

# Application Guide



## Alteon OS™ 21.0

Layer 2-3 GbE Switch Module  
for IBM @server BladeCenter

**Version 1.1**

Part Number: 24R9731, September 2005

**NORTEL**  
NETWORKS™

4655 Great America Parkway  
Santa Clara, CA 95054  
[www.nortelnetworks.com](http://www.nortelnetworks.com)  
Reference: 217024-B

Copyright 2004 Nortel Networks, Inc., 4655 Great America Parkway, Santa Clara, California 95054, USA. All rights reserved. Part Number: 24R9731.

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 Nortel Networks, 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).

Nortel Networks, Inc. reserves the right to change any products described herein at any time, and without notice. Nortel Networks, Inc. assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by Nortel Networks, 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 Nortel Networks, Inc.

Alteon OS and ACEswitch are trademarks of Nortel Networks, 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. Check Point® and FireWall-1® are trademarks or registered trademarks of Check Point Software Technologies Ltd. Any other trademarks appearing in this manual are owned by their respective companies.

Originated in the U.S.A.

# Contents

---

## **Preface 13**

- Who Should Use This Guide 13
- What You'll Find in This Guide 13
- Typographic Conventions 15
- How to Get Help 16

## **Part 1: Basic Switching 17**

---

### **Chapter 1: Accessing the Switch 19**

- Management module setup 20
  - Factory-Default vs. MM assigned IP Addresses 20
  - Default Gateway 21
  - Configuring management module for switch access 21
- Using Telnet 23
  - Connect to the Switch via SSH 24
  - BOOTP Relay Agent 24
  - DHCP Relay Agent 26
- Using the Browser-Based Interface 27
  - Configuring BBI Access via HTTP 27
  - Configuring BBI Access via HTTPS 28
- Using SNMP 31
  - SNMP v1.0 31
  - SNMP v3.0 31
  - Configuring SNMP Trap Hosts 36
- Securing Access to the Switch 39
  - Setting Allowable Source IP Address Ranges 40
  - RADIUS Authentication and Authorization 41
  - TACACS+ Authentication 45
  - Secure Shell and Secure Copy 49

## **Chapter 2: Port-based Network Access Control 57**

- Extensible Authentication Protocol over LAN 57
  - 802.1x Authentication Process 59
  - 802.1x Port States 61
  - Supported RADIUS Attributes 62
  - Configuration Guidelines 63

## **Chapter 3: VLANs 65**

- Overview 66
- VLANs and Port VLAN ID Numbers 66
  - VLAN Numbers 66
  - PVID Numbers 67
- VLAN Tagging 68
- VLAN Topologies and Design Considerations 72
  - VLAN configuration rules 73
  - Example 1: Multiple VLANs with Tagging Adapters 73
- VLANs and Default Gateways 75
  - Segregating VLAN Traffic 75
  - Configuring the Local Network 77
  - Configuring Gateways per VLAN 77

## **Chapter 4: Ports and Trunking 81**

- Overview 82
  - Statistical Load Distribution 83
  - Built-In Fault Tolerance 83
  - Before you configure static trunks 83
  - Trunk group configuration rules 84
- Port Trunking Example 85
- Configurable Trunk Hash Algorithm 87
- Link Aggregation Control Protocol 88
  - Configuring LACP 90

## **Chapter 5: Spanning Tree Group 91**

- Overview 92
- Bridge Protocol Data Units (BPDUs) 93
  - Determining the Path for Forwarding BPDUs 93
- Spanning Tree Group configuration guidelines 94

Multiple Spanning Trees	96
Default Spanning Tree configuration	96
Why Do We Need Multiple Spanning Trees?	97
Switch-Centric Spanning Tree Group	97
VLAN Participation in Spanning Tree Groups	98
Configuring Multiple Spanning Tree Groups	99
Port Fast Forwarding	100
Configuring Port Fast Forwarding	100
Fast Uplink Convergence	101
Configuration Guidelines	101
Configuring Fast Uplink Convergence	101

## **Chapter 6: Rapid Spanning Tree Protocol/Multiple Spanning Tree Protocol 103**

Rapid Spanning Tree Protocol	104
Port State Changes	104
Port Type and Link Type	105
RSTP Configuration Guidelines	105
RSTP Configuration Example	106
Multiple Spanning Tree Protocol	107
MSTP Region	107
Common Internal Spanning Tree	107
MSTP Configuration Guidelines	108
MSTP Configuration Example	108

## **Chapter 7: Quality of Service 109**

Overview	110
Using ACL Filters	111
Summary of packet classifiers	111
Summary of ACL Actions	113
Using ACL Blocks and ACL Groups	113
Port-Based QoS Actions	115
Viewing ACL Statistics	116
ACL Configuration Examples	117
Using DSCP Values to Provide QoS	119
Differentiated Services Concepts	119
Using 802.1p Priorities to Provide QoS	123
802.1p Configuration Example	124

Queuing and Scheduling 125

## **Part 2: IP Routing 127**

---

### **Chapter 8: Basic IP Routing 129**

IP Routing Benefits 130

Routing Between IP Subnets 130

Example of Subnet Routing 133

Defining IP Address Ranges for the Local Route Cache 137

Dynamic Host Configuration Protocol 138

    DHCP Relay Agent 139

    DHCP Relay Agent Configuration 140

### **Chapter 9: Routing Information Protocol 141**

Distance Vector Protocol 141

Stability 141

Routing Updates 142

RIPv1 142

RIPv2 142

RIPv2 in RIPv1 compatibility mode 143

RIP Features 143

RIP Configuration Example 144

### **Chapter 10: IGMP Snooping 147**

Overview 148

FastLeave 149

IGMP Filtering 149

Static Multicast Router 150

IGMP Snooping Configuration Example 151

### **Chapter 11: Border Gateway Protocol 155**

Internal Routing Versus External Routing 156

Forming BGP Peer Routers 157

What is a Route Map? 157

    Incoming and Outgoing Route Maps 158

    Precedence 159

    Configuration Overview 159

Aggregating Routes 161

Redistributing Routes	161
BGP Attributes	162
Local Preference Attribute	162
Metric (Multi-Exit Discriminator) Attribute	162
Selecting Route Paths in BGP	163
BGP Failover Configuration	164
Default Redistribution and Route Aggregation Example	167

## **Chapter 12: OSPF 169**

OSPF Overview	170
Types of OSPF Areas	170
Types of OSPF Routing Devices	172
Neighbors and Adjacencies	173
The Link-State Database	173
The Shortest Path First Tree	174
Internal Versus External Routing	174
OSPF Implementation in Alteon OS	175
Configurable Parameters	175
Defining Areas	176
Interface Cost	178
Electing the Designated Router and Backup	178
Summarizing Routes	178
Default Routes	179
Virtual Links	180
Router ID	181
Authentication	181
Host Routes for Load Balancing	184
OSPF Features Not Supported in This Release	185
OSPF Configuration Examples	185
Example 1: Simple OSPF Domain	186
Example 2: Virtual Links	188
Example 3: Summarizing Routes	192
Verifying OSPF Configuration	194

## **Part 3: High Availability Fundamentals 195**

---

## **Chapter 13: High Availability 197**

Layer 2 Failover	198
VLAN Monitor	198
Setting the Failover Limit	199
L2 Failover with Other Features	199
Configuration Guidelines	199
L2 Failover Configurations	200
Configuring Trunk Failover	203
VRRP Overview	204
VRRP Components	204
VRRP Operation	206
Selecting the Master VRRP Router	206
Failover Methods	207
Active-Active Redundancy	208
Hot-Standby Redundancy	209
Alteon OS extensions to VRRP	210
Tracking VRRP Router Priority	210
Virtual Router Deployment Considerations	211
Assigning VRRP Virtual Router ID	211
Configuring the Switch for Tracking	211
High Availability Configurations	213
Active-Active Configuration	213
Hot-Standby Configuration	218

## **Part 4: Appendices 223**

---

### **Appendix A: Troubleshooting 225**

Monitoring Ports	226
Port Mirroring behavior	227
Configuring Port Mirroring	231

### **Appendix B: RADIUS Server Configuration Notes 233**

### **Glossary 235**

### **Index 237**



# Figures

---

Figure 1-1:	Switch management on the BladeCenter management module	22
Figure 1-2:	BOOTP Relay Agent Configuration	24
Figure 1-3:	DHCP Relay Agent Configuration	26
Figure 2-1:	Authenticating a Port Using EAPoL	59
Figure 3-1:	Default VLAN settings	69
Figure 3-2:	Port-based VLAN assignment	70
Figure 3-3:	802.1Q tagging (after port-based VLAN assignment)	71
Figure 3-4:	802.1Q tag assignment	71
Figure 3-5:	802.1Q tagging (after 802.1Q tag assignment)	72
Figure 3-6:	Example 1: Multiple VLANs with VLAN-Tagged Gigabit Adapters	73
Figure 3-7:	Default Gateways per VLAN	75
Figure 4-1:	Port Trunk Group	82
Figure 4-2:	Port Trunk Group Configuration Example	85
Figure 5-1:	Using Multiple Instances of Spanning Tree Group	97
Figure 5-2:	Implementing Multiple Spanning Tree Groups	98
Figure 7-1:	QoS Model	110
Figure 7-2:	Layer 3 IPv4 packet	119
Figure 7-3:	Layer 2 802.1q/802.1p VLAN tagged packet	124
Figure 8-1:	The Router Legacy Network	131
Figure 8-2:	Switch-Based Routing Topology	132
Figure 8-3:	DHCP Relay Agent Configuration	140
Figure 11-1:	iBGP and eBGP	156
Figure 11-2:	Distributing Network Filters in Access Lists and Route Maps	158
Figure 11-3:	BGP Failover Configuration Example	164
Figure 11-4:	Route Aggregation and Default Route Redistribution	167
Figure 12-1:	OSPF Area Types	171
Figure 12-2:	OSPF Domain and an Autonomous System	172
Figure 12-3:	Injecting Default Routes	179
Figure 12-4:	OSPF Authentication	182
Figure 12-5:	A Simple OSPF Domain	186
Figure 12-6:	Configuring a Virtual Link	188
Figure 12-7:	Summarizing Routes	192
Figure 13-1:	Basic Layer 2 Trunk Failover	200
Figure 13-2:	Two trunks, each in a different Failover Trigger	201

Figure 13-3:	Two trunks, one Failover Trigger	202
Figure 13-4:	A Non-VRRP, Hot-Standby Configuration	207
Figure 13-5:	Active-Active Redundancy	208
Figure 13-6:	Hot-Standby Redundancy	209
Figure 13-7:	Active-Active High-Availability Configuration	213
Figure 13-8:	Hot-Standby Configuration	219

# Tables



---

Table 1-1:	GbESM IP addresses, based on switch-module bay numbers	20
Table 1-2:	User Access Levels	44
Table 1-3:	Alteon OS-proprietary Attributes for RADIUS	44
Table 1-4:	Alteon OS-Proprietary Attributes for TACACS+	46
Table 3-1:	Route Cache Example	76
Table 4-1:	Actor vs. Partner LACP configuration	89
Table 5-1:	Ports, Trunk Groups, and VLANs	92
Table 7-1:	Well-Known Protocol Types	112
Table 7-2:	Well-Known Application Ports	112
Table 7-3:	Well-Known TCP flag values	113
Table 7-4:	Default QoS Service Levels	121
Table 8-1:	Subnet Routing Example: IP Address Assignments	133
Table 8-2:	Subnet Routing Example: IP Interface Assignments	133
Table 8-3:	Subnet Routing Example: Optional VLAN Ports	135
Table 8-4:	Local Routing Cache Address Ranges	137
Table 13-1:	VRRP Tracking Parameters	210





# Preface

---

The *Alteon OS 21.0 Application Guide* describes how to configure and use the Alteon OS software on the Layer 2-3 GbE Switch Module for IBM eServer BladeCenter. For documentation on installing the switch physically, see the *Installation Guide* for your GbE Switch Module (GbESM).

## Who Should Use This Guide

---

This *Application 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 Alteon OS 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 GbE Switch Module to configure, view information and run statistics on the switch. This chapter also discusses different methods to manage the switch for remote administrators using specific IP addresses, RADIUS authentication, Secure Shell (SSH), and Secure Copy (SCP).
- [Chapter 2, “Port-based Network Access Control,”](#) describes how to authenticate 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 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 how Jumbo frames can be used to ease server processing overhead.

- [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, “Rapid Spanning Tree Protocol/Multiple Spanning Tree Protocol,”](#) describes Rapid Spanning Tree and Multiple Spanning Tree configurations.
- [Chapter 7, “Quality of Service,”](#) discusses Quality of Service features, including IP filtering using Access Control Lists, Differentiated Services, and IEEE 802.1p priority values.

## Part 2: IP Routing

- [Chapter 8, “Basic IP Routing,”](#) describes how to configure the GbE Switch Module for IP routing using IP subnets, and DHCP Relay.
- [Chapter 9, “Routing Information Protocol,”](#) describes how the Alteon OS software implements standard RIP for exchanging TCP/IP route information with other routers.
- [Chapter 10, “IGMP Snooping,”](#) describes how the Alteon OS software implements IGMP Snooping to handle multicast traffic efficiently.
- [Chapter 11, “Border Gateway Protocol,”](#) describes BGP concepts and BGP features supported in Alteon OS.
- [Chapter 12, “OSPF,”](#) describes OSPF concepts, how OSPF is implemented in Alteon OS, and examples of how to configure your switch for OSPF support.

## Part 3: High Availability Fundamentals

- [Chapter 13, “High Availability,”](#) describes how to use the Virtual Router Redundancy Protocol (VRRP) to ensure that network resources remain available if one GbE Switch Module is removed for service.

## Part 4: Appendices

- [Appendix A, “Troubleshooting,”](#) discusses two tools for troubleshooting your switch—monitoring ports and filtering session dumps.
- [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 GbE Switch Module.

# Typographic Conventions

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

**Table 1** Typographic Conventions

Typeface 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#
<b>AaBbCc123</b>	This bold type appears in command examples. It shows text that must be typed in exactly as shown.	Main# <b>sys</b>
<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# <b>telnet</b> <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# <b>ls</b> [-a]

## How to Get Help

---

If you need help, service, or technical assistance, see the "Getting help and technical assistance" appendix in the Nortel Networks *Layer 2-3 GbE Switch Module for IBM eserver BladeCenter Installation Guide* on the IBM *BladeCenter Documentation CD*.



# 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
- Quality of Service





## CHAPTER 1

# Accessing the Switch

---

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

- “Management module setup” on page 20
- “Using Telnet” on page 23
- “Using the Browser-Based Interface” on page 27
- “Using SNMP” on page 31
- “Securing Access to the Switch” on page 39
  - “Setting Allowable Source IP Address Ranges” on page 40
  - “RADIUS Authentication and Authorization” on page 41
  - “TACACS+ Authentication” on page 45
  - “Secure Shell and Secure Copy” on page 49

## Management module setup

---

The BladeCenter GbE Switch Module 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 100-Mbps Ethernet port on the management module to configure and manage the GbE Switch Module. The GbE Switch Module communicates with the management module through its internal port 15 (MGT1) and port 16 (MGT2), which you can access through the 100 Mbps Ethernet port on the management module. The factory default settings will permit *only* management and control access to the switch module through the 10/100 Mbps Ethernet port on the management module, or the built-in serial port. You can use the six external 10/100/1000 Mbps Ethernet ports 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 on the IBM *BladeCenter Documentation* CD for more information).

### Factory-Default vs. MM assigned IP Addresses

Each GbE Switch Module 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 10.90.90.9x, where x corresponds to the number of the bay into which the GbE Switch Module is installed. For additional information, see the *Installation Guide*). The management module assigns an IP address of 192.168.70.1xx, where xx corresponds to the number of the bay into which each GbE Switch Module is installed, as shown in the following table:

**Table 1-1** 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** – Switch Modules installed in Bay 1 and Bay 2 connect to server NICs 1 and 2, respectively. However, Windows operating systems show that Switch Modules 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 GbE Switch Module from outside your local network, use the management module to assign a default Gateway address to the GbE Switch Module. Choose **I/O Module Tasks > Management** 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* on the *IBM BladeCenter Documentation CD*. The management module provides the appropriate IP addresses for network access (see the applicable *BladeCenter Installation and User's Guide* publications on the *IBM BladeCenter Documentation CD* for more information).**
3. **Select Management on the I/O Module Tasks menu on the left side of the BladeCenter Management Module window. See [Figure 1-1](#).**

The screenshot shows the BladeCenter Management Module interface. The top navigation bar includes the IBM logo, the text "BladeCenter Management Module", and the "e.server" logo. The left sidebar contains a navigation menu with the following items: Bay 1: WMN315716344, Monitors (System Status, Event Log, LEDs, Hardware VPD, Firmware VPD), Blade Tasks (Power/Restart, Remote Control, Firmware Update, Configuration), I/O Module Tasks (Power/Restart, Management), MM Control (General Settings, Login Profiles, Alerts, Network Interfaces, Network Protocols, Security, Configuration File, Firmware Update, Restore Defaults, Restart MM), and Log Off. The main content area is titled "Switch Management" and contains the following text: "Use the following links to jump down to different sections on this page." followed by links for Bay 1, Bay 2, Bay 3, and Bay 4. Below this is a section for "Bay 1 (Ethernet SM)" with a "Current IP Configuration" table showing: Configuration method: Static, IP address: 192.168.70.127, Subnet mask: 255.255.255.0, and Gateway address: 192.168.70.125. A "New Static IP Configuration" section shows the status as "Enabled" and includes a note: "To change the IP configuration for this switch module, fill in the following fields and click 'Save'. This will save and enable the new IP configuration." Below the note are three input fields: IP address (192.168.70.127), Subnet mask (255.255.255.0), and Gateway address (192.168.70.125). At the bottom of the main content area, there is a link for "Advanced Management" and a "Save" button.

**Figure 1-1** Switch management on the BladeCenter management module

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 (available in Director release 4.21)

---

**NOTE – If you change the IP address of the GbE Switch Module, 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 (Switch Tasks > Management > Advanced Management):**

- External Ports
- External management over all ports (required if you want to access the management network through the six external ports on the GbE Switch Module)

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

---

**NOTE** – In the switch management 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 GbE Switch Module.

## Using Telnet

---

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

Once that you have configured the GbE Switch Module 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, minus certain telnet and management commands.

