffic is treated. The GbESM QoS actions include the following:

- Pass or Drop
- Re-mark a new DiffServ Code Point (DSCP)
- Re-mark the 802.1p field
- Set the COS queue

ACL Order of Precedence

When multiple ACLs are assigned to a port, the order in which the ACLs are evaluated depends on the order in which the ACLs are assigned to the port. The first ACL added to the port is processed first and has the highest precedence. The last ACL added has the lowest precedence (top to bottom). For example, if ACL 3, ACL 1, and ACL 2 are assigned to a port in that order, then ACL 3 has highest precedence, followed by ACL 1, and finally ACL 2.

When a packet matches an ACL, the configured action is taken. The remaining ACLs are then considered, in order of precedence. If any ACL action contradicts the action of a preceding ACL (one with higher-precedence), the action of the lower-precedence ACL is ignored.

However, each port supports a maximum of eight precedence levels. To extend the number of ACLs that can be assigned to ports, related ACLs (those that use the same parameters for network mask and packet format) can be grouped together into ACL Blocks (see [“ACL Blocks and ACL Groups” on page 135](#)). Each ACL Block can contain up to eight ACLs, all of which share the same order of precedence.

ACL Blocks and ACL Groups

ACLs allow you to classify packets according to a particular content in the packet header, such as the source address, destination address, source port number, destination port number, and others. Once classified, packet flows can be identified for more processing.

To assist in organizing multiple ACLs and assigning them to ports, you can place ACLs into ACL Blocks and ACL Groups, thereby defining complex traffic profiles.

Each port accepts up to 127 ACL entities, comprised as follows:

■ Individual ACLs

The GbESM supports up to 4096 ACLs. Each ACL defines one filter rule for matching traffic criteria, and can include an action (such as permit or deny). For example:

ACL 1:
 VLAN = 1
 SIP = 10.10.10.1 (255.255.255.0)
 Action = permit

ACLs can be assigned to a port by itself or in combination with other ACLs, ACL Blocks, or ACL Groups. Each individually assigned ACL counts as one ACL filter entity, and occupies one of the eight precedence levels available on the port.

■ Access Control List Blocks

An Access Control List Block (ACL Block) is a collection of up to eight related ACLs. All ACLs within an ACL Block share the same set of filter criteria for network range and packet format type. For example:

ACL Block 1	
ACL 1: VLAN = 1 SIP = 10.10.10.1 (255.255.255.0) Action = permit	ACL 2: VLAN = 2 SIP = 10.10.10.2 (255.255.255.0) Action = deny
ACL Block 2	
ACL 3: PRI = 0 SIP = 20.20.20.1 (255.255.0.0) Action = deny	ACL 4: PRI = 1 SIP = 20.20.20.2 (255.255.0.0) Action = deny

Each ACL Block assigned to a port counts as a single filter entity and occupies one precedence level, regardless of how many ACLs it contains. Use ACL Blocks to group ACLs that do not require different precedence levels, thereby extending the total number of ACLs that can operate on a port.

The GbESM supports up to 4096 ACL Blocks.

■ Access Control List Groups

An Access Control List Group (ACL Group) is a collection of ACLs and ACL Blocks used for assigning a more complex traffic profile to a port. For example:

ACL Group 1	
ACL Block 1	
ACL 1: VLAN = 1 SIP = 10.10.10.1 (255.255.255.0) Action = permit	ACL 2: VLAN = 2 SIP = 10.10.10.2 (255.255.255.0) Action = deny
ACL 3: Priority = 7 DIP = 10.10.10.3 (255.255.255.0) Action = permit	

Each ACL and ACL Block in an ACL Group has its own precedence level, based on its position in the ACL Group. In the example above, ACL Group 1 has two precedence levels. The two ACLs in ACL Block 1 share the highest precedence level, while ACL 3 has a lower precedence, based on their positions in the ACL Group.

The GbESM supports up to 4096 ACL Groups.

Assigning ACLs to a Port

Once you configure an ACL, you must assign the ACL to a port. Each port can accept multiple ACLs. Note that higher-priority ACLs are considered first, and their action takes precedence over lower-priority ACLs.

When you assign an ACL to a port, the ACL acts only upon ingress traffic, not egress traffic.

To assign an ACL to a port, use the following command:

```
# /cfg/port <x>/aclqos/add acl 120
```

To assign an ACL Block to a port, use the following command:

```
# /cfg/port <x>/aclqos/add blk 2
```

To assign an ACL Group to a port, use the following command:

```
# /cfg/port <x>/aclqos/add grp 20
```

ACL Metering and Re-Marking

You can define a profile for the aggregate traffic flowing through the GbESM by configuring a QoS meter (if desired) and assigning ACL Groups to ports. When you add ACL Groups to a port, make sure they are ordered correctly in terms of precedence.

For example, consider two ACL Groups, ACL Group 1 and ACL Group 2, where each includes three levels of precedence. If you add both ACL Groups to a port, then the port utilizes a total of six precedence levels. A maximum of eight is permitted.

Actions taken by an ACL are called *In-Profile* actions. You can configure additional In-Profile and Out-of-Profile actions on a port. Data traffic can be metered, and re-marked to ensure that the traffic flow provides certain levels of service in terms of bandwidth for different types of network traffic.

Metering

QoS metering provides different levels of service to data streams through user-configurable parameters. A meter is used to measure the traffic stream against a traffic profile which you create. Thus, creating meters yields In-Profile and Out-of-Profile traffic for each ACL, as follows:

- **In-Profile**—If there is no meter configured or if the packet conforms to the meter, the packet is classified as In-Profile.
- **Out-of-Profile**—If a meter is configured and the packet does not conform to the meter (exceeds the committed rate or maximum burst rate of the meter), the packet is classified as Out-of-Profile.

Using meters, you set a Committed Rate in Kbps (1000 bits per second in each Kbps). All traffic within this Committed Rate is In-Profile. Additionally, you can set a Maximum Burst Size that specifies an allowed data burst larger than the Committed Rate for a brief period. These parameters define the In-Profile traffic.

Meters keep the sorted packets within certain parameters. You can configure a meter on an ACL, and perform actions on metered traffic, such as packet re-marking.

Re-Marking

Re-marking allows for the treatment of packets to be reset based on new network specifications or desired levels of service. You can configure the ACL to re-mark a packet as follows:

- Change the DSCP value of a packet, used to specify the service level traffic should receive.
- Change the 802.1p priority of a packet.

Viewing ACL Statistics

ACL statistics display how many packets hit (matched) each ACL. Use ACL statistics to check filter performance, and debug the ACL filters.

You must enable statistics (`cfg/acl/acl <x>/stats ena`) for each ACL that you want to monitor.

ACL Configuration Examples

ACL Example 1

Use this configuration to block traffic to a specific host. All traffic that ingresses on port EXT1 is denied if it is destined for the host at IP address 100.10.1.1

1. Configure an Access Control List.

```
>> Main# cfg/acl/acl 1 (Define ACL 1)
>> ACL 1# ipv4/dip 100.10.1.1
Enter destination IP address mask (default 255.255.255.255):
>> Filtering IPv4# ..
>> ACL 1# action deny
```

2. Add ACL 1 to port EXT1.

```
>> Main# cfg/port ext1/aclqos (Select port EXT 1 to assign ACLs)
>> Port EXT1 ACL# add acl 1 (Assign ACL 1 to the port)
```

3. Apply and save the configuration.

```
>> Port EXT1 ACL# apply
>> Port EXT1 ACL# save
```

ACL Example 2

Use this configuration to block traffic from a network destined for a specific host address. All traffic that ingresses in port EXT2 with source IP from the class 100.10.1.0/24 and destination IP 200.20.2.2 is denied.

1. Configure an Access Control List.

```
>> Main# cfg/acl/acl 2 (Define ACL 2)
>> ACL 2# ipv4/sip 100.10.1.0 255.255.255.0
>> Filtering IPv4# dip 200.20.2.2 255.255.255.255
>> Filtering IPv4# ..
>> ACL 2# action deny
```

2. Add ACL 2 to port EXT2.

```
>> ACL 2# /cfg/port ext2/aclqos (Select port EXT2 to assign ACLs)
>> Port EXT2 ACL# add acl 2 (Assign ACL 2 to the port)
```

3. Apply and save the configuration.

```
>> Port EXT2 ACL# apply
>> Port EXT2 ACL# save
```

ACL Example 3

This configuration blocks traffic from a network that is destined for a specific egress port. All traffic that ingresses port EXT1 from the network 100.10.1.0/24 and is destined for port INT1 is denied.

1. Configure an Access Control List.

```
>> Main# cfg/acl/acl 3 (Define ACL 3)
>> ACL 3# ipv4/sip 100.10.1.0 255.255.255.0
>> Filtering IPv4# ..
>> ACL 3# egrport int1
>> ACL 3# action deny
```

2. Add ACL 3 to port EXT1.

```
>> Main# cfg/port ext1/aclqos (Select port EXT1 to assign ACLs)
>> Port EXT1 ACL# add acl 3 (Assign ACL 3 to the port)
```

3. Apply and save the configuration.

```
>> Port EXT2 ACL# apply
>> Port EXT2 ACL# save
```

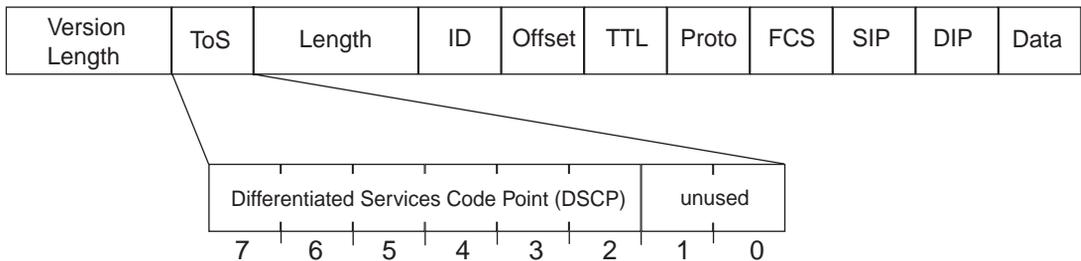
Using DSCP Values to Provide QoS

The six most significant bits in the TOS byte of the IP header are defined as DiffServ Code Points (DSCP). Packets are marked with a certain value depending on the type of treatment the packet must receive in the network device. DSCP is a measure of the Quality of Service (QoS) level of the packet.

Differentiated Services Concepts

To differentiate between traffic flows, packets can be classified by their DSCP value. The Differentiated Services (DS) field in the IP header is an octet, and the first six bits, called the DS Code Point (DSCP), can provide QoS functions. Each packet carries its own QoS state in the DSCP. There are 64 possible DSCP values (0-63).

Figure 16 Layer 3 IPv4 Packet



The GbESM can perform the following actions to the DSCP:

- Read the DSCP value of ingress packets
- Re-mark the DSCP value to a new value
- Map the DSCP value to an 802.1p priority

Once the DSCP value is marked, the GbESM can use it to direct traffic prioritization.

Per-Hop Behavior

The DSCP value determines the Per Hop Behavior (PHB) of each packet. The PHB is the forwarding treatment given to packets at each hop. QoS policies are built by applying a set of rules to packets, based on the DSCP value, as they hop through the network.

The GbESM default settings are based on the following standard PHBs, as defined in the IEEE standards:

- Expedited Forwarding (EF)—This PHB has the highest egress priority and lowest drop precedence level. EF traffic is forwarded ahead of all other traffic. EF PHB is described in RFC 2598.
- Assured Forwarding (AF)—This PHB contains four service levels, each with a different drop precedence, as shown below. Routers use drop precedence to determine which packets to discard last when the network becomes congested. AF PHB is described in RFC 2597.

Drop Precedence	Class 1	Class 2	Class 3	Class 4
Low	AF11 (DSCP 10)	AF21 (DSCP 18)	AF31 (DSCP 26)	AF41 (DSCP 34)
Medium	AF12 (DSCP 12)	AF22 (DSCP 20)	AF32 (DSCP 28)	AF42 (DSCP 36)
High	AF13 (DSCP 14)	AF23 (DSCP 22)	AF33 (DSCP 30)	AF43 (DSCP 38)

- Class Selector (CS)—This PHB has eight priority classes, with CS7 representing the highest priority, and CS0 representing the lowest priority, as shown below. CS PHB is described in RFC 2474.

Priority	Class Selector	DSCP
Highest	CS7	56
	CS6	48
	CS5	40
	CS4	32
	CS3	24
	CS2	16
	CS1	8
Lowest	CS0	0

QoS Levels

Table 15 shows the default service levels provided by the GbESM, listed from highest to lowest importance:

Table 15 Default QoS Service Levels

Service Level	Default PHB	802.1p Priority
Critical	CS7	7
Network Control	CS6	6
Premium	EF, CS5	5
Platinum	AF41, AF42, AF43, CS4	4
Gold	AF31, AF32, AF33, CS3	3
Silver	AF21, AF22, AF23, CS2	2
Bronze	AF11, AF12, AF13, CS1	1
Standard	DF, CS0	0

DSCP Re-Marking and Mapping

DSCP Re-Marking Overview

The GbESM can re-mark the DSCP value of ingress packets to a new value, and set the 802.1p priority value, based on the DSCP value. You can view the settings by using the following command:

```
>> # /cfg/qos/dscp/cur
Current DSCP Remarking Configuration: OFF
```

DSCP	New DSCP	New 802.1p Prio
0	0	0
1	1	0
...		
51	51	0
52	52	0
53	53	0
54	54	0
55	55	0
56	56	7
57	57	0
58	58	0
59	59	0
60	60	0
61	61	0
62	62	0
63	63	0

Use the `cfg/qos/dscp/on` command to turn on DSCP re-marking globally. Then you must enable DSCP re-marking (`cfg/port <x>/dscpmrk/ena`) on any port that you wish to perform this function.

Note – If an ACL meter is configured for DSCP re-marking, the meter function takes precedence over QoS re-marking.

DSCP Re-Marking Configuration Example

1. Turn DSCP re-marking on globally, and define the DSCP-DSCP-802.1p mapping. You can use the default mapping, as shown in the `cfg/qos/dscp/cur` command output.

```

>> Main# cfg/qos/dscp/on                                (Turn on DSCP re-marking)
>> DSCP Remark# dscp 8                                  (Define DSCP re-marking)
Current DSCP remark (for DSCP 8): 8
Enter new DSCP remark (for DSCP 8) [0-63]: 10
>> DSCP Remark# prio                                    (Define DSCP-to-802.1p mapping)
Enter DSCP [0-63]: 10
Current prio (for DSCP 10): 1
Enter new prio (for DSCP 10) [0-7]: 2
>> DSCP Remark# apply

```

2. Enable DSCP re-marking on a port.

```

>> Main# cfg/port EXT1                                  (Select port)
>> Port EXT1# dscpmrk ena                               (Enable DSCP re-marking)
Current DSCP remarking: disabled
New DSCP remarking:      enabled
>> Port EXT1# apply

```

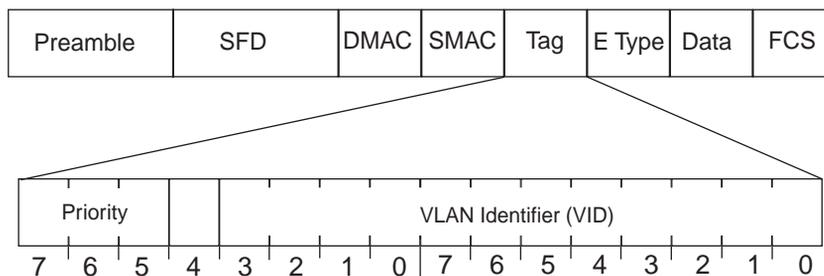
Using 802.1p Priorities to Provide QoS

802.1p Overview

BLADEOS provides Quality of Service functions based on the priority bits in a packet's VLAN header. (The priority bits are defined by the 802.1p standard within the IEEE 802.1q VLAN header.) The 802.1p bits, if present in the packet, specify the priority that should be given to packets during forwarding. Packets with a numerically higher (non-zero) priority are given forwarding preference over packets with lower priority bit value.

The IEEE 802.1p standard uses eight levels of priority (0-7). Priority 7 is assigned to highest priority network traffic, such as OSPF or RIP routing table updates, priorities 5-6 are assigned to delay-sensitive applications such as voice and video, and lower priorities are assigned to standard applications. A value of 0 (zero) indicates a "best effort" traffic prioritization, and this is the default when traffic priority has not been configured on your network. The GbESM can filter packets based on the 802.1p values, and it can assign or overwrite the 802.1p value in the packet.

Figure 17 Layer 2 802.1q/802.1p VLAN Tagged Packet



Ingress packets receive a priority value, as follows:

- **Tagged packets**—GbESM reads the 802.1p priority in the VLAN tag.
- **Untagged packets**—GbESM tags the packet and assigns an 802.1p priority, based on the port's default priority (`/cfg/port <x>/8021ppri`).

Egress packets are placed in a COS queue based on the priority value, and scheduled for transmission based on the scheduling weight of the COS queue.

802.1p Configuration Example

1. Configure a port's default 802.1p priority.

```
>> Main# cfg/port EXT1                                (Select port)
>> Port EXT1# 8021ppri                                (Set port's default 802.1p priority)
Current 802.1p priority: 0
Enter new 802.1p priority [0-7]: 1
>> Port EXT1# ena
>> Port EXT1# apply
```

2. Map the 802.1p priority value to a COS queue and set the COS queue scheduling weight.

```
>> Port EXT1# /cfg/qos/8021p                            (Select 802.1p menu)
>> 802.1p# priq                                         (Set COS queue assignments)
Enter priority [0-7]: 1
Current COS queue (for priority 1): 0
Enter new COS queue (for priority 1) [0-1]: 1
>> 802.1p# qweight                                    (Set COS queue weights)
Enter COS queue [0-1]: 1
Current weight (for COS queue 1): 3
Enter new weight (for COS queue 1) [0-15]: 10
>> 802.1p# apply
```

Queuing and Scheduling

The GbESM can be configured to have either 2 or 8 output Class of Service (COS) queues per port, into which each packet is placed. Each packet's 802.1p priority determines its COS queue, except when an ACL action sets the COS queue of the packet.

Each COS queue uses Weighted Round Robin (WRR) scheduling, with user configurable weight from 1 to 15. The weight of 0 (zero) indicates strict priority, which might starve the low priority queues.

You can configure the following attributes for COS queues:

- Map 802.1p priority value to a COS queue
- Define the scheduling weight of each COS queue

Use the 802.1p menu (`/cfg/qos/8021p`) to configure COS queues.

Part 2: IP Routing

This section discusses Layer 3 switching functions. In addition to switching traffic at near line rates, the application switch can perform multi-protocol routing. This section discusses basic routing and advanced routing protocols:

- Basic Routing
- Routing Information Protocol (RIP)
- Internet Group Management Protocol (IGMP)
- Border Gateway Protocol (BGP)
- Open Shortest Path First (OSPF)
- Remote Monitoring (RMON)

CHAPTER 9

Basic IP Routing

This chapter provides configuration background and examples for using the GbE Switch Module (GbESM) to perform IP routing functions. The following topics are addressed in this chapter:

- [“IP Routing Benefits” on page 151](#)
- [“Routing Between IP Subnets” on page 152](#)
- [“Subnet Routing Example” on page 154](#)
- [“Dynamic Host Configuration Protocol” on page 158](#)

IP Routing Benefits

The GbESM uses a combination of configurable IP switch interfaces and IP routing options. The switch IP routing capabilities provide the following benefits:

- Connects the server IP subnets to the rest of the backbone network.
- Provides the ability to route IP traffic between multiple Virtual Local Area Networks (VLANs) configured on the switch.

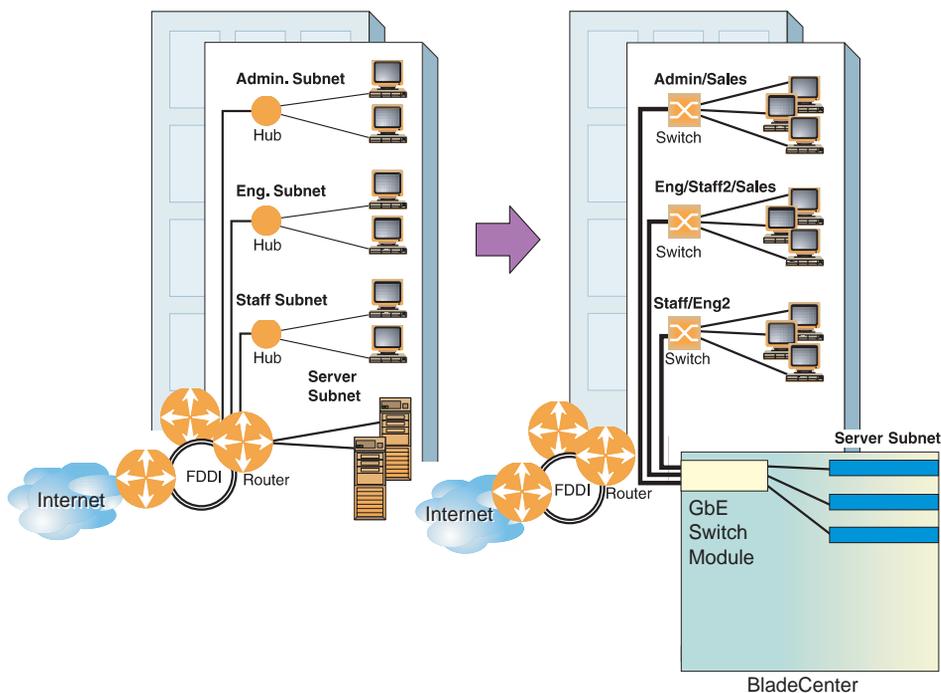
Routing Between IP Subnets

The physical layout of most corporate networks has evolved over time. Classic hub/router topologies have given way to faster switched topologies, particularly now that switches are increasingly intelligent. GbESMs are intelligent and fast enough to perform routing functions on par with wire-speed Layer 2 switching.

The combination of faster routing and switching in a single device provides another service—it allows you to build versatile topologies that account for legacy configurations.

For example, consider the following topology migration:

Figure 18 The Router Legacy Network



In this example, a corporate campus has migrated from a router-centric topology to a faster, more powerful, switch-based topology. As is often the case, the legacy of network growth and redesign has left the system with a mix of illogically distributed subnets.

This is a situation that switching alone cannot cure. Instead, the router is flooded with cross-subnet communication. This compromises efficiency in two ways:

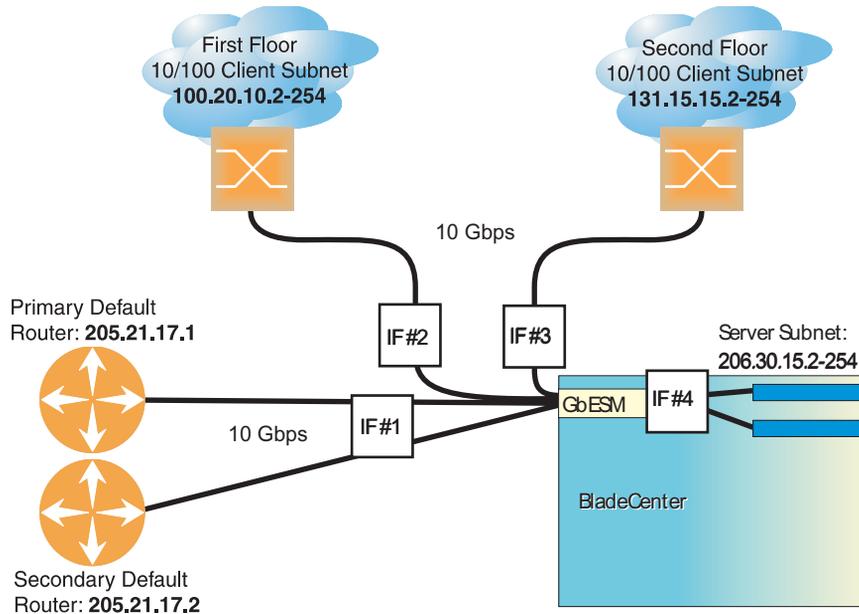
- Routers can be slower than switches. The cross-subnet side trip from the switch to the router and back again adds two hops for the data, slowing throughput considerably.
- Traffic to the router increases, increasing congestion.

Even if every end-station could be moved to better logical subnets (a daunting task), competition for access to common server pools on different subnets still burdens the routers.

This problem is solved by using GbESMs with built-in IP routing capabilities. Cross-subnet LAN traffic can now be routed within the switches with wire speed Layer 2 switching performance. This not only eases the load on the router but saves the network administrators from reconfiguring each and every end-station with new IP addresses.