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

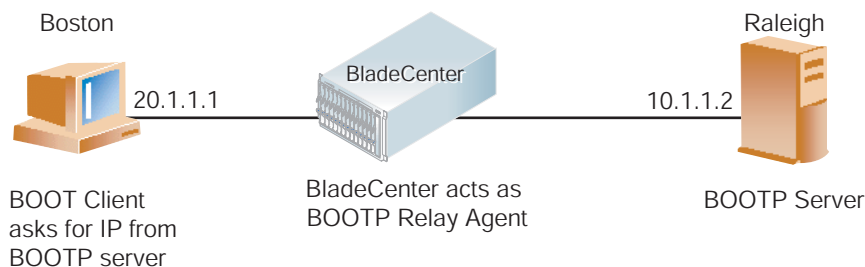
The SSH (Secure Shell) protocol enables you to securely log into another computer over a network to execute commands remotely. As a secure alternative to using Telnet to manage switch configuration, SSH ensures that all data sent over the network is encrypted and secure. For more information, see “[Secure Shell and Secure Copy](#)” on page 49. For more information on the CLI, see the *Alteon OS Command Reference*.

## BOOTP Relay Agent

The GbE Switch Module 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 1-2 figure shows a basic BOOTP network example.



**Figure 1-2** 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.

## Configuring the BOOTP Relay Agent

To enable the GbE Switch Module 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 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 Alteon Application 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 command on the switch IP interface closest to the client so that the DHCP server knows from which IP subnet the newly allocated IP address should come.

The following figure shows a basic DHCP network example:

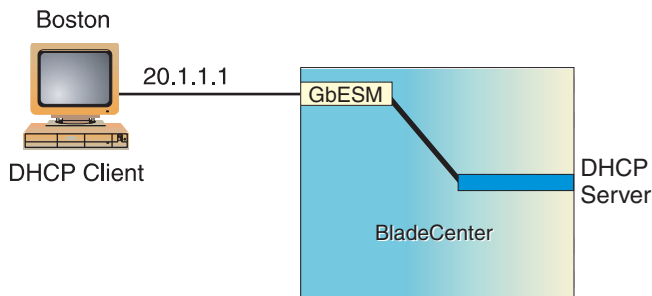


Figure 1-3 DHCP Relay Agent Configuration

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

The Browser-based Interface (BBI) provides access to the common configuration, management and operation features of the GbE Switch Module 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 BBI access on the switch via HTTP, use the following command:

```
/cfg/sys/access/http ena
```

To change the HTTP web server port from the default port 80, use the following command:

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

To access 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.

To enable BBI Access on the switch via HTTPS, use the following command:

```
/cfg/sys/access/https/https ena
```

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

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

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
```

The certificate can be saved to flash for use if the switch is rebooted by using the `apply` and `save` commands.

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 *BBI Quick Guide*.

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
  - Radius
  - TACACS+
  - NTP
  - Boot
  - Syslog/Trap Features
  - Config/Image Control
  - Mgmt
- **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
  - Virtual LANs
  - Spanning Tree Groups
  - MSTP/RSTP
  - Trunk Groups
  - Trunk Hash
  - LACP
  - Uplink Fast
- **Layer 3** – configure Layer 3 features for the switch.
  - IP Interfaces
  - Network Routes

- Route Maps
- Border Gateway Protocol
- 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 SNMP

---

Alteon OS provides SNMP v1.0 and SNMP v3.0 support for access through any network management software, such as IBM Director or HP-OpenView.

### SNMP v1.0

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 traps sent out by the SNMP agent on the switch, the trap host on the switch should be configured with the following command:

```
/cfg/sys/ssnmp/<trap1/trap2>
```

### SNMP v3.0

SNMPv3 is an enhanced version of the Simple Network Management Protocol, approved by the Internet Engineering Steering Group in March, 2002. SNMP v3.0 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 ensures that the client can use SNMPv3 to query the MIBs, mainly for security.

To access the SNMP v3.0 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 *Alteon OS Command Reference*.

## Default configuration

Alteon OS has two users by default. Both the users 'adminmd5' and 'adminsha' have access to all the MIBs supported by the switch.

- 1) username 1: adminmd5/password adminmd5. Authentication used is MD5.
- 2) username adminsha/password 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. Currently we support two authentication algorithms: MD5 and SHA. These can be specified using the command  
 /c/sys/ssnmp/snmpv3/usm <x>/auth md5|sha.

1. **To configure a user with name 'test,' authentication type MD5, and authentication password of 'test,' privacy option DES with privacy password of 'test,' use the following CLI commands.**

```
>> # /cfg/sys/ssnmp/snmpv3/usm 5
>> SNMPv3 usmUser 5 # name "test"
>> SNMPv3 usmUser 5 # auth md5
>> SNMPv3 usmUser 5 # authpw test
>> SNMPv3 usmUser 5 # priv des
>> SNMPv3 usmUser 5 # privpw test
```



2. Once a user is configured you need to specify the access level for this user along with the views the user is allowed access to. This is specified in the access table.

```
>> # /cfg/sys/ssnmp/snmpv3/access 5
>> SNMPv3 vacmAccess 5 # name "testgrp"
>> SNMPv3 vacmAccess 5 # level authPriv
>> SNMPv3 vacmAccess 5 # rview "iso"
>> SNMPv3 vacmAccess 5 # wview "iso"
>> SNMPv3 vacmAccess 5 # nview "iso"
```

3. The group table links the user to a particular access group.

```
>> # /cfg/sys/ssnmp/snmpv3/group 5
>> SNMPv3 vacmSecurityToGroup 5 # uname test
>> SNMPv3 vacmSecurityToGroup 5 # gname testgrp
```

If you want to allow user access only to certain MIBs, see the 'View based Configuration' section.

## View based Configurations

### *CLI User equivalent*

To configure an SNMP user equivalent to the CLI 'user,' use the following configuration:

```

/c/sys/ssnmp/snmpv3/usm 4
name "usr"
/c/sys/ssnmp/snmpv3/access 3
name "usrgrp"
rview "usr"
wview "usr"
nview "usr"
/c/sys/ssnmp/snmpv3/group 4
uname usr
gname usrgrp
/c/sys/ssnmp/snmpv3/view 6
name "usr"
tree "1.3.6.1.4.1.1872.2.5.1.2"
/c/sys/ssnmp/snmpv3/view 7
name "usr"
tree "1.3.6.1.4.1.1872.2.5.1.3"
/c/sys/ssnmp/snmpv3/view 8
name "usr"
tree "1.3.6.1.4.1.1872.2.5.2.2"
/c/sys/ssnmp/snmpv3/view 9
name "usr"
tree "1.3.6.1.4.1.1872.2.5.2.3"
/c/sys/ssnmp/snmpv3/view 10
name "usr"
tree "1.3.6.1.4.1.1872.2.5.3.2"
/c/sys/ssnmp/snmpv3/view 11
name "usr"
tree "1.3.6.1.4.1.1872.2.5.3.3"
/c/sys/ssnmp/snmpv3/view 12
name "usr"
tree "1.3.6.1.4.1.1872.2.5.4.2"
/c/sys/ssnmp/snmpv3/view 13
name "usr"
tree "1.3.6.1.4.1.1872.2.5.4.3"
/c/sys/ssnmp/snmpv3/view 14
name "usr"
tree "1.3.6.1.4.1.1872.2.5.5.2"
/c/sys/ssnmp/snmpv3/view 15
name "usr"
tree "1.3.6.1.4.1.1872.2.5.5.3"
/c/sys/ssnmp/snmpv3/view 16
name "usr"
tree "1.3.6.1.4.1.1872.2.5.6.2"

```

*CLI oper equivalent*

```

/c/sys/ssnmp/snmpv3/usm 5
name "oper"
/c/sys/ssnmp/snmpv3/access 4
name "opergrp"
rview "oper"
wview "oper"
nview "oper"
/c/sys/ssnmp/snmpv3/group 4
uname oper
gname opergrp
/c/sys/ssnmp/snmpv3/view 20
name "oper"
tree "1.3.6.1.4.1.1872.2.5.1.2"
/c/sys/ssnmp/snmpv3/view 21
name "oper"
tree "1.3.6.1.4.1.1872.2.5.1.3"
/c/sys/ssnmp/snmpv3/view 22
name "oper"
tree "1.3.6.1.4.1.1872.2.5.2.2"
/c/sys/ssnmp/snmpv3/view 23
name "oper"
tree "1.3.6.1.4.1.1872.2.5.2.3"
/c/sys/ssnmp/snmpv3/view 24
name "oper"
tree "1.3.6.1.4.1.1872.2.5.3.2"
/c/sys/ssnmp/snmpv3/view 25
name "oper"
tree "1.3.6.1.4.1.1872.2.5.3.3"
/c/sys/ssnmp/snmpv3/view 26
name "oper"
tree "1.3.6.1.4.1.1872.2.5.4"
/c/sys/ssnmp/snmpv3/view 27
name "oper"
tree "1.3.6.1.4.1.1872.2.5.4.1"
type excluded
/c/sys/ssnmp/snmpv3/view 28
name "oper"
tree "1.3.6.1.4.1.1872.2.5.5.2"
/c/sys/ssnmp/snmpv3/view 29
name "oper"
tree "1.3.6.1.4.1.1872.2.5.5.3"
/c/sys/ssnmp/snmpv3/view 30
name "oper"
tree "1.3.6.1.4.1.1872.2.5.6.2"

```

## Configuring SNMP Trap Hosts

### SNMPv1 trap host

1. **Configure a user with no authentication and password.**

```
/c/sys/ssnmp/snmpv3/usm 10
name "vltrap"
```

2. **Configure an access group and group table entries for the user. The command `/c/sys/ssnmp/snmpv3/access <x>/nview` can be used to specify which traps can be received by the user. In the example below the user will receive the traps sent by the switch.**

```
/c/sys/ssnmp/snmpv3/access 10
name "vltrap"
model snmpv1
nview "iso"
/c/sys/ssnmp/snmpv3/group 10
model snmpv1
uname vltrap
gname vltrap
```

3. **Configure an entry in the notify table.**

```
/c/sys/ssnmp/snmpv3/notify 10
name vltrap
tag vltrap
```

4. **Specify the IP address and other trap parameters in the targetAddr and targetParam tables. The `c/sys/ssnmp/snmpv3/tparam <x>/uname` command is used to specify the user name used with this targetParam table.**

```
/c/sys/ssnmp/snmpv3/taddr 10
name vltrap
addr 47.80.23.245
taglist vltrap
pname vlparam
/c/sys/ssnmp/snmpv3/tparam 10
name vlparam
mpmodel snmpv1
uname vltrap
model snmpv1
```

## 5. The community string used in the traps is specified using the community table.

```
/c/sys/ssnmp/snmpv3/comm 10
index v1trap
name public
uname v1trap
```

## SNMPv2 trap host configuration

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

```
c/sys/ssnmp/snmpv3/usm 10
name "v2trap"
/c/sys/ssnmp/snmpv3/access 10
    name "v2trap"
    model snmpv2
    nview "iso"
/c/sys/ssnmp/snmpv3/group 10
    model snmpv2
    uname v2trap
    gname v2trap
/c/sys/ssnmp/snmpv3/taddr 10
    name v2trap
    addr 47.81.25.66
    taglist v2trap
    pname v2param
/c/sys/ssnmp/snmpv3/tparam 10
    name v2param
    mpmodel snmpv2c
    uname v2trap
    model snmpv2
/c/sys/ssnmp/snmpv3/notify 10
    name v2trap
    tag v2trap
/c/sys/ssnmp/snmpv3/comm 10
    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 command `/c/sys/ssnmp/snmpv3/access <x>/level`, and `/c/sys/ssnmp/snmpv3/tparam <x>`. The user in the user table should be configured 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:

```
/c/sys/ssnmp/snmpv3/usm 11
    name "v3trap"
    auth md5
    authpw v3trap
/c/sys/ssnmp/snmpv3/access 11
    name "v3trap"
    level authNoPriv
    nview "iso"
/c/sys/ssnmp/snmpv3/group 11
    uname v3trap
    gname v3trap
/c/sys/ssnmp/snmpv3/taddr 11
    name v3trap
    addr 47.81.25.66
    taglist v3trap
    pname v3param
/c/sys/ssnmp/snmpv3/tparam 11
    name v3param
    uname v3trap
    level authNoPriv
/c/sys/ssnmp/snmpv3/notify 11
    name v3trap
    tag v3trap
```

## 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:

- Limiting management users to a specific IP address range. See [“Setting Allowable Source IP Address Ranges” on page 40](#)
- Authentication and authorization of remote administrators: see [“RADIUS Authentication and Authorization” on page 41](#)
- Encryption of management information exchanged between the remote administrator and the switch: see [“Secure Shell and Secure Copy” on page 49](#)

The following sections are addressed in this section:

- [“Setting Allowable Source IP Address Ranges” on page 40](#)
- [“RADIUS Authentication and Authorization” on page 41](#)
- [“TACACS+ Authentication” on page 45](#)
- [“Secure Shell and Secure Copy” on page 49](#)

## Setting Allowable Source IP Address Ranges

To limit access to the switch without having to configure filters for each switch port, you can set a source IP address (or range) that will be allowed to connect to the switch IP interface through Telnet, SSH, SNMP, or the Alteon OS Browser-Based Interface (BBI). This will also help to prevent spoofing or attacks on the switch's TCP/IP stack.

When an IP packet reaches the switch, the source IP address is checked against the range of addresses defined by the management network and mask, (`mnet` and `mmask`) . If the source IP address of the host or hosts are within this range, they are allowed to attempt to log in. Any packet addressed to a switch IP interface with a source IP address outside this range is discarded.

### Configuring an IP Address Range for the Management Network

Configure Management network IP address and mask from the System menu in the Command Line Interface (CLI).

```
>> Main# /cfg/sys/access/mnet 192.192.192.0
Current management network:      0.0.0.0
New pending management network: 192.192.192.0
>> System Access# mmask 255.255.255.128
Current management netmask:      0.0.0.0
New pending management netmask: 255.255.255.128
```

In this example, the management network is set to 192.192.192.0 and management mask is set to 255.255.255.128. This defines the following range of allowed IP addresses: 192.192.192.1 to 192.192.192.126. The following source IP addresses are granted or not granted access to the switch:

- A host with a source IP address of 192.192.192.21 falls within the defined range and would be allowed to access the switch.

---

**NOTE** – A host with a source IP address of 192.192.192.192 falls outside the defined range and is not granted access. To make this source IP address valid, you would need to shift the host to an IP address within the valid range specified by the `mnet` and `mmask` or modify the `mnet` to be 192.192.192.128 and the `mmask` to be 255.255.255.128. This would put the 192.192.192.192 host within the valid range allowed by the `mnet` and `mmask` (192.192.192.128-255).

---



## RADIUS Authentication and Authorization

Alteon OS 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 GbE Switch Module—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 GbE Switch Module. For more information, see [Appendix B, “RADIUS Server Configuration Notes.”](#)

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>
```



**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 TCP 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]: <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# time
Current RADIUS server timeout: 3
Enter new RADIUS server timeout [1-10]: 10 (Enter the timeout period in minutes)
```

## RADIUS Authentication Features in Alteon OS

Alteon OS supports the following RADIUS authentication features:

- Supports RADIUS client on the switch, based on the protocol definitions in RFC 2138 and RFC 2866.
- Allows RADIUS secret password up to 32 bytes and less than 16 octets.
- 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 in 1-3 retries. The switch will also automatically retry connecting to the RADIUS server 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 1-2](#) can be defined in the RADIUS server dictionary file.

**Table 1-2** 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	The Operator manages all functions of the switch. The Operator can reset ports or the entire switch.	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 Alteon OS 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. The default is `disable` for Telnet access. 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 Alteon OS user privileges levels:

**Table 1-3** Alteon OS-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>	250

## TACACS+ Authentication

Alteon OS supports authentication and authorization with networks using the Cisco Systems TACACS+ protocol. The GbE Switch Module 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 GbE Switch Module 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 41](#).

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 Alteon OS

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. Alteon OS 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 mapping between TACACS+ authorization levels and Alteon OS management access levels is shown in [Table 1-4](#). The authorization levels must be defined on the TACACS+ server.

**Table 1-4** Alteon OS-Proprietary Attributes for TACACS+

Alteon OS User Access Level	TACACS+ level
user	0
oper	3
admin	6

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. The default is `disable` for Telnet access. Backdoor access is always enabled on the console port.

---

**NOTE** – To obtain the TACACS+ backdoor password for your GbESM, 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 GbE Switch Module 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`), Alteon OS configuration commands are sent to the TACACS+ server for authorization. When TACACS+ Command Logging is enabled (`/cfg/sys/tacacs/clog ena`), Alteon OS configuration commands are logged on the TACACS+ server.

The following examples illustrate the format of Alteon OS 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/l3/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 Alteon OS, 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
```

## 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 retry attempts for contacting the TACACS+ server, and the time-out 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]: 10(Enter the timeout period in minutes)

```

5. **Apply and save the configuration.**

## Secure Shell and Secure Copy

Secure Shell (SSH) and Secure Copy (SCP) use secure tunnels to encrypt and secure messages between a remote administrator and the switch. Telnet does not provide this level of security. The Telnet method of managing an GbE Switch Module does not provide a secure connection.

**SSH** is a protocol that enables remote administrators to log securely into the GbE Switch Module 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 GbE Switch Module, SCP is used to download and upload the switch configuration via secure channels.

The benefits of using SSH and SCP are listed below:

- Authentication of remote administrators
- Identifying the administrator using Name/Password
- 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 Alteon OS 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

Before you can use SSH commands, use the following commands to turn on SSH/SCP. SSH and SCP are disabled by default.

### *To enable or disable the SSH feature:*

Begin a Telnet session from the management module 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)
SSHd# apply                                           (Apply the 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 and save)

```

## Configuring the SCP Administrator Password

To configure the `scpadm` (SCP Administrator) password, first connect to the switch via the management module. For security reasons, the `scpadm` password may only be configured when connected through the management module.

To configure the password, enter the following command via the CLI. At factory default settings, the current SCP administrator password is `admin`.

```

>> /cfg/sys/sshd/scpadm
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 <switch IP address>:getcfg <local filename>
```

Example:

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

*To upload the configuration to the switch:*

Syntax:

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

Example:

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

## To apply and save the configuration

The `apply` and `save` commands are still needed after the last command (`scp ad4.cfg 205.178.15.157:putcfg`). Or, instead, you can use the following commands:

```
>> # scp ad4.cfg 205.178.15.157:putcfg_apply
>> # scp ad4.cfg 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 GbE Switch Module. The server key is 768 bits and is used to make it impossible to decipher a captured session by breaking into the GbE Switch Module 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 GbE Switch Module through the console port or the management module (commands are not available via external Telnet connection), and enter the following commands to generate them manually.

---

```
>> # /cfg/sys/ssh/hkeygen           (Generates the host key)
>> # /cfg/sys/ssh/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/ssh/intrval <number of hours (0-24)>
```

A value of 0 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, or if another client has logged in immediately prior. 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.

---





## 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 GbESM that connect to blade servers.

The following topics are discussed in this section:

- “Extensible Authentication Protocol over LAN” on page 57
- “802.1x Authentication Process” on page 59
- “802.1x Port States” on page 61
- “Supported RADIUS Attributes” on page 62
- “Configuration Guidelines” on page 63

## Extensible Authentication Protocol over LAN

---

Alteon OS 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.

## 802.1x 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 2-1 shows a typical message exchange initiated by the client.

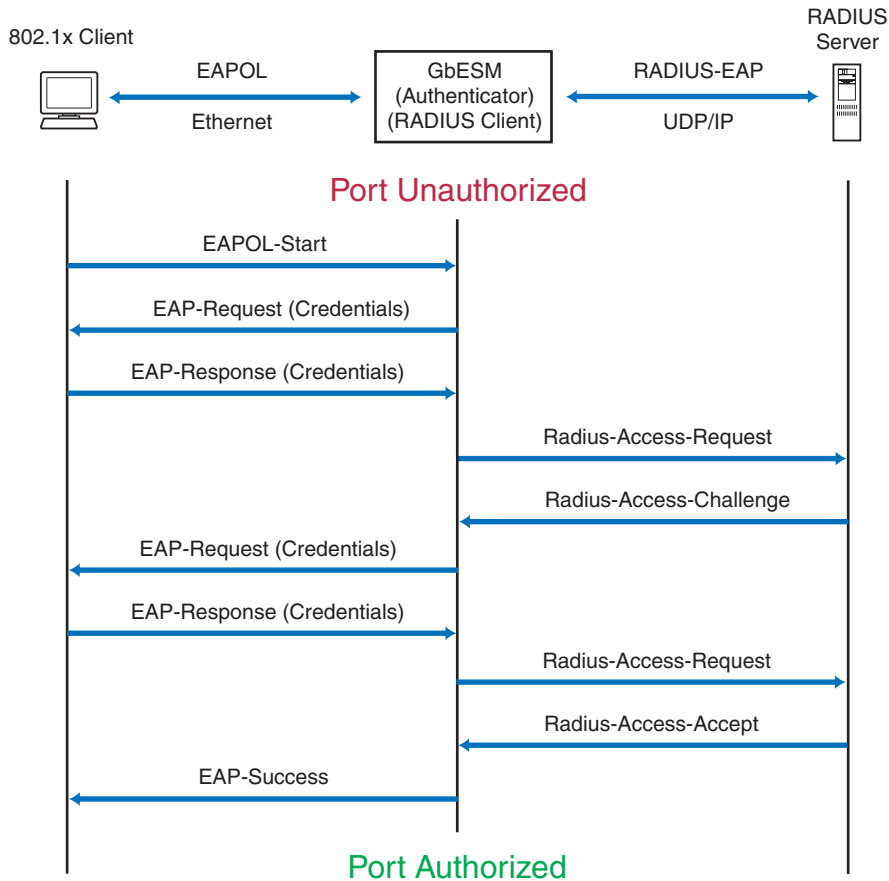


Figure 2-1 Authenticating a Port Using EAPoL

## 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.

---

## 802.1x 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.

## Supported RADIUS Attributes

The Alteon 802.1x Authenticator relies on external RADIUS servers for authentication with EAP. Table 2 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 2** 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.

## 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.
- 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 tunneling 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 66
- “VLAN Tagging” on page 68
- “VLAN Topologies and Design Considerations” on page 72

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.

- “VLANs and Default Gateways” on page 75

---

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

---

## 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 supports jumbo frames, up to 9,216 bytes.

## VLANs and Port VLAN ID Numbers

---

### VLAN Numbers

Alteon OS 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 used by the management network, which includes the management ports and (by default) the internal blade ports. This configuration allows Serial over LAN (SoL) management, a feature available on certain server blades.

### Viewing VLANs

- VLAN information menu:

```
>> Main# info/l2/vlan
```

VLAN	Name	Status	Ports
1	Default VLAN	ena	EXT1 EXT4 INT1-INT14
2	VLAN 2	ena	EXT2 EXT3 EXT5 EXT6
4095	Mgmt Vlan	ena	MGT1 MGT2

**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.

## Viewing and Configuring PVIDs

Use the following CLI commands to view PVIDs:

- Port information:

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

Alias	Port	Tag	FAST	PVID	NAME	VLAN (s)
INT1	1	y	n	1	INT1	1 4095
INT2	2	y	n	1	INT2	1 4095
INT3	3	y	n	1	INT3	1 4095
INT4	4	y	n	1	INT4	1 4095
INT5	5	y	n	1	INT5	1 4095
INT6	6	y	n	1	INT6	1 4095
INT7	7	y	n	1	INT7	1 4095
INT8	8	y	n	1	INT8	1 4095
INT9	9	y	n	1	INT9	1 4095
INT10	10	y	n	1	INT10	1 4095
INT11	11	y	n	1	INT11	1 4095
INT12	12	y	n	1	INT12	1 4095
INT13	13	y	n	1	INT13	1 4095
INT14	14	y	n	1	INT14	1 4095
MGT1	15	y	n	4095	MGT1	4095
MGT2	16	y	n	4095	MGT2	4095
EXT1	17	n	n	1	EXT1	1
EXT2	18	n	n	1	EXT2	1
EXT3	19	n	n	1	EXT3	1
EXT4	20	n	n	1	EXT4	1
EXT5	20	n	n	1	EXT5	1
EXT6	20	n	n	1	EXT6	1

**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.

### ■ Port Configuration:

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

>> Port INT7#
```

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 68).

## VLAN Tagging

---

Alteon OS 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 must also enable tagging on that port.

Since tagging fundamentally changes the format of frames transmitted on a tagged port, you must carefully plan network designs to prevent tagged frames from being transmitted to devices that do not support 802.1Q VLAN tags, or devices where tagging is not enabled.

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.

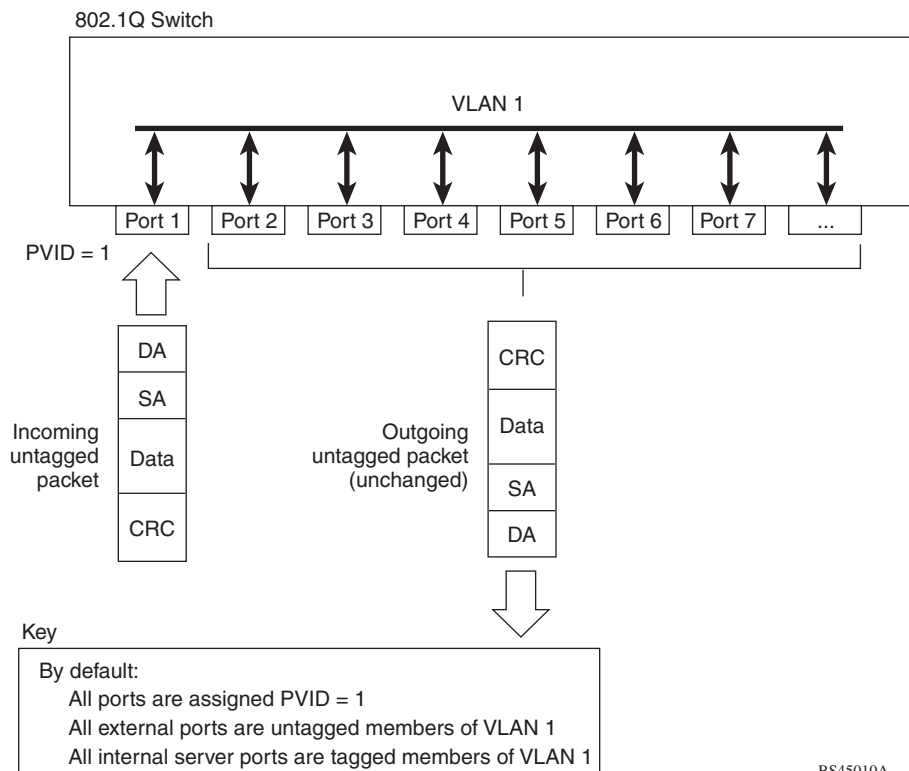
- **Untagged member**—a port that has been configured as an untagged member of a specific VLAN. When an untagged frame exits the switch through an untagged member port, the frame header remains unchanged. When a tagged frame exits the switch through an untagged member port, the tag is stripped and the 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 3-1** Default VLAN settings



BS45010A

---

**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 3-2](#) through [Figure 3-5](#)).

In [Figure 3-1](#), 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.

The default configuration settings for GbE Switch Modules 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 3-1 on page 69](#), all incoming packets are assigned to VLAN 1 by the default port VLAN identifier (PVID = 1).

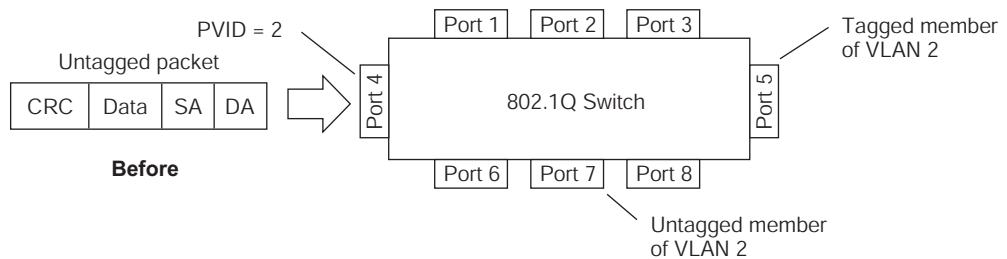
[Figure 3-2](#) through [Figure 3-5](#) illustrate generic examples of VLAN tagging.

---

**NOTE** – The port assignments in the following figures are not meant to match the GbE Switch Module.

---

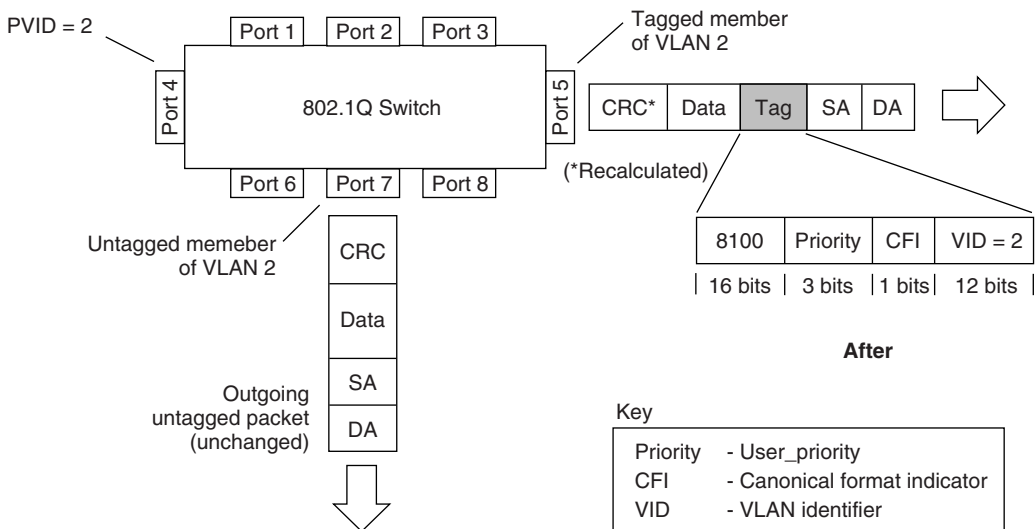
**Figure 3-2** Port-based VLAN assignment



BS45011A

As shown in Figure 3-3, 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.

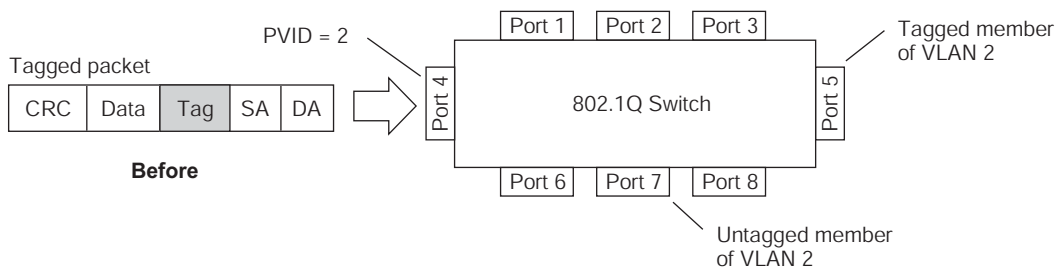
**Figure 3-3** 802.1Q tagging (after port-based VLAN assignment)



BS45012A

In Figure 3-4, 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.

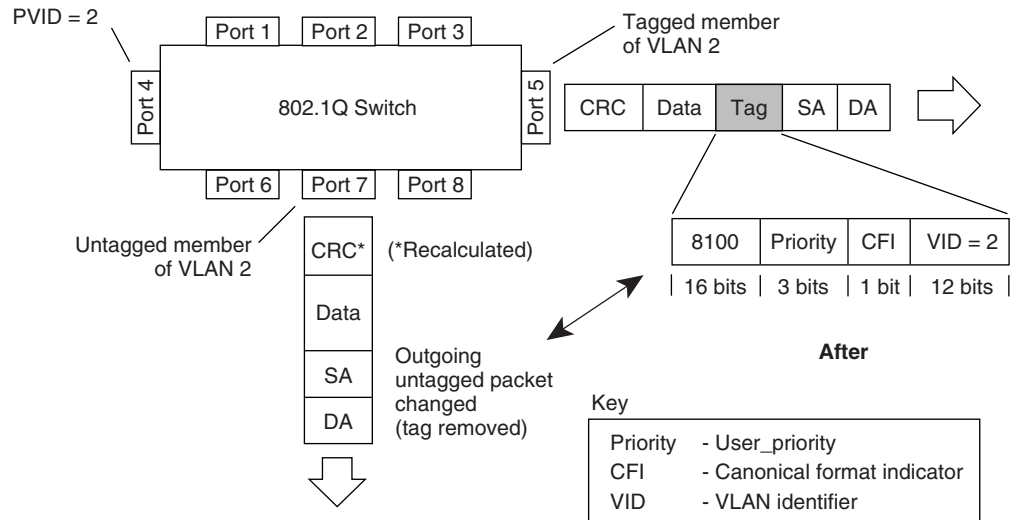
**Figure 3-4** 802.1Q tag assignment



BS45013A

As shown in Figure 3-5, 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 3-5** 802.1Q tagging (after 802.1Q tag assignment)



BS45014A

**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 Alteon OS software is configured so that tagging is disabled on all external ports, and enabled for all internal ports.
- By default, the Alteon OS software is configured so that all internal ports are members of VLAN 1. These ports are also members of VLAN 4095 (the management VLAN), to allow Serial over LAN (SoL) management, a feature of certain server blades.
- By default, the Alteon OS software is configured so that the management ports are members of VLAN 4095 (the management VLAN).



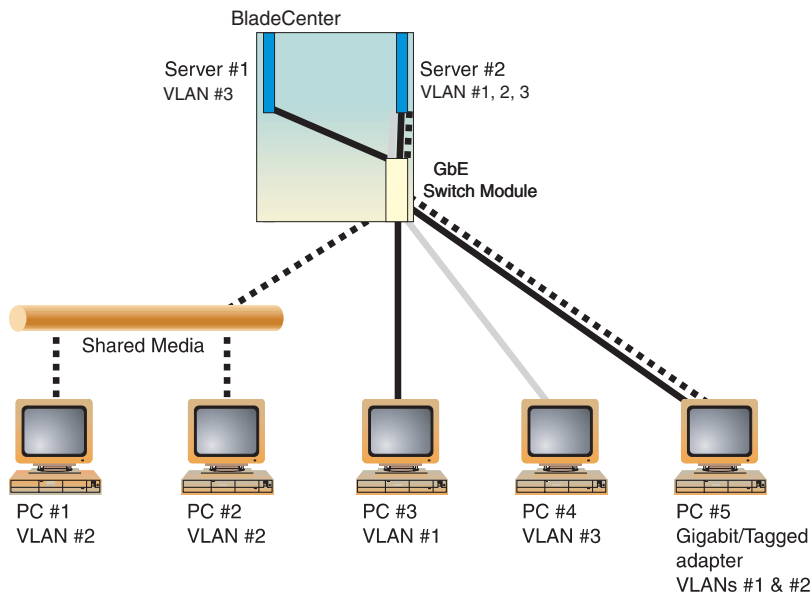
- If configuring Spanning Tree Groups (STG), note that Spanning Tree Groups 2-32 may contain only one VLAN.

## 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 85](#).
- 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” on page 226](#).

### Example 1: Multiple VLANs with Tagging Adapters



**Figure 3-6** Example 1: 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 GbE Switch Modules or on ports that connect to tag-capable end-stations, such as servers with VLAN-tagging adapters.

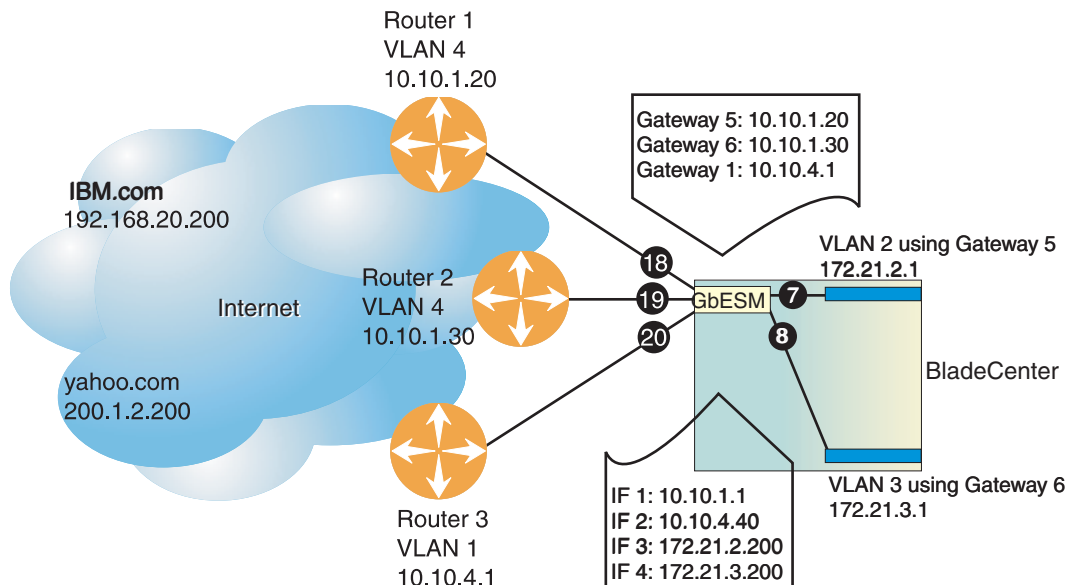
## VLANs and Default Gateways

Alteon OS allows you to assign different gateways for each VLAN. You can effectively map multiple customers to specific gateways on a single switch. The benefits of segregating customers to different default gateways are:

- Resource optimization
- Enhanced customer segmentation
- Improved service differentiation

### Segregating VLAN Traffic

Deploy this feature in an environment where you want to segregate VLAN traffic to a configured default gateway. In [Figure 3-7](#), VLANs 2 and 3 have different routing requirements. VLAN 2 is required to route traffic through default gateway 5 and VLAN 3 is required to route traffic through default gateway 6.



**Figure 3-7** Default Gateways per VLAN

You can configure up to 132 gateways with one gateway per VLAN with values starting from 5 through 132. If the gateways per VLAN fail, then traffic is directed to the default gateway. The default gateway is used as a backup when a specific gateway that has been assigned to a VLAN is down.

In the example shown in [Figure 3-7](#), if gateways 5 or 6 fail, then traffic is directed to default gateway 1, which is configured with IP address 10.10.4.1. If the default gateways is not configured on the switch, then packets from VLAN 2 and VLAN 3 are discarded.

The route cache table on the switch records each session request by mapping the destination IP address with the MAC address of the default gateway. The command `/info/13/arp/dump` on the switch command line will display the entries in the route cache similar to those shown in [Table 3-1](#). The destination IP addresses (see the last two rows) are associated with the MAC addresses of the gateways.

**Table 3-1** Route Cache Example

Destination IP address	Flags	MAC address	VLAN	Port	Referenced SPs
10.10.1.1	P	00:60:cf:46:48:60	4		1-4
10.10.1.20		00:60:cf:44:cd:a0	4	2 (Gig)	empty
10.10.1.30		00:60:cf:42:3b:40	4	3 (Gig)	empty
10.10.4.1		00:60:cf:42:77:e0	1	4 (Gig)	empty
10.10.4.40	P	00:60:cf:46:48:60	1		1-4
172.21.2.27		00:50:da:17:c8:05	2	7	1
172.21.2.200	P	00:60:cf:46:48:60	2		1-4
172.21.3.14		00:c0:4f:09:3e:56	3	8	2
172.21.2.200	P	00:60:cf:46:48:60	3		1-4
192.168.20.200	R	00:60:cf:44:cd:a0	4	5	7
200.1.2.200	R	00:60:cf:42:3b:40	4	6	8

As shown in [Table 3-1](#), traffic from VLAN 2 uses gateway 5 to access destination IP address 192.168.20.200. If traffic from VLAN 3 requests the same destination address, then traffic is routed via gateway 5 instead of gateway 6, because 192.168.20.200 in the route cache is mapped to gateway 5. If the requested route is not in the route cache, then the switch reads the routing table. If the requested route is not in the routing table, then the switch looks at the configured default gateway.

## Configuring the Local Network

To completely segregate VLAN traffic to its own default gateway, you can configure the local network addresses of the VLAN. This will ensure that all traffic from VLAN 2 is forwarded to gateway 5 and all traffic from VLAN 3 is forwarded to gateway 6.

Typically, the switch routes traffic based on the routes in the routing table. The routing table will contain an entry of the configured local network with the default gateway. The route cache will not contain the route entry. This configuration provides a more secure environment, but affects performance if the routing table is close to its maximum capacity.

## Configuring Gateways per VLAN

Follow this procedure to configure the example shown in [Figure 3-7 on page 75](#):

1. Assign an IP address for each router and client workstation.
2. Assign an IP interface for each subnet attached to the switch.