Take a closer look at the BladeCenter's GbESM in the following configuration example:

Figure 19 Switch-Based Routing Topology



The GbESM connects the Gigabit Ethernet and Fast Ethernet trunks from various switched subnets throughout one building. Common servers are placed on another subnet attached to the switch. A primary and backup router are attached to the switch on yet another subnet.

Without Layer 3 IP routing on the switch, cross-subnet communication is relayed to the default gateway (in this case, the router) for the next level of routing intelligence. The router fills in the necessary address information and sends the data back to the switch, which then relays the packet to the proper destination subnet using Layer 2 switching.

With Layer 3 IP routing in place on the GbESM, routing between different IP subnets can be accomplished entirely within the switch. This leaves the routers free to handle inbound and outbound traffic for this group of subnets.

Subnet Routing Example

Prior to configuring, you must be connected to the switch 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 *BLADEOS 5.1 Command Reference*.

1. Assign an IP address (or document the existing one) for each router and client workstation.

In the example topology in [Figure 19 on page 153](#), the following IP addresses are used:

Table 16 Subnet Routing Example: IP Address Assignments

Subnet	Devices	IP Addresses
1	Primary and Secondary Default Routers	205.21.17.1 and 205.21.17.2
2	First Floor Client Workstations	100.20.10.2-254
3	Second Floor Client Workstations	131.15.15.2-254
4	Common Servers	206.30.15.2-254

2. Assign an IP interface for each subnet attached to the switch.

Since there are four IP subnets connected to the switch, four IP interfaces are needed:

Table 17 Subnet Routing Example: IP Interface Assignments

Interface	Devices	IP Interface Address
IF 1	Primary and Secondary Default Routers	205.21.17.3
IF 2	First Floor Client Workstations	100.20.10.1
IF 3	Second Floor Client Workstations	131.15.15.1
IF 4	Common Servers	206.30.15.1

IP interfaces are configured using the following commands at the CLI:

```
>> # /cfg/13/if 1                (Select IP interface 1)
>> IP Interface 1# addr 205.21.17.3 (Assign IP address for the interface)
>> IP Interface 1# ena            (Enable IP interface 1)
>> IP Interface 1# ../if 2       (Select IP interface 2)
>> IP Interface 2# addr 100.20.10.1 (Assign IP address for the interface)
>> IP Interface 2# ena            (Enable IP interface 2)
>> IP Interface 2# ../if 3       (Select IP interface 3)
>> IP Interface 3# addr 131.15.15.1 (Assign IP address for the interface)
>> IP Interface 3# ena            (Enable IP interface 3)
>> IP Interface 3# ../if 4       (Select IP interface 4)
>> IP Interface 4# addr 206.30.15.1 (Assign IP address for the interface)
>> IP Interface 4# ena            (Enable IP interface 4)
```

3. Set each server and workstation's default gateway to the appropriate switch IP interface (the one in the same subnet as the server or workstation).
4. Configure the default gateways to the routers' addresses.

Configuring the default gateways allows the switch to send outbound traffic to the routers:

```
>> IP Interface 5# ../gw 1        (Select primary default gateway)
>> Default gateway 1# addr 205.21.17.1 (Assign IP address for primary router)
>> Default gateway 1# ena        (Enable primary default gateway)
>> Default gateway 1# ../gw 2    (Select secondary default gateway)
>> Default gateway 2# addr 205.21.17.2 (Assign address for secondary router) >>
Default gateway 2# ena          (Enable secondary default gateway)
```

5. Apply and verify the configuration.

```
>> Default gateway 2# # apply    (Make your changes active)
>> Default gateway 2# /cfg/13/cur (View current IP settings)
```

Examine the resulting information. If any settings are incorrect, make the appropriate changes.

6. Save your new configuration changes.

```
>> IP# save                       (Save for restore after reboot)
```

Using VLANs to Segregate Broadcast Domains

In the previous example, devices that share a common IP network are all in the same broadcast domain. If you want to limit the broadcasts on your network, you could use VLANs to create distinct broadcast domains. For example, as shown in the following procedure, you could create one VLAN for the client trunks, one for the routers, and one for the servers.

In this example, you are adding to the previous configuration.

1. Determine which switch ports and IP interfaces belong to which VLANs.

The following table adds port and VLAN information:

Table 18 Subnet Routing Example: Optional VLAN Ports

VLAN	Devices	IP Interface	Switch Port	VLAN #
1	First Floor Client Workstations	2	EXT1	1
	Second Floor Client Workstations	3	EXT2	1
2	Primary Default Router	1	EXT3	2
	Secondary Default Router	1	EXT4	2
3	Common Servers 1	4	INT5	3
	Common Servers 2	4	INT6	3

2. Add the switch ports to their respective VLANs.

The VLANs shown in [Table 18](#) are configured as follows:

```

>> # /cfg/12/vlan 1                               (Select VLAN 1)
>> VLAN 1# add ext1                               (Add port for 1st floor to VLAN 1)
>> VLAN 1# add ext2                               (Add port for 2nd floor to VLAN 1)
>> VLAN 1# ena                                    (Enable VLAN 1)
>> VLAN 1# ../vlan 2                               (Select VLAN 2)
>> VLAN 2# add ext3                               (Add port for default router 1)
>> VLAN 2# add ext4                               (Add port for default router 2)
>> VLAN 2# ena                                    (Enable VLAN 2)
>> VLAN 2# ../vlan 3                               (Select VLAN 3)
>> VLAN 3# add int5                               (Add port for default router 3)
>> VLAN 3# add int6                               (Add port for common server 1)
>> VLAN 3# ena                                    (Enable VLAN 3)

```

Each time you add a port to a VLAN, you may get the following prompt:

```
Port 4 is an untagged port and its current PVID is 1.
Confirm changing PVID from 1 to 2 [y/n]?
```

Enter **y** to set the default Port VLAN ID (PVID) for the port.

3. Add each IP interface to the appropriate VLAN.

Now that the ports are separated into three VLANs, the IP interface for each subnet must be placed in the appropriate VLAN. From [Table 18 on page 156](#), the settings are made as follows:

```
>> VLAN 3# /cfg/l3/if 1                (Select IP interface 1 for def. routers)
>> IP Interface 1# vlan 2              (Set to VLAN 2)
>> IP Interface 1# ../if 2            (Select IP interface 2 for first floor)
>> IP Interface 2# vlan 1              (Set to VLAN 1)
>> IP Interface 2# ../if 3            (Select IP interface 3 for second floor)
>> IP Interface 3# vlan 1              (Set to VLAN 1)
>> IP Interface 3# ../if 4            (Select IP interface 4 for servers)
>> IP Interface 4# vlan 3              (Set to VLAN 3)
```

4. Apply and verify the configuration.

```
>> IP Interface 5# apply                (Make your changes active)
>> IP Interface 5# /info/l2/vlan        (View current VLAN information)
>> Information# port                    (View current port information)
```

Examine the resulting information. If any settings are incorrect, make the appropriate changes.

5. Save your new configuration changes.

```
>> Information# save                    (Save for restore after reboot)
```

Dynamic Host Configuration Protocol

Dynamic Host Configuration Protocol (DHCP) is a transport protocol that provides a framework for automatically assigning IP addresses and configuration information to other IP hosts or clients in a large TCP/IP network. Without DHCP, the IP address must be entered manually for each network device. DHCP allows a network administrator to distribute IP addresses from a central point and automatically send a new IP address when a device is connected to a different place in the network.

DHCP is an extension of another network IP management protocol, Bootstrap Protocol (BOOTP), with an additional capability of being able to dynamically allocate reusable network addresses and configuration parameters for client operation.

Built on the client/server model, DHCP allows hosts or clients on an IP network to obtain their configurations from a DHCP server, thereby reducing network administration. The most significant configuration the client receives from the server is its required IP address; (other optional parameters include the “generic” file name to be booted, the address of the default gateway, and so forth).

DHCP relay agent eliminates the need to have DHCP/BOOTP servers on every subnet. It allows the administrator to reduce the number of DHCP servers deployed on the network and to centralize them. Without the DHCP relay agent, there must be at least one DHCP server deployed at each subnet that has hosts needing to perform the DHCP request.

DHCP Relay Agent

DHCP is described in RFC 2131, and the DHCP relay agent supported on GbESMs is described in RFC 1542. DHCP uses UDP as its transport protocol. The client sends messages to the server on port 67 and the server sends messages to the client on port 68.

DHCP defines the methods through which clients can be assigned an IP address for a finite lease period and allowing reassignment of the IP address to another client later. Additionally, DHCP provides the mechanism for a client to gather other IP configuration parameters it needs to operate in the TCP/IP network.

In the DHCP environment, the GbESM acts as a relay agent. The DHCP relay feature (`/cfg/13/bootp`) enables the switch to forward a client request for an IP address to two BOOTP servers with IP addresses that have been configured on the switch.

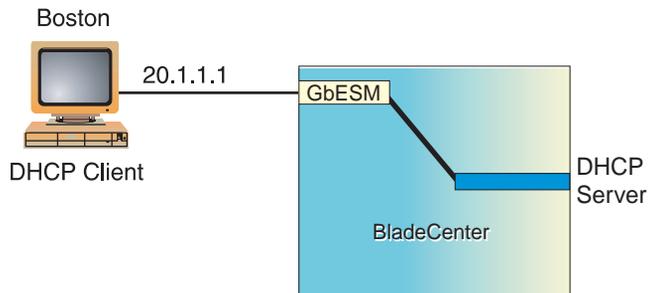
When a switch receives a UDP broadcast on port 67 from a DHCP client requesting an IP address, the switch acts as a proxy for the client, replacing the client source IP (SIP) and destination IP (DIP) addresses. The request is then forwarded as a UDP Unicast MAC layer message to two BOOTP servers whose IP addresses are configured on the switch. The servers respond as a UDP Unicast message back to the switch, with the default gateway and IP address for the client. The destination IP address in the server response represents the interface address on the switch that received the client request. This interface address tells the switch on which VLAN to send the server response to the client.

DHCP Relay Agent Configuration

To enable the GbESM to be the BOOTP forwarder, you need to configure the DHCP/BOOTP server IP addresses on the switch. Generally, you should configure the switch IP interface on the client side to match the client's subnet, and configure VLANs to separate client and server subnets. The DHCP server knows from which IP subnet the newly allocated IP address should come.

The following figure shows a basic DHCP network example:

Figure 20 DHCP Relay Agent Configuration



In GbESM implementation, there is no need for primary or secondary servers. The client request is forwarded to the BOOTP servers configured on the switch. The use of two servers provide failover redundancy. However, no health checking is supported.

Use the following commands to configure the switch as a DHCP relay agent:

```
>> # /cfg/13/bootp
>> Bootstrap Protocol Relay# addr           (Set IP address of BOOTP server)
>> Bootstrap Protocol Relay# addr2        (Set IP address of 2nd BOOTP server)
>> Bootstrap Protocol Relay# on           (Globally turn BOOTP relay on)
>> Bootstrap Protocol Relay# off         (Globally turn BOOTP relay off)
>> Bootstrap Protocol Relay# cur         (Display current configuration)
```

Additionally, DHCP Relay functionality can be assigned on a per interface basis. Use the following command to enable the Relay functionality:

```
>> # /cfg/13/if <interface number>/relay ena
```

CHAPTER 10

Routing Information Protocol

In a routed environment, routers communicate with one another to keep track of available routes. Routers can learn about available routes dynamically using the Routing Information Protocol (RIP). BLADEOS software supports RIP version 1 (RIPv1) and RIP version 2 (RIPv2) for exchanging TCP/IP route information with other routers.

Distance Vector Protocol

RIP is known as a distance vector protocol. The vector is the network number and next hop, and the distance is the cost associated with the network number. RIP identifies network reachability based on metric, and metric is defined as hop count. One hop is considered to be the distance from one switch to the next, which typically is 1.

When a switch receives a routing update that contains a new or changed destination network entry, the switch adds 1 to the metric value indicated in the update and enters the network in the routing table. The IP address of the sender is used as the next hop.

Stability

RIP includes a number of other stability features that are common to many routing protocols. For example, RIP implements the split horizon and hold-down mechanisms to prevent incorrect routing information from being propagated.

RIP prevents routing loops from continuing indefinitely by implementing a limit on the number of hops allowed in a path from the source to a destination. The maximum number of hops in a path is 15. The network destination network is considered unreachable if increasing the metric value by 1 causes the metric to be 16 (that is infinity). This limits the maximum diameter of a RIP network to less than 16 hops.

RIP is often used in stub networks and in small autonomous systems that do not have many redundant paths.

Routing Updates

RIP sends routing-update messages at regular intervals and when the network topology changes. Each router “advertises” routing information by sending a routing information update every 30 seconds. If a router doesn’t receive an update from another router for 180 seconds, those routes provided by that router are declared invalid. The routes are removed from the routing table, but they remain in the RIP routes table (`/info/13/rip/routes`). After another 120 seconds without receiving an update for those routes, the routes are removed from regular updates.

When a router receives a routing update that includes changes to an entry, it updates its routing table to reflect the new route. The metric value for the path is increased by 1, and the sender is indicated as the next hop. RIP routers maintain only the best route (the route with the lowest metric value) to a destination.

For more information see *The Configuration Menu, Routing Information Protocol Configuration* (`/cfg/13/rip`) in the *BLADEOS 5.1 Command Reference*.

RIPv1

RIP version 1 uses broadcast User Datagram Protocol (UDP) data packets for the regular routing updates. The main disadvantage is that the routing updates do not carry subnet mask information. Hence, the router cannot determine whether the route is a subnet route or a host route. It is of limited usage after the introduction of RIPv2. For more information about RIPv1 and RIPv2, refer to RFC 1058 and RFC 2453.

RIPv2

RIPv2 is the most popular and preferred configuration for most networks. RIPv2 expands the amount of useful information carried in RIP messages and provides a measure of security. For a detailed explanation of RIPv2, refer to RFC 1723 and RFC 2453.

RIPv2 improves efficiency by using multicast UDP (address 224.0.0.9) data packets for regular routing updates. Subnet mask information is provided in the routing updates. A security option is added for authenticating routing updates, by using a shared password. BLADEOS supports using clear password for RIPv2.

RIPv2 in RIPv1 Compatibility Mode

BLADEOS allows you to configure RIPv2 in RIPv1 compatibility mode, for using both RIPv2 and RIPv1 routers within a network. In this mode, the regular routing updates use broadcast UDP data packet to allow RIPv1 routers to receive those packets. With RIPv1 routers as recipients, the routing updates have to carry natural or host mask. Hence, it is not a recommended configuration for most network topologies.

Note – When using both RIPv1 and RIPv2 within a network, use a single subnet mask throughout the network.

RIP Features

BLADEOS provides the following features to support RIPv1 and RIPv2:

Poison Reverse

Simple split horizon in RIP omits routes learned from one neighbor in updates sent to that neighbor. That is the most common configuration used in RIP, with the Poison Reverse feature disabled. Split horizon with poisoned reverse enabled includes such routes in updates, but sets their metrics to 16. The disadvantage of using this feature is the increase of size in the routing updates.

Triggered Updates

Triggered updates are an attempt to speed up convergence. When Triggered Updates is enabled (`/cfg/13/rip/if <x>/trigg/e`), whenever a router changes the metric for a route, it sends update messages almost immediately, without waiting for the regular update interval. It is recommended to enable Triggered Updates.

Multicast

RIPv2 messages use IP multicast address (224.0.0.9) for periodic updates. Multicast RIPv2 updates are not processed by RIPv1 routers. IGMP is not needed since these are inter-router messages which are not forwarded.

To configure RIPv2 in RIPv1 compatibility mode, set multicast to `disable`, and set version to `both`.

Default Route

The RIP router can listen and supply a default route, usually represented as 0.0.0.0 in the routing table. When a router does not have an explicit route to a destination network in its routing table, it uses the default route to forward those packets.

Metric

The metric field contains a configurable value between 1 and 15 (inclusive) which specifies the current metric for the interface. The metric value typically indicates the total number of hops to the destination. The metric value of 16 represents an unreachable destination.

Authentication

RIPv2 authentication uses plain text password for authentication. If configured using Authentication password, then it is necessary to enter an authentication key value.

The following method is used to authenticate a RIP message:

- If the router is not configured to authenticate RIPv2 messages, then RIPv1 and unauthenticated RIPv2 messages are accepted; authenticated RIPv2 messages are discarded.
- If the router is configured to authenticate RIPv2 messages, then RIPv1 and RIPv2 messages which pass authentication testing are accepted; unauthenticated and failed authentication RIPv2 messages are discarded.

For maximum security, RIPv1 messages are ignored when authentication is enabled (`cfg/13/rip/if <x>/auth/password`); otherwise, the routing information from authenticated messages is propagated by RIPv1 routers in an unauthenticated manner.

RIP Configuration Example

Note – An interface RIP disabled uses all the default values of the RIP, no matter how the RIP parameters are configured for that interface. RIP sends out RIP regular updates to include an UP interface, but not a DOWN interface.

1. Add VLANs for routing interfaces.

```
>> Main# /cfg/12/vlan 2/ena           (Enable VLAN 2)
>> VLAN 2# add ext2                 (Add port EXT2 to VLAN 2)
Port EXT2 is an UNTAGGED port and its current PVID is 1.
Confirm changing PVID from 1 to 2 [y/n]: y
>> VLAN 2# /cfg/12/vlan 3/ena       (Enable VLAN 3)
>> VLAN 3# add ext3                 (Add port EXT3 to VLAN 3)
Port EXT3 is an UNTAGGED port and its current PVID is 1.
Confirm changing PVID from 1 to 3 [y/n]: y
```

2. Add IP interfaces to VLANs.

```
>> VLAN 3# /cfg/13/if 2/ena         (Enable interface 2)
>> IP Interface 2# addr 102.1.1.1   (Define IP address for interface 2)
>> IP Interface 2# vlan 2           (Add interface 2 to VLAN 2)
>> IP Interface 2# /cfg/13/if 3/ena (Enable interface 3)
>> IP Interface 3# addr 103.1.1.1   (Define IP address for interface 3)
>> IP Interface 3# vlan 3           (Add interface 3 to VLAN 3)
```

3. Turn on RIP globally and enable RIP for each interface.

```
>> IP Interface 3# /cfg/13/rip/on    (Turn on RIP globally)
>> Routing Information Protocol# if 2/ena (Enable RIP on IP interface 2)
>> RIP Interface 2# ..
>> Routing Information Protocol# if 3/ena (Enable RIP on IP interface 3)
>> RIP Interface 3# apply           (Apply your changes)
>> RIP Interface 3# save            (Save the configuration)
```

Use the `/maint/route/dump` command to check the current valid routes in the routing table of the switch.

For those RIP learnt routes within the garbage collection period, that are routes phasing out of the routing table with metric 16, use the `/info/13/rip/routes` command. Locally configured static routes do not appear in the RIP Routes table.

CHAPTER 11

IGMP

Internet Group Management Protocol (IGMP) is used by IP Multicast routers to learn about the existence of host group members on their directly attached subnet (see RFC 2236). The IP Multicast routers get this information by broadcasting IGMP Membership Queries and listening for IP hosts reporting their host group memberships. This process is used to set up a client/server relationship between an IP Multicast source that provides the data streams and the clients that want to receive the data.

The GbE Switch Module (GbESM) can perform IGMP Snooping, or act as an IGMP Relay (proxy) device.

The following topics are discussed in this chapter:

- [“IGMP Snooping” on page 168](#)
- [“IGMP Relay” on page 172](#)
- [“Additional IGMP Features” on page 174](#)

IGMP Snooping

IGMP Snooping allows the switch to forward multicast traffic only to those ports that request it. IGMP Snooping prevents multicast traffic from being flooded to all ports. The switch learns which server hosts are interested in receiving multicast traffic, and forwards it only to ports connected to those servers.

IGMP Snooping conserves bandwidth. With IGMP Snooping, the switch learns which ports are interested in receiving multicast data, and forwards multicast data only to those ports. In this way, other ports are not burdened with unwanted multicast traffic.

The switch can sense IGMP Membership Reports from attached clients and act as a proxy to set up a dedicated path between the requesting host and a local IP Multicast router. After the pathway is established, the switch blocks the IP Multicast stream from flowing through any port that does not connect to a host member, thus conserving bandwidth.

The client-server path is set up as follows:

- An IP Multicast Router (Mrouter) sends *Membership Queries* to the switch, which forwards them to all ports in a given VLAN.
- Hosts that want to receive the multicast data stream send *Membership Reports* to the switch, which sends a proxy Membership Report to the Mrouter.
- The switch sets up a path between the Mrouter and the host, and blocks all other ports from receiving the multicast.
- Periodically, the Mrouter sends Membership Queries to ensure that the host wants to continue receiving the multicast. If a host fails to respond with a Membership Report, the Mrouter stops sending the multicast to that path.
- The host can send a *Leave Report* to the switch, which sends a proxy Leave Report to the Mrouter. The multicast path is terminated immediately.

IGMPv3

IGMPv3 includes new membership report messages to extend IGMP functionality. The GbESM provides snooping capability for all types of IGMP version 3 (IGMPv3) Membership Reports, as described in RFC 3376.

IGMPv3 supports Source-Specific Multicast (SSM). SSM identifies session traffic by both source and group addresses. The GbESM uses *source filtering*, which allows hosts to report interest in receiving multicast packets only from specific source addresses, or from all but specific source addresses.

The GbESM supports the following IGMPv3 filter modes:

- **INCLUDE** mode: The host requests membership to a multicast group and provides a list of IP addresses from which it wants to receive traffic.
- **EXCLUDE** mode: The host requests membership to a multicast group and provides a list of IP addresses from which it *does not* want to receive traffic. This indicates that the host wants to receive traffic only from sources that are not part of the Exclude list. To disable snooping on EXCLUDE mode reports, use the following command:

```
/cfg/13/igmp/snoop/igmpv3/exclude dis
```

By default, the GbESM snoops the first eight sources listed in the IGMPv3 Group Record. Use the following command to change the number of snooping sources:

```
/cfg/13/igmp/snoop/igmpv3/sources <1-64>
```

IGMPv3 Snooping is compatible with IGMPv1 and IGMPv2 Snooping. You can disable snooping on version 1 and version 2 reports, using the following command:

```
/cfg/13/igmp/snoop/igmpv3/v1v2 dis
```

IGMP Snooping Configuration Example

This section provides steps to configure IGMP Snooping on the GbESM, using the Command-Line Interface (CLI).

1. Configure port and VLAN membership on the switch.
2. Turn on IGMP.

```
>> /cfg/13/igmp/on
```

```
(Turn on IGMP)
```

3. Add VLANs to IGMP Snooping and enable the feature.

```
>> /cfg/13/igmp/snoop (Access IGMP Snoop menu)
>> IGMP Snoop# add 1 (Add VLAN 1 to IGMP snooping)
>> IGMP Snoop# ena (Enable IGMP Snooping)
```

4. Enable IGMPv3 Snooping (optional).

```
>> IGMP Snoop# igmpv3 (Access IGMPv3 menu)
>> IGMP V3 Snoop# ena (Enable IGMPv3 Snooping)
```

5. Apply and save the configuration.

```
>> IGMP V3 Snoop# apply (Apply the configuration)
>> IGMP V3 Snoop# save (Save your changes)
```

6. View dynamic IGMP information.

To display information about IGMP Groups:

```
>> /info/13/igmp/dump
Note: Local groups (224.0.0.x) are not snooped/relayed and do not appear.
Source      Group      VLAN      Port      Version    Mode      Expires    Fwd
-----
10.1.1.1    232.1.1.1  2         EXT4      V3         INC       4:16       Yes
10.1.1.5    232.1.1.1  2         EXT4      V3         INC       4:16       Yes
*           232.1.1.1  2         EXT4      V3         INC       -          No
10.10.10.43 235.0.0.1  9         EXT1      V3         INC       2:26       Yes
*           236.0.0.1  9         EXT1      V3         EXC       -          Yes
```

To display information about Mrouters learned by the switch:

```
>> /info/13/igmp/mrouter/dump
VLAN      Port      Version    Expires    Max Query Resp. Time    QRV    QQIC
-----
1         EXT4      V2         static     unknown           -      -
2         EXT3      V3         4:09      128                2      125
```

Note – If IGMP Snooping v1/v2 is enabled and IGMPv3 Snooping is disabled, the output of IGMPv3 reports and queries show some items as IGMPv3 (V3), though they retain v2 behavior. For example, the Source IP address is not relevant for v2 entries.

Static Multicast Router

A static multicast router (Mrouter) can be configured for a particular port on a particular VLAN. A static Mrouter does not have to be learned through IGMP Snooping.

A total of 16 static Mrouters can be configured on the GbESM. Both internal and external ports can accept a static Mrouter.

Note – When static Mrouters are used, the switch will continue learning dynamic Mrouters via IGMP snooping. However, dynamic Mrouters may not replace static Mrouters. If a dynamic Mrouter has the same port and VLAN combination as a static Mrouter, the dynamic Mrouter will not be learned.

Following is an example of configuring a static multicast router:

1. Configure a port to which the static Multicast Router is connected, and enter the appropriate VLAN.

>> / cfg/13/igmp/mrouter	<i>(Select IGMP Mrouter menu)</i>
>> Static Multicast Router# add EXT4	<i>(Add EXT4 Static Mrouter port)</i>
Enter VLAN number: (1-4094) 1	<i>(Enter the VLAN number)</i>
Enter the version number of mrouter [1 2 3]: 2	<i>(Enter IGMP version number)</i>

2. Apply, verify, and save the configuration.

>> Static Multicast Router# apply	<i>(Apply the configuration)</i>
>> Static Multicast Router# cur	<i>(View the configuration)</i>
>> Static Multicast Router# save	<i>(Save your changes)</i>

IGMP Relay

The GbESM can act as an IGMP Relay (or IGMP Proxy) device that relays IGMP multicast messages and traffic between an Mrouter and end stations. IGMP Relay allows the GbESM to participate in network multicasts with no configuration of the various multicast routing protocols, so you can deploy it in the network with minimal effort.

To an IGMP host connected to the GbESM, IGMP Relay appears to be an IGMP multicast router (Mrouter). IGMP Relay sends Membership Queries to hosts, which respond by sending an IGMP response message. A host can also send an unsolicited Join message to the IGMP Relay.

To a multicast router, IGMP Relay appears as a host. The Mrouter sends IGMP host queries to IGMP Relay, and IGMP Relay responds by forwarding IGMP host reports and unsolicited join messages from its attached hosts.

IGMP Relay also forwards multicast traffic between the Mrouter and end stations, similar to IGMP Snooping.

You can configure up to two Mrouters to use with IGMP Relay. One Mrouter acts as the primary Mrouter, and one is the backup Mrouter. The GbESM uses health checks to select the primary Mrouter.

Configuration Guidelines

Consider the following guidelines when you configure IGMP Relay:

- IGMP Relay and IGMP Snooping are mutually exclusive—if you enable IGMP Relay, you must turn off IGMP Snooping.
- Upstream Mrouters must be connected to external ports (EXTx).
- Add the upstream Mrouter VLAN to the IGMP Relay list, using the following command:

```
/cfg/l3/igmp/relay/add <VLAN number>
```
- If IGMP hosts reside on different VLANs, you must disable IGMP flooding (`/cfg/l3/igmp/adv/flood dis`) and enable CPU forwarding (`/cfg/l3/igmp/adv/cpu ena`) to ensure that multicast data is forwarded across the VLANs.

Configure IGMP Relay

Use the following procedure to configure IGMP Relay.

1. Configure an IP interface and assign VLANs.

```
>> /cfg/l3/if 2                               (Select IP interface 2)
>> IP Interface 2# addr 10.10.1.1             (Configure IP address for IF 2)
>> IP Interface 2# mask 255.255.255.0        (Configure mask for IF 2)
>> IP Interface 2# vlan 2                     (Assign VLAN 2 to IF 2)
>> /cfg/l3/if 3                               (Select IP interface 3)
>> IP Interface 3# addr 10.10.2.1             (Configure IP address for IF 3)
>> IP Interface 3# mask 255.255.255.0        (Configure mask for IF 3)
>> IP Interface 3# vlan 3                     (Assign VLAN 3 to IF 3)
```

2. Turn IGMP on.

```
>> /cfg/l3/igmp/on                             (Turn on IGMP)
```

3. Enable IGMP Relay and add VLANs to the downstream network.

```
>> /cfg/l3/igmp/relay/ena                     (Enable IGMP Relay)
>> IGMP Relay# add 2                           (Add VLAN 2 to IGMP Relay)
Vlan 2 added.
>> IGMP Relay# add 3                           (Add VLAN 3 to IGMP Relay)
Vlan 3 added.
```

4. Configure the upstream Mrouters.

```
>> IGMP Relay# mrtr 1/addr 100.0.1.2/ena
Current IP address:    0.0.0.0
New pending IP address: 100.0.1.2
Current status: disabled
New status:    enabled
>> Multicast router 1# ..
>> IGMP Relay# mrtr 2/addr 100.0.2.4/ena
Current IP address:    0.0.0.0
New pending IP address: 100.0.2.4
Current status: disabled
New status:    enabled
```

5. Apply and save the configuration.

```
>> Multicast router 2# apply                   (Apply the configuration)
>> Multicast router 2# save                     (Save the configuration)
```

Additional IGMP Features

The following topics are discussed in this section:

- [“FastLeave” on page 174](#)
- [“IGMP Filtering” on page 175](#)

FastLeave

In normal IGMP operation, when the switch receives an IGMPv2 leave message, it sends a Group-Specific Query to determine if any other devices in the same group (and on the same port) are still interested in the specified multicast group traffic. The switch removes the affiliated port from that particular group, if the following conditions apply:

- If the switch does not receive an IGMP Membership Report within the query-response-interval.
- If no multicast routers have been learned on that port.

With FastLeave enabled on the VLAN, a port can be removed immediately from the port list of the group entry when the IGMP Leave message is received, unless a multicast router was learned on the port.

Enable FastLeave only on VLANs that have only one host connected to each physical port.

IGMP Filtering

With IGMP Filtering, you can allow or deny a port to send and receive multicast traffic to certain multicast groups. Unauthorized users are restricted from streaming multicast traffic across the network.

If access to a multicast group is denied, IGMP Membership Reports from the port are dropped, and the port is not allowed to receive IP multicast traffic from that group. If access to the multicast group is allowed, Membership Reports from the port are forwarded for normal processing.

To configure IGMP Filtering, you must globally enable IGMP filtering, define an IGMP filter, assign the filter to a port, and enable IGMP Filtering on the port. To define an IGMP filter, you must configure a range of IP multicast groups, choose whether the filter will allow or deny multicast traffic for groups within the range, and enable the filter.

Configuring the Range

Each IGMP Filter allows you to set a start and end point that defines the range of IP addresses upon which the filter takes action. Each IP address in the range must be between 224.0.1.0 and 239.255.255.255.

Configuring the Action

Each IGMP filter can allow or deny IP multicasts to the range of IP addresses configured. If you configure the filter to deny IP multicasts, then IGMP Membership Reports from multicast groups within the range are dropped. You can configure a secondary filter to allow IP multicasts to a small range of addresses within a larger range that a primary filter is configured to deny. The two filters work together to allow IP multicasts to a small subset of addresses within the larger range of addresses.

Note – Lower-numbered filters take precedence over higher-number filters. For example, the action defined for IGMP Filter 1 supersedes the action defined for IGMP Filter 2.

Configure IGMP Filtering

1. Enable IGMP Filtering on the switch.

```
>> /cfg/l3/igmp/igmpflt (Select IGMP Filtering menu)
>> IGMP Filter# ena (Enable IGMP Filtering)
Current status: disabled
New status: enabled
```

2. Define an IGMP filter.

```
>> /cfg/l3/igmp/igmpflt (Select IGMP Filtering menu)
>>IGMP Filter# filter 1 (Select Filter 1 Definition menu)
>>IGMP Filter 1 Definition# range 224.0.1.0(Enter first IP address of the range)
Current multicast address2:
Enter new multicast address2: 226.0.0.0 (Enter second IP address)
Current multicast address1:
New pending multicast address1: 224.0.1.0
Current multicast address2:
New pending multicast address2: 226.0.0.0
>>IGMP Filter 1 Definition# action deny (Deny multicast traffic)
>>IGMP Filter 1 Definition# ena (Enable the filter)
```

3. Assign the IGMP filter to a port.

```
>> /cfg/l3/igmp/igmpflt (Select IGMP Filtering menu)
>>IGMP Filter# port EXT3 (Select port EXT3)
>>IGMP Port EXT3# filt ena (Enable IGMP Filtering on the port)
Current port EXT3 filtering: disabled
New port EXT3 filtering: enabled
>>IGMP Port EXT3# add 1 (Add IGMP Filter 1 to the port)
>>IGMP Port EXT3# apply (Make your changes active)
```

CHAPTER 12

Border Gateway Protocol

Border Gateway Protocol (BGP) is an Internet protocol that enables routers on a network to share and advertise routing information with each other about the segments of the IP address space they can access within their network and with routers on external networks. BGP allows you to decide what is the “best” route for a packet to take from your network to a destination on another network rather than simply setting a default route from your border router(s) to your upstream provider(s). BGP is defined in RFC 1771.

GbE Switch Modules (GbESMs) can advertise their IP interfaces and IP addresses using BGP and take BGP feeds from as many as 16 BGP router peers. This allows more resilience and flexibility in balancing traffic from the Internet.

The following topics are discussed in this section:

- [“Internal Routing Versus External Routing” on page 177](#)
- [“Forming BGP Peer Routers” on page 179](#)
- [“What is a Route Map?” on page 179](#)
- [“Aggregating Routes” on page 183](#)
- [“Redistributing Routes” on page 183](#)
- [“BGP Attributes” on page 184](#)
- [“Selecting Route Paths in BGP” on page 185](#)
- [“BGP Failover Configuration” on page 186](#)
- [“Default Redistribution and Route Aggregation Example” on page 188](#)

Internal Routing Versus External Routing

To ensure effective processing of network traffic, every router on your network needs to know how to send a packet (directly or indirectly) to any other location/destination in your network. This is referred to as *internal routing* and can be done with static routes or using active, internal dynamic routing protocols, such as RIP, RIPv2, and OSPF.

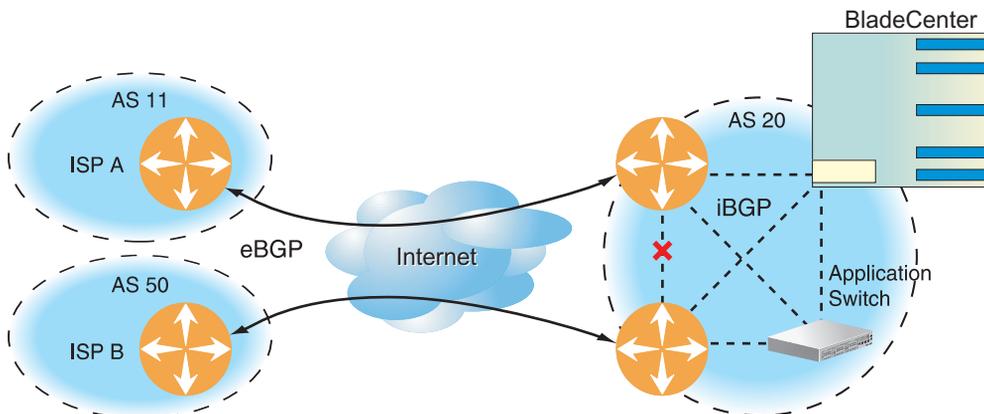
Static routes should have a higher degree of precedence than dynamic routing protocols. If the destination route is not in the route cache, then the packets are forwarded to the default gateway which may be incorrect if a dynamic routing protocol is enabled.

It is also useful to tell routers outside your network (upstream providers or *peers*) about the routes you can access in your network. External networks (those outside your own) that are under the same administrative control are referred to as *autonomous systems* (AS). Sharing of routing information between autonomous systems is known as *external routing*.

External BGP (eBGP) is used to exchange routes between different autonomous systems whereas internal BGP (iBGP) is used to exchange routes within the same autonomous system. An iBGP is a type of internal routing protocol you can use to do active routing inside your network. It also carries AS path information, which is important when you are an ISP or doing BGP transit.

The iBGP peers have to maintain reciprocal sessions to every other iBGP router in the same AS (in a full-mesh manner) in order to propagate route information throughout the AS. If the iBGP session shown between the two routers in AS 20 was not present (as indicated in Figure 21), the top router would not learn the route to AS 50, and the bottom router would not learn the route to AS 11, even though the two AS 20 routers are connected via the BladeCenter and the Application Switch.

Figure 21 iBGP and eBGP



Typically, an AS has one or more *border routers*—peer routers that exchange routes with other ASs—and an internal routing scheme that enables routers in that AS to reach every other router and destination within that AS. When you *advertise* routes to border routers on other autonomous systems, you are effectively committing to carry data to the IP space represented in the route being advertised. For example, if you advertise 192.204.4.0/24, you are declaring that if another router sends you data destined for any address in 192.204.4.0/24, you know how to carry that data to its destination.

Forming BGP Peer Routers

Two BGP routers become peers or neighbors once you establish a TCP connection between them. For each new route, if a peer is interested in that route (for example, if a peer would like to receive your static routes and the new route is static), an update message is sent to that peer containing the new route. For each route removed from the route table, if the route has already been sent to a peer, an update message containing the route to withdraw is sent to that peer.

For each Internet host, you must be able to send a packet to that host, and that host has to have a path back to you. This means that whoever provides Internet connectivity to that host must have a path to you. Ultimately, this means that they must “hear a route” which covers the section of the IP space you are using; otherwise, you will not have connectivity to the host in question.

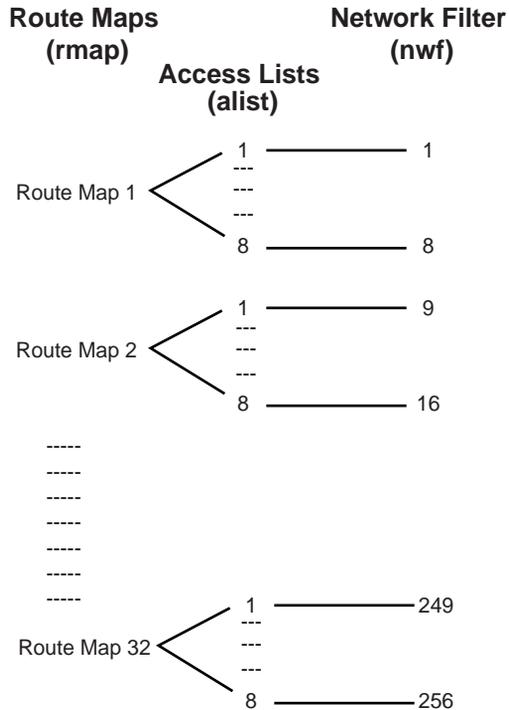
What is a Route Map?

A route map is used to control and modify routing information. Route maps define conditions for redistributing routes from one routing protocol to another or controlling routing information when injecting it in and out of BGP. Route maps are used by OSPF only for redistributing routes. For example, a route map is used to set a preference value for a specific route from a peer router and another preference value for all other routes learned via the same peer router. For example, the following command is used to define a route map:

```
>> # /cfg/13/rmap 1 (Select a route map)
```

A route map allows you to match attributes, such as metric, network address, and AS number. It also allows users to overwrite the local preference metric and to append the AS number in the AS route. See [“BGP Failover Configuration” on page 186](#).

BLADEOS allows you to configure 32 route maps. Each route map can have up to eight access lists. Each access list consists of a network filter. A network filter defines an IP address and subnet mask of the network that you want to include in the filter. [Figure 22](#) illustrates the relationship between route maps, access lists and network filters.

Figure 22 Distributing Network Filters in Access Lists and Route Maps

Incoming and Outgoing Route Maps

You can have two types of route maps: incoming and outgoing. A BGP peer router can be configured to support up to eight route maps in the incoming route map list and outgoing route map list.

If a route map is not configured in the incoming route map list, the router imports all BGP updates. If a route map is configured in the incoming route map list, the router ignores all unmatched incoming updates. If you set the action to **deny**, you must add another route map to permit all unmatched updates.

Route maps in an outgoing route map list behave similar to route maps in an incoming route map list. If a route map is not configured in the outgoing route map list, all routes are advertised or permitted. If a route map in the outgoing route map list is set to **permit**, matched routes are advertised and unmatched routes are ignored.

Precedence

You can set a priority to a route map by specifying a precedence value with the following command:

```
>> /cfg/l3/rmap <x>/pre (Specify a precedence)
```

The smaller the value the higher the precedence. If two route maps have the same precedence value, the smaller number has higher precedence.

Configuration Overview

To configure route maps, you need to do the following:

1. Define network filter.

```
>> # /cfg/l3/nwf 1 (Specify a network filter number)
>> IP Network Filter 1# addr <IP address> (Specify network address)
>> IP Network Filter 1# mask <IP mask> (Specify network mask)
>> IP Network Filter 1# ena (Enable network filter)
```

Enter a filter number from 1 to 256. Specify the IP address and subnet mask of the network that you want to match. Enable the network filter. You can distribute up to 256 network filters among 32 route maps each containing eight access lists.

2. (Optional) Define the criteria for the access list and enable it.

Specify the access list and associate the network filter number configured in Step 1.

```
>> # /cfg/l3/rmap 1 (Specify a route map number)
>> IP Route Map 1# alist 1 (Specify the access list number)
>> IP Access List 1# nwf 1 (Specify the network filter number)
>> IP Access List 1# metric (Define a metric)
>> IP Access List 1# action deny (Specify action for the access list)
>> IP Access List 1# ena (Enable the access list)
```

Steps 2 and 3 are optional, depending on the criteria that you want to match. In Step 2, the network filter number is used to match the subnets defined in the network filter. In Step 3, the autonomous system number is used to match the subnets. Or, you can use both (Step 2 and Step 3) criteria: access list (network filter) and access path (AS filter) to configure the route maps.

3. (Optional) Configure the attributes in the AS filter menu.

```
>> # cfg/l3/rmap 1/aspath 1           (Specify the attributes in the filter)
>> AS Filter 1# as 1                 (Specify the AS number)
>> AS Filter 1# action deny          (Specify the action for the filter)
>> AS Filter 1# ena                  (Enable the AS filter)
```

4. Set up the BGP attributes.

If you want to overwrite the attributes that the peer router is sending, then define the following BGP attributes:

- Specify the AS numbers that you want to prepend to a matched route and the local preference for the matched route.
- Specify the metric [Multi Exit Discriminator (MED)] for the matched route.

```
>> # cfg/l3/rmap 1                   (Specify a route map number)
>> IP Route Map 1# ap                 (Specify the AS numbers to prepend)
>> IP Route Map 1# lp                 (Specify the local preference)
>> IP Route Map 1# metric            (Specify the metric)
```

5. Enable the route map.

```
>> # cfg/l3/rmap 1/en                (Enable the route map)
```

6. Turn BGP on.

```
>> # cfg/l3/bgp/on                  (Globally turn BGP on)
```

7. Assign the route map to a peer router.

Select the peer router and then add the route map to the incoming route map list,

```
>> Border Gateway Protocol# peer 1/addi (Add to the incoming route map)
```

or to the outgoing route map list.

```
>> Border Gateway Protocol# peer 1/addo (Add to the outgoing route map)
```

8. Apply and save the configuration.

```
>> Border Gateway Protocol# apply     (Apply the configuration)
>> Border Gateway Protocol# save      (Save your changes)
```

Aggregating Routes

Aggregation is the process of combining several different routes in such a way that a single route can be advertised, which minimizes the size of the routing table. You can configure aggregate routes in BGP either by redistributing an aggregate route into BGP or by creating an aggregate entry in the BGP routing table.

To define an aggregate route in the BGP routing table, use the following commands:

```
>> # cfg/13/bgp                                (Specify BGP)
>> Border Gateway Protocol# aggr 1             (Specify aggregate list number)
>> BGP aggr 1 # addr                             (Enter aggregation network address)
>> BGP aggr 1 # mask                             (Enter aggregation network mask)
>> BGP aggr 1 # ena                             (Enable aggregation)
```

An example of creating a BGP aggregate route is shown in [“Default Redistribution and Route Aggregation Example”](#) on page 188.

Redistributing Routes

In addition to running multiple routing protocols simultaneously, BLADEOS software can redistribute information from one routing protocol to another. For example, you can instruct the switch to use BGP to re-advertise static routes. This applies to all of the IP-based routing protocols.

You can also conditionally control the redistribution of routes between routing domains by defining a method known as route maps between the two domains. For more information on route maps, see [“What is a Route Map?”](#) on page 179. Redistributing routes is another way of providing policy control over whether to export OSPF routes, fixed routes, and static routes. For an example configuration, see [“Default Redistribution and Route Aggregation Example”](#) on page 188.

Default routes can be configured using the following methods:

- Import
- Originate—The router sends a default route to peers if it does not have any default routes in its routing table.
- Redistribute—Default routes are either configured through the default gateway or learned via other protocols and redistributed to peer routers. If the default routes are from the default gateway, enable the static routes because default routes from the default gateway are static routes. Similarly, if the routes are learned from another routing protocol, make sure you enable that protocol for redistribution.
- None

BGP Attributes

The following two BGP attributes are discussed in this section: Local preference and metric (Multi-Exit Discriminator).

Local Preference Attribute

When there are multiple paths to the same destination, the local preference attribute indicates the preferred path. The path with the higher preference is preferred (the default value of the local preference attribute is 100). Unlike the weight attribute, which is only relevant to the local router, the local preference attribute is part of the routing update and is exchanged among routers in the same AS.

The local preference attribute can be set in one of two ways:

- **`/cfg/13/bgp/pref`**

This command uses the BGP default local preference method, affecting the outbound direction only.

- **`/cfg/13/rmap/lp`**

This command uses the route map local preference method, which affects both inbound and outbound directions.

Metric (Multi-Exit Discriminator) Attribute

This attribute is a hint to external neighbors about the preferred path into an AS when there are multiple entry points. A lower metric value is preferred over a higher metric value. The default value of the metric attribute is 0.

Unlike local preference, the metric attribute is exchanged between ASs; however, a metric attribute that comes into an AS does not leave the AS.

When an update enters the AS with a certain metric value, that value is used for decision making within the AS. When BGP sends that update to another AS, the metric is reset to 0.

Unless otherwise specified, the router compares metric attributes for paths from external neighbors that are in the same AS.

Selecting Route Paths in BGP

BGP selects only one path as the best path. It does not rely on metric attributes to determine the best path. When the same network is learned via more than one BGP peer, BGP uses its policy for selecting the best route to that network. The BGP implementation on the GbESM uses the following criteria to select a path when the same route is received from multiple peers.

1. Local fixed and static routes are preferred over learned routes.
2. With iBGP peers, routes with higher local preference values are selected.
3. In the case of multiple routes of equal preference, the route with lower AS path weight is selected.
AS path weight = $128 \times$ AS path length (number of autonomous systems traversed).
4. In the case of equal weight and routes learned from peers that reside in the same AS, the lower metric is selected.

Note – A route with a metric is preferred over a route without a metric.

5. The lower cost to the next hop of routes is selected.
6. In the case of equal cost, the eBGP route is preferred over iBGP.
7. If all routes are from eBGP, the route with the lower router ID is selected.

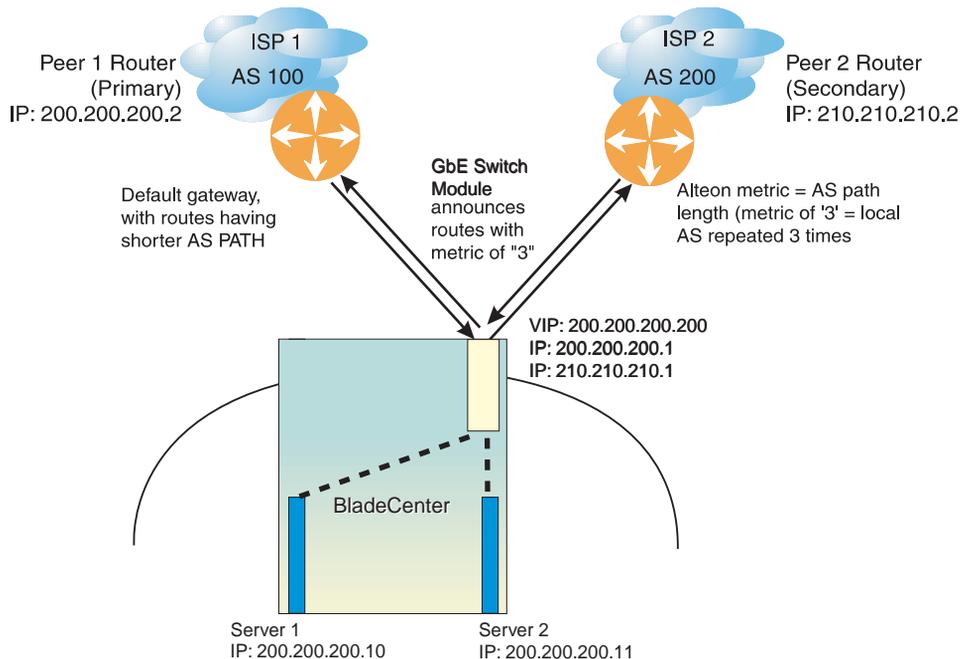
When the path is selected, BGP puts the selected path in its routing table and propagates the path to its neighbors.