```

>> /cfg/l3/if 1                                     (Select IP interface 1 for gateway 5 &
                                                    6 subnet)
>> IP Interface 1# addr 10.10.1.1                 (Assign IP address for interface 1)
>> IP Interface 1# mask 255.255.255.0            (Assign mask for IF 1)
>> IP Interface 1# vlan 4                         (Assign VLAN 4 to IF 1)
>> IP Interface 1# ../if 2                       (Select IP interface 2 for gateway 1)
>> IP Interface 2# addr 10.10.4.40                (Assign IP address for interface 2)
>> IP Interface 2# mask 255.255.255.0            (Assign mask for IF 2)
>> IP Interface 2# vlan 1                         (Assign VLAN 1 to IF 2)
>> IP Interface 2# ../if 3                       (Select IP interface 3 for VLAN 2
                                                    subnet)
>> IP Interface 3# addr 172.21.2.200             (Assign IP address for interface 3)
>> IP Interface 3# mask 255.255.255.0            (Assign mask for IF 3)
>> IP Interface 3# vlan 2                         (Assign VLAN 2 to IF 3)
>> IP Interface 3# ../if 4                       (Select IP interface 4 for VLAN 3)
                                                    subnet)
>> IP Interface 4# addr 172.21.3.200             (Assign IP address for interface 4)
>> IP Interface 4# mask 255.255.255.0            (Assign mask for IF 4)
>> IP Interface 4# vlan 3                         (Assign VLAN 3 to IF 4)

```

### 3. Configure the default gateways.

Configuring gateways 5 and 6 for VLANs 2 and 3 respectively. Configure default gateway 1 as a backup when gateways 5 and 6 fail.

```
>> /cfg/13/gw 5                               (Select gateway 5)
>> Default gateway 5# addr 10.10.1.20        (Assign IP address for gateway 5)
>> Default gateway 5# ../gw 6                (Select default gateway 6)
>> Default gateway 6# addr 10.10.1.30        (Assign IP address for gateway 6)
>> Default gateway 6# ../gw 1                (Select default gateway 1)
>> Default gateway 1# addr 10.10.4.1        (Assign IP address for gateway 1)
```

---

**NOTE** – The IP address for the default gateway must be unique. IP addresses for default gateways 5 to 132 can be set to the same IP address as the other gateways (including default gateway 1). For example, you can configure two gateways with the same IP address for two different VLANs.

---

### 4. Add the VLANs to the gateways and enable them.

```
>> /cfg/13/gw 5                               (Select gateway 5)
>> Default gateway 5# vlan 2                  (Add VLAN 2 for default gateway 5)
>> Default gateway 5# ena                     (Enable gateway 5)
>> Default gateway 5# ../ gw 6                (Select gateway 6)
>> Default gateway 6# vlan 3                  (Add VLAN 3 for default gateway 6)
>> Default gateway 6# ena                     (Enable gateway 6)
>> Default gateway 6# ../gw 1                (Select default gateway 1)
>> Default gateway 1# ena                     (Enable gateway 1 for all VLAN s)
```

### 5. Apply and verify your configuration.

```
>> Default gateway 1# ../cur                   (View current IP settings)
```

Examine the results under the gateway section. If any settings are incorrect, make the appropriate changes.

**6. (Optional) Configure the local networks to ensure that the VLANs use the configured default gateways.**

```
>> Layer 3# frwd/local (Select the local network Menu)
>> IP Forwarding# add 10.10.0.0 (Specify the network for routers 1, 2,
& 3)
>> IP Forwarding# mask 255.255.0.0 (Add the mask for the routers)
>> IP Forwarding# add 172.21.2.0 (Specify the network for VLAN 2)
>> IP Forwarding# mask 255.255.255.0 (Add the mask for VLAN 2 network)
>> IP Forwarding# add 172.21.3.0 (Specify the network for VLAN 3)
>> IP Forwarding# mask 255.255.255.0 (Add the mask for VLAN 3)
```

**7. Apply and save your new configuration changes.**

```
>> IP Forwarding# apply
>> IP Forwarding# save
```







## CHAPTER 4

# Ports and Trunking

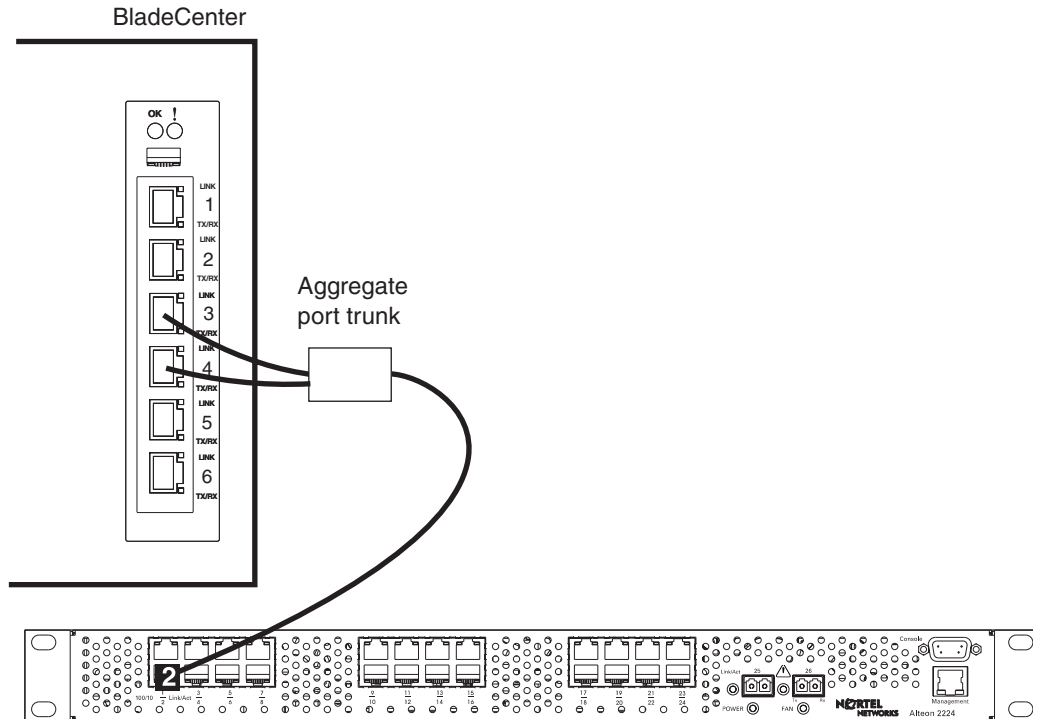
---

Trunk groups can provide super-bandwidth, multi-link connections between GbE Switch Modules 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:

- “[Overview](#)” on this page
- “[Port Trunking Example](#)” on page 85
- “[Configurable Trunk Hash Algorithm](#)” on page 87
- “[Link Aggregation Control Protocol](#)” on page 88

## Overview

When using port trunk groups between two switches, as shown in [Figure 4-1](#), you can create a virtual link between the switches, operating up to 4 Gig per second, depending on how many physical ports are combined. Each GbESM supports up to 13 trunk groups, consisting of one to six ports in each group. Both internal (INT1-INT14) and external ports (EXT1-EXT6) can become members of a trunk group.



**Figure 4-1** Port Trunk Group

Trunk groups are also useful for connecting a GbE Switch Module 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. Nortel Networks 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 Alteon OS-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.

## 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 next section, “[Trunk group configuration rules](#)” on page 84.**
2. **Determine which switch ports (up to six) 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 [Chapter 5, “Spanning Tree Group”](#) for Spanning Tree Group configuration guidelines.**
4. **Consider how existing VLANs will be affected by the addition of a trunk.**

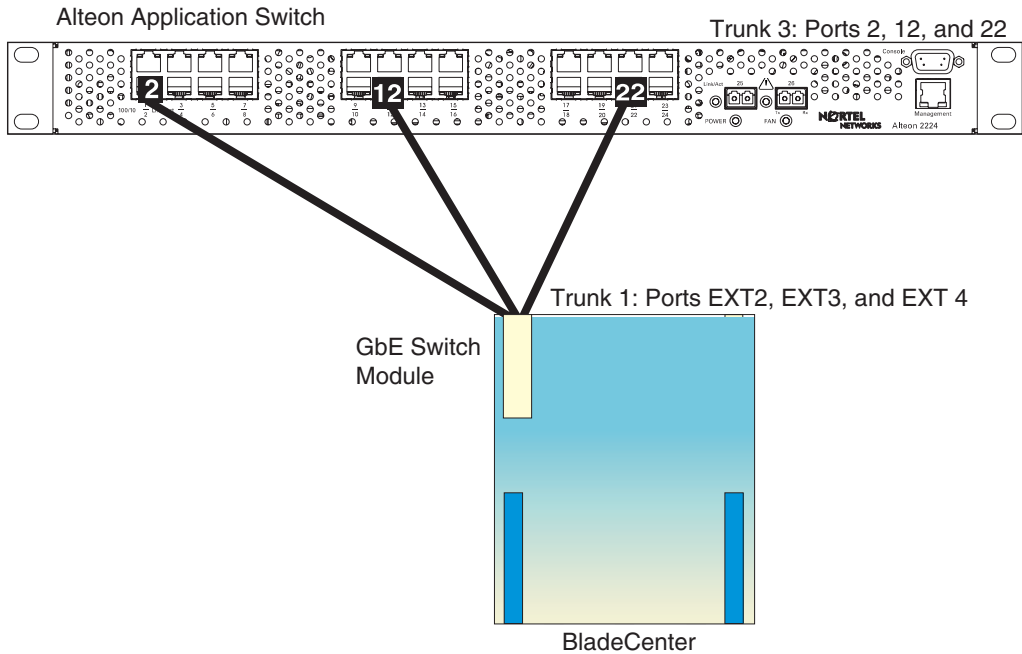
## Trunk group configuration rules

The trunking feature operates according to specific configuration rules. When creating trunks, consider the following rules that determine how the 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.
- 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 `/cfg/12/trunk/ena` command. The Spanning Tree parameters for the port then change to reflect the new trunk settings.
- All trunk members must be in the same Spanning Tree Group (STG) and can belong to only one Spanning Tree Group (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.
- 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.
- To guarantee proper trunking behavior, all ports in static trunks must be configured for full-duplex mode (`cfg/port x/gig/mode full`).

## Port Trunking Example

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



**Figure 4-2** 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 *Alteon OS 21.0 Command Reference*.

---

1. Connect the switch ports that will be members in the trunk group.
2. Follow these steps on the GbESM:

- (a) Define a trunk group.