BGP Failover Configuration

Use the following example to create redundant default gateways for a GbESM at a Web Host/ISP site, eliminating the possibility, should one gateway go down, that requests will be forwarded to an upstream router unknown to the switch.

As shown in [Figure 23](#), the switch is connected to ISP 1 and ISP 2. The customer negotiates with both ISPs to allow the switch to use their peer routers as default gateways. The ISP peer routers will then need to announce themselves as default gateways to the GbESM.

Figure 23 BGP Failover Configuration Example



On the GbESM, one peer router (the secondary one) is configured with a longer AS path than the other, so that the peer with the shorter AS path will be seen by the switch as the primary default gateway. ISP 2, the secondary peer, is configured with a metric of “3,” thereby appearing to the switch to be three router *hops* away.

1. Define the VLANs.

For simplicity, both default gateways are configured in the same VLAN in this example. The gateways could be in the same VLAN or different VLANs.

```
>> # /cfg/12/vlan 1                                (Select VLAN 1)
>> vlan 1# add <port number>                       (Add a port to the VLAN membership)
```

2. Define the IP interfaces.

The switch will need an IP interface for each default gateway to which it will be connected. Each interface must be placed in the appropriate VLAN. These interfaces will be used as the primary and secondary default gateways for the switch.

```
>> # /cfg/l3/if 1                               (Select interface 1)
>> IP Interface 1# ena                          (Enable switch interface 1)
>> IP Interface 1# addr 200.200.200.1          (Configure IP address of interface 1)
>> IP Interface 1# mask 255.255.255.0        (Configure IP subnet address mask)
>> IP Interface 1# ../if 2                     (Select interface 2)
>> IP Interface 2# ena                          (Enable switch interface 2)
>> IP Interface 2# addr 210.210.210.1        (Configure IP address of interface 2)
>> IP Interface 2# mask 255.255.255.0        (Configure IP subnet address mask)
```

3. Enable IP forwarding.

IP forwarding is turned on by default and is used for VLAN-to-VLAN (non-BGP) routing. Make sure IP forwarding is on if the default gateways are on different subnets or if the switch is connected to different subnets and those subnets need to communicate through the switch (which they almost always do).

```
>> # /cfg/l3/frwd/on                            (Enable IP forwarding)
```

Note – To help eliminate the possibility for a Denial of Service (DoS) attack, the forwarding of directed broadcasts is disabled by default.

4. Configure BGP peer router 1 and 2.

```
>> # /cfg/l3/bgp/peer 1                        (Select BGP peer router 1)
>> BGP Peer 1# ena                            (Enable this peer configuration)
>> BGP Peer 1# addr 200.200.200.2            (Set IP address for peer router 1)
>> BGP Peer 1# ras 100                       (Set remote AS number)
>> BGP Peer 1# /cfg/l3/bgp/peer 2           (Select BGP peer router 2)
>> BGP Peer 2# ena                            (Enable this peer configuration)
>> BGP Peer 2# addr 210.210.210.2          (Set IP address for peer router 2)
>> BGP Peer 2# ras 200                       (Set remote AS number)
```

5. On the switch, apply and save your configuration changes.

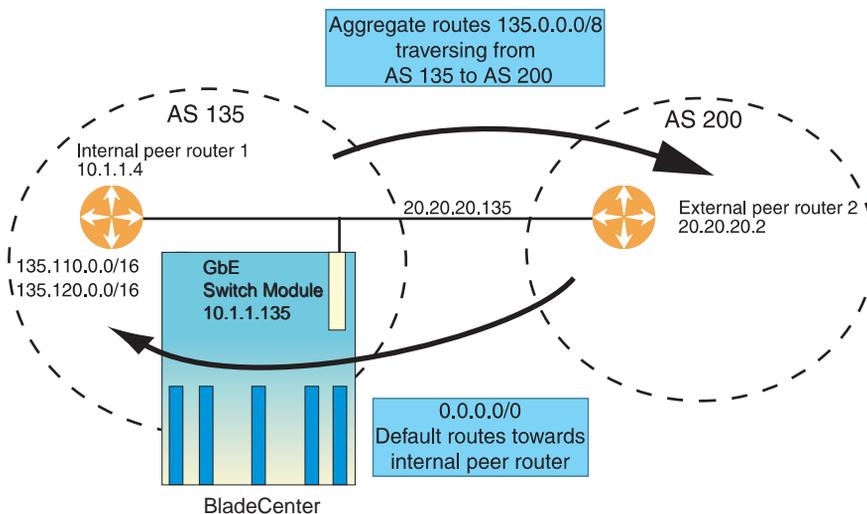
```
>> BGP Peer 2# apply                          (Make your changes active)
>> save                                       (Save for restore after reboot)
```

Default Redistribution and Route Aggregation Example

This example shows you how to configure the switch to redistribute information from one routing protocol to another and create an aggregate route entry in the BGP routing table to minimize the size of the routing table.

As illustrated in [Figure 24](#), you have two peer routers: an internal and an external peer router. Configure the GbESM to redistribute the default routes from AS 200 to AS 135. At the same time, configure for route aggregation to allow you to condense the number of routes traversing from AS 135 to AS 200.

Figure 24 Route Aggregation and Default Route Redistribution



1. Configure the IP interface.
2. Configure the AS number (AS 135) and router ID number (10.1.1.135).

```
>> # /cfg/13/bgp (Select BGP menu)
>> Border Gateway Protocol# as 135 (Specify an AS number)
>> Border Gateway Protocol# .. /rtrid 10.1.1.135(Specify a router ID number)
```

3. Configure internal peer router 1 and external peer router 2.

```

>> # /cfg/l3/bgp/peer 1                (Select internal peer router 1)
>> BGP Peer 1# ena                      (Enable this peer configuration)
>> BGP Peer 1# addr 10.1.1.4           (Set IP address for peer router 1)
>> BGP Peer 1# ras 135                 (Set remote AS number)
>> BGP Peer 1# ../peer 2              (Select external peer router 2)
>> BGP Peer 2# ena                      (Enable this peer configuration)
>> BGP Peer 2# addr 20.20.20.2        (Set IP address for peer router 2)
>> BGP Peer 2# ras 200                 (Set remote AS number)

```

4. Configure redistribution for Peer 1.

```

>> # /cfg/l3/bgp/peer 1/redist         (Select redistribute)
>> BGP Peer 1# default redistribute    (Set default to redistribute)
>> BGP Peer 1# fixed ena              (Enable fixed routes)

```

5. Configure aggregation policy control.

Configure the routes that you want aggregated.

```

>> # /cfg/l3/bgp/aggr 1                (Set aggregation number)
>> BGP aggr 1# addr 135.0.0.0          (Add IP address to aggregate 1)
>> BGP Peer 1# mask 255.0.0.0         (Add IP mask to aggregate 1)

```


CHAPTER 13

OSPF

BLADEOS supports the Open Shortest Path First (OSPF) routing protocol. The BLADEOS implementation conforms to the OSPF version 2 specifications detailed in Internet RFC 1583. The following sections discuss OSPF support for the GbE Switch Module (GbESM):

- [“OSPF Overview” on page 191](#). This section provides information on OSPF concepts, such as types of OSPF areas, types of routing devices, neighbors, adjacencies, link state database, authentication, and internal versus external routing.
- [“OSPF Implementation in BLADEOS” on page 196](#). This section describes how OSPF is implemented in BLADEOS, such as configuration parameters, electing the designated router, summarizing routes, defining route maps and so forth.
- [“OSPF Configuration Examples” on page 205](#). This section provides step-by-step instructions on configuring different configuration examples:
 - Creating a simple OSPF domain
 - Creating virtual links
 - Summarizing routes

OSPF Overview

OSPF is designed for routing traffic within a single IP domain called an Autonomous System (AS). The AS can be divided into smaller logical units known as *areas*.

All routing devices maintain link information in their own Link State Database (LSDB). The LSDB for all routing devices within an area is identical but is not exchanged between different areas. Only routing updates are exchanged between areas, thereby significantly reducing the overhead for maintaining routing information on a large, dynamic network.

The following sections describe key OSPF concepts.

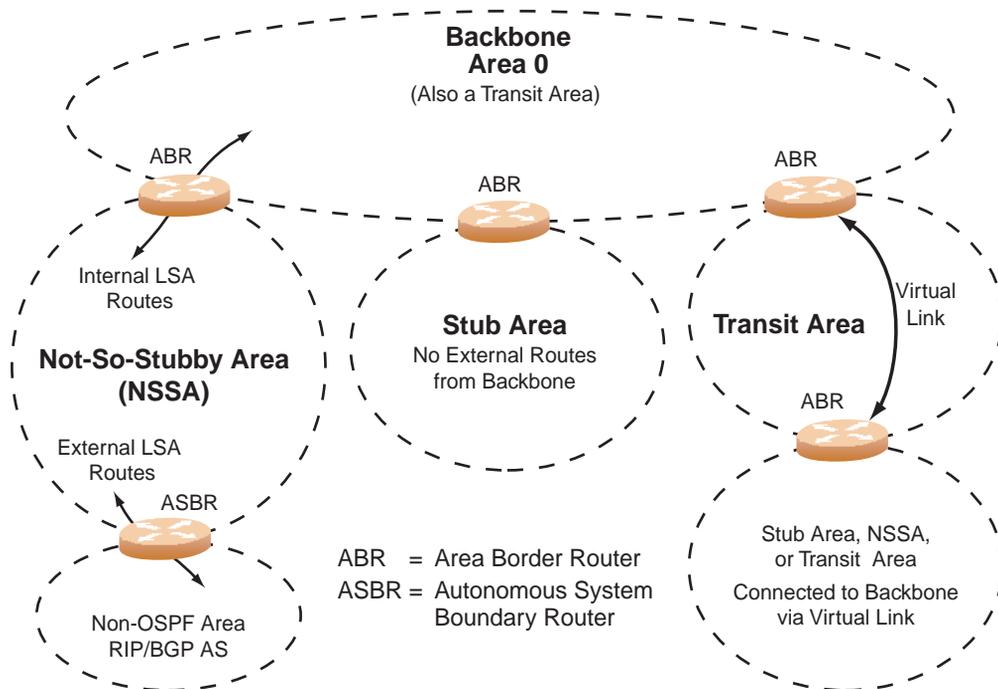
Types of OSPF Areas

An AS can be broken into logical units known as *areas*. In any AS with multiple areas, one area must be designated as area 0, known as the *backbone*. The backbone acts as the central OSPF area. All other areas in the AS must be connected to the backbone. Areas inject summary routing information into the backbone, which then distributes it to other areas as needed.

As shown in [Figure 25](#), OSPF defines the following types of areas:

- **Stub Area**—an area that is connected to only one other area. External route information is not distributed into stub areas.
- **Not-So-Stubby-Area (NSSA)**—similar to a stub area with additional capabilities. Routes originating from within the NSSA can be propagated to adjacent transit and backbone areas. External routes from outside the AS can be advertised within the NSSA but are not distributed into other areas.
- **Transit Area**—an area that allows area summary information to be exchanged between routing devices. The backbone (area 0), any area that contains a virtual link to connect two areas, and any area that is not a stub area or an NSSA are considered transit areas.

Figure 25 OSPF Area Types

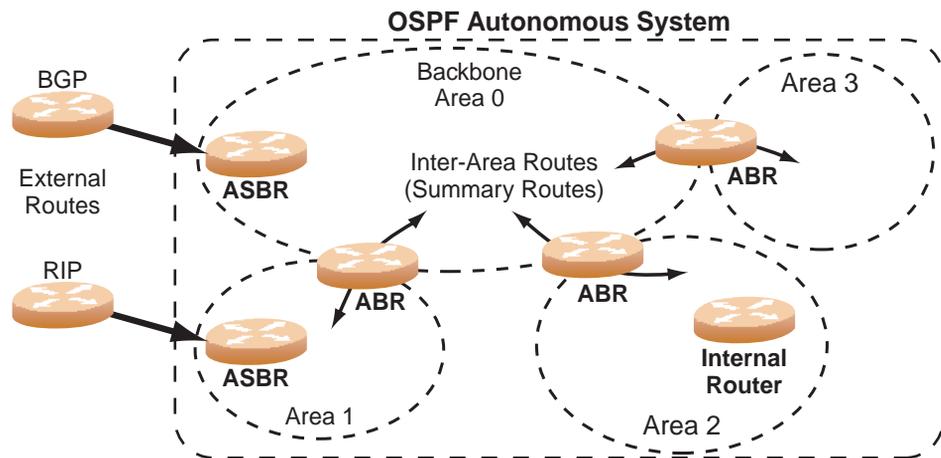


Types of OSPF Routing Devices

As shown in [Figure 26](#), OSPF uses the following types of routing devices:

- Internal Router (IR)—a router that has all of its interfaces within the same area. IRs maintain LSDBs identical to those of other routing devices within the local area.
- Area Border Router (ABR)—a router that has interfaces in multiple areas. ABRs maintain one LSDB for each connected area and disseminate routing information between areas.
- Autonomous System Boundary Router (ASBR)—a router that acts as a gateway between the OSPF domain and non-OSPF domains, such as RIP, BGP, and static routes.

Figure 26 OSPF Domain and an Autonomous System



Neighbors and Adjacencies

In areas with two or more routing devices, *neighbors* and *adjacencies* are formed.

Neighbors are routing devices that maintain information about each others' health. To establish neighbor relationships, routing devices periodically send hello packets on each of their interfaces. All routing devices that share a common network segment, appear in the same area, and have the same health parameters (hello and dead intervals) and authentication parameters respond to each other's hello packets and become neighbors. Neighbors continue to send periodic hello packets to advertise their health to neighbors. In turn, they listen to hello packets to determine the health of their neighbors and to establish contact with new neighbors.

The hello process is used for electing one of the neighbors as the area's Designated Router (DR) and one as the area's Backup Designated Router (BDR). The DR is adjacent to all other neighbors and acts as the central contact for database exchanges. Each neighbor sends its database information to the DR, which relays the information to the other neighbors.

The BDR is adjacent to all other neighbors (including the DR). Each neighbor sends its database information to the BDR just as with the DR, but the BDR merely stores this data and does not distribute it. If the DR fails, the BDR will take over the task of distributing database information to the other neighbors.

The Link-State Database

OSPF is a link-state routing protocol. A *link* represents an interface (or routable path) from the routing device. By establishing an adjacency with the DR, each routing device in an OSPF area maintains an identical Link-State Database (LSDB) describing the network topology for its area.

Each routing device transmits a Link-State Advertisement (LSA) on each of its *active* interfaces. LSAs are entered into the LSDB of each routing device. OSPF uses *flooding* to distribute LSAs between routing devices. Interfaces may also be *passive*. Passive interfaces send LSAs to active interfaces, but do not receive LSAs, hello packets, or any other OSPF protocol information from active interfaces. Passive interfaces behave as stub networks, allowing OSPF routing devices to be aware of devices that do otherwise participate in OSPF (either because they do not support it, or because the administrator chooses to restrict OSPF traffic exchange or transit).

When LSAs result in changes to the routing device's LSDB, the routing device forwards the changes to the adjacent neighbors (the DR and BDR) for distribution to the other neighbors.

OSPF routing updates occur only when changes occur, instead of periodically. For each new route, if an adjacency is interested in that route (for example, if configured to receive static routes and the new route is indeed static), an update message containing the new route is sent to the adjacency. For each route removed from the route table, if the route has already been sent to an adjacency, an update message containing the route to withdraw is sent.

The Shortest Path First Tree

The routing devices use a link-state algorithm (Dijkstra's algorithm) to calculate the shortest path to all known destinations, based on the cumulative *cost* required to reach the destination.

The cost of an individual interface in OSPF is an indication of the overhead required to send packets across it. The cost is inversely proportional to the bandwidth of the interface. A lower cost indicates a higher bandwidth.

Internal Versus External Routing

To ensure effective processing of network traffic, every routing device on your network needs to know how to send a packet (directly or indirectly) to any other location/destination in your network. This is referred to as *internal routing* and can be done with static routes or using active internal routing protocols, such as OSPF, RIP, or RIPv2.

It is also useful to tell routers outside your network (upstream providers or *peers*) about the routes you have access to in your network. Sharing of routing information between autonomous systems is known as *external routing*.

Typically, an AS will have one or more border routers (peer routers that exchange routes with other OSPF networks) as well as an internal routing system enabling every router in that AS to reach every other router and destination within that AS.

When a routing device *advertises* routes to boundary routers on other autonomous systems, it is effectively committing to carry data to the IP space represented in the route being advertised. For example, if the routing device advertises 192.204.4.0/24, it is declaring that if another router sends data destined for any address in the 192.204.4.0/24 range, it will carry that data to its destination.

OSPF Implementation in BLADEOS

BLADEOS supports a single instance of OSPF and up to 4K routes on the network. The following sections describe OSPF implementation in BLADEOS:

- [“Configurable Parameters” on page 196](#)
- [“Defining Areas” on page 197](#)
- [“Interface Cost” on page 199](#)
- [“Electing the Designated Router and Backup” on page 199](#)
- [“Summarizing Routes” on page 199](#)
- [“Default Routes” on page 200](#)
- [“Virtual Links” on page 201](#)
- [“Router ID” on page 201](#)
- [“Authentication” on page 202](#)

Configurable Parameters

In BLADEOS, OSPF parameters can be configured through the Command Line Interfaces (CLI/ISCLI), Browser-Based Interface (BBI), or through SNMP. For more information, see [“Accessing the Switch” on page 23.](#)”

The CLI supports the following parameters: interface output cost, interface priority, dead and hello intervals, retransmission interval, and interface transmit delay.

In addition to the above parameters, you can also specify the following:

- Shortest Path First (SPF) interval—Time interval between successive calculations of the shortest path tree using the Dijkstra’s algorithm.
- Stub area metric—A stub area can be configured to send a numeric metric value such that all routes received via that stub area carry the configured metric to potentially influence routing decisions.
- Default routes—Default routes with weight metrics can be manually injected into transit areas. This helps establish a preferred route when multiple routing devices exist between two areas. It also helps route traffic to external networks.
- Passive—When enabled, the interface sends LSAs to upstream devices, but does not otherwise participate in OSPF protocol exchanges.
- Point-to-Point—For LANs that have only two OSPF routing agents (the GbESM and one other device), this option allows the switch to significantly reduce the amount of routing information it must carry and manage.

Defining Areas

If you are configuring multiple areas in your OSPF domain, one of the areas must be designated as area 0, known as the *backbone*. The backbone is the central OSPF area and is usually physically connected to all other areas. The areas inject routing information into the backbone which, in turn, disseminates the information into other areas.

Since the backbone connects the areas in your network, it must be a contiguous area. If the backbone is partitioned (possibly as a result of joining separate OSPF networks), parts of the AS will be unreachable, and you will need to configure *virtual links* to reconnect the partitioned areas (see “[Virtual Links](#)” on page 201).

Up to three OSPF areas can be connected to the GbESM with BLADEOS software. To configure an area, the OSPF number must be defined and then attached to a network interface on the switch. The full process is explained in the following sections.

An OSPF area is defined by assigning *two* pieces of information—an *area index* and an *area ID*. The command to define an OSPF area is as follows:

```
>> # /cfg/13/ospf/aindex <area index>/areaid <n.n.n.n>
```

Note – The `aindex` option above is an arbitrary index used only on the switch and does not represent the actual OSPF area number. The actual OSPF area number is defined in the `areaid` portion of the command as explained in the following sections.

Assigning the Area Index

The `aindex <area index>` option is actually just an arbitrary index (0-2) used only by the GbESM. This index does not necessarily represent the OSPF area number, though for configuration simplicity, it should where possible.

For example, both of the following sets of commands define OSPF area 0 (the backbone) and area 1 because that information is held in the area ID portion of the command. However, the first set of commands is easier to maintain because the arbitrary area indexes agree with the area IDs:

- Area index and area ID agree

```
/cfg/13/ospf/aindex 0/areaid 0.0.0.0 (Use index 0 to set area 0 in ID octet format)
/cfg/13/ospf/aindex 1/areaid 0.0.0.1 (Use index 1 to set area 1 in ID octet format)
```

- Area index set to an arbitrary value

```
/cfg/13/ospf/aindex 1/areaid 0.0.0.0 (Use index 1 to set area 0 in ID octet format)
/cfg/13/ospf/aindex 2/areaid 0.0.0.1 (Use index 2 to set area 1 in ID octet format)
```

Using the Area ID to Assign the OSPF Area Number

The OSPF area number is defined in the `areaid <IP address>` option. The octet format is used in order to be compatible with two different systems of notation used by other OSPF network vendors. There are two valid ways to designate an area ID:

- Placing the area number in the last octet (0.0.0.*n*)

Most common OSPF vendors express the area ID number as a single number. For example, the Cisco IOS-based router command “network 1.1.1.0 0.0.0.255 area 1” defines the area number simply as “area 1.” On the GbESM, using the last octet in the area ID, “area 1” is equivalent to “areaid 0.0.0.1”.
- Multi-octet (*IP address*)

Some OSPF vendors express the area ID number in multi-octet format. For example, “area 2.2.2.2” represents OSPF area 2 and can be specified directly on the GbESM as “areaid 2.2.2.2”.

Note – Although both types of area ID formats are supported, be sure that the area IDs are in the same format throughout an area.

Attaching an Area to a Network

Once an OSPF area has been defined, it must be associated with a network. To attach the area to a network, you must assign the OSPF area index to an IP interface that participates in the area. The format for the command is as follows:

```
>> # /cfg/l3/ospf/if <interface number>/aindex <area index>
```

For example, the following commands could be used to configure IP interface 14 for a presence on the 10.10.10.1/24 network, to define OSPF area 1, and to attach the area to the network:

```
>> # /cfg/l3/if 14                                     (Select menu for IP interface 14)
>> IP Interface 14# addr 10.10.10.1                    (Define IP address on backbone)
>> IP Interface 14# mask 255.255.255.0                (Define IP mask on backbone)
>> IP Interface 14# ena                                 (Enable IP interface 14)
>> IP Interface 14# ../ospf/aindex 1                   (Select menu for area index 1)
>> OSPF Area (index) 1 # areaid 0.0.0.1                (Define area ID as OSPF area 1)
>> OSPF Area (index) 1 # ena                            (Enable area index 1)
>> OSPF Area (index) 1 # ../if 14                       (Select OSPF menu for interface 14)
>> OSPF Interface 14# aindex 1                          (Attach area to interface 14 network)
>> OSPF Interface 14# enable                            (Enable interface 14 for area index 1)
```

Interface Cost

The OSPF link-state algorithm (Dijkstra's algorithm) places each routing device at the root of a tree and determines the cumulative *cost* required to reach each destination. Usually, the cost is inversely proportional to the bandwidth of the interface. Low cost indicates high bandwidth. You can manually enter the cost for the output route with the following command:

```
>> # /cfg/13/ospf/if <OSPF interface number>/cost <cost value (1-65535)>
```

Electing the Designated Router and Backup

In any area with more than two routing devices, a Designated Router (DR) is elected as the central contact for database exchanges among neighbors, and a Backup Designated Router (BDR) is elected in case the DR fails.

DR and BDR elections are made through the hello process. The election can be influenced by assigning a priority value to the OSPF interfaces on the GbESM. The command is as follows:

```
>> # /cfg/13/ospf/if <OSPF interface number>/prio <priority value (0-255)>
```

A priority value of 255 is the highest, and 1 is the lowest. A priority value of 0 specifies that the interface cannot be used as a DR or BDR. In case of a tie, the routing device with the highest router ID wins. Interfaces configured as *passive* do not participate in the DR or BDR election process:

```
>> # /cfg/13/ospf/if <OSPF interface number>/passive enable
```

Summarizing Routes

Route summarization condenses routing information. Without summarization, each routing device in an OSPF network would retain a route to every subnet in the network. With summarization, routing devices can reduce some sets of routes to a single advertisement, reducing both the load on the routing device and the perceived complexity of the network. The importance of route summarization increases with network size.

Summary routes can be defined for up to 16 IP address ranges using the following command:

```
>> # /cfg/13/ospf/range <range number>/addr <IP address>/mask <mask>
```

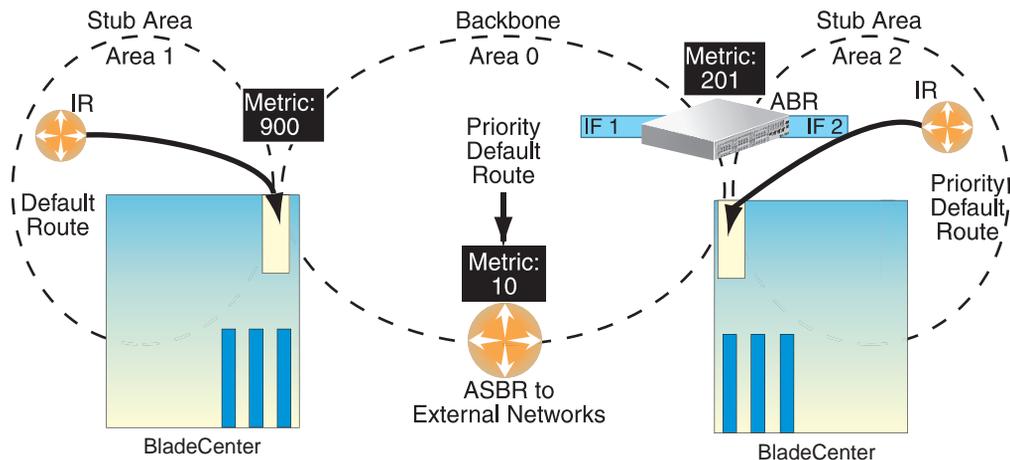
where *<range number>* is a number 1 to 16, *<IP address>* is the base IP address for the range, and *<mask>* is the IP address mask for the range. For a detailed configuration example, see [“Example 3: Summarizing Routes”](#) on page 212.

Default Routes

When an OSPF routing device encounters traffic for a destination address it does not recognize, it forwards that traffic along the *default route*. Typically, the default route leads upstream toward the backbone until it reaches the intended area or an external router.

Each GbESM acting as an ABR automatically inserts a default route into each attached area. In simple OSPF stub areas or NSSAs with only one ABR leading upstream (see Area 1 in [Figure 27](#)), any traffic for IP address destinations outside the area is forwarded to the switch's IP interface, and then into the connected transit area (usually the backbone). Since this is automatic, no further configuration is required for such areas.

Figure 27 Injecting Default Routes



If the switch is in a transit area and has a configured default gateway, it can inject a default route into rest of the OSPF domain. Use the following command to configure the switch to inject OSPF default routes:

```
>> # /cfg/13/ospf/default <metric value> <metric type (1 or 2)>
```

In the command above, *<metric value>* sets the priority for choosing this switch for default route. The value *none* sets no default and 1 sets the highest priority for default route. Metric type determines the method for influencing routing decisions for external routes.

When the switch is configured to inject a default route, an AS-external LSA with link state ID 0.0.0.0 is propagated throughout the OSPF routing domain. This LSA is sent with the configured metric value and metric type.

The OSPF default route configuration can be removed with the command:

```
>> # /cfg/13/ospf/default none
```

Virtual Links

Usually, all areas in an OSPF AS are physically connected to the backbone. In some cases where this is not possible, you can use a *virtual link*. Virtual links are created to connect one area to the backbone through another non-backbone area (see [Figure 25 on page 192](#)).

The area which contains a virtual link must be a transit area and have full routing information. Virtual links cannot be configured inside a stub area or NSSA. The area type must be defined as `transit` using the following command:

```
>> # /cfg/13/ospf/aindex <area index>/type transit
```

The virtual link must be configured on the routing devices at each endpoint of the virtual link, though they may traverse multiple routing devices. To configure a GbESM as one endpoint of a virtual link, use the following command:

```
>> # /cfg/13/ospf/virt <link number>/aindex <area index>/nbr <router ID>
```

where `<link number>` is a value between 1 and 3, `<area index>` is the OSPF area index of the transit area, and `<router ID>` is the IP address of the virtual neighbor (nbr), the routing device at the target endpoint. Another router ID is needed when configuring a virtual link in the other direction. To provide the GbESM with a router ID, see the following section, [Router ID](#).

For a detailed configuration example on Virtual Links, see [“Example 2: Virtual Links” on page 208](#).

Router ID

Routing devices in OSPF areas are identified by a router ID, expressed in IP address format. The router ID is not required to be part of any IP interface range or in any OSPF area, and may even use the GbESM loopback interface (see [“Loopback Interfaces in OSPF” on page 204](#)).

The router ID can be configured in one of the following two ways:

- Dynamically (the default)—OSPF protocol configures the router ID as the lowest IP loopback interface IP address, if available, or else the lowest IP interface IP address, if available. Once dynamically configured, the router ID does not normally undergo further updates.
- Statically—Use the following command to manually configure the router ID:

```
>> # /cfg/13/rtrid <IP address>
```

To change the router ID from static to dynamic, set the router ID to 0.0.0.0, save the configuration, and reboot the GbESM. To view the router ID, enter:

```
>> # /info/13/ospf/gen
```

Authentication

OSPF protocol exchanges can be authenticated so that only trusted routing devices can participate. This ensures less processing on routing devices that are not listening to OSPF packets.

OSPF allows packet authentication and uses IP multicast when sending and receiving packets. Routers participate in routing domains based on pre-defined passwords. BLADEOS supports simple password (type 1 plain text passwords) and MD5 cryptographic authentication. This type of authentication allows a password to be configured per area.

Figure 28 shows authentication configured for area 0 with the password `test`. Simple authentication is also configured for the virtual link between area 2 and area 0. Area 1 is not configured for OSPF authentication.

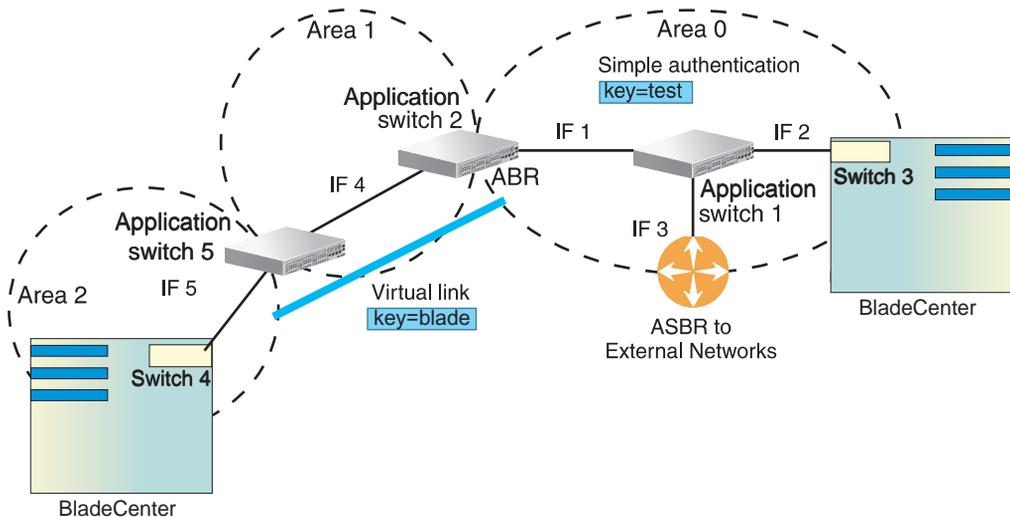


Figure 28 OSPF Authentication

Configuring Plain Text OSPF Passwords

To configure plain text OSPF passwords as shown in Figure 28 use the following commands:

1. Enable OSPF authentication for Area 0 on switches 1, 2, and 3.

```
>> # /cfg/13/ospf/aindex 0/auth password (Turn on password authentication)
```

2. Configure a simple text password up to eight characters for each OSPF IP interface in Area 0 on switches 1, 2, and 3.

```
>> # /cfg/13/ospf/if 1/key test
>> OSPF Interface 1 # ../if 2/key test
>> OSPF Interface 2 # ../if 3/key test
```

3. Enable OSPF authentication for Area 2 on switch 4.

```
>> # /cfg/13/ospf/aindex 2/auth password      (Turn on password authentication)
```

4. Configure a simple text password up to eight characters for the virtual link between Area 2 and Area 0 on switches 2 and 4.

```
>> # /cfg/13/ospf/virt 1/key blade
```

Configuring MD5 Authentication

Use the following commands to configure MD5 authentication on the switches shown in [Figure 28](#):

1. Enable OSPF MD5 authentication for Area 0 on switches 1, 2, and 3.

```
>> # /cfg/13/ospf/aindex 0/auth md5          (Turn on MD5 authentication)
```

2. Configure MD5 key ID for Area 0 on switches 1, 2, and 3.

```
>> # /cfg/13/ospf/md5key 1/key test
```

3. Assign MD5 key ID to OSPF interfaces on switches 1, 2, and 3.

```
>> # /cfg/13/ospf/if 1/mdkey 1
>> OSPF Interface 1 # ../if 2/mdkey 1
>> OSPF Interface 2 # ../if 3/mdkey 1
```

4. Enable OSPF MD5 authentication for Area 2 on switch 4.

```
>> # /cfg/13/ospf/aindex 2/auth md5
```

5. Configure MD5 key for the virtual link between Area 2 and Area 0 on switches 2 and 4.

```
>> # /cfg/13/ospf/md5key 2/key blade
```

6. Assign MD5 key ID to OSPF virtual link on switches 2 and 4.

```
>> # /cfg/13/ospf/virt 1/mdkey 2
```

Host Routes for Load Balancing

BLADEOS implementation of OSPF includes host routes. Host routes are used for advertising network device IP addresses to external networks, accomplishing the following goals:

- ABR Load Sharing

As a form of load balancing, host routes can be used for dividing OSPF traffic among multiple ABRs. To accomplish this, each switch provides identical services but advertises a host route for a different IP address to the external network. If each IP address serves a different and equal portion of the external world, incoming traffic from the upstream router should be split evenly among ABRs.

- ABR Failover

Complementing ABR load sharing, identical host routes can be configured on each ABR. These host routes can be given different costs so that a different ABR is selected as the preferred route for each server and the others are available as backups for failover purposes.

- Equal Cost Multipath (ECMP)

With equal cost multipath, a router potentially has several available next hops towards any given destination. ECMP allows separate routes to be calculated for each IP Type of Service. All paths of equal cost to a given destination are calculated, and the next hops for all equal-cost paths are inserted into the routing table.

If redundant routes via multiple routing processes (such as OSPF, RIP, BGP, or static routes) exist on your network, the switch defaults to the OSPF-derived route.

Loopback Interfaces in OSPF

Because loopback interfaces are always available on the switch, loopback interfaces may present an advantage when used as the router ID.

If dynamic router ID selection is used (see “[Router ID](#)” on page 201), loopback interfaces can be used to force router ID selection. If a loopback interface is configured, its IP address is automatically selected as the router ID, even if other IP interfaces have lower IP addresses. If more than one loopback interface is configured, the lowest loopback interface IP address is selected.

Loopback interfaces can be advertised into the OSPF domain by specifying an OSPF host route with the loopback interface IP address.

Note – Loopback interfaces are not advertised via the OSPF route redistribution of fixed routes.

OSPF Features Not Supported in This Release

The following OSPF features are not supported in this release:

- Summarizing external routes
- Filtering OSPF routes
- Using OSPF to forward multicast routes
- Configuring OSPF on non-broadcast multi-access networks (such as frame relay, X.25, or ATM)

OSPF Configuration Examples

A summary of the basic steps for configuring OSPF on the GbESM is listed here. Detailed instructions for each of the steps is covered in the following sections:

1. Configure IP interfaces.

One IP interface is required for each desired network (range of IP addresses) being assigned to an OSPF area on the switch.

2. (Optional) Configure the router ID.

The router ID is required only when configuring virtual links on the switch.

3. Enable OSPF on the switch.

4. Define the OSPF areas.

5. Configure OSPF interface parameters.

IP interfaces are used for attaching networks to the various areas.

6. (Optional) Configure route summarization between OSPF areas.

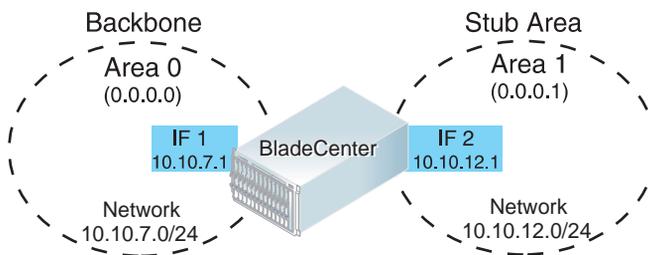
7. (Optional) Configure virtual links.

8. (Optional) Configure host routes.

Example 1: Simple OSPF Domain

In this example, two OSPF areas are defined—one area is the backbone and the other is a stub area. A stub area does not allow advertisements of external routes, thus reducing the size of the database. Instead, a default summary route of IP address 0.0.0.0 is automatically inserted into the stub area. Any traffic for IP address destinations outside the stub area will be forwarded to the stub area's IP interface, and then into the backbone.

Figure 29 A Simple OSPF Domain



Follow this procedure to configure OSPF support as shown in [Figure 29](#):

1. Configure IP interfaces on each network that will be attached to OSPF areas.

In this example, two IP interfaces are needed:

- Interface 1 for the backbone network on 10.10.7.0/24
- Interface 2 for the stub area network on 10.10.12.0/24

```
>> # /cfg/l3/if 1                                     (Select menu for IP interface 1)
>> IP Interface 1 # addr 10.10.7.1                   (Set IP address on backbone)
>> IP Interface 1 # mask 255.255.255.0              (Set IP mask on backbone)
>> IP Interface 1 # enable                           (Enable IP interface 1)
>> IP Interface 1 # ../if 2                           (Select menu for IP interface 2)
>> IP Interface 2 # addr 10.10.12.1                  (Set IP address on stub area)
>> IP Interface 2 # mask 255.255.255.0              (Set IP mask on stub area)
>> IP Interface 2 # enable                           (Enable IP interface 2)
```

2. Enable OSPF.

```
>> IP Interface 2 # /cfg/l3/ospf/on                 (Enable OSPF on the switch)
```

3. Define the backbone.

The backbone is always configured as a transit area using `areaid 0.0.0.0`.

```
>> Open Shortest Path First # aindex 0           (Select menu for area index 0)
>> OSPF Area (index) 0 # areaid 0.0.0.0         (Set the ID for backbone area 0)
>> OSPF Area (index) 0 # type transit           (Define backbone as transit type)
>> OSPF Area (index) 0 # enable                 (Enable the area)
```

4. Define the stub area.

```
>> OSPF Area (index) 0 # ../aindex 1           (Select menu for area index 1)
>> OSPF Area (index) 1 # areaid 0.0.0.1         (Set the area ID for OSPF area 1)
>> OSPF Area (index) 1 # type stub             (Define area as stub type)
>> OSPF Area (index) 1 # enable                 (Enable the area)
```

5. Attach the network interface to the backbone.

```
>> OSPF Area 1 # ../if 1                       (Select OSPF menu for interface 1)
>> OSPF Interface 1 # aindex 0                 (Attach network to backbone index)
>> OSPF Interface 1 # enable                   (Enable the backbone interface)
```

6. Attach the network interface to the stub area.

```
>> OSPF Interface 1 # ../if 2                 (Select OSPF menu for interface 2)
>> OSPF Interface 2 # aindex 1                 (Attach network to stub area index)
>> OSPF Interface 2 # enable                   (Enable the stub area interface)
```

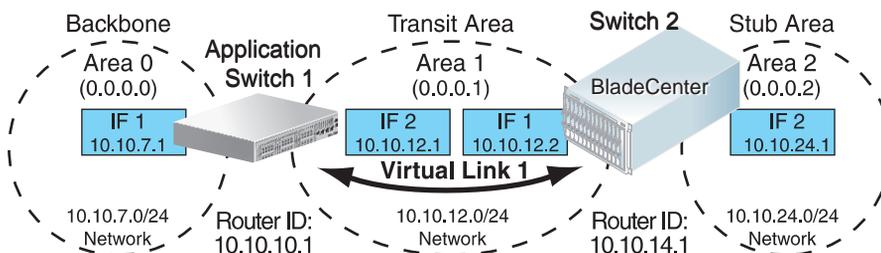
7. Apply and save the configuration changes.

```
>> OSPF Interface 2 # apply                     (Apply all configuration changes)
>> OSPF Interface 2 # save                       (Save applied changes)
```

Example 2: Virtual Links

In the example shown in [Figure 30](#), area 2 is not physically connected to the backbone as is usually required. Instead, area 2 will be connected to the backbone via a virtual link through area 1. The virtual link must be configured at each endpoint.

Figure 30 Configuring a Virtual Link



Configuring OSPF for a Virtual Link on Switch #1

1. Configure IP interfaces on each network that will be attached to the switch.

In this example, two IP interfaces are needed:

- Interface 1 for the backbone network on 10.10.7.0/24
- Interface 2 for the transit area network on 10.10.12.0/24

```
>> # /cfg/13/if 1                                     (Select menu for IP interface 1)
>> IP Interface 1 # addr 10.10.7.1                    (Set IP address on backbone)
>> IP Interface 1 # mask 255.255.255.0              (Set IP mask on backbone)
>> IP Interface 1 # enable                            (Enable IP interface 1)
>> IP Interface 1 # ../if 2                           (Select menu for IP interface 2)
>> IP Interface 2 # addr 10.10.12.1                  (Set IP address on transit area)
>> IP Interface 2 # mask 255.255.255.0              (Set IP mask on transit area)
>> IP Interface 2 # enable                            (Enable interface 2)
```

2. Configure the router ID.

A router ID is required when configuring virtual links. Later, when configuring the other end of the virtual link on Switch 2, the router ID specified here will be used as the target virtual neighbor (nbr) address.

```
>> IP Interface 2 # /cfg/13/rtrid 10.10.10.1        (Set static router ID on switch 1)
```

3. Enable OSPF.

```
>> IP # /cfg/13/ospf/on                             (Enable OSPF on switch 1)
```

4. Define the backbone.

```
>> Open Shortest Path First # aindex 0           (Select menu for area index 0)
>> OSPF Area (index) 0 # areaid 0.0.0.0         (Set area ID for backbone area 0)
>> OSPF Area (index) 0 # type transit           (Define backbone as transit type)
>> OSPF Area (index) 0 # enable                 (Enable the area)
```

5. Define the transit area.

The area that contains the virtual link must be configured as a transit area.

```
>> OSPF Area (index) 0 # ../aindex 1           (Select menu for area index 1)
>> OSPF Area (index) 1 # areaid 0.0.0.1         (Set the area ID for OSPF area 1)
>> OSPF Area (index) 1 # type transit           (Define area as transit type)
>> OSPF Area (index) 1 # enable                 (Enable the area)
```

6. Attach the network interface to the backbone.

```
>> OSPF Area (index) 1 # ../if 1               (Select OSPF menu for interface 1)
>> OSPF Interface 1 # aindex 0                 (Attach network to backbone index)
>> OSPF Interface 1 # enable                   (Enable the backbone interface)
```

7. Attach the network interface to the transit area.

```
>> OSPF Interface 1 # ../if 2                 (Select OSPF menu for interface 2)
>> OSPF Interface 2 # aindex 1                 (Attach network to transit area)
>> OSPF Interface 2 # enable                   (Enable the transit area interface)
```

8. Configure the virtual link.

The nbr router ID configured in this step must be the same as the router ID that will be configured for Switch #2 in [Step 2 on page 210](#).

```
>> OSPF Interface 2 # ../virt 1               (Specify a virtual link number)
>> OSPF Virtual Link 1 # aindex 1             (Set transit area for virtual link)
>> OSPF Virtual Link 1 # nbr 10.10.14.1         (Set the router ID of the recipient)
>> OSPF Virtual Link 1 # enable                 (Enable the virtual link)
```

9. Apply and save the configuration changes.

```
>> OSPF Interface 2 # apply                   (Apply all configuration changes)
>> OSPF Interface 2 # save                     (Save applied changes)
```

Configuring OSPF for a Virtual Link on Switch #2

1. Configure IP interfaces on each network that will be attached to OSPF areas.

In this example, two IP interfaces are needed:

- Interface 1 for the transit area network on 10.10.12.0/24
- Interface 2 for the stub area network on 10.10.24.0/24

```
>> # /cfg/l3/if 1                               (Select menu for IP interface 1)
>> IP Interface 1 # addr 10.10.12.2             (Set IP address on transit area)
>> IP Interface 1 # mask 255.255.255.0         (Set IP mask on transit area)
>> IP Interface 1 # enable                      (Enable IP interface 1)
>> IP Interface 1 # ../if 2                    (Select menu for IP interface 2)
>> IP Interface 2 # addr 10.10.24.1            (Set IP address on stub area)
>> IP Interface 2 # mask 255.255.255.0         (Set IP mask on stub area)
>> IP Interface 2 # enable                      (Enable IP interface 2)
```

2. Configure the router ID.

A router ID is required when configuring virtual links. This router ID should be the same one specified as the target virtual neighbor (nbr) on switch 1 in [Step 8 on page 209](#).

```
>> IP Interface 2 # /cfg/l3/rtrid 10.10.14.1   (Set static router ID on switch 2)
```

3. Enable OSPF.

```
>> IP# /cfg/l3/ospf/on                          (Enable OSPF on switch 2)
```

4. Define the backbone.

This version of BLADEOS requires that a backbone index be configured on the non-backbone end of the virtual link as follows:

```
>> Open Shortest Path First # aindex 0         (Select the menu for area index 0)
>> OSPF Area (index) 0 # areaid 0.0.0.0       (Set the area ID for OSPF area 0)
>> OSPF Area (index) 0 # enable                (Enable the area)
```

5. Define the transit area.

```
>> OSPF Area (index) 0 # ../aindex 1          (Select menu for area index 1)
>> OSPF Area (index) 1 # areaid 0.0.0.1       (Set the area ID for OSPF area 1)
>> OSPF Area (index) 1 # type transit         (Define area as transit type)
>> OSPF Area (index) 1 # enable                (Enable the area)
```

6. Define the stub area.

```
>> OSPF Area (index) 1 # ../aindex 2           (Select the menu for area index 2)
>> OSPF Area (index) 2 # areaid 0.0.0.2       (Set the area ID for OSPF area 2)
>> OSPF Area (index) 2 # type stub           (Define area as stub type)
>> OSPF Area (index) 2 # enable             (Enable the area)
```

7. Attach the network interface to the backbone.

```
>> OSPF Area (index) 2 # ../if 1             (Select OSPF menu for interface 1)
>> OSPF Interface 1 # aindex 1             (Attach network to transit area)
>> OSPF Interface 1 # enable             (Enable the transit area interface)
```

8. Attach the network interface to the transit area.

```
>> OSPF Interface 1 # ../if 2             (Select OSPF menu for interface 2)
>> OSPF Interface 2 # aindex 2             (Attach network to stub area index)
>> OSPF Interface 2 # enable             (Enable the stub area interface)
```

9. Configure the virtual link.

The nbr router ID configured in this step must be the same as the router ID that was configured for switch #1 in [Step 2 on page 208](#).

```
>> OSPF Interface 2 # ../virt 1           (Specify a virtual link number)
>> OSPF Virtual Link 1 # aindex 1         (Set transit area for virtual link)
>> OSPF Virtual Link 1 # nbr 10.10.10.1   (Set the router ID of the recipient)
>> OSPF Virtual Link 1 # enable         (Enable the virtual link)
```

10. Apply and save the configuration changes.

```
>> OSPF Interface 2 # apply             (Apply all configuration changes)
>> OSPF Interface 2 # save             (Save applied changes)
```

Other Virtual Link Options

- You can use redundant paths by configuring multiple virtual links.
- Only the endpoints of the virtual link are configured. The virtual link path may traverse multiple routers in an area as long as there is a routable path between the endpoints.

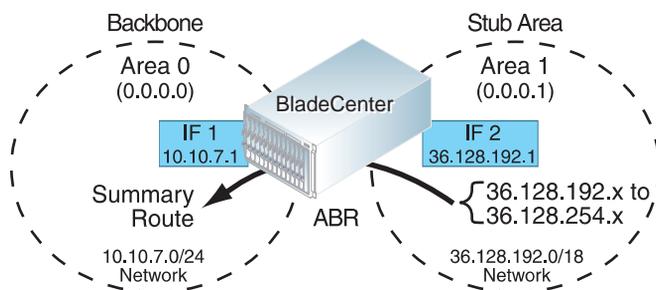
Example 3: Summarizing Routes

By default, ABRs advertise all the network addresses from one area into another area. Route summarization can be used for consolidating advertised addresses and reducing the perceived complexity of the network.

If the network IP addresses in an area are assigned to a contiguous subnet range, you can configure the ABR to advertise a single summary route that includes all the individual IP addresses within the area.