```

>> # /cfg/12/trunk 1                (Select trunk group 1)
>> Trunk group 1# add EXT2           (Add port EXT2 to trunk group 1)
>> Trunk group 1# add EXT3           (Add port EXT3 to trunk group 1)
>> Trunk group 1# add EXT4           (Add port EXT4 to trunk group 1)
>> Trunk group 1# ena                (Enable trunk group 1)
```

- (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 Alteon Application Switch).

---

**NOTE** – In this example, a GbE Switch Module 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.

The following restrictions apply:

- Any physical switch port can belong to only one trunk group.
- Up to six ports can belong to the same trunk group.
- Best performance is achieved when all ports in any given trunk group are configured for the same speed.
- Trunking from third-party devices must comply with Cisco® EtherChannel® technology.

## Configurable Trunk Hash Algorithm

---

This feature allows you to configure the particular parameters for the GbESM Trunk Hash algorithm instead of having to utilize the defaults. You can configure new default behavior for Layer 2 traffic and Layer 3 traffic using the CLI menu `cfg/12/thash`. You can select a minimum of one 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)

# Link Aggregation Control Protocol

---

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.

The 802.3ad standard 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 re-assigned dynamically to the remaining link/s of the dynamic trunk group.

---

**NOTE** – LACP implementation in Alteon OS 21.0 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 Market 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 4-1](#).

**Table 4-1** 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)	
Port EXT3 (admin key = 200)		Port 3 (admin key = 60)
Port EXT4 (admin key = 200)		Port 4 (admin key = 60)

In the configuration shown in [Table 4-1](#), Actor switch ports EXT1 and EXT2 aggregate to form an LACP trunk group with Partner switch ports 1 and 2. At the same time, Actor switch ports EXT3 and EXT4 form a different LACP trunk group with a different partner.

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.

---

**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/12/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/12/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:

- “Overview” on page 92
- “Bridge Protocol Data Units (BPDUs)” on page 93
- “Multiple Spanning Trees” on page 96
- “Port Fast Forwarding” on page 100
- “Fast Uplink Convergence” on page 101

## 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.

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 5-1](#).

**Table 5-1** 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 GbE Switch Module to inter-operate well within Rapid Spanning Tree networks.

---

## Bridge Protocol Data Units (BPDUs)

---

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 the root bridge, configure the bridge priority lower than all other switches and bridges on your network. The lower the value, the higher the bridge priority. The bridge priority is configured using the `/cfg/l2/stg x/brg/prio` command in the CLI.

#### Port Priority

The port priority helps determine which bridge port becomes the designated port. In a network topology that has multiple bridge ports connected to a single segment, the port with the lowest port priority becomes the designated port for the segment. The port priority is configured using the `/cfg/l2/stg x/port x/prio` command in the CLI.

## 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 indicates that the default cost will be computed for an auto-negotiated link speed.

# Spanning Tree Group configuration 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 94](#) for information on adding ports to VLANs.
- Add the VLAN to the STG using the `/cfg/12/stg <stg-#>/add <vlan-number>` command.

---

**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
- You cannot delete or move VLAN 1 from STG 1.

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 5-1](#).

## Multiple Spanning Trees

---

Each GbE Switch Module supports a maximum of 32 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 GbE Switch Modules 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 32 STGs running simultaneously. The default STG 1 may contain an unlimited number of VLANs. All other STGs 2-32 may contain only one VLAN each.

---

**NOTE** – When you configure multiple spanning trees, Layer 3 forwarding must be turned off (`/cfg/l3/frwd off`).

---

## 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. Also you cannot delete the default VLAN (VLAN 1) from STG 1.

All other STGs, except the default STG and the management STG (32), are empty and VLANs must be added by the user. However, you cannot assign ports directly to an STG. Add ports to a VLAN and add the VLAN to the STG. Each STG is enabled by default, and assigned an ID number from 2 to 32.

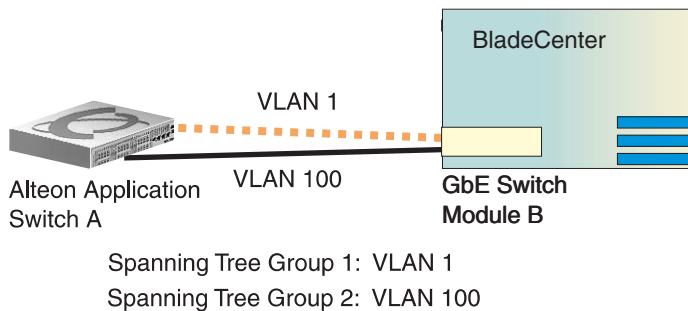
By default, the spanning tree on the management ports is turned off in both STP/PVST+ mode and in MSTP/RSTP mode.



## Why Do We Need Multiple Spanning Trees?

Figure 5-1 shows a simple example of why we need multiple Spanning Trees. Two VLANs, VLAN 1 and VLAN 100 exist between application switch A and GbE Switch Module 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.

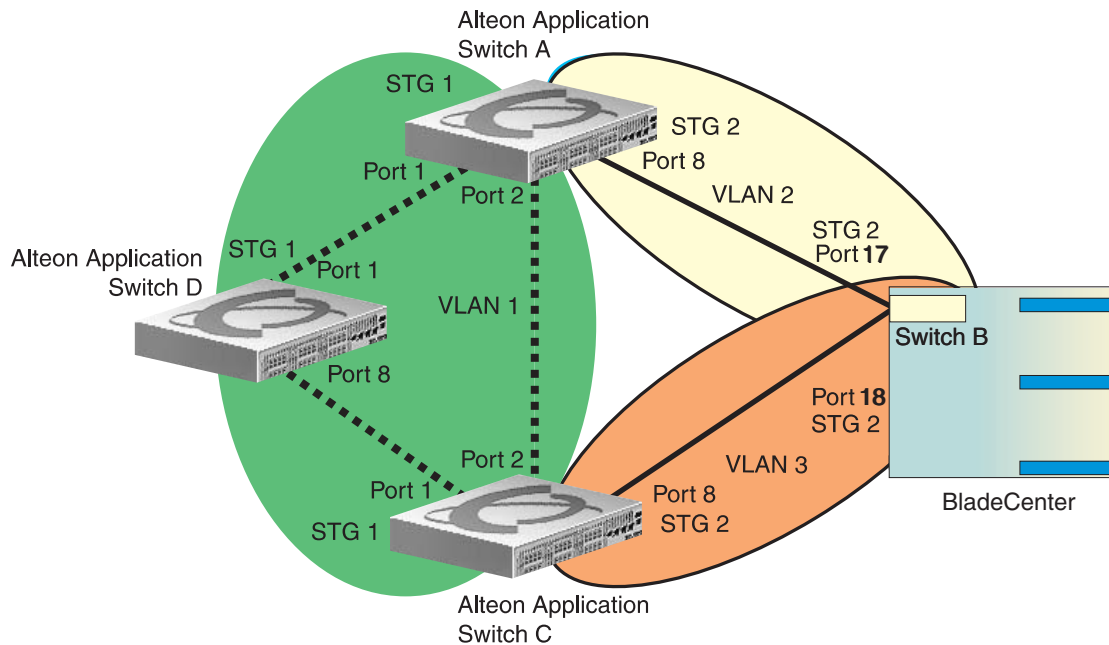
If VLAN 1 and VLAN 100 belong to different Spanning Tree Groups, then the two instances of Spanning Tree separate the topology without forming a loop. Both VLANs can forward packets between the switches without losing connectivity.



**Figure 5-1** Using Multiple Instances of Spanning Tree Group

## Switch-Centric Spanning Tree Group

In Figure 5-2 on page 98, VLAN 2 is shared by application switch A and GbE Switch Module B on ports 8 and 17 respectively. Application Switch A identifies VLAN 2 in Spanning Tree Group 2 and GbE Switch Module B identifies VLAN 2 in Spanning Tree Group 1. 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 5-2** Implementing Multiple Spanning Tree Groups

## VLAN Participation in Spanning Tree Groups

The VLAN participation for each Spanning Tree Group in [Figure 5-2 on page 98](#) 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. GbE Switch Module B receives this BPDU on its port 1. Port 1 on GbE Switch Module B is on VLAN 2, Spanning Tree Group 1. Because switch B has no additional ports participating in Spanning Tree Group 1, this BPDU is not be forwarded to any additional ports and application switch A remains the designated root.

## ■ VLAN 3 Participation

For VLAN 3 you can have GbE Switch Module 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 2. 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 GbE Switch Module 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 5-2 on page 98](#).

By default Spanning Trees 2-32 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 GbE Switch Module B:

Add port 1 to VLAN 2, port 2 to VLAN 3 and define Spanning Tree Groups 2 for VLAN 3.

```
>> # /cfg/12/vlan2                (Select VLAN 2 menu)
>> VLAN 2# add 1                  (Add port 1)
>> VLAN 2# ../vlan3              (Select VLAN 3 menu)
>> VLAN 3# add 2                  (Add port 2)
>> 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** – 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/vlan3	<i>(Select VLAN 3 menu)</i>
>> VLAN 3# add 8	<i>(Add port 8)</i>
>> VLAN 3# ../stg 2	<i>(Select Spanning Tree Group 2)</i>
>> Spanning Tree Group 2# add 3	<i>(Add VLAN 3)</i>

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 GbE Switch Module to interoperate well within Rapid Spanning Tree (RSTP) networks.

### Configuring Port Fast Forwarding

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

>> # /cfg/port ext1	<i>(Select port EXT 1)</i>
>> Port EXT1# fastfwd ena	<i>(Enable Port Fast Forwarding)</i>
>> Port EXT1# apply	<i>(Make your changes active)</i>
>> Port EXT1# save	<i>(Save for restore after reboot)</i>

## 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 sends multicasts of 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, Alteon OS automatically makes the following configuration changes:

- Increases 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 commands to enable Fast Uplink Convergence on external ports.

```
>> # /cfg/l2/upfast ena           (Enable Fast Uplink convergence)
>> Layer 2# apply                 (Make your changes active)
>> Layer 2# save                   (Save for restore after reboot)
```



## CHAPTER 6

# Rapid Spanning Tree Protocol/ Multiple Spanning Tree Protocol

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 104
  - “Port State Changes” on page 104
  - “Port Type and Link Type” on page 105
  - “RSTP Configuration Guidelines” on page 105
  - “RSTP Configuration Example” on page 106
- “Multiple Spanning Tree Protocol” on page 107
  - “MSTP Region” on page 107
  - “Common Internal Spanning Tree” on page 107
  - “MSTP Configuration Guidelines” on page 108
  - “MSTP Configuration Example” on page 108

## 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 [Chapter 5, “Spanning Tree Group.”](#)

RSTP parameters are configured in Spanning Tree Group 1. STP Groups 2-32 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 3](#) compares the port states between 802.1d Spanning Tree and 802.1w Rapid Spanning Trees.

**Table 3** 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 1-32 (`/cfg/12/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 1-16 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. A full-duplex link is point-to-point (**p2p**), while a half-duplex link should be configured as **shared**. If you select **auto** as the link type, the port dynamically configures the link type.

## RSTP Configuration Guidelines

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

- When RSTP is turned on, STP parameters apply only to STP Group 1.
- When RSTP is turned on, STG 32 is automatically added to STG 1, and STG 2-31 are automatically turned off.
- When RSTP is turned on, all VLANs (including the management VLAN 4095) are moved to Spanning Tree Group 1.

## RSTP Configuration Example

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

### Configure Rapid Spanning Tree

1. **Configure port and VLAN membership on the switch.**
2. **Disable and clear STP groups 2 through 32.**

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

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

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

4. **Configure STP Group 1 parameters.**

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

# 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 [Chapter 5, “Spanning Tree Group.”](#)

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.

---

**NOTE** – When MSTP is turned on, Layer 3 forwarding must be turned off (`/cfg/l3/frwd off`).

---

## 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
- Version 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 1-32. 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 32.
- For enabling MSTP, Region Name must be configured, 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).

### Configure Multiple Spanning Tree Protocol

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 xxxxxx (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)
```

4. **Turn off Layer 3 forwarding.**

```
>> /cfg/13/frwd off (Turn Layer 3 forwarding off)
>> IP Forwarding# apply (Apply the configuration)
```



## CHAPTER 7

# Quality of Service

---

Quality of Service 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 Quality of Service (QoS) level.

The following topics are discussed in this section:

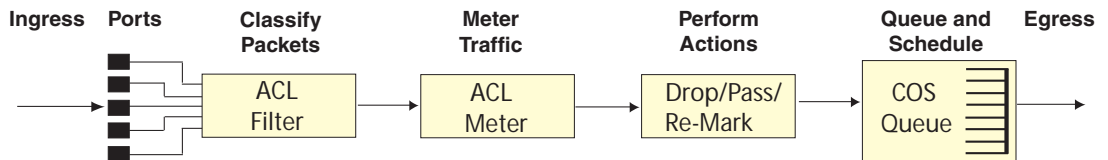
- [“Overview” on page 110](#)
- [“Using ACL Filters” on page 111](#)
- [“Using DSCP Values to Provide QoS” on page 119](#)
- [“Using 802.1p Priorities to Provide QoS” on page 123](#)
- [“Queuing and Scheduling” on page 125](#)

## Overview

QoS helps you allocate guaranteed bandwidth to the 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 cannot tolerate delay by 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 7-1 shows the basic QoS model used by the GbESM.



**Figure 7-1** QoS Model

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

With DiffServ, you can establish policies to direct 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 eight COS queues
  - Schedule transmission based on the COS queue weight

## Using ACL Filters

---

Access Control Lists 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 7-1](#).

**Table 7-1** 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 7-2](#) on page 112
  - TCP/UDP application destination port as shown in [Table 7-2](#) on page 112
  - TCP/UDP flag value as shown in [Table 7-3](#) on page 113

**Table 7-2** Well-Known Application Ports

Number	TCP/UDP Application	Number	TCP/UDP Application	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 7-3** 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 packets

## 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

## Using ACL Blocks and ACL Groups

Access Control Lists (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. Packet classifiers identify flows for more processing.

You can define a traffic profile by compiling a number of ACLs into an ACL Group, and assigning the ACL Group to a port. Each port can accept up to 127 filters, including ACL Groups.

Each ACL (or ACL Block) in an ACL Group has a precedence level, based on its position in the ACL Group. The first ACL added to the ACL Group has the highest precedence, the last ACL added has the lowest precedence (top to bottom). If a packet matches the highest precedence ACL, the following ACLs are also considered, in order of precedence.

ACL Groups are assigned and enabled on a per-port basis. Each ACL can be used by itself or in combination with other ACLs, ACL Blocks, or ACL Groups on a given switch port. When multiple ACLs are grouped together on a port, each ACL's position determines its order of precedence: the first or topmost ACL is checked first. When traffic is encountered at the switch port, if the ACL criteria matches, its configured action takes place. Other ACLs are checked in order of precedence.

ACLs can be grouped in the following manner:

#### ■ Access Control Lists

Access Control Lists (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. Packet classifiers identify flows for more processing.

The GbESM supports up to 4096 ACLs. Each ACL defines one filter rule. Each filter rule is a collection of matching criteria, and can include an action (permit or deny the packet). For example:

<p><b>ACL 1:</b>  VLAN = 1  SIP = 10.10.10.1 (255.255.255.0)  Action = permit</p>
---

#### ■ Access Control Blocks

An Access Control Block (ACL Block) is a collection of ACLs that define filtering criteria on the same set of fields and within the same mask. Each ACL in an ACL Block must match the same parameters and the same mask, range, and packet format type. For example:

<b>ACL Block 1</b>	
<p><b>ACL 1:</b>  VLAN = 1  SIP = 10.10.10.1 (255.255.255.0)  Action = permit</p>	<p><b>ACL 2:</b>  VLAN = 2  SIP = 10.10.10.2 (255.255.255.0)  Action = deny</p>
<b>ACL Block 2</b>	
<p><b>ACL 3:</b>  PRI = 0  SIP = 20.20.20.1 (255.255.0.0)  Action = deny</p>	<p><b>ACL 4:</b>  PRI = 1  DIP = 20.20.20.2 (255.255.0.0)  Action = deny</p>

All ACLs within an ACL Block remain at the same precedence level. Use ACL Blocks to define a number of ACLs that do not require different precedence levels and therefore, utilize only one of the eight available precedence levels. The GbESM supports up to 4096 ACL Blocks.

### ■ Access Control Groups

An Access Control Group (ACL Group) is a collection of ACLs and ACL Blocks. For example:

<b>ACL Group 1</b>	
<b>ACLBLK 1:</b>	
<b>ACL 1:</b> VLAN = 1 SIP = 10.10.10.1 (255.255.255.0) Action = permit	<b>ACL 2:</b> VLAN = 2 SIP = 10.10.10.2 (255.255.255.0) Action = deny
<b>ACL 3:</b> Priority = 7 DIP = 10.10.10.3 (255.255.255.0) Action = permit	

In the example above, ACLBLK 1 can contain several ACLs, with each ACL defining a filter rule. ACLBLK 1 has a higher precedence than ACL 3, based on its position in the ACL Group.

Use ACL Groups to create a traffic profile by gathering ACLs and ACL Blocks into an ACL Group, and assigning the ACL Group to a port. The GbESM supports up to 4096 ACL Groups.

## Port-Based QoS Actions

You can define a profile for the aggregate traffic flowing through the GbESM, by assigning an ACL Group to the port, and configuring a QoS meter (if desired). 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. Each contains three levels of precedence. If you add ACL Group 1 to a port, then add ACL Group 2 to the port, the port's ACL filters contain a total of six precedence levels. ACL Group 1 has precedence over ACL Group 2.

Each port supports up to seven precedence levels.

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, 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 Kb/s (1000 bits per second in each Kb/s). All traffic within this Committed Rate is In-Profile. Additionally, you 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 assign an ACL to a meter, or perform actions on metered traffic, such as packet re-marking. Each port can accept up to 127 meters.

## 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

### Configure Access Control Lists and Groups

#### 1. Configure Access Control Lists (ACLs).

```

>> Main# cfg/acl/acl 1 (Define ACL 1)
>> ACL 1# ethernet/dmac 00:60:af:00:02:20
>> Enter Destination MAC mask (default FF:FF:FF:FF:FF:FF):
>> Filtering Ethernet# ..
>> ACL 1# action permit
>> ACL 1# ..

>> ACL# acl 2 (Define ACL 2)
>> ACL 2# ethernet/dmac 00:60:af:00:02:30
Enter Destination MAC mask (default FF:FF:FF:FF:FF:FF):
>> Filtering Ethernet# ..
>> ACL 2# action setcos
Enter COS to remap packet to [0-7]: 2
>>ACL 2# ..
>>ACL# ..

>>ACL# acl 3 (Define ACL 3)
>>ACL 3# ipv4/sip 10.10.10.1
Enter Destination MAC mask (default 255.255.255.255):
>>Filtering IPv4# ..
>>ACL 3# action permit
>>ACL 3# apply

```

#### 2. Configure ACL Blocks and ACL Groups. In this example, ACL Block 1 takes precedence over ALC 3.

```

>> Main# cfg/acl/block 1 (Define ACL Block 1)
>> ACL Block 1# addacl 1
>> ACL Block 1# addacl 2
>> ACL Block 1# ..

>> ACL# group 1 (Define ACL Group 1)
>> ACL Group 1# addblk 1
>> ACL Group 1# addacl 3
>> ACL Group 1# apply

```

## Assign ACL Groups to Ports, and Configure Port-Based QoS Actions

### 1. Add ACL Group 1 to port EXT 1.

```
>> Main# cfg/port ext1/aclqos (Select port and assign ACL)
>> Port EXT1 ACL# /add 1
>> Port EXT1 ACL# apply
```

### 2. Configure QoS metering.

```
>> Main# cfg/port ext1/aclqos (Select port and define meter)
>> Port EXT1 ACL# meter 1
>> Metering# cir 6400
>> Metering# mbsize 256
>> Metering# assign grp 1
>> Metering# dpass p
>> Metering# enable
Current Metering: disabled
Enter new Metering [d/e]: e
>> Metering# apply
```

### 3. Configure Re-marking.

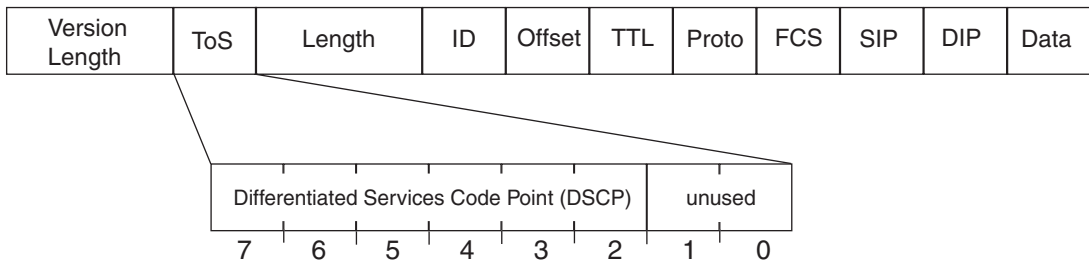
```
>> Main# cfg/port ext1/aclqos (Select port and define re-marking)
>> Port EXT1 ACL# re-mark 1
>> Re-mark Menu# inprof/up1p (Define in-profile re-marking)
>> Update User Priority# value
Enter update precedence [0-7]: 7
>> Update User Priority# ..
>> Re-marking - In Profile# ..
>> Re-mark Menu# outprof (Define out-of-profile re-marking)
>> Re-marking - Out Of Profile# updscp
Enter update dscp [0-63]: 15
>> Re-marking - Out Of Profile# apply
```

## 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 7-2** 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 7-4 shows the default service levels provided by the GbESM, listed from highest to lowest importance:

**Table 7-4** 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

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 default settings by using the `cfg/qos/dscp/cur` command, as shown below.

```
>> DSCP Remark# 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
```

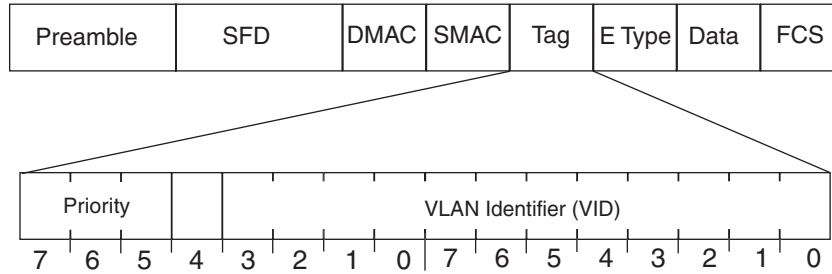
2. Enable DSCP re-marking on a port.

```
>> 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

Alteon OS 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 7-3** 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.

```

>> Main# 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): 1
Enter new COS queue (for priority 1) [0-7]: 2
>> 802.1p# qweight                                    (Set COS queue weights)
Enter COS queue [0-7]: 2
Current weight (for COS queue 2): 3
Enter new weight (for COS queue 2) [0-15]: 10
>> 802.1p# apply

```

## Queuing and Scheduling

The GbESM has eight 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)
- IGMP Snooping
- Border Gateway Protocol (BGP)
- Open Shortest Path First (OSPF)







## CHAPTER 8

# Basic IP Routing

---

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

- “IP Routing Benefits” on page 130
- “Routing Between IP Subnets” on page 130
- “Example of Subnet Routing” on page 133
- “Defining IP Address Ranges for the Local Route Cache” on page 137
- “Dynamic Host Configuration Protocol” on page 138

## IP Routing Benefits

---

The GbE Switch Module 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 another means to invisibly introduce Jumbo frame technology into the server-switched network by automatically fragmenting UDP Jumbo frames when routing to non-Jumbo frame VLANs or subnets.
- Provides the ability to route IP traffic between multiple Virtual Local Area Networks (VLANs) configured on the switch.

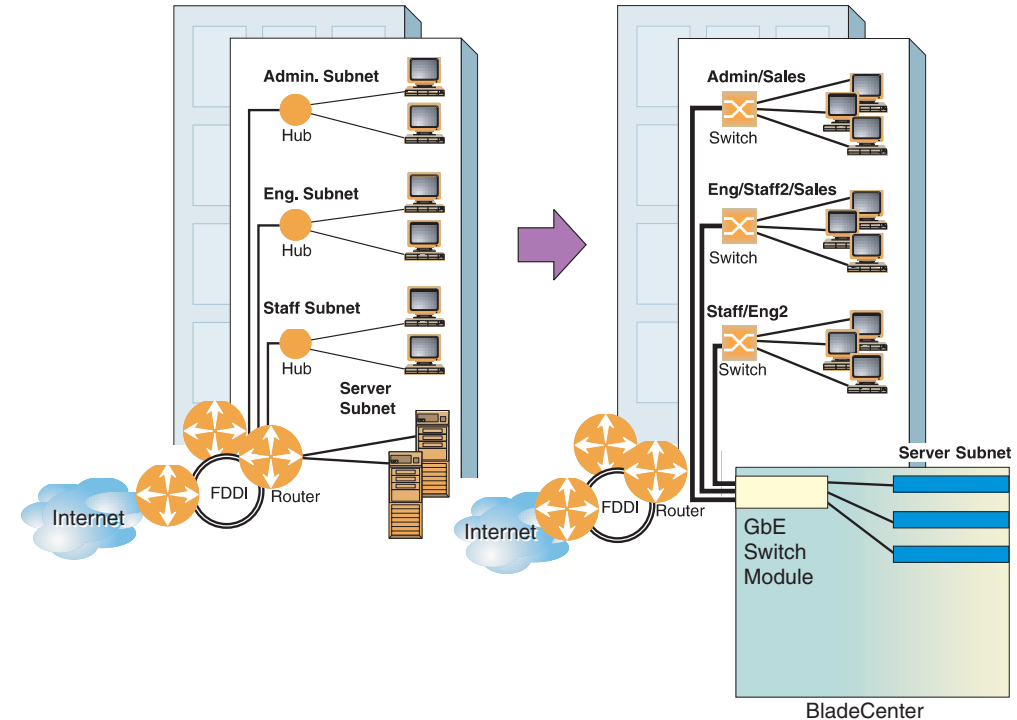
## Routing Between IP Subnets

---

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. GbE Switch Modules are intelligent and fast enough to perform routing functions on a 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 8-1** 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.

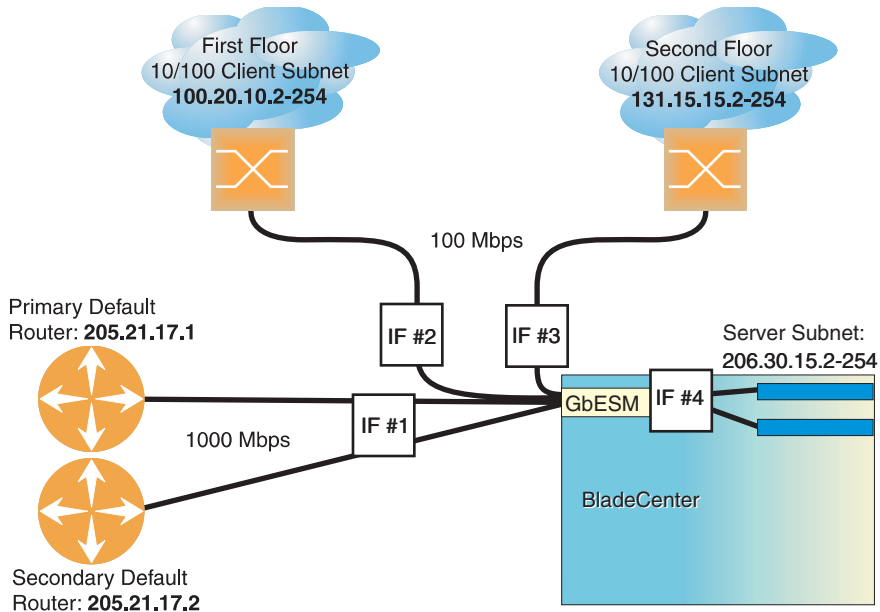
This is a situation that switching alone cannot cure. Instead, the router is flooded with cross-subnet communication. This compromises efficiency in two ways:

- 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 GbE Switch Modules 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 GbE Switch Module in the following configuration example:



**Figure 8-2** Switch-Based Routing Topology

The GbE Switch Module 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 GbE Switch Module, 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.

To make implementation even easier, UDP Jumbo frame traffic is automatically fragmented to regular Ethernet frame sizes when routing to non-Jumbo frame VLANs or subnets. This automatic frame conversion allows servers to communicate using Jumbo frames, all transparently to the user.

## Example of Subnet Routing

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 *Alteon OS Command Reference*.

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

In the example topology in [Figure 8-2 on page 132](#), the following IP addresses are used:

**Table 8-1** 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 8-2** 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 5)
```

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. Enable, apply, and verify the configuration.

```
>> Default gateway 2# ../fwrdd                  (Select the IP Forwarding Menu)
>> IP Forwarding# on                            (Turn IP forwarding on)
>> IP Forwarding# apply                         (Make your changes active)
>> IP Forwarding# /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 8-3** 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 8-3](#) are configured as follows:

```

>> # /cfg/12/vlan 1                               (Select VLAN 1)
>> VLAN 1# add port EXT1                          (Add port for 1st floor to VLAN 1)
>> VLAN 1# add port 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 port EXT3                          (Add port for default router 1)
>> VLAN 2# add port EXT4                          (Add port for default router 2)
>> VLAN 2# ena                                    (Enable VLAN 2)
>> VLAN 2# ../VLAN 3                              (Add port for default router 3)
>> VLAN 3# add port INT5                          (Select VLAN 3)
>> VLAN 3# add port INT6                          (Select 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 8-3 on page 135](#), the settings are made as follows:

```
>> VLAN 3# /cfg/13/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/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)
```



## Defining IP Address Ranges for the Local Route Cache

---

A local route cache lets you use switch resources more efficiently. The local network address and local network mask parameters (accessed via the `/cfg/13/frwd/local/add` command) define a range of addresses that will be cached on the switch. The *local network address* is used to define the base IP address in the range that will be cached. The *local network mask* is applied to produce the range. To determine if a route should be added to the memory cache, the destination address is masked (bit-wise AND) with the local network mask and checked against the local network address.

By default, the local network address and local network mask are both set to 0.0.0.0. This produces a range that includes all Internet addresses for route caching: 0.0.0.0 through 255.255.255.255.

To limit the route cache to your local hosts, you could configure the parameters as shown in the following example:

**Table 8-4** Local Routing Cache Address Ranges

Local Host Address Range	Local Network Address	Local Network Mask
0.0.0.0 - 127.255.255.255	0.0.0.0	128.0.0.0
128.0.0.0 - 128.255.255.255	128.0.0.0	128.0.0.0 or 255.0.0.0
205.32.0.0 - 205.32.255.255	205.32.0.0	255.255.0.0

---

**NOTE** – Static routes must be configured within the configured range. All other addresses that fall outside the defined range are forwarded to the default gateway.

---

# 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).

Nortel Networks 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 GbE Switch Modules 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.

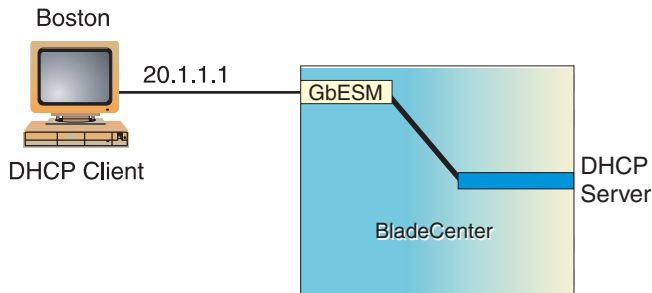
In the DHCP environment, the GbE Switch Module acts as a relay agent. The DHCP relay feature (`/cfg/13/bootp`) enables the switch to forward a client request for an IP address to two BOOTP servers with IP addresses that have been configured on the switch.

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 GbE Switch Module to be the BOOTP forwarder, you need to configure the DHCP/BOOTP server IP addresses on the switch. Generally, you should configure the command on the switch IP interface closest to the client so that 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 8-3** DHCP Relay Agent Configuration

In GbE Switch Module 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 9

# 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). Alteon OS 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 cost, and cost is defined as hop count. One hop is considered to be the distance from one switch to the next which is typically 1. This cost or hop count is known as the metric.

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. After another 120 seconds without receiving an update for those routes, the routes are removed from the routing table and respective 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 *Alteon OS Command Reference*.

### RIPv1

RIP version 1 use 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. Alteon OS supports using clear password for RIPv2.

## RIPv2 in RIPv1 compatibility mode

Alteon OS 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

Alteon OS provides the following features to support RIPv1 and RIPv2:

### Poison

Simple split horizon in RIP scheme omits routes learned from one neighbor in updates sent to that neighbor. That is the most common configuration used in RIP, that is setting this Poison to DISABLE. Split horizon with poisoned reverse 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/l3/rip 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 broadcasts. Multicast RIPv2 announcements 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.

## Default

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 plaintext 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 messages 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/l3/rip 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.

```

>> Main# 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.

```

>> Main# 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 10

# 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.

The following topics are discussed in this chapter:

- “Overview” on page 148
- “FastLeave” on page 149
- “IGMP Filtering” on page 149
- “Static Multicast Router” on page 150
- “IGMP Snooping Configuration Example” on page 151

## Overview

---

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.

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 currently supports snooping for IGMP version 1 and version 2.

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.

## FastLeave

In normal operation, when the switch with IGMP Snooping enabled 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 learnt 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 learnt 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. If you choose **any** as the start point, then the filter acts upon all addresses between 224.0.0.0 and the address entered as the end point. If you enter **any** as the end point, then the filter acts upon all addresses between the address entered as the start point 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.

## 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.

When you configure a static MRouter on a VLAN, it replaces any dynamic MRouters learnt through IGMP Snooping.

## IGMP Snooping Configuration Example

This section provides steps to configure IGMP Snooping on the GbESM, using the Command-Line Interface (CLI).

### Configure IGMP Snooping

1. Configure port and VLAN membership on the switch.
2. Add VLANs to IGMP Snooping and enable the feature.

```
>> /cfg/13/igmp/snoop                (Select IGMP Snoop menu)
>> IGMP Snoop# add 1                 (Add VLAN 1 to IGMP snooping)
>> IGMP Snoop# ena                   (Enable IGMP Snooping)
```

3. View dynamic IGMP information.

```
>> /info/13/igmp                    (Select IGMP information menu)
>> IGMP Multicast Group# dump       (Show IGMP Group information)

>> IGMP Multicast# dump
Note: Local groups (224.0.0.x) are not snooped and will not appear.
  Group          VLAN    Port    Version    Expires
  -----
  238.1.0.0      1      EXT1    V2         4:17
  238.1.0.4      1      EXT1    V2         4:18

>> /info/13/igmp/mrouter           (Select MRouter information menu)
>> Mrouter# dump                   (Show IGMP Group information)

>> IGMP Multicast Router# dump
VLAN      Port          Version    Expires    Max Query Resp. Time
-----
  1        EXT4          V2         static     unknown
```

These commands display information about IGMP Groups and MRouters learnt through IGMP Snooping.

## Configure IGMP Filtering

### 1. Enable IGMP Filtering on the switch.

```
>> //cfg/13/igmp/igmpflt (Select IGMP Filtering menu)
>> IGMP Filter# ena (Enable IGMP Filtering)
Current status: disabled
New status: enabled
```

### 2. Define an IGMP filter.

```
>> //cfg/13/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: any
Enter new multicast address2 or any: 226.0.0.0 (Enter second IP address)
Current multicast address1: any
New pending multicast address1: 224.0.1.0
Current multicast address2: any
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/13/igmp/igmpflt (Select IGMP Filtering menu)
>>IGMP Filter# port EXT6 (Select port EXT6)
>>IGMP Port EXT6# filt ena (Enable IGMP Filtering on the port)
Current port EXT6 filtering: disabled
New port EXT6 filtering: enabled
>>IGMP Port EXT6# add 1 (Add IGMP Filter 1 to the port)
>>IGMP Port EXT6# apply (Make your changes active)
```

## Configure a Static Multicast Router (CLI example)

### 1. Configure a port to which the static Multicast Router is connected, and enter the appropriate VLAN.