The following example shows one summary route from area 1 (stub area) injected into area 0 (the backbone). The summary route consists of all IP addresses from 36.128.192.0 through 36.128.254.255 except for the routes in the range 36.128.200.0 through 36.128.200.255.

Figure 31 Summarizing Routes



Note – You can specify a range of addresses to prevent advertising by using the `hide` option. In this example, routes in the range 36.128.200.0 through 36.128.200.255 are kept private.

Follow this procedure to configure OSPF support as shown in [Figure 31](#):

1. Configure IP interfaces for each network which will be attached to OSPF areas.

```
>> # /cfg/13/if 1                               (Select menu for IP interface 1)
>> IP Interface 1 # addr 10.10.7.1               (Set IP address on backbone)
>> IP Interface 1 # mask 255.255.255.0          (Set IP mask on backbone)
>> IP Interface 1 # ena                           (Enable IP interface 1)
>> IP Interface 1 # ../if 2                       (Select menu for IP interface 2)
>> IP Interface 2 # addr 36.128.192.1           (Set IP address on stub area)
>> IP Interface 2 # mask 255.255.192.0          (Set IP mask on stub area)
>> IP Interface 2 # ena                           (Enable IP interface 2)
```

2. Enable OSPF.

```
>> IP Interface 2 # /cfg/13/ospf/on             (Enable OSPF on the switch)
```

3. Define the backbone.

```
>> Open Shortest Path First # aindex 0           (Select menu for area index 0)
>> OSPF Area (index) 0 # areaid 0.0.0.0         (Set the ID for backbone area 0)
>> OSPF Area (index) 0 # type transit           (Define backbone as transit type)
>> OSPF Area (index) 0 # enable                 (Enable the area)
```

4. Define the stub area.

```
>> OSPF Area (index) 0 # ../aindex 1           (Select menu for area index 1)
>> OSPF Area (index) 1 # areaid 0.0.0.1         (Set the area ID for OSPF area 1)
>> OSPF Area (index) 1 # type stub             (Define area as stub type)
>> OSPF Area (index) 1 # enable                 (Enable the area)
```

5. Attach the network interface to the backbone.

```
>> OSPF Area (index) 1 # ../if 1               (Select OSPF menu for interface 1)
>> OSPF Interface 1 # aindex 0                 (Attach network to backbone index)
>> OSPF Interface 1 # enable                   (Enable the backbone interface)
```

6. Attach the network interface to the stub area.

```
>> OSPF Interface 1 # ../if 2                 (Select OSPF menu for interface 2)
>> OSPF Interface 2 # aindex 1                 (Attach network to stub area index)
>> OSPF Interface 2 # enable                   (Enable the stub area interface)
```

7. Configure route summarization by specifying the starting address and mask of the range of addresses to be summarized.

```
>> OSPF Interface 2 # ../range 1              (Select menu for summary range)
>> OSPF Summary Range 1 # addr 36.128.192.0    (Set base IP address of range)
>> OSPF Summary Range 1 # mask 255.255.192.0   (Set mask address for range)
>> OSPF Summary Range 1 # aindex 0            (Add summary route to backbone)
>> OSPF Summary Range 1 # enable              (Enable summary range)
```

8. Use the hide command to prevent a range of addresses from advertising to the backbone.

```
>> OSPF Interface 2 # ../range 2              (Select menu for summary range)
>> OSPF Summary Range 2 # addr 36.128.200.0    (Set base IP address)
>> OSPF Summary Range 2 # mask 255.255.255.0   (Set mask address)
>> OSPF Summary Range 2 # hide enable          (Hide the range of addresses)
```

9. Apply and save the configuration changes.

```
>> OSPF Summary Range 2 # apply                (Apply all configuration changes)
>> OSPF Summary Range 2 # save                 (Save applied changes)
```

Verifying OSPF Configuration

Use the following commands to verify the OSPF configuration on your switch:

- `/info/13/ospf/general`
- `/info/13/ospf/nbr`
- `/info/13/ospf/dbase/dbsum`
- `/info/13/ospf/route`
- `/stats/13/route`

Refer to the *BLADEOS Command Reference* for information on the above commands.

CHAPTER 14

Remote Monitoring

Remote Monitoring (RMON) allows network devices to exchange network monitoring data.

RMON performs the following major functions:

- Gathers cumulative statistics for Ethernet interfaces
- Tracks a history of statistics for Ethernet interfaces
- Creates and triggers alarms for user-defined events

RMON Overview

The RMON MIB provides an interface between the RMON agent on the switch and an RMON management application. The RMON MIB is described in RFC 1757.

The RMON standard defines objects that are suitable for the management of Ethernet networks. The RMON agent continuously collects statistics and proactively monitors switch performance. RMON allows you to monitor traffic flowing through the switch.

The switch supports the following RMON Groups, as described in RFC 1757:

- Group 1: Statistics
- Group 2: History
- Group 3: Alarms
- Group 9: Events

RMON Group 1—Statistics

The switch supports collection of Ethernet statistics as outlined in the RMON statistics MIB, in reference to `etherStatsTable`. You can enable RMON statistics on a per-port basis, and you can view them using the following command: `/stat/port <x>/rmon`. RMON statistics are sampled every second, and new data overwrites any old data on a given port.

Note – RMON port statistics must be enabled for the port before you can view RMON statistics.

To configure RMON Statistics:

1. Enable RMON on each port where you wish to collect RMON statistics.

```
>> # /cfg/port 23/rmon                (Select Port 23 RMON)
>> Port 23 RMON# ena                  (Enable RMON)
>> Port 23 RMON# apply                (Make your changes active)
>> Port 23 RMON# save                 (Save for restore after reboot)
```

2. View RMON statistics for the port.

```
>> # /stats/port 23                    (Select Port 23 Stats)
>> Port Statistics# rmon
-----
RMON statistics for port 23:
etherStatsDropEvents:                  NA
etherStatsOctets:                      7305626
etherStatsPkts:                        48686
etherStatsBroadcastPkts:               4380
etherStatsMulticastPkts:               6612
etherStatsCRCAlignErrors:              22
etherStatsUndersizePkts:                0
etherStatsOversizePkts:                0
etherStatsFragments:                   2
etherStatsJabbers:                     0
etherStatsCollisions:                  0
etherStatsPkts64Octets:                 27445
etherStatsPkts65to127Octets:           12253
etherStatsPkts128to255Octets:          1046
etherStatsPkts256to511Octets:           619
etherStatsPkts512to1023Octets:         7283
etherStatsPkts1024to1518Octets:         38
```

RMON Group 2—History

The RMON History Group allows you to sample and archive Ethernet statistics for a specific interface during a specific time interval.

Note – RMON port statistics must be enabled for the port before an RMON history group can monitor the port.

Data is stored in buckets, which store data gathered during discreet sampling intervals. At each configured interval, the history instance takes a sample of the current Ethernet statistics, and places them into a bucket. History data buckets reside in dynamic memory. When the switch is re-booted, the buckets are emptied.

Requested buckets (`/cfg/rmon/hist <x>/rbnum`) are the number of buckets, or data slots, requested by the user for each History Group. Granted buckets (`/info/rmon/hist <x>/gbnum`) are the number of buckets granted by the system, based on the amount of system memory available. The system grants a maximum of 50 buckets.

Use an SNMP browser to view History samples.

History MIB Objects

The type of data that can be sampled must be of an `ifIndex` object type, as described in RFC1213 and RFC1573. The most common data type for the history sample is as follows:

```
1.3.6.1.2.1.2.2.1.1.<x> -mgmt.interfaces.ifTable.ifIndex.interface
```

The last digit (*x*) represents the interface on which to monitor, which corresponds to the switch port number. History sampling is done per port, by utilizing the interface number to specify the port number.

Configuring RMON History

This example configuration creates an RMON History Group to monitor port 23. It takes a data sample every two minutes, and places the data into one of the 30 requested buckets. After 30 samples are gathered, the new samples overwrite the previous samples, beginning with the first bucket.

1. Enable RMON on each port where you wish to collect RMON History.

```
>> # /cfg/port 23/rmon                (Select Port 23 RMON)
>> Port 23# ena                      (Enable RMON)
>> Port 23 RMON# apply                (Make your changes active)
>> Port 23 RMON# save                 (Save for restore after reboot)
```

2. Configure the RMON History parameters.

```
>> # /cfg/rmon/hist 1                 (Select RMON History 1)
>> RMON History 1# ifoid 1.3.6.1.2.1.2.2.1.1.23
>> RMON History 1# rbnum 30
>> RMON History 1# intrval 120
>> RMON History 1# owner "Owner_History_1"
```

3. Apply and save the configuration.

```
>> RMON History 1# apply              (Make your changes active)
>> RMON History 1# save               (Save for restore after reboot)
```

Use SNMP to view the data.

RMON Group 3—Alarms

The RMON Alarm Group allows you to define a set of thresholds used to determine network performance. When a configured threshold is crossed, an alarm is generated. For example, you can configure the switch to issue an alarm if more than 1,000 CRC errors occur during a 10-minute time interval.

Each Alarm index consists of a variable to monitor, a sampling time interval, and parameters for rising and falling thresholds. The Alarm group can be used to track rising or falling values for a MIB object. The object must be a counter, gauge, integer, or time interval.

Use the `/cfg/rmon/alarm <x>/revtidx` command or the `/cfg/rmon/alarm <x>/fevtidx` command to correlate an alarm index to an event index. When the alarm threshold is reached, the corresponding event is triggered.

Alarm MIB Objects

The most common data types used for alarm monitoring are `ifStats`: errors, drops, bad CRCs, and so on. These MIB Object Identifiers (OIDs) correlate to the ones tracked by the History group. An example of an ICMP stat is as follows:

```
1.3.6.1.2.1.5.1.<x> - mgmt.icmp.icmpInMsgs
```

where *x* represents the interface on which to monitor, which corresponds to the switch interface number or port number, as follows:

- 1 through 128 = Switch interface number
- 129 = Switch port 1
- 130 = Switch port 2
- 131 = Switch port 3, and so on.

This value represents the alarm's MIB OID, as a string. Note that for non-tables, you must supply a `.0` to specify an end node.

Configuring RMON Alarms

Alarm Example 1

This example configuration creates an RMON alarm that checks `ifInOctets` on port 20 once every hour. If the statistic exceeds two billion, an alarm is generated that triggers event index 6.

1. Configure the RMON Alarm parameters to track the number of packets received on a port.

```
>> # /cfg/rmon/alarm 6 (Select RMON Alarm 6)
>> RMON Alarm 6# oid 1.3.6.1.2.1.2.2.1.10.276
>> RMON Alarm 6# intrval 3600
>> RMON Alarm 6# almtypе rising
>> RMON Alarm 6# rlimit 2000000000
>> RMON Alarm 6# revtidх 6
>> RMON Alarm 6# sample abs
>> RMON Alarm 6# owner "Alarm_for_ifInOctets"
```

2. Apply and save the configuration.

```
>> RMON Alarm 6# apply (Make your changes active)
>> RMON Alarm 6# save (Save for restore after reboot)
```

Alarm Example 2

This example configuration creates an RMON alarm that checks `icmpInEchos` on the switch once every minute. If the statistic exceeds 200 within a 60 second interval, an alarm is generated that triggers event index 5.

1. Configure the RMON Alarm parameters to track ICMP messages.

```
>> # /cfg/rmon/alarm 5 (Select RMON Alarm 5)
>> RMON Alarm 5# oid 1.3.6.1.2.1.5.8.0
>> RMON Alarm 5# intrval 60
>> RMON Alarm 5# almtypе rising
>> RMON Alarm 5# rlimit 200
>> RMON Alarm 5# revtidх 5
>> RMON Alarm 5# sample delta
>> RMON Alarm 5# owner "Alarm_for_icmpInEchos"
```

2. Apply and save the configuration.

```
>> RMON Alarm 5# apply (Make your changes active)
>> RMON Alarm 5# save (Save for restore after reboot)
```

RMON Group 9—Events

The RMON Event Group allows you to define events that are triggered by alarms. An event can be a log message, an SNMP trap message, or both.

When an alarm is generated, it triggers a corresponding event notification. Use the `/cfg/rmon/alarm <x>/revidx` and `/cfg/rmon/alarm <x>/fevidx` commands to correlate an event index to an alarm.

RMON events use SNMP and system logs to send notifications. Therefore, an SNMP trap host must be configured for trap event notification to work properly.

RMON uses a syslog host to send syslog messages. Therefore, an existing syslog host (`/cfg/sys/syslog`) must be configured for event log notification to work properly. Each log event generates a system log message of type RMON that corresponds to the event.

Configuring RMON Events

This example configuration creates an RMON event that sends a SYSLOG message each time it is triggered by an alarm.

1. Configure the RMON Event parameters.

```
>> # /cfg/rmon/event 5                               (Select RMON Event 5)
>> RMON Event 5# descn "SYSLOG_generation_event"
>> RMON Event 5# type log
>> RMON Event 5# owner "Owner_event_5"
```

2. Apply and save the configuration.

```
>> RMON Alarm 5# apply                               (Make your changes active)
>> RMON Alarm 5# save                               (Save for restore after reboot)
```


Part 3: High Availability Fundamentals

Internet traffic consists of myriad services and applications which use the Internet Protocol (IP) for data delivery. However, IP is not optimized for all the various applications. High Availability goes beyond IP and makes intelligent switching decisions to provide redundant network configurations.

CHAPTER 15

High Availability

BNT GbE Switch Modules (GbESMs) support high-availability network topologies through an enhanced implementation of the Virtual Router Redundancy Protocol (VRRP).

The following topics are discussed in this chapter:

- [“Layer 2 Failover” on page 226](#). This section discusses trunk failover without using VRRP.
- [“VRRP Overview” on page 234](#). This section discusses VRRP operation and BLADEOS redundancy configurations.
- [“Failover Methods” on page 237](#). This section describes the three modes of high availability.
- [“BLADEOS Extensions to VRRP” on page 240](#). This section describes VRRP enhancements implemented in BLADEOS.
- [“Virtual Router Deployment Considerations” on page 241](#). This section describes issues to consider when deploying virtual routers.
- [“High Availability Configurations” on page 242](#). This section discusses the more useful and easily deployed redundant configurations.
 - [“Active-Active Configuration” on page 242](#)
 - [“Hot-Standby Configuration” on page 247](#)

Layer 2 Failover

The primary application for Layer 2 Failover is to support Network Adapter Teaming. With Network Adapter Teaming, all the NICs on each server share the same IP address, and are configured into a team. One NIC is the primary link, and the other is a standby link. For more details, refer to the NetXen 10 Gb Ethernet Adapter documentation.

Note – Only two links per server blade can be used for Layer 2 Trunk Failover (one primary and one backup). Network Adapter Teaming allows only one backup NIC for each server blade.

Auto Monitoring Trunk Links

Layer 2 Failover can be enabled on any trunk group in the GbESM, including LACP trunks. Trunks can be added to failover trigger groups. Then, if some specified number of trigger links fail, the switch disables all the internal ports in the switch (unless VLAN Monitor is turned on). When the internal ports are disabled, it causes the NIC team on the affected server blades to failover from the primary to the backup NIC. This process is called a failover event.

When the appropriate number of links in a trigger group return to service, the switch enables the internal ports. This causes the NIC team on the affected server blades to fail back to the primary switch (unless Auto-Fallback is disabled on the NIC team). The backup switch processes traffic until the primary switch's internal links come up, which can take up to five seconds.

VLAN Monitor

The VLAN Monitor allows Layer 2 Failover to discern different VLANs. With VLAN Monitor turned on:

- If enough links in a trigger fail (see [“Setting the Failover Limit” on page 229](#)), the switch disables all internal ports that reside in the same VLAN membership as the trunk(s) in the trigger.
- When enough links in the trigger return to service, the switch enables the internal ports that reside in the same VLAN membership as the trunk(s) in the trigger.

If you turn off the VLAN Monitor (`/cfg/l2/failovr/vlan/off`), only one failover trigger is allowed. When a link failure occurs on the trigger, the switch disables all internal server-blade ports.

Auto Monitor Configurations

Figure 32 is a simple example of Layer 2 Failover. One GbESM is the primary, and the other is used as a backup. In this example, all external ports on the primary switch belong to a single trunk group, with Layer 2 Failover enabled, and Failover Limit set to 2. If two or fewer links in trigger 1 remain active, the switch temporarily disables all internal server-blade ports that reside in VLAN 1. This action causes a failover event on Server 1 and Server 2.

Figure 32 Basic Layer 2 Failover

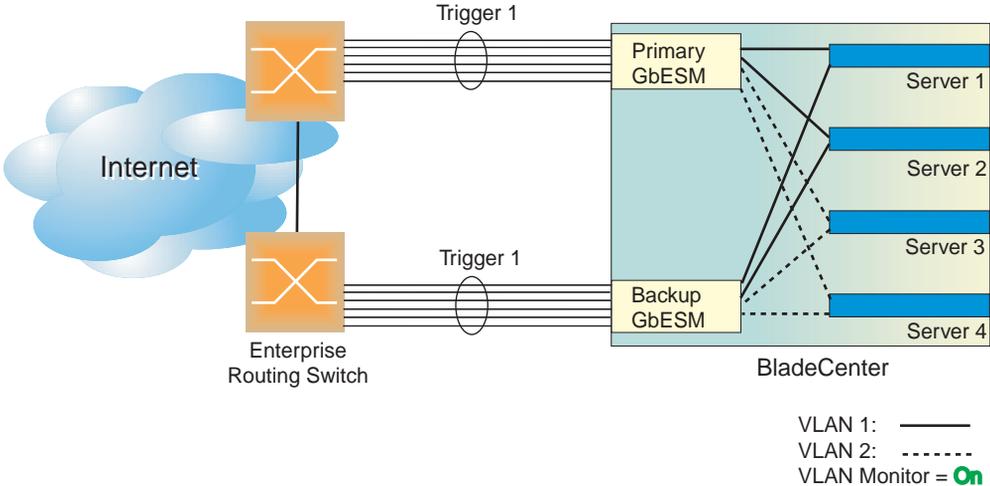


Figure 33 shows a configuration with two trunks, each in a different Failover Trigger. GbESM 1 is the primary switch for Server 1 and Server 2. GbESM 2 is the primary switch for Server 3 and Server 4. VLAN Monitor is turned on. STP is turned off.

If all links go down in trigger 1, GbESM 1 disables all internal ports that reside in VLAN 1. If all links in trigger 2 go down, GbESM 1 disables all internal ports that reside in VLAN 2.

Figure 33 Two trunks, each in a different Failover Trigger

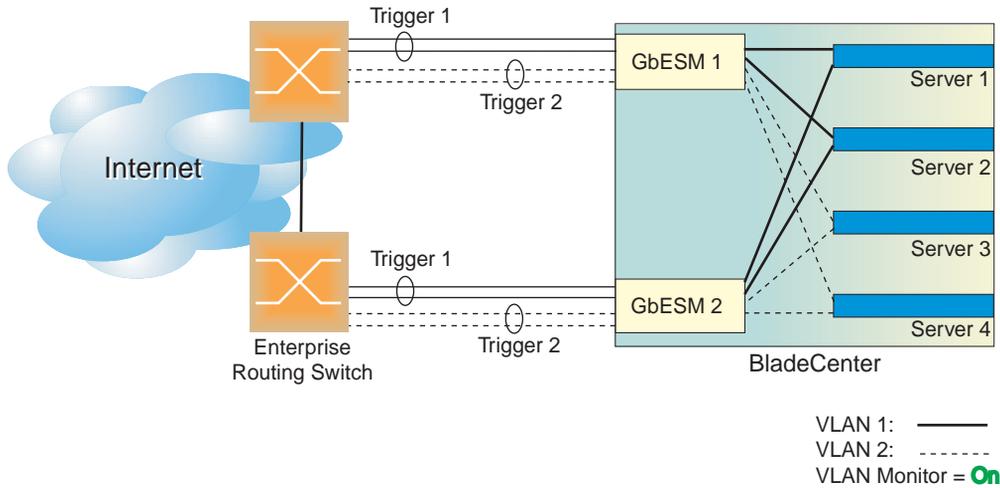
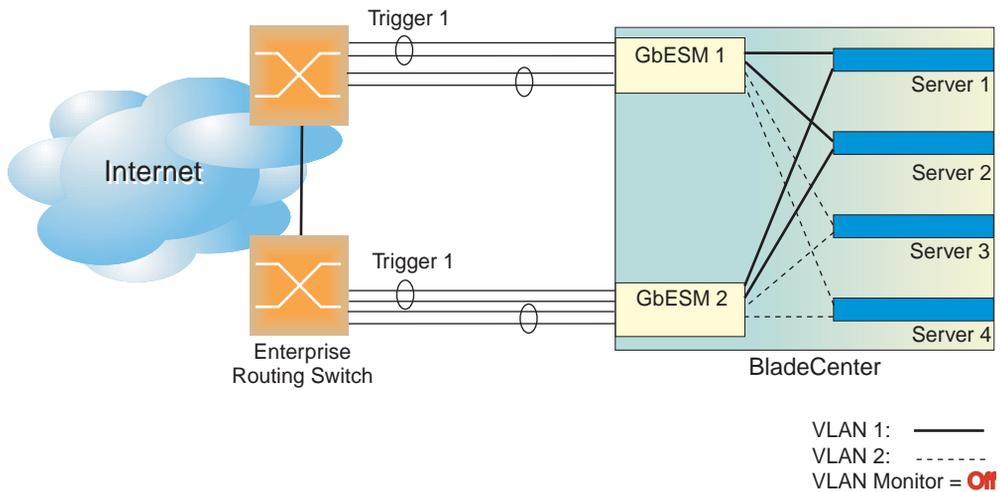


Figure 34 shows a configuration with two trunks. VLAN Monitor is turned off, so only one Failover Trigger is configured on each switch. GbESM 1 is the primary switch for Server 1 and Server 2. GbESM 2 is the primary switch for Server 3 and Server 4. STP is turned off.

If all links in trigger 1 go down, GbESM 1 disables all internal links to server blades.

Figure 34 Two trunks, one Failover Trigger



Setting the Failover Limit

The failover limit lets you specify the minimum number of operational links required within each trigger before the trigger initiates a failover event. For example, if the limit is two (`/cfg/l2/failover/trigger <x>/limit 2`), a failover event occurs when the number of operational links in the trigger is two or fewer. When you set the limit to zero, the switch triggers a failover event only when no links in the trigger are operational.

Manually Monitoring Port Links

The Manual Monitor allows you to configure a set of ports and/or trunks to monitor for link failures (a monitor list), and another set of ports and/or trunks to disable when the trigger limit is reached (a control list). When the switch detects a link failure on the monitor list, it automatically disables the items in control list. When server ports are disabled, the corresponding server's network adapter can detect the disabled link, and trigger a network-adapter failover to another port or trunk on the switch, or another switch in the chassis.

The switch automatically enables the control list items when the monitor list items return to service.

Monitor Port State

A monitor port is considered operation as long as the following conditions are true:

- The port must be in the `Link Up` state.
- If STP is enabled, the port must be in the `Forwarding` state.
- If the port is part of an LACP trunk, the port must be in the `Aggregated` state.

If any of the above conditions is false, the monitor port is considered to have failed.

Control Port State

A control port is considered Operational if the monitor trigger is up. As long as the trigger is up, the port is considered operational from a teaming perspective, even if the port itself is actually in the `Down` state, `Blocking` state (if STP is enabled on the port), or `Not Aggregated` state (if part of an LACP trunk).

A control port is considered to have failed on if the monitor trigger is in the `Down` state.

To view the state of any port, use one of the following commands:

<code>>> # /info/link</code>	<i>(View port link status)</i>
<code>>> # /info/l2/stp</code>	<i>(View port STP status)</i>
<code>>> # /info/l2/lacp/dump</code>	<i>(View port LACP status)</i>

L2 Failover with Other Features

L2 Failover works together with Link Aggregation Control Protocol (LACP) and with Spanning Tree Protocol (STP), as described below.

LACP

Link Aggregation Control Protocol allows the switch to form dynamic trunks. You can use the *admin key* to add up to two LACP trunks to a failover trigger using automatic monitoring. When you add an *admin key* to a trigger (`/cfg/l2/failovr/trigger <x>/amon/addkey`), any LACP trunk with that *admin key* becomes a member of the trigger.

Spanning Tree Protocol

If Spanning Tree Protocol (STP) is enabled on the ports in a failover trigger, the switch monitors the port STP state rather than the link state. A port failure results when STP is not in a Forwarding state (that is, Listening, Learning, Blocking, or No Link). The switch automatically disables the appropriate internal ports, based on the VLAN monitor.

When the switch determines that ports in the trigger are in STP Forwarding state, then it automatically enables the appropriate internal ports, based on the VLAN monitor. The switch *fails back* to normal operation.