```
>> //cfg/13/igmp/mrouter (Select IGMP MRouter menu)
>> Static Multicast Router# add EXT4 (Add port EXT4 as Static MRouter port)
Enter VLAN number: (1-4095) 1 (Enter the VLAN number)
Enter the version number of mrouter [1|2]: 2 (Enter the IGMP version number)
```



## 2. Apply, verify, and save the configuration.

>> Static Multicast Router# <b>apply</b>	<i>(Apply the configuration)</i>
>> Static Multicast Router# <b>cur</b>	<i>(View the configuration)</i>
>> Static Multicast Router# <b>save</b>	<i>(Save the configuration)</i>





## CHAPTER 11

# 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 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 156](#)
- [“Forming BGP Peer Routers” on page 157](#)
- [“What is a Route Map?” on page 157](#)
- [“Aggregating Routes” on page 161](#)
- [“Redistributing Routes” on page 161](#)
- [“BGP Attributes” on page 162](#)
- [“Selecting Route Paths in BGP” on page 163](#)
- [“BGP Failover Configuration” on page 164](#)
- [“Default Redistribution and Route Aggregation Example” on page 167](#)

## 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.

**NOTE** – The iBGP peers must be part of a fully meshed network, as shown in Figure 11-1.

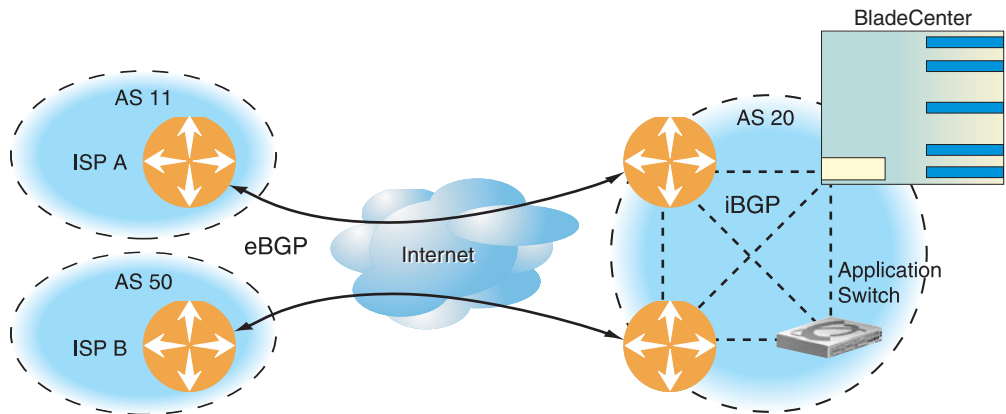


Figure 11-1 iBGP and eBGP

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?

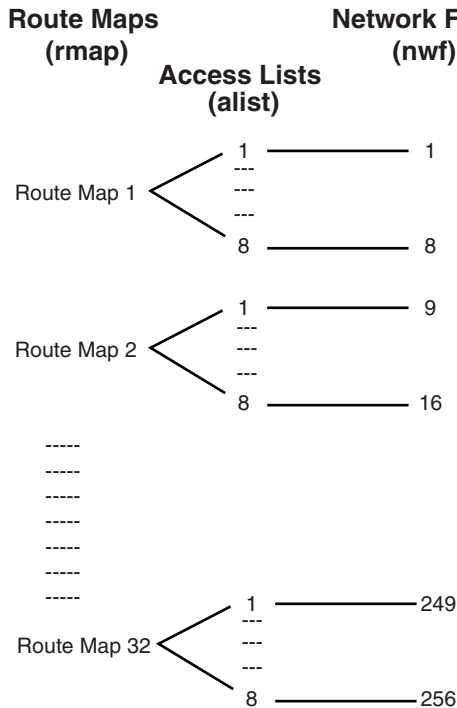
---

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 164.

Alteon OS 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 11-2 illustrates the relationship between route maps, access lists and network filters.



**Figure 11-2** 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.

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 is configured in the outgoing route map list, 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/13/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/13/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/13/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/13/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/13/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# met                 (Specify the metric)
```

### 5. Enable the route map.

```
>> # cfg/13/rmap 1/en                   (Enable the route map)
```

### 6. Assign the route map to a peer router.

Select the peer router and then add the route map to the incoming route map list,

```
>> # cfg/13/bgp/peer 1/addi           (Add to the incoming route map)
```

or to the outgoing route map list.

```
>> # cfg/13/bgp/peer 1/addo           (Add to the outgoing route map)
```



## 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.

When a subnet is redistributed from an Interior Gateway Protocol (IGP) into BGP, only the network route is injected into the BGP table. By default, this automatic summarization is disabled. To define the route to aggregate, use the following commands:

>> # <b>cfg/13/bgp</b>	<i>(Specify BGP)</i>
>> Border Gateway Protocol# <b>aggr 1</b>	<i>(Specify aggregate list number)</i>
>> BGP aggr 1 # <b>addr</b>	<i>(Enter aggregation network address)</i>
>> BGP aggr 1 # <b>mask</b>	<i>(Enter aggregation network mask)</i>
>> BGP aggr 1 # <b>ena</b>	<i>(Enable aggregation)</i>

An example of creating a BGP aggregate route is shown in [“Default Redistribution and Route Aggregation Example”](#) on page 167.

## Redistributing Routes

---

In addition to running multiple routing protocols simultaneously, Alteon OS software can redistribute information from one routing protocol to another. For example, you can instruct the switch to use BGP to readvertise 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 157. Redistributing routes is another way of providing policy control over whether to export OSPF routes, fixed routes, static routes, and virtual IP address routes. For an example configuration, see [“Default Redistribution and Route Aggregation Example”](#) on page 167.

Default routes can be configured using the following methods:

- Import
- Originate—The router sends a default route to peers even though 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 metrics 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 GbE Switch Module 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 x AS path length (number of autonomous systems transversed).

- 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.

---

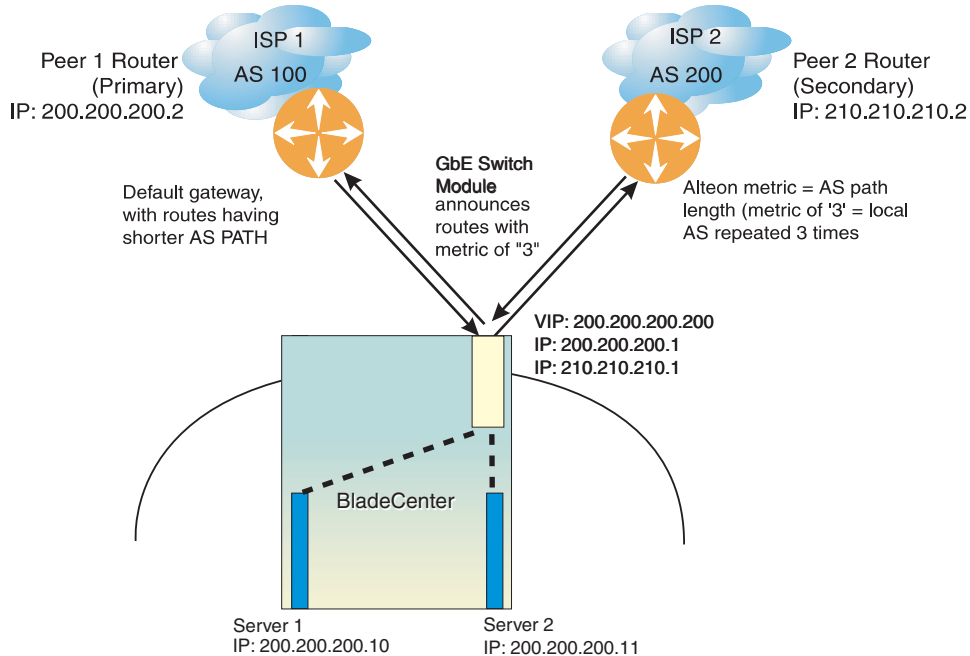
- 5. The lower cost to the next hop of routes is selected.**
- 6. In the case of equal cost, the eBGP route is preferred over iBGP.**
- 7. If all routes are from eBGP, the route with the lower router ID is selected.**

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 GbE Switch Module 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 11-3](#), 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 GbE Switch Module.



**Figure 11-3** BGP Failover Configuration Example

On the GbE Switch Module, 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/l2/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 will need to be placed in the appropriate VLAN. These interfaces will be used as the primary and secondary default gateways for the switch.

```
>> IP# metric strict                (Set metric for default gateway)
>> IP# if 1                          (Select default gateway 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# ../ip/if 2        (Select default gateway 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 disabled by default and is used for VLAN-to-VLAN (non-BGP) routing. Make sure IP forwarding is enabled 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.

Peer 1 is the primary gateway router. Peer 2 is configured with a metric of “3.” The `metric` option is key to ensuring gateway traffic is directed to Peer 1, as it will make Peer 2 appear to be three router hops away from the switch. Thus, the switch should never use it unless Peer 1 goes down.

```

>> # /cfg/13/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# if 200.200.200.1       (Set IP interface for peer router 1)
>> BGP Peer 1# ras 100                (Set remote AS number)
>> BGP Peer 1# /cfg/13/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# if 210.210.210.1     (Set IP interface for peer router 2)
>> BGP Peer 2# ras 200                (Set remote AS number)
>> BGP Peer 2# metric 3              (Set AS path length to 3 router hops)

```

The `metric` command in the peer menu tells the GbE Switch Module to create an AS path of “3” when advertising via BGP.

#### 5. On the switch, apply and save your configuration changes.

```

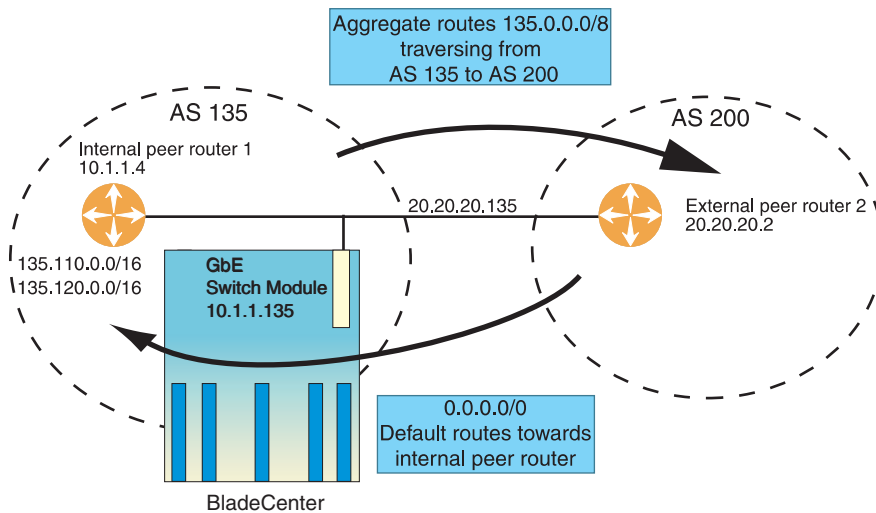
>> BGP Peer 2# apply                  (Make your changes active)
>> save                                (Save for restore after reboot)

```

## 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 11-4](#), you have two peer routers: an internal and an external peer router. Configure the GbE Switch Module 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 11-4** 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 0.0.0.1    (Specify a router ID number)
```

### 3. Configure internal peer router 1 and external peer router 2.

```

>> # /cfg/13/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/13/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/13/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 12

# OSPF

---

Alteon OS supports the Open Shortest Path First (OSPF) routing protocol. The Alteon OS implementation conforms to the OSPF version 2 specifications detailed in Internet RFC 1583. The following sections discuss OSPF support for the GbE Switch Module:

- [“OSPF Overview” on page 170](#). 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 Alteon OS” on page 175](#). This section describes how OSPF is implemented in Alteon OS, such as configuration parameters, electing the designated router, summarizing routes, defining route maps and so forth.
- [“OSPF Configuration Examples” on page 185](#). 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.

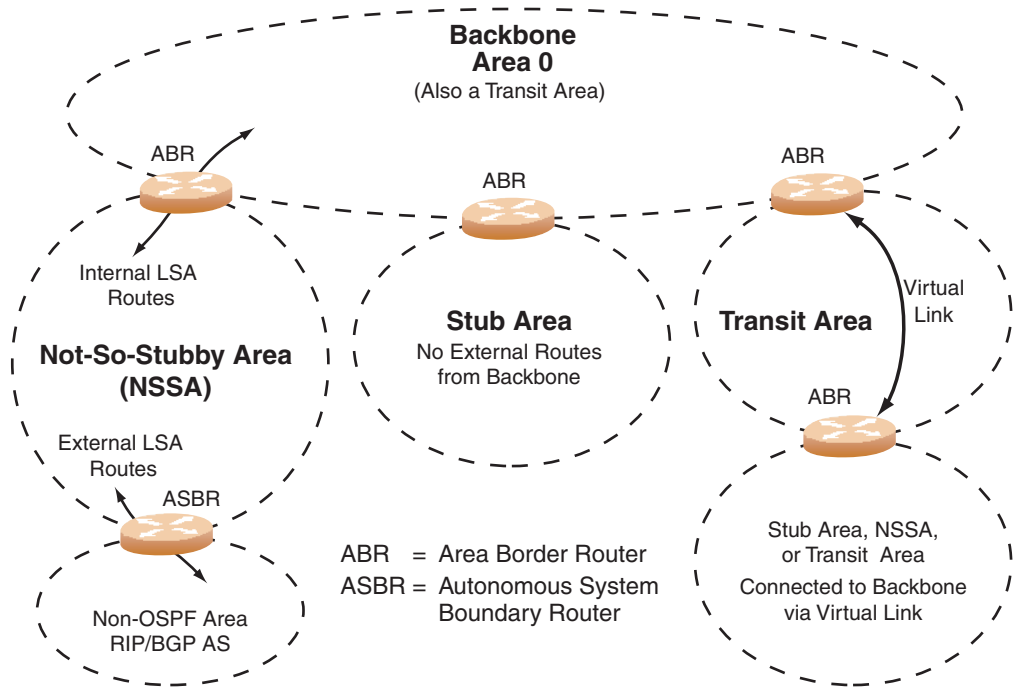
## Types of OSPF Areas

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 12-1](#), 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.

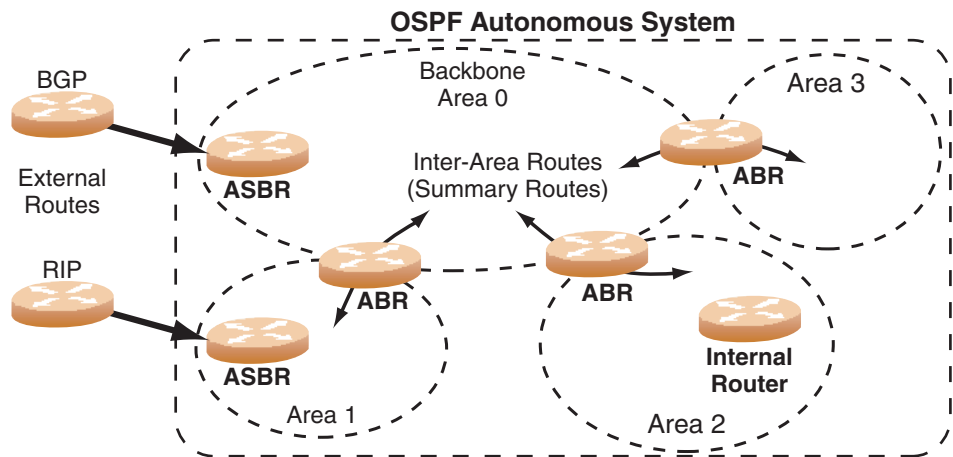


**Figure 12-1** OSPF Area Types

## Types of OSPF Routing Devices

As shown in [Figure 12-2](#), 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 12-2** 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 interfaces. LSAs are entered into the LSDB of each routing device. OSPF uses *flooding* to distribute LSAs between routing devices.

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 Alteon OS

---

Alteon OS supports a single instance of OSPF and up to 4 K routes on the network. The following sections describe OSPF implementation in Alteon OS:

- “Configurable Parameters” on page 175
- “Defining Areas” on page 176
- “Interface Cost” on page 178
- “Electing the Designated Router and Backup” on page 178
- “Summarizing Routes” on page 178
- “Default Routes” on page 179
- “Virtual Links” on page 180
- “Router ID” on page 181
- “Authentication” on page 181

## Configurable Parameters

In Alteon OS, OSPF parameters can be configured through the Command Line Interface (CLI), Browser-Based Interface (BBI) for GbE Switch Modules, or through SNMP. For more information, see [Chapter 1, “Accessing the Switch.”](#)

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.

## 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 180).

Up to three OSPF areas can be connected to the GbE Switch Module with Alteon OS 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 GbE Switch Module. 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 GbE Switch Module, 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 GbE Switch Module 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
                                                    network)
>> 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 network on interface
                                                    14)
>> 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/l3/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 GbE Switch Module. The command is as follows:

```
>> # /cfg/l3/ospf/if <OSPF interface number>/prio <priority value (0-127)>
```

A priority value of 127 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 lowest router ID wins.

## 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:

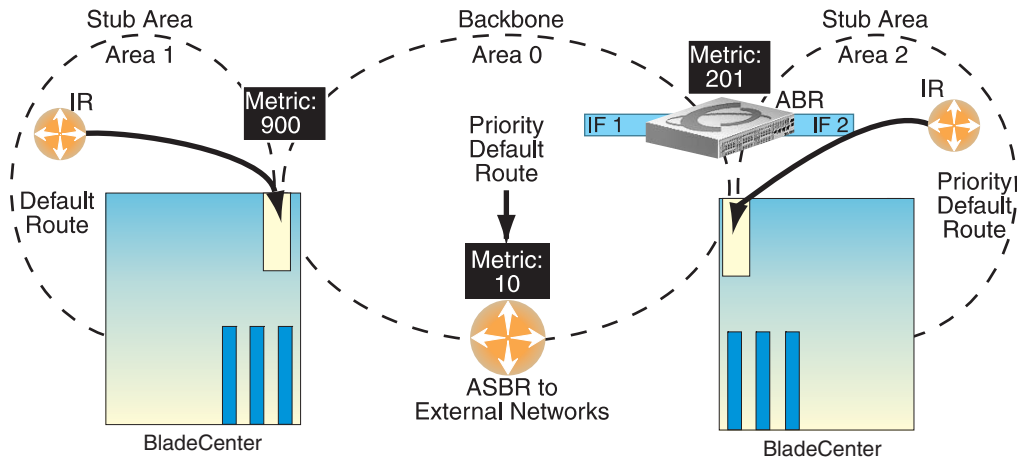
```
>> # /cfg/l3/ospf/range <range number>/addr <IP address>/mask
<mask>
```

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 192](#).

## 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 GbE Switch Module 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 12-3](#)), 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 12-3** Injecting Default Routes

In more complex OSPF areas with multiple ABRs or ASBRs (such as area 0 and area 2 in [Figure 12-3](#)), there are multiple routes leading from the area. In such areas, traffic for unrecognized destinations cannot tell which route leads upstream without further configuration.

To resolve the situation and select one default route among multiple choices in an area, you can manually configure a metric value on each ABR. The metric assigns a priority to the ABR for its selection as the priority default route in an area. The following command is used for setting the metric value:

```
>> # /cfg/l3/ospf/default <metric value> <metric type (1 or 2)>
```

where *<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.

To clear a default route metric from the switch, use the following command:

```
>> # /cfg/l3/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 12-1 on page 171](#)).