Configuration Guidelines

This section provides important information about configuring Layer 2 Failover.

Note – Auto Monitor and Manual Monitor are mutually exclusive. They cannot both be configured on the switch.

Auto Monitor Guidelines

- Any specific failover trigger may monitor static trunks only or LACP trunks only, but not both.
- All external ports in all static or LACP trunks added to any specific failover trigger must belong to the same VLAN.
- A maximum of two LACP keys can be added per trigger.
- When VLAN Monitor is on, the following additional guidelines apply:
 - All external ports in all static or LACP trunks added to a specific failover trigger must belong to the same VLAN and have the same PVID.
 - Different triggers are not permitted to operate on the same VLAN.
 - Different triggers are not permitted to operate on the same internal port.
 - For each port in each trunk in a specific failover trigger, the trigger will monitor the STP state on only the default PVID.

Manual Monitor Guidelines

- A Manual Monitor can monitor only external ports.
- Any specific failover trigger can monitor external ports only, static trunks only, or LACP trunks only. The different types cannot be combined in the same trigger.
- A maximum of two LACP keys can be added per trigger.
- Port membership for different triggers should not overlap. Any specific port should be a member of only one trigger.

Configuring Layer 2 Failover

Auto Monitor Example

The following procedure pertains to the configuration shown in [Figure 32](#).

1. Configure Network Adapter Teaming on the servers.
2. Define a trunk group on the GbESM.

```
>> # /cfg/l2/trunk 1                (Select trunk group 1)
>> Trunk group 1# add EXT1          (Add port EXT1 to 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# ena                (Enable trunk group 1)
```

3. Configure Failover parameters.

```
>> # /cfg/l2/failovr/on             (Turn Failover on)
>> Failover# trigger 1              (Select trigger group 1)
>> Trigger 1# ena                   (Enable trigger group 1)
>> Trigger 1# limit 2               (Set Failover limit to 2 links)
>> Trigger 1# amon                  (Select Auto Monitor menu)
>> Auto Monitor# addtrnk 1          (Add trunk group 1)
```

4. Apply and verify the configuration.

```
>> Auto Monitor# apply              (Make your changes active)
>> Auto Monitor# cur                (View current trunking configuration)
```

5. Save the configuration.

```
>> Auto Monitor# save                (Save for restore after reboot)
```

Manual Monitor Example

Use the following procedure to configure a Layer 2 Failover Manual Monitor.

1. Configure Network Adapter Teaming on the servers.
2. Configure general Layer 2 Failover parameters.

```
>> # /cfg/l2/failovr/on           (Turn Failover on)
>> Failover# trigger 2          (Select trigger 2)
>> Trigger 2# ena              (Enable trigger 2)
>> Trigger 2# limit 2          (Set Failover limit to 2 links)
```

3. Specify the links to monitor.

```
>> Trigger 2# mmon/monitor      (Select Manual Monitor, Monitor menu)
>> Monitor# addport EXT4       (Add port EXT4)
>> Monitor# addport EXT5       (Add port EXT5)
>> Monitor# addport EXT6       (Add port EXT6)
>> Monitor# ..
```

4. Specify the links to disable when the failover limit is reached.

```
>> Manual Monitor# control      (Select Manual Monitor - Control menu)
>> Control# addport INT13       (Add port INT13)
>> Control# addport INT14       (Add port INT14)
```

5. Apply and verify the configuration.

```
>> Control# apply              (Make your changes active)
>> # /cfg/l2/failovr/cur       (View current Failover configuration)
```

6. Save the configuration.

```
>> Failover# save              (Save for restore after reboot)
```

VRRP Overview

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. This means that your network will remain in service despite the failure of any single device. To achieve this usually requires redundancy for all vital network components.

VRRP enables 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. One of the virtual routers 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, one of the backup virtual routers will take control of the virtual router IP address and actively process traffic addressed to it.

With VRRP, Virtual Interface Routers (VIR) allow two VRRP routers to share an IP interface across the routers. VIRs provide a single Destination IP (DIP) for upstream routers to reach various servers, and provide a virtual default Gateway for the server blades.

VRRP Components

Each physical router running VRRP is known as a *VRRP router*.

Virtual Router

Two or more VRRP routers can be configured to form a *virtual router* (RFC 2338). Each VRRP router may participate in one or more virtual routers. Each virtual router consists of a user-configured *virtual router identifier* (VRID) and an IP address.

Virtual Router MAC Address

The VRID is used to build the *virtual router MAC Address*. The five highest-order octets of the virtual router MAC Address are the standard MAC prefix (00-00-5E-00-01) defined in RFC 2338. The VRID is used to form the lowest-order octet.

Owners and Renters

Only one of the VRRP routers in a virtual router may be configured as the IP address owner. This router has the virtual router's IP address as its real interface address. This router responds to packets addressed to the virtual router's IP address for ICMP pings, TCP connections, and so on.

There is no requirement for any VRRP router to be the IP address owner. Most VRRP installations choose not to implement an IP address owner. For the purposes of this chapter, VRRP routers that are not the IP address owner are called *renters*.

Master and Backup Virtual Router

Within each virtual router, one VRRP router is selected to be the virtual router master. See [“Selecting the Master VRRP Router” on page 236](#) for an explanation of the selection process.

Note – If the IP address owner is available, it will always become the virtual router master.

The virtual router master forwards packets sent to the virtual router. It also responds to Address Resolution Protocol (ARP) requests sent to the virtual router's IP address. Finally, the virtual router master sends out periodic advertisements to let other VRRP routers know it is alive and its priority.

Within a virtual router, the VRRP routers not selected to be the master are known as virtual router backups. Should the virtual router master fail, one of the virtual router backups becomes the master and assumes its responsibilities.

Virtual Interface Router

At Layer 3, a Virtual Interface Router (VIR) allows two VRRP routers to share an IP interface across the routers. VIRs provide a single Destination IP (DIP) for upstream routers to reach various destination networks, and provide a virtual default Gateway.

Note – Every VIR must be assigned to an IP interface, and every IP interface must be assigned to a VLAN. If no port in a VLAN has link up, the IP interface of that VLAN is down, and if the IP interface of a VIR is down, that VIR goes into INIT state.

VRRP Operation

Only the virtual router master responds to ARP requests. Therefore, the upstream routers only forward packets destined to the master. The master also responds to ICMP ping requests. The backup does not forward any traffic, nor does it respond to ARP requests.

If the master is not available, the backup becomes the master and takes over responsibility for packet forwarding and responding to ARP requests.

Selecting the Master VRRP Router

Each VRRP router is configured with a priority between 1–254. A bidding process determines which VRRP router is or becomes the master—the VRRP router with the highest priority.

The master periodically sends advertisements to an IP multicast address. As long as the backups receive these advertisements, they remain in the backup state. If a backup does not receive an advertisement for three advertisement intervals, it initiates a bidding process to determine which VRRP router has the highest priority and takes over as master.

If, at any time, a backup determines that it has higher priority than the current master does, it can preempt the master and become the master itself, unless configured not to do so. In preemption, the backup assumes the role of master and begins to send its own advertisements. The current master sees that the backup has higher priority and will stop functioning as the master.

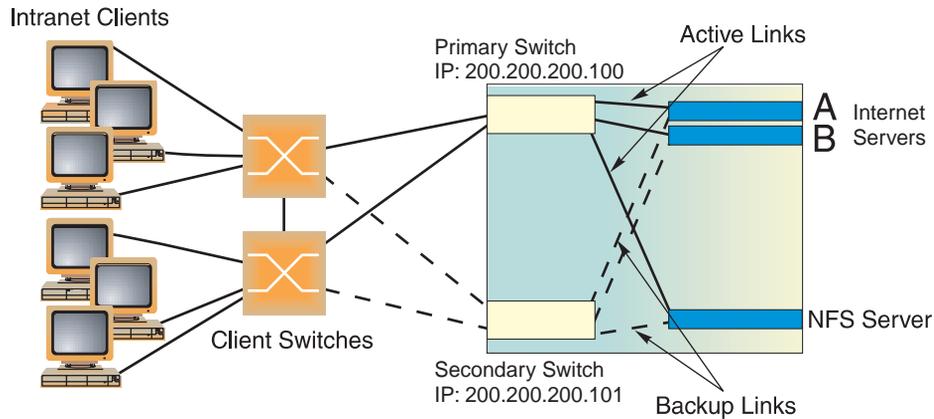
A backup router can stop receiving advertisements for one of two reasons—the master can be down, or all communications links between the master and the backup can be down. If the master has failed, it is clearly desirable for the backup (or one of the backups, if there is more than one) to become the master.

Note – If the master is healthy but communication between the master and the backup has failed, there will then be two masters within the virtual router. To prevent this from happening, configure redundant links to be used between the switches that form a virtual router.

Failover Methods

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. Traditionally, these configurations have been *hot-standby* configurations, where one switch is active and the other is in a standby mode. A non-VRRP hot-standby configuration is shown in the figure below:

Figure 35 A Non-VRRP, Hot-Standby Configuration



While hot-standby configurations increase site availability by removing single points-of-failure, service providers increasingly view them as an inefficient use of network resources because one functional application switch sits by idly until a failure calls it into action. Service providers now demand that vendors' 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.

BLADEOS high availability configurations are based on VRRP. The BLADEOS implementation of VRRP includes proprietary extensions.

The BLADEOS implementation of VRRP supports the following modes of high availability:

- **Active-Active**—based on proprietary BLADEOS extensions to VRRP
- **Hot-Standby**—supports Network Adapter Teaming on your server blades

Virtual Router Group

The virtual router group ties all virtual routers on the switch together as a single entity. By definition, hot-standby requires that all virtual routers failover as a group, and not individually. As members of a group, all virtual routers on the switch (and therefore the switch itself), are in either a master or standby state.

The virtual router group cannot be used for active-active configurations or any other configuration that require shared interfaces.

A VRRP group has the following characteristics:

- When enabled, all virtual routers behave as one entity, and all group settings override any individual virtual router settings.
- All individual virtual routers, once the VRRP group is enabled, assume the group's tracking and priority.
- When one member of a VRRP group fails, the priority of the group decreases, and the state of the entire switch changes from Master to Standby.

Each VRRP advertisement can include up to 128 addresses. All virtual routers are advertised within the same packet, conserving processing and buffering resources.

BLADEOS Extensions to VRRP

This section describes VRRP enhancements that are implemented in BLADEOS.

BLADEOS supports a tracking function that dynamically modifies the priority of a VRRP router, based on its current state. The objective of tracking is to have, whenever possible, the master bidding processes for various virtual routers in a LAN converge on the same switch. Tracking ensures that the selected switch is the one that offers optimal network performance. For tracking to have any effect on virtual router operation, preemption must be enabled.

BLADEOS can track the attributes listed in [Table 19](#):

Table 19 VRRP Tracking Parameters

Parameter	Description
Number of IP interfaces on the switch that are active (“up”) <code>/cfg/l3/vrrp/track/ifs</code>	Helps elect the virtual routers with the most available routes as the master. (An IP interface is considered active when there is at least one active port on the same VLAN.) This parameter influences the VRRP router's priority in virtual interface routers.
Number of active ports on the same VLAN <code>/cfg/l3/vrrp/track/ports</code>	Helps elect the virtual routers with the most available ports as the master. This parameter influences the VRRP router's priority in virtual interface routers. Note: In a hot-standby configuration, only external ports are tracked.
Number of virtual routers in master mode on the switch <code>/cfg/l3/vrrp/track/vrs</code>	Useful for ensuring that traffic for any particular client/server pair is handled by the same switch, increasing routing efficiency. This parameter influences the VRRP router's priority in virtual interface routers.

Each tracked parameter has a user-configurable weight associated with it. As the count associated with each tracked item increases (or decreases), so does the VRRP router's priority, subject to the weighting associated with each tracked item. If the priority level of a standby is greater than that of the current master, then the standby can assume the role of the master.

See [“Configuring the Switch for Tracking” on page 241](#) for an example on how to configure the switch for tracking VRRP priority.

Virtual Router Deployment Considerations

Assigning VRRP Virtual Router ID

During the software upgrade process, VRRP virtual router IDs will be automatically assigned if failover is enabled on the switch. When configuring virtual routers at any point after upgrade, virtual router ID numbers (`/cfg/13/vrrp/vr #/vr id`) must be assigned. The virtual router ID may be configured as any number between 1 and 128.

Configuring the Switch for Tracking

Tracking configuration largely depends on user preferences and network environment. Consider the configuration shown in [Figure 36 on page 238](#). Assume the following behavior on the network:

- Switch 1 is the master router upon initialization.
- If switch 1 is the master and it has one fewer active servers than switch 2, then switch 1 remains the master.
This behavior is preferred because running one server down is less disruptive than bringing a new master online and severing all active connections in the process.
- If switch 1 is the master and it has two or more active servers fewer than switch 2, then switch 2 becomes the master.
- If switch 2 is the master, it remains the master even if servers are restored on switch 1 such that it has one fewer or an equal number of servers.
- If switch 2 is the master and it has one active server fewer than switch 1, then switch 1 becomes the master.

The user can implement this behavior by configuring the switch for tracking as follows:

1. Set the priority for switch 1 to 101.
2. Leave the priority for switch 2 at the default value of 100.
3. On both switches, enable tracking based on ports (**ports**), interfaces (**ifs**), or virtual routers (**vr**). You can choose any combination of tracking parameters, based on your network configuration.

Note – There is no shortcut to setting tracking parameters. The goals must first be set and the outcomes of various configurations and scenarios analyzed to find settings that meet the goals.

High Availability Configurations

GbESMs offer flexibility in implementing redundant configurations. This section discusses the more useful and easily deployed configurations:

- “Active-Active Configuration” on page 242
- “Hot-Standby Configuration” on page 247

Active-Active Configuration

Figure 38 shows an example configuration where two GbESMs are used as VRRP routers in an active-active configuration. In this configuration, both switches respond to packets.

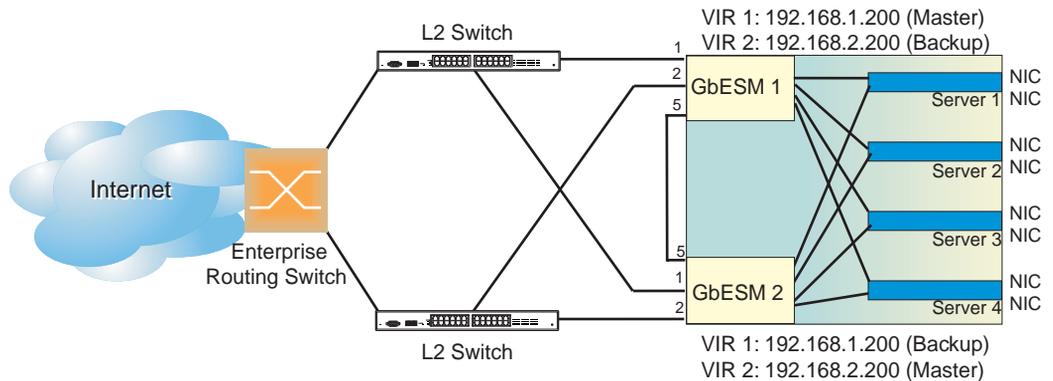


Figure 38 Active-Active High-Availability Configuration

Although this example shows only two switches, there is no limit on the number of switches used in a redundant configuration. It is possible to implement an active-active configuration across all the VRRP-capable switches in a LAN.

Each VRRP-capable switch in an active-active configuration is autonomous. Switches in a virtual router need not be identically configured.

In the scenario illustrated in Figure 38, traffic destined for IP address 10.0.1.1 is forwarded through the Layer 2 switch at the top of the drawing, and ingresses GbESM 1 on port EXT1. Return traffic uses default gateway 1 (192.168.1.1). If the link between GbESM 1 and the Layer 2 switch fails, GbESM 2 becomes the Master because it has a higher priority. Traffic is forwarded to GbESM 2, which forwards it to GbESM 1 through port EXT5. Return traffic uses default gateway 2 (192.168.2.1), and is forwarded through the Layer 2 switch at the bottom of the drawing.

To implement the active-active example, perform the following switch configuration.

Task 1: Configure GbESM 1

1. Configure client and server interfaces.

```

/cfg/l3/if 1                               (Select interface 1)
>> IP Interface 1# addr 192.168.1.100    (Define IP address for interface 1)
>> IP Interface 1# vlan 10                (Assign VLAN 10 to interface 1)
>> IP Interface 1# ena                    (Enable interface 1)
>> IP Interface 1# ..
>> Layer 3# if 2                           (Select interface 2)
>> IP Interface 2# addr 192.168.2.101    (Define IP address for interface 2)
>> IP Interface 2# vlan 20                (Assign VLAN 20 to interface 2)
>> IP Interface 2# ena                    (Enable interface 2)
>> IP Interface 2# ..
>> Layer 3# if 3                           (Select interface 3)
>> IP Interface 3# addr 10.0.1.100        (Define IP address for interface 3)
>> IP Interface 3# mask 255.255.255.0    (Define subnet mask for interface 3)
>> IP Interface 3# ena                    (Enable interface 3)
>> IP Interface 3# ..
>> Layer 3# if 4                           (Select interface 4)
>> IP Interface 4# addr 10.0.2.101        (Define IP address for interface 4)
>> IP Interface 4# mask 255.255.255.0    (Define subnet mask for interface 4)
>> IP Interface 4# ena                    (Enable interface 4)

```

2. Configure the default gateways. Each default gateway points to a Layer 3 router.

```

/cfg/l3/gw 1                               (Select default gateway 1)
>> Default gateway 1# addr 192.168.1.1  (Point gateway to the first L3 router)
>> Default gateway 1# ena                (Enable the default gateway)
>> Default gateway 1# ..
>> Layer 3# gw 2                           (Select default gateway 2)
>> Default gateway 2# addr 192.168.2.1  (Point gateway to the second router)
>> Default gateway 2# ena                (Enable the default gateway)

```

3. Turn on VRRP and configure two Virtual Interface Routers.

```

/cfg/l3/vrrp/on                            (Turn VRRP on)
>> Virtual Router Redundancy Protocol# vr 1 (Select virtual router 1)
>> VRRP Virtual Router 1# vrid 1          (Set VRID to 1)
>> VRRP Virtual Router 1# if 1            (Set interface 1)
>> VRRP Virtual Router 1# addr 192.168.1.200 (Define IP address)
>> VRRP Virtual Router 1# ena              (Enable virtual router 1)
>> VRRP Virtual Router 1# ..                (Enable virtual router 1)
>> Virtual Router Redundancy Protocol# vr 2 (Select virtual router 2)
>> VRRP Virtual Router 2# vrid 2          (Set VRID to 2)
>> VRRP Virtual Router 2# if 2            (Set interface 2)
>> VRRP Virtual Router 2# addr 192.168.2.200 (Define IP address)
>> VRRP Virtual Router 2# ena              (Enable virtual router 2)

```

4. Enable tracking on ports. Set the priority of Virtual Router 1 to 101, so that it becomes the Master.

```

/cfg/l3/vrrp/vr 1 (Select VRRP virtual router 1)
>> VRRP Virtual Router 1# track/ports/ena (Set tracking on ports)
>> VRRP Virtual Router 1 Priority Tracking# ..
>> VRRP Virtual Router 1# prio 101 (Set the VRRP priority)
>> VRRP Virtual Router 1# ..
>> Virtual Router Redundancy Protocol# vr 2(Select VRRP virtual router 2)
>> VRRP Virtual Router 1# track/ports/ena (Set tracking on ports)

```

5. Configure ports.

```

/cfg/l2/vlan 10 (Select VLAN 10)
>> VLAN 10# ena (Enable VLAN 10)
>> VLAN 10# add ext1 (Add port EXT 1 to VLAN 10)
>> VLAN 10# ..
>> Layer 2# vlan 20 (Select VLAN 20)
>> VLAN 20# ena (Enable VLAN 20)
>> VLAN 20# add ext2 (Add port EXT 2 to VLAN 20)

```

6. Turn off Spanning Tree Protocol globally.

```

/cfg/l2/stg 1/off (Turn off STG)
>> Spanning Tree Group 1# apply
>> Spanning Tree Group 1# save

```

Task 2: Configure GbESM 2

1. Configure client and server interfaces.

```

/cfg/l3/if 1                               (Select interface 1)
>> IP Interface 1# addr 192.168.1.101     (Define IP address for interface 1)
>> IP Interface 1# vlan 10                 (Assign VLAN 10 to interface 1)
>> IP Interface 1# ena                     (Enable interface 1)
>> IP Interface 1# ..
>> Layer 3# if 2                           (Select interface 2)
>> IP Interface 2# addr 192.168.2.100     (Define IP address for interface 2)
>> IP Interface 2# vlan 20                 (Assign VLAN 20 to interface 2)
>> IP Interface 2# ena                     (Enable interface 2)
>> IP Interface 2# ..
>> Layer 3# if 3                           (Select interface 3)
>> IP Interface 3# addr 10.0.1.101        (Define IP address for interface 3)
>> IP Interface 3# mask 255.255.255.0     (Define subnet mask for interface 3)
>> IP Interface 3# ena                     (Enable interface 3)
>> IP Interface 3# ..
>> Layer 3# if 4                           (Select interface 4)
>> IP Interface 4# addr 10.0.2.100        (Define IP address for interface 4)
>> IP Interface 4# mask 255.255.255.0     (Define subnet mask for interface 4)
>> IP Interface 4# ena                     (Enable interface 4)

```

2. Configure the default gateways. Each default gateway points to a Layer 3 router.

```

/cfg/l3/gw 1                               (Select default gateway 1)
>> Default gateway 1# addr 192.168.2.1   (Point gateway to the first L3 router)
>> Default gateway 1# ena                 (Enable the default gateway)
>> Default gateway 1# ..
>> Layer 3# gw 2                           (Select default gateway 2)
>> Default gateway 2# addr 192.168.1.1   (Point gateway to the second router)
>> Default gateway 2# ena                 (Enable the default gateway)

```

3. Turn on VRRP and configure two Virtual Interface Routers.

```

/cfg/l3/vrrp/on                            (Turn VRRP on)
>> Virtual Router Redundancy Protocol# vr 1 (Select virtual router 1)
>> VRRP Virtual Router 1# vrid 1          (Set VRID to 1)
>> VRRP Virtual Router 1# if 1           (Set interface 1)
>> VRRP Virtual Router 1# addr 192.168.1.200 (Define IP address)
>> VRRP Virtual Router 1# ena             (Enable virtual router 1)
>> VRRP Virtual Router 1# ..             (Enable virtual router 1)
>> Virtual Router Redundancy Protocol# vr 2 (Select virtual router 2)
>> VRRP Virtual Router 2# vrid 2          (Set VRID to 2)
>> VRRP Virtual Router 2# if 2           (Set interface 2)
>> VRRP Virtual Router 2# addr 192.168.2.200 (Define IP address)
>> VRRP Virtual Router 2# ena             (Enable virtual router 2)

```

4. Enable tracking on ports. Set the priority of Virtual Router 2 to 101, so that it becomes the Master.

```

/cfg/l3/vrrp/vr 1 (Select VRRP virtual router 1)
>> VRRP Virtual Router 1# track/ports/ena (Set tracking on ports)
>> VRRP Virtual Router 1 Priority Tracking# ..
>> VRRP Virtual Router 1# ..
>> Virtual Router Redundancy Protocol# vr 2(Select VRRP virtual router 2)
>> VRRP Virtual Router 2# track/ports/ena (Set tracking on ports)
>> VRRP Virtual Router 2 Priority Tracking# ..
>> VRRP Virtual Router 2# prio 101 (Set the VRRP priority)

```

5. Configure ports.

```

/cfg/l2/vlan 10 (Select VLAN 10)
>> VLAN 10# ena (Enable VLAN 10)
>> VLAN 10# add ext1 (Add port EXT 1 to VLAN 10)
>> VLAN 10# ..
>> Layer 2# vlan 20 (Select VLAN 20)
>> VLAN 20# ena (Enable VLAN 20)
>> VLAN 20# add ext2 (Add port EXT 2 to VLAN 20)

```

6. Turn off Spanning Tree Protocol globally. Apply and save changes.

```

/cfg/l2/stg 1/off (Turn off STG)
>> Spanning Tree Group 1# apply
>> Spanning Tree Group 1# save

```

Hot-Standby Configuration

The primary application for VRRP-based hot-standby is to support Network Adapter Teaming on your server blades. With Network Adapter Teaming, the NICs on each server share the same IP address, and are configured into a team. One NIC is the primary link, and the others are backup links. For more details, refer to the NetXen 10 Gb Ethernet Adapter documentation.

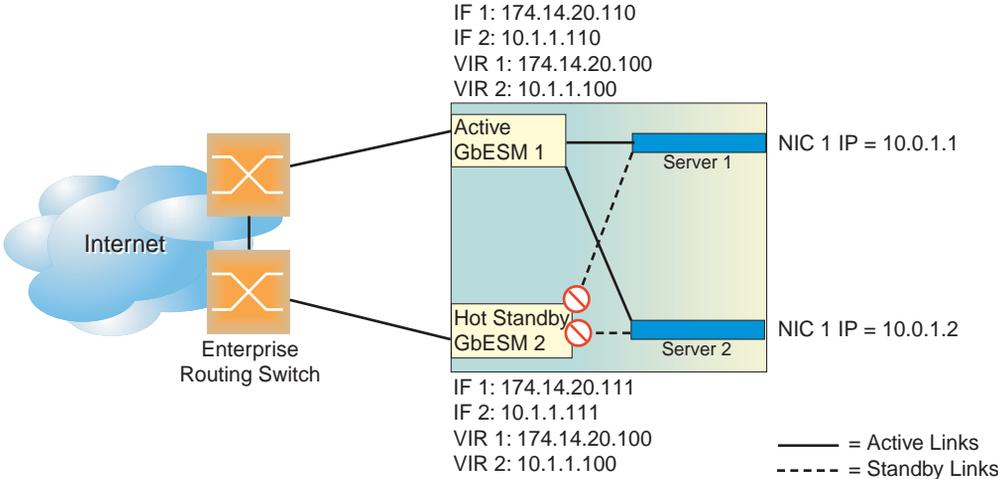
A hot-standby configuration allows all processes to failover to a standby switch if any type of failure should occur. All Virtual Interface Routers (VIRs) are bundled into one Virtual Router group, and then they failover together. When there is a failure that causes the VRRP Master to failover to the Standby, then the original primary switch temporarily disables the internal server links, which, in turn, causes the NIC teams to failover as well.

Note – When using hot-standby redundancy, peer switches should have an equal number of connected ports.

If hot-standby is implemented in a looped environment, the hot-standby feature automatically disables the hot-standby ports on the VRRP Standby. If the Master switch should failover to the Standby switch, it would change the hot-standby ports from *disabled* to *forwarding*, without relying on Spanning Tree or manual intervention. Therefore, Spanning Tree must be disabled.

Figure 39 illustrates a common hot-standby implementation on a single blade server. Notice that the BladeCenter server NICs are configured into a team that shares the same IP address across both NICs. Because only one link can be active at a time, the hot-standby feature controls the NIC failover by having the Standby switch disable its internal ports (holding down the server links).

Figure 39 Hot-Standby Configuration



Task 1: Configure GbESM 1

1. On GbESM 1, configure the interfaces for clients (174.14.20.110) and servers (10.1.1.110).

```

/cfg/13/if 1
>> IP Interface 1# addr 174.14.20.110      (Define IP address for interface 1)
>> IP Interface 1# ena                    (Enable interface 1)
>> IP Interface 1# ..
>> Layer 3# if 2
>> IP Interface 2# addr 10.1.1.110        (Define IP address for interface 2)
>> IP Interface 2# ena                    (Enable interface 2)

```

2. Configure Virtual Interface Routers.

```

/cfg/13/vrrp/on                                (Turn on VRRP)
>> Virtual Router Redundancy Protocol# vr 1(Select Virtual Router 1)
>> VRRP Virtual Router 1# ena              (Enable VR 1)
>> VRRP Virtual Router 1# vrid 1          (Select the Virtual Router ID)
>> VRRP Virtual Router 1# if 1            (Select interface for VR 1)
>> VRRP Virtual Router 1# addr 174.14.20.100(Define IP address for VR 1)
>> VRRP Virtual Router 1# ..
>> Virtual Router Redundancy Protocol# vr 2(Select Virtual Router 2)
>> VRRP Virtual Router 2# ena              (Enable VR 2)
>> VRRP Virtual Router 2# vrid 2          (Select the Virtual Router ID)
>> VRRP Virtual Router 2# if 2            (Select interface for VR 2)
>> VRRP Virtual Router 2# addr 10.1.1.100 (Define IP address for VR 2)

```

3. Enable VRRP Hot Standby.

```

/cfg/13/vrrp/hotstan ena                    (Enable Hot Standby)

```

4. Configure VRRP Group parameters. Set the VRRP priority to 101, so that this switch is the Master.

```

/cfg/13/vrrp/group
>> VRRP Virtual Router Group# ena          (Enable Virtual Router Group)
>> VRRP Virtual Router Group# vrid 1      (Set Virtual Router ID for Group)
>> VRRP Virtual Router Group# if 1        (Set interface for Group)
>> VRRP Virtual Router Group# prio 101    (Set VRRP priority to 101)
>> VRRP Virtual Router Group# track/ports ena(Enable tracking on ports)

```

5. Turn off Spanning Tree Protocol globally. Apply and save changes.

```

/cfg/12/stg 1/off                              (Turn off Spanning Tree)
>> Spanning Tree Group 1# apply           (Apply changes)
>> Spanning Tree Group 1# save

```

Task 2: Configure GbESM 2

1. On GbESM 2, configure the interfaces for clients (174.14.20.111) and servers (10.1.1.111).

```

/cfg/13/if 1
>> IP Interface 1# addr 174.14.20.111      (Define IP address for interface 1)
>> IP Interface 1# ena                    (Enable interface 1)
>> IP Interface 1# ..
>> Layer 3# if 2
>> IP Interface 2# addr 10.1.1.111        (Define IP address for interface 2)
>> IP Interface 2# ena                    (Enable interface 2)

```

2. Configure Virtual Interface Routers.

```

/cfg/13/vrrp/on                                (Turn on VRRP)
>> Virtual Router Redundancy Protocol# vr 1(Select Virtual Router 1)
>> VRRP Virtual Router 1# ena              (Enable VR 1)
>> VRRP Virtual Router 1# vrid 1          (Select the Virtual Router ID)
>> VRRP Virtual Router 1# if 1            (Select interface for VR 1)
>> VRRP Virtual Router 1# addr 174.14.20.100(Define IP address for VR 1)
>> VRRP Virtual Router 1# ..
>> Virtual Router Redundancy Protocol# vr 2(Select Virtual Router 2)
>> VRRP Virtual Router 2# ena              (Enable VR 2)
>> VRRP Virtual Router 2# vrid 2          (Select the Virtual Router ID)
>> VRRP Virtual Router 2# if 2            (Select interface for VR 2)
>> VRRP Virtual Router 2# addr 10.1.1.100 (Define IP address for VR 2)

```

3. Enable VRRP Hot Standby.

```

/cfg/13/vrrp/hotstan ena                    (Enable Hot Standby)

```

4. Configure VRRP Group parameters. Use the default VRRP priority of 100, so that this switch is the Standby.

```

/cfg/13/vrrp/group
>> VRRP Virtual Router Group# ena          (Enable Virtual Router Group)
>> VRRP Virtual Router Group# vrid 1      (Set Virtual Router ID for Group)
>> VRRP Virtual Router Group# if 1        (Set interface for Group)
>> VRRP Virtual Router Group# track/ports ena(Enable tracking on ports)

```

5. Turn off Spanning Tree Protocol globally. Apply and save changes.

```

/cfg/12/stg 1/off                              (Turn off Spanning Tree)
>> Spanning Tree Group 1# apply           (Apply changes)
>> Spanning Tree Group 1# save

```


Part 4: Appendices

This section describes the following topics:

- Port Monitoring for Troubleshooting
- RADIUS Server Configuration Notes
- Glossary

APPENDIX A

Monitoring Ports for Troubleshooting

BLADEOS provides two main methods for monitoring GbESM network traffic:

- sFlow—Statistical flow monitoring for high-speed data ports.
- Port Mirroring—Complete port data capture.

sFlow

The GbESM supports sFlow technology for monitoring traffic in data networks. The switch includes an embedded sFlow agent which can be configured to provide continuous monitoring information to a central sFlow analyzer.

Note – The switch is responsible only for forwarding sFlow information. A separate sFlow analyzer is required elsewhere on the network in order to interpret sFlow data.

sFlow Statistical Counters

The GbESM can be configured to send network statistics to an sFlow analyzer at regular intervals. For each port, a polling interval of 5 to 60 seconds can be configured, or 0 (the default) to disable this feature.

When polling is enabled, at the end of each configured polling interval, the GbESM reports general port statistics (as found in the output of the `/stats/port <x>/if` command) and port Ethernet statistics (as found in the output of the `/stats/port <x>/ether` command).

sFlow Example Configuration

1. Specify the location of the sFlow analyzer (the server and optional port to which the sFlow information will be sent):

```
>> # /cfg/sys/sflow/saddress <IP address>      (sFlow server)
>> sFlow# sport <service port>                 (Set the optional service port)
>> sFlow# ena                                   (Enable sFlow features)
```

By default, the switch uses established sFlow service port 6343.

To disable sFlow features across all ports, use the `/cfg/sys/sflow/dis` command.

2. On a per-port basis, define the statistics polling rate:

```
>> sFlow# port <port number or alias>          (Select the port)
>> sFlow Port# polling <polling rate>         (Statistics polling rate)
```

Specify a polling rate between 5 and 60 seconds, or 0 to disable. By default, polling is 0 (disabled) for each port.

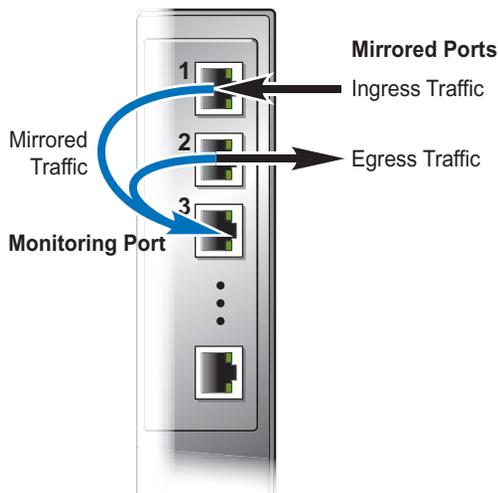
3. Apply and save the configuration.

Port Mirroring

The port mirroring feature in the BLADEOS allows you to attach a sniffer to a monitoring port that is configured to receive a copy of all packets that are forwarded from the mirrored port. BLADEOS enables you to mirror port traffic for all layer 2 and layer 3. Port mirroring can be used as a troubleshooting tool or to enhance the security of your network. For example, an IDS server can be connected to the monitor port to detect intruders attacking the network.

As shown in [Figure 40](#), port EXT3 is monitoring *ingress* traffic (traffic entering the switch) on port EXT1 and *egress* traffic (traffic leaving the switch) on port EXT2. You can attach a device to port EXT3 to monitor the traffic on ports EXT1 and EXT2.

Figure 40 Mirroring Ports



[Figure 40](#) shows two mirrored ports being monitored by a single port. The following configurations are also supported:

- One mirrored port to one monitored port
- More than two mirrored ports to one monitored port

BLADEOS does not support mirroring traffic from any specific port to multiple monitoring ports. For instance, EXT1 traffic cannot be monitored by both EXT3 and EXT4 simultaneously, nor can EXT1 ingress traffic be monitored by a different port than its egress traffic.

Ingress and egress traffic is duplicated and sent to the monitor port after processing.

Note – The GbE Switch Module (GbESM) cannot mirror LACPDU packets. Also, traffic on management VLANs is not mirrored to the external ports.

Port Mirroring Behavior

This section describes the composition of monitored packets in the GbESM, based on the configuration of the ports.

- Packets mirrored at port egress are mirrored prior to VLAN tag processing and may have a different PVID than packets that egress the port toward their actual network destination.
- Packets mirrored at port ingress are not modified.

Configuring Port Mirroring

To configure port mirroring for the example shown in [Figure 40](#),

1. Specify the monitoring port.

```
>> # /cfg/pmirr/monport EXT3           (Select port EXT3 for monitoring)
```

2. Select the ports that you want to mirror.

```
>> Port EXT3 # add EXT1                 (Select port EXT1 to mirror)
>> Enter port mirror direction [in, out, or both]: in
                                           (Monitor ingress traffic on port EXT1)
>> Port EXT3 # add EXT2                 (Select port EXT2 to mirror)
>> Enter port mirror direction [in, out, or both]: out
                                           (Monitor egress traffic on port EXT2)
```

3. Enable port mirroring.

```
>> # /cfg/pmirr/mirr ena               (Enable port mirroring)
```

4. Apply and save the configuration.

```
>> PortMirroring# apply                 (Apply the configuration)
>> PortMirroring# save                  (Save the configuration)
```

5. View the current configuration.

```
>> PortMirroring# cur (Display the current settings)  
Port mirroring is enabled  
Monitoring Ports      Mirrored Ports  
INT1                  none  
INT2                  none  
INT3                  none  
INT4                  none  
...  
EXT1                  none  
EXT2                  none  
EXT3                  (EXT1, in) (EXT2, out)  
EXT4                  none  
...
```


APPENDIX B

RADIUS Server Configuration Notes

Use the following information to modify your RADIUS configuration files for the Nortel Networks BaySecure Access Control RADIUS server, to provide authentication for users of the GbE Switch Module (GbESM).

1. Create a dictionary file called `blade.dct`, with the following content:

```
#####
# blade.dct - RADLINX BLADE dictionary
#
# (See README.DCT for more details on the format of this file)
#####
#
# Use the Radius specification attributes in lieu of the
# RADLINX BLADE ones
#
@radius.dct

#
# Define additional RADLINX BLADE parameters
# (add RADLINX BLADE specific attributes below)

ATTRIBUTE   Radlinx-Vendor-Specific  26  [vid=648 data=string]  R

#####
# blade.dct - RADLINX BLADE dictionary
#####

#Define GbE Switch Module dictionary
#@radius.dct

@blade.dct
    VALUE           Service-Type      user      255
    VALUE           Service-Type      oper      252
#####
```

2. Open the `dictiona.dcm` file, and add the following line (as in the example): `@blade.dct`

```
#####
# dictiona.dcm
#####
# Generic Radius

@radius.dct

#
# Specific Implementations (vendor specific)
#
@pprtl213.dct
@acc.dct
@accessbd.dct
@blade.dct
.
.
.
#####
# dictiona.dcm
#####
```

3. Open the vendor file (`vendor.ini`), and add the following data to the Vendor-Product identification list:

```
vendor-product    = BLADE Blade-server module
dictionary        = blade
ignore-ports     = no
help-id          = 0
```

Glossary

DIP	The destination IP address of a frame.
Dport	The destination port (application socket: for example, http-80/https-443/DNS-53)
NAT	Network Address Translation Any time an IP address is changed from one source IP or destination IP address to another address, network address translation can be said to have taken place. In general, half NAT is when the destination IP or source IP address is changed from one address to another. Full NAT is when both addresses are changed from one address to another. No NAT is when neither source nor destination IP addresses are translated.
Preemption	In VRRP, preemption will cause a Virtual Router that has a lower priority to go into backup should a peer Virtual Router start advertising with a higher priority.
Priority	In VRRP, the value given to a Virtual Router to determine its ranking with its peer(s). Minimum value is 1 and maximum value is 254. Default is 100. A higher number will win out for master designation.
Proto (Protocol)	The protocol of a frame. Can be any value represented by a 8-bit value in the IP header adherent to the IP specification (for example, TCP, UDP, OSPF, ICMP, and so on.)
SIP	The source IP address of a frame.
SPort	The source port (application socket: for example, HTTP-80/HTTPS-443/DNS-53).
Tracking	<p>In VRRP, a method to increase the priority of a virtual router and thus master designation (with preemption enabled). Tracking can be very valuable in an active/active configuration.</p> <p>You can track the following:</p> <ul style="list-style-type: none">■ Active IP interfaces on the Web switch (increments priority by 2 for each)■ Active ports on the same VLAN (increments priority by 2 for each)■ Number of virtual routers in master mode on the switch
VIR	Virtual Interface Router A VRRP address for that is an IP interface address shared between two or more virtual routers.

Virtual Router A shared address between two devices utilizing VRRP, as defined in RFC 2338. One virtual router is associated with an IP interface. This is one of the IP interfaces that the switch is assigned. All IP interfaces on the GbESMs must be in a VLAN. If there is more than one VLAN defined on the Web switch, then the VRRP broadcasts will only be sent out on the VLAN of which the associated IP interface is a member.

VRID **Virtual Router Identifier**

In VRRP, a numeric ID is used by each virtual router to create its MAC address and identify its peer for which it is sharing this VRRP address. The VRRP MAC address as defined in the RFC is 00-00-5E-00-01-*<VRID>*. For a virtual router with a VRID greater than 255, the following block of MAC addresses is allocated: 00:0F:6A:9A:40:00 through 00:0F:6A:9A:47:FF.

If you have a VRRP address that two switches are sharing, then the VRID number needs to be identical on both switches so each virtual router on each switch knows with whom to share.

VRRP **Virtual Router Redundancy Protocol**

A protocol that acts very similarly to Cisco's proprietary HSRP address sharing protocol. The reason for both of these protocols is so devices have a next hop or default gateway that is always available. Two or more devices sharing an IP interface are either advertising or listening for advertisements. These advertisements are sent via a broadcast message to an address such as 224.0.0.18.

With VRRP, one switch is considered the master and the other the backup. The master is always advertising via the broadcasts. The backup switch is always listening for the broadcasts. Should the master stop advertising, the backup will take over ownership of the VRRP IP and MAC addresses as defined by the specification. The switch announces this change in ownership to the devices around it by way of a Gratuitous ARP, and advertisements. If the backup switch didn't do the Gratuitous ARP the Layer 2 devices attached to the switch would not know that the MAC address had moved in the network. For a more detailed description, refer to RFC 2338.

CNA **Converged Network Adapter**

A device used for I/O consolidation such as that in Converged Enhanced Ethernet (CEE) environments implementing Fibre Channel over Ethernet (FCoE). The CNA performs the duties of both a Network Interface Card (NIC) for Local Area Networks (LANs) and a Host Bus Adapter (HBA) for Storage Area Networks (SANs).

HBA **Host Bus Adapter**

An adapter or card that interfaces with device drivers in the host operating system and the storage target in a Storage Area Network (SAN). It is equivalent to a Network Interface Controller (NIC) from a Local Area Network (LAN).

Index

Symbols

[]..... 20

Numerics

802.1Q VLAN tagging..... 74

A

accessing the switch

 Browser-based Interface 31

 LDAP 51

 RADIUS authentication..... 42

 security..... 42

active-active redundancy 238

administrator account..... 45

aggregating routes 183

 example..... 188

anycast address, IPv6..... 273

application ports 134

authenticating, in OSPF 202

autoconfiguration, IPv6..... 273

autonomous systems (AS) 195

B

BBI. *See* Browser-Based Interface

Border Gateway Protocol (BGP)..... 177

 attributes..... 184

 failover configuration 186

 route aggregation 183

 route maps 179

 selecting route paths 185

Bridge Protocol Data Unit (BPDU) 104

broadcast domains 71, 156

Browser-Based Interface 196

C

Cisco EtherChannel..... 95

CIST 119

Class of Service queue 147

command conventions 20

Command Line Interface 196

Community VLAN..... 88

configuration rules

 port mirroring 95

 spanning tree..... 95

 Trunking 95

 VLANs 95

configuring

 BGP failover 186

 IP routing 154

 OSPF 205

 port trunking 97

 spanning tree groups 110

D

default gateway..... 153

 configuration example..... 155

default password 45

default route, OSPF 200

Differentiated Services Code Point (DSCP) 141

E

EAPoL..... 64

End user access control, configuring 59

EtherChannel 92

 as used with port trunking 95

Extensible Authentication Protocol over LAN 64

external routing..... 177, 195

F

failover.....	226
overview.....	237
fault tolerance with port trunking	93
frame size	72
frame tagging. <i>See</i> VLANs tagging.	

G

gateway. *See* default gateway.

H

high-availability	225
host routes, OSPF.....	204
Hot Links	113
hot-standby redundancy	238
HP-OpenView.....	35

I

IBM Director	35
ICMP	133
IEEE standards	
802.1D	103
802.1s	115, 119
802.1x	64
IGMP	133, 167
IGMP Relay.....	172
IGMP Snooping	168
IGMPv3	169
incoming route maps.....	180
internal routing.....	177, 195
Internet Group Management Protocol (IGMP).....	167
Inter-Switch Link, ISL	94
IP address routing example.....	154
IP interfaces, example configuration.....	154, 157
IP routing	
cross-subnet example	152
default gateway configuration.....	155
IP interface configuration.....	154, 157
IP subnets	152
network diagram.....	152
subnet configuration example.....	154
switch-based topology	153
IP subnets	153
routing.....	152, 153
VLANs	71
IPv6 addressing	271
ISL Trunking	92
Isolated VLAN.....	88

J

jumbo frames.....	72
-------------------	----

L

LACP	99
Layer 2 Failover.....	226
LDAP authentication	51
Link Aggregation Control Protocol	99
logical segment. <i>See</i> IP subnets.	
LSAs	194

M

management module.....	24
manual style conventions	20
Maximum Transmission Unit	72
meter	138
mirroring ports.....	255
monitoring ports.....	255
MSTP	119
MTU.....	72
multi-links between switches using port trunking	91
multiple spanning tree groups	108
Multiple Spanning Tree Protocol	119

N

Neighbor Discovery, IPv6.....	274
network management.....	35

O

OSPF	
area types	192
authentication.....	202
configuration examples	206 to 213
default route.....	200
external routes.....	205
filtering criteria	133
host routes	204
link state database.....	194
neighbors.....	194
overview	191
redistributing routes	179, 183
route maps	179, 181
route summarization	199
router ID	201
virtual link.....	201
outgoing route maps	180

P

packet size	72
password	
administrator account	45
default	45
user account	45
payload size	72
Per Hop Behavior (PHB).....	142
port mirroring	255
configuration rules	95
port trunking	93
configuration example	96
EtherChannel	92
fault tolerance.....	93
ports	
for services	134
monitoring	255
physical. <i>See</i> switch ports.	
priority value (802.1p)	146
Private VLANs	88
promiscuous port.....	89
protocol types	133
PVID (port VLAN ID).....	73
PVLAN	83

R

RADIUS	
authentication	42
port 1812 and 1645	134
port 1813	134
SSH/SCP	57
Rapid Spanning Tree Protocol (RSTP)	116
redistributing routes.....	179, 183, 188
redundancy	
active-active	238
hot-standby	238
re-mark.....	138
RIP (Routing Information Protocol)	
advertisements.....	162
distance vector protocol	161
hop count.....	161
TCP/IP route information.....	18, 161
version 1	161
route aggregation.....	183, 188
route maps	179
configuring	181
incoming and outgoing	180
route paths in BGP	185
Router ID, OSPF	201

routers.....	152, 155
border	195
peer	195
port trunking	92
switch-based routing topology	153
routes, advertising	195
routing	177
internal and external.....	195
Routing Information Protocol. <i>See</i> RIP	
RSA keys	57
RSTP	116

S

SecurID.....	58
security	
LDAP authentication	51
port mirroring	255
RADIUS authentication.....	42
VLANs	71
segmentation. <i>See</i> IP subnets.	
segments. <i>See</i> IP subnets.	
service ports	134
SNMP	35, 196
HP-OpenView	35
Source-Specific Multicast	169
Spanning Tree Protocol	
configuration rules	95
Spanning-Tree Protocol	
multiple instances	108
SSH/SCP	
configuring	54
RSA host and server keys	57
statistical load distribution	93
summarizing routes	199
switch failover	237
switch ports VLANs membership	73

T

TACACS+	46
tagging. <i>See</i> VLANs tagging.	
TCP	133
technical terms	
port VLAN identifier (PVID)	74
tagged frame	74
tagged member.....	74
untagged frame	74
untagged member	74
VLAN identifier (VID)	74

text conventions	20
thash	98
Trunk Hash algorithm	98
Trunking configuration rules	95
typographic conventions	20

U

UDP.....	133
user account.....	45

V

virtual interface router (VIR)	234
virtual link, OSPF.....	201
Virtual Local Area Networks. <i>See</i> VLANs.	
virtual router group.....	239
virtual router ID numbering	241
VLANs	
broadcast domains	71, 156
configuration rules	95
default PVID	73
example showing multiple VLANs	78
ID numbers	72
IP interface configuration.....	157
multiple spanning trees	103
multiple VLANs	74
port members	73
PVID	73
routing.....	156
security.....	71
Spanning-Tree Protocol	103
tagging	73 to 79
topologies	78
VRRP (Virtual Router Redundancy Protocol)	
active-active redundancy	238
hot-standby redundancy	238
overview.....	234
virtual interface router	234
virtual router ID numbering.....	241
vrid	234