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/l3/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 GbE Switch Module as one endpoint of a virtual link, use the following command:

```
>> # /cfg/l3/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 GbE Switch Module 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 188](#).

## Router ID

Routing devices in OSPF areas are identified by a router ID. The router ID is expressed in IP address format. The IP address of the router ID is not required to be included in any IP interface range or in any OSPF area.

The router ID can be configured in one of the following two ways:

- Dynamically—OSPF protocol configures the lowest IP interface IP address as the router ID. This is the default.
- Statically—Use the following command to manually configure the router ID:

```
>> # /cfg/l3/rtrid <IP address>
```

- To modify the router ID from static to dynamic, set the router ID to 0.0.0.0, save the configuration, and reboot the GbE Switch Module. To view the router ID, enter:

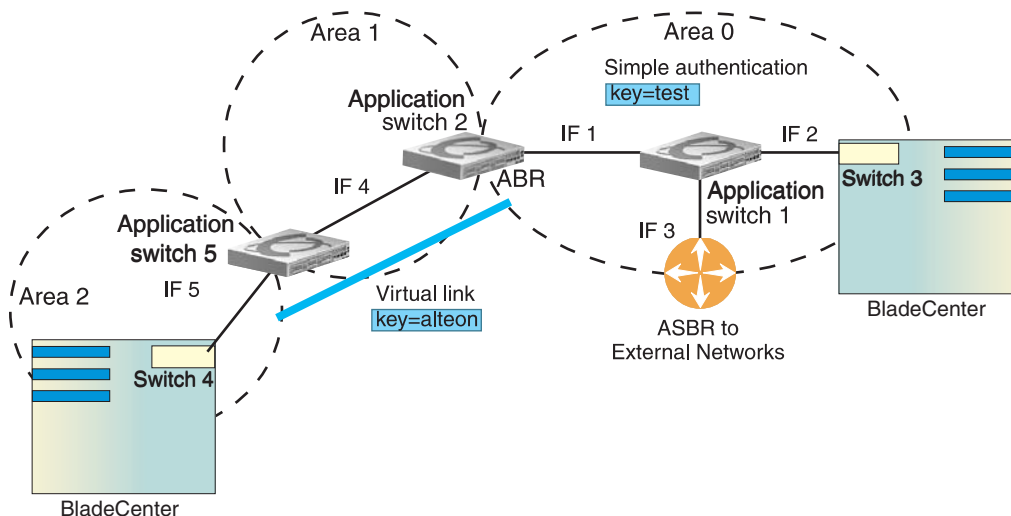
```
>> # /info/l3/ospf/gen
```

## 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 predefined passwords. Alteon OS 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 12-4 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 12-4** OSPF Authentication

To configure simple plain text OSPF passwords on the switches shown in Figure 12-4 use the following commands:

1. **Enable OSPF authentication for Area 0 on switches 1, 2, and 3.**

```
>> # /cfg/l3/ospf/aindex 0/auth password
                                     (Turn on OSPF 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/l3/ospf/if 1
>> OSPF Interface 1 # key test
>> OSPF Interface 1 # ../if 2
>> OSPF Interface 2 # key test
>> OSPF Interface 1 # ../if 3
>> OSPF Interface 3 # key test
```

**3. Enable OSPF authentication for Area 2 on switch 4.**

```
>> # /cfg/l3/ospf/aindex 2/auth password
                                     (Turn on OSPF 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/l3/ospf/virt 1/key alteon
```

Use the following commands to configure MD5 authentication on the switches shown in [Figure 12-4](#):

**1. Enable OSPF MD5 authentication for Area 0 on switches 1, 2, and 3.**

```
>> # /cfg/l3/ospf/aindex 0/auth md5   (Turn on MD5 authentication)
```

**2. Configure MD5 key ID for Area 0 on switches 1, 2, and 3.**

```
>> # /cfg/l3/ospf/md5key 1/key test
```

**3. Assign MD5 key ID to OSPF interfaces on switches 1, 2, and 3.**

```
>> # /cfg/l3/ospf/if 1
>> OSPF Interface 1 # mdkey 1
>> OSPF Interface 1 # ../if 2
>> OSPF Interface 2 # mdkey 1
>> OSPF Interface 1 # ../if 3
>> OSPF Interface 3 # mdkey 1
```

**4. Enable OSPF MD5 authentication for Area 2 on switch 4.**

```
>> # /cfg/l3/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/l3/ospf/md5key 2/key alteon
```

## 6. Assign MD5 key ID to OSPF virtual link on switches 2 and 4.

```
>> # /cfg/l3/ospf/virt 1/mdkey 2
```

## Host Routes for Load Balancing

Alteon OS 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.



## 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, and ATM)

## OSPF Configuration Examples

---

A summary of the basic steps for configuring OSPF on the GbE Switch Module 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.

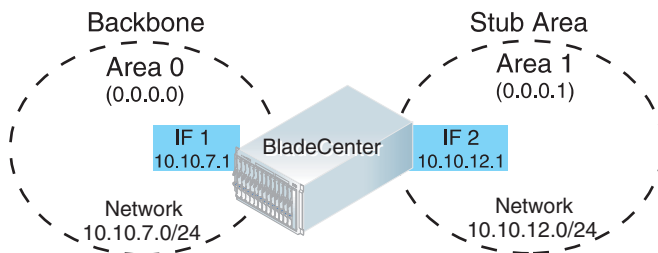
**6. (Optional) Configure route summarization between OSPF areas.**

**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 12-5** A Simple OSPF Domain

Follow this procedure to configure OSPF support as shown in [Figure 12-5](#):

### 1. Configure IP interfaces on each network that will be attached to OSPF areas.

In this example, two IP interfaces are needed: one for the backbone network on 10.10.7.0/24 and one for the stub 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 network)
>> IP Interface 1 # mask 255.255.255.0 (Set IP mask on backbone network)
>> 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 network)
>> IP Interface 2 # mask 255.255.255.0 (Set IP mask on stub area network)
>> IP Interface 2 # enable (Enable IP interface 2)
```

### 2. Enable OSPF.

```
>> IP Interface 2 # /cfg/13/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 IP 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 IP 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                (Global command to apply all changes)
>> OSPF Interface 2 # save                (Global command to save all changes)
```

## Example 2: Virtual Links

In the example shown in Figure 12-6, 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.

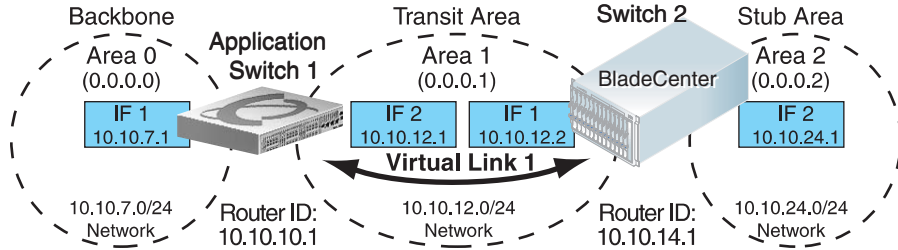


Figure 12-6 Configuring a Virtual Link

### 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 on Switch #1: one for the backbone network on 10.10.7.0/24 and one for the transit 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 network)
>> IP Interface 1 # mask 255.255.255.0              (Set IP mask on backbone network)
>> 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 network)
>> IP Interface 2 # mask 255.255.255.0              (Set IP mask on transit area network)
>> 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/l3/rtrid 10.10.10.1        (Set static router ID on switch 1)
```

#### 3. Enable OSPF.

```
>> IP # /cfg/l3/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 the 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 IP 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 IP interface 2)
>> OSPF Interface 2 # aindex 1           (Attach network to transit area index)
>> 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 190](#).

```
>> OSPF Interface 2 # ../virt 1          (Specify a virtual link number)
>> OSPF Virtual Link 1 # aindex 1        (Specify the transit area for the virtual link)
>> OSPF Virtual Link 1 # nbr 10.10.14.1  (Specify 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             (Global command to apply all changes)
>> OSPF Interface 2 # save             (Global command to save all changes)
```

## Configuring OSPF for a Virtual Link on Switch #2

### 1. Configure IP interfaces on each network that will be attached to OSPF areas.

Two IP interfaces are needed on Switch #2: one for the transit area network on 10.10.12.0/24 and one for the stub area network on 10.10.24.0/24.

```
>> # /cfg/13/if 1 (Select menu for IP interface 1)
>> IP Interface 1 # addr 10.10.12.2 (Set IP address on transit area network)
>> IP Interface 1 # mask 255.255.255.0 (Set IP mask on transit area network)
>> 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 network)
>> IP Interface 2 # mask 255.255.255.0 (Set IP mask on stub area network)
>> 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 189](#).

```
>> IP Interface 2 # /cfg/13/rtrid 10.10.14.1 (Set static router ID on switch 2)
```

### 3. Enable OSPF.

```
>> IP# /cfg/13/ospf/on (Enable OSPF on switch 2)
```

### 4. Define the backbone.

This version of Alteon OS 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 IP interface 1)
>> OSPF Interface 1 # aindex 1           (Attach network to transit area index)
>> 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 IP 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 188](#).

```
>> OSPF Interface 2 # ../virt 1           (Specify a virtual link number)
>> OSPF Virtual Link 1 # aindex 1         (Specify the transit area for the virtual link)
>> OSPF Virtual Link 1 # nbr 10.10.10.1   (Specify 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               (Global command to apply all changes)
>> OSPF Interface 2 # save                (Global command to save all changes)
```

## Other Virtual Link Options

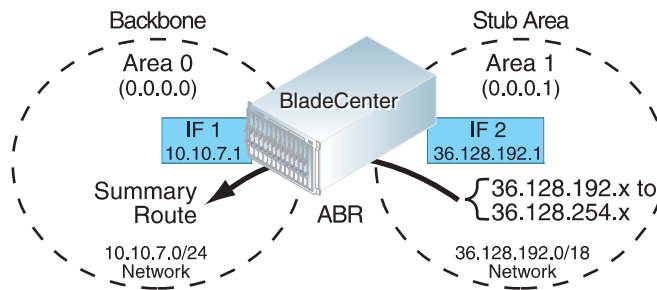
- You can use redundant paths by configuring multiple virtual links.
- Only the endpoints of the virtual link are configured. The virtual link path may traverse multiple routers in an area as long as there is a routable path between the endpoints.

### 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 12-7** 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 12-7](#):

### 1. Configure IP interfaces for each network which will be attached to OSPF areas.

```
>> # /cfg/l3/if 1 (Select menu for IP interface 1)
>> IP Interface 1 # addr 10.10.7.1 (Set IP address on backbone network)
>> IP Interface 1 # mask 255.255.255.0 (Set IP mask on backbone network)
>> 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 network)
>> IP Interface 2 # mask 255.255.192.0 (Set IP mask on stub area network)
>> IP Interface 2 # ena (Enable IP interface 2)
```

### 2. Enable OSPF.

```
>> IP Interface 2 # /cfg/l3/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 IP 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 IP 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 summary range)
>> OSPF Summary Range 1 # mask 255.255.192.0 (Set mask address for summary range)
>> OSPF Summary Range 1 # aindex 0         (Inject summary route into 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           (Global command to apply all changes)
>> OSPF Summary Range 2 # save           (Global command to save all changes)
```

## Verifying OSPF Configuration

Use the following commands to verify the OSPF configuration on your switch:

- /info/l3/ospf/general
- /info/l3/ospf/nbr
- /info/l3/ospf/dbase/dbsum
- /info/l3/ospf/route
- /stats/l3/route

Refer to the *Alteon OS Command Reference* for information on the above commands.

# 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.

- High Availability





## CHAPTER 13

# High Availability

---

GbE Switch Modules 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 198](#). This section discusses trunk failover without using VRRP.
- [“VRRP Overview” on page 204](#). This section discusses VRRP operation and Alteon OS redundancy configurations.
- [“Failover Methods” on page 207](#). This section describes the three modes of high availability.
- [“Alteon OS extensions to VRRP” on page 210](#). This section describes VRRP enhancements implemented in Alteon OS.
- [“Virtual Router Deployment Considerations” on page 211](#). This section describes issues to consider when deploying virtual routers.
- [“High Availability Configurations” on page 213](#). This section discusses the more useful and easily deployed redundant configurations.
  - [“Active-Active Configuration” on page 213](#)
  - [“Hot-Standby Configuration” on page 218](#)

## Layer 2 Failover

---

The primary application for Layer 2 Failover is to support Network Adapter Teaming. With Network Adapter Teaming, the NICs on each server all 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 “Configuring Teaming” in the Broadcom NetXtreme™ Gigabit Ethernet Adapter User Guide (<http://www.ibm.com/pc/support/site.wss/MIGR-43152.html>).

---

**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.

---

Layer 2 Failover can be enabled on any trunk group in the GbE Switch Module, including LACP trunks. Trunks can be added to failover trigger groups such that if some (or all) of the links fail in a trigger, the switch disables all 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 takes up to five seconds.

## VLAN Monitor

The VLAN Monitor allows L2 Failover to discern different VLANs. With VLAN Monitor turned on:

- If enough links in a trigger go down (see “Setting the Failover Limit” on page 199), 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/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.

## 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 four (`/cfg/l2/failovr/trigger x/limit 4`), a failover event occurs when the number of operational links in the trigger is four or fewer. When you set the limit to zero, the switch triggers a failover event only when no links in the trigger are operational.

## 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 LACP trunks to a failover trigger. When you add an *admin key* to a trigger (`/cfg/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

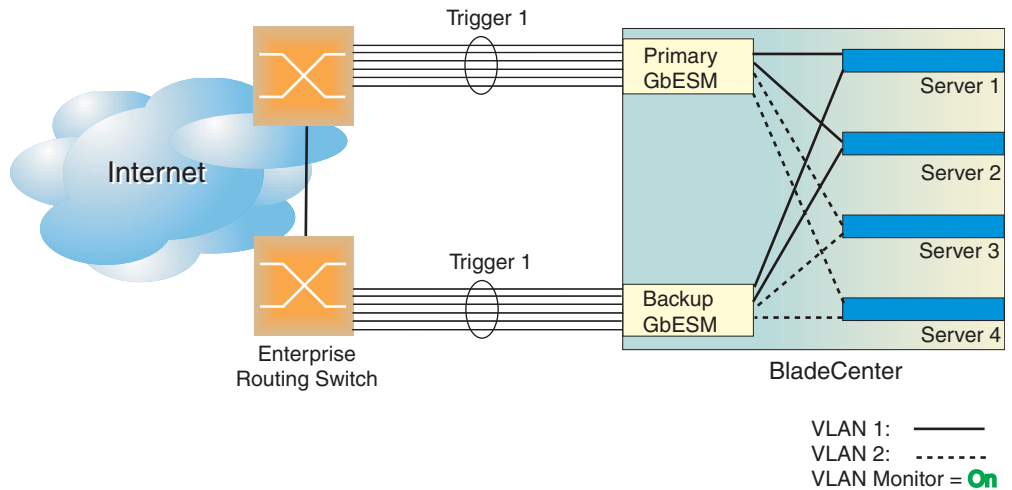
This section provides important information about configuring L2 Failover:

- A failover trigger can monitor multiple static trunks or a single LACP key, but not both.
- With VLAN Monitor on, the following additional guidelines apply:
  - All external ports in all trunks that are added to a single failover trigger must have the same VLAN membership and have the same PVID.
  - Each failover trigger must operate on a different VLAN membership.
  - Multiple failover triggers cannot operate on the same internal port.

- For each port in each trunk in a failover trigger, the trigger monitors the STP state only on the default PVID.

## L2 Failover Configurations

Figure 13-1 is a simple example of Layer 2 Failover. One GbE Switch Module 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.

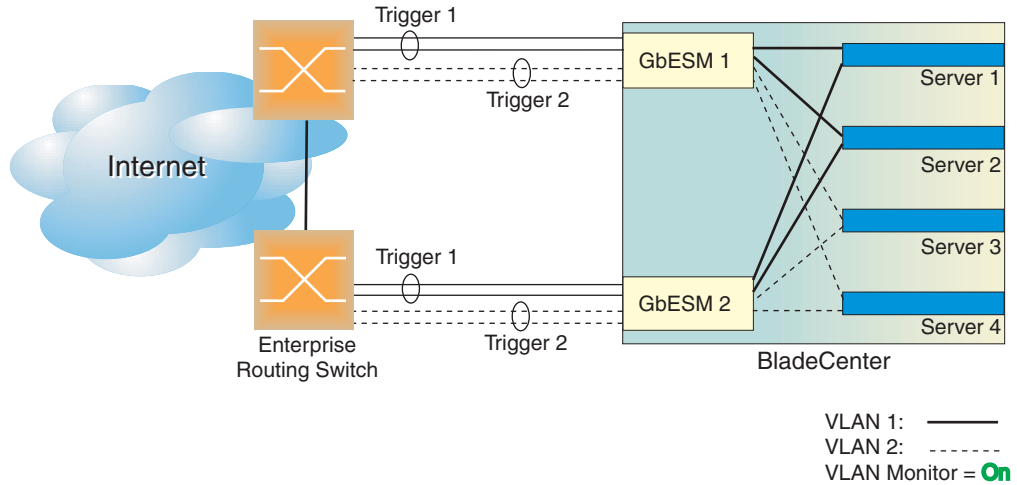


**Figure 13-1** Basic Layer 2 Trunk Failover



Figure 13-2 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 13-2** Two trunks, each in a different Failover Trigger

Figure 13-3 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.

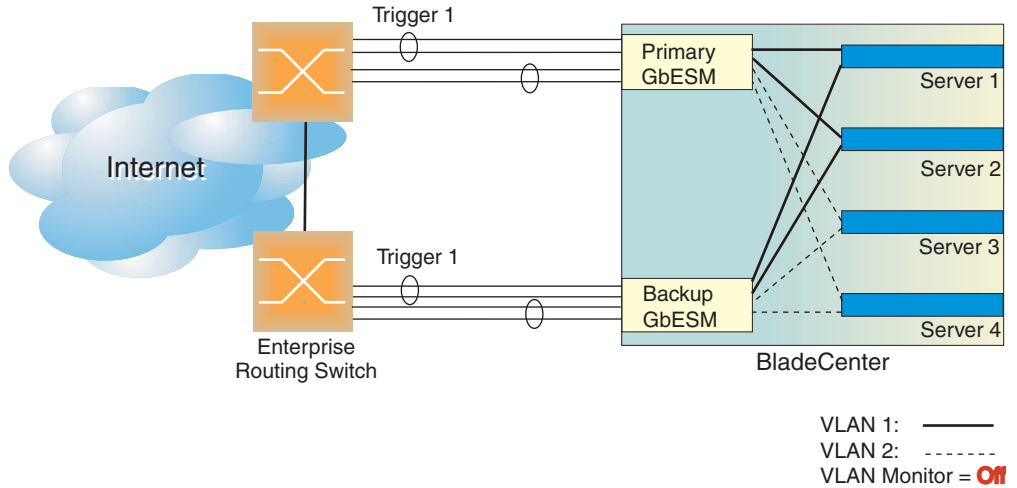


Figure 13-3 Two trunks, one Failover Trigger

## Configuring Trunk Failover

The following procedure pertains to example 1, as shown in [Figure 13-1](#).

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# add EXT4          (Add port EXT4 to trunk group 1)
>> Trunk group 1# add EXT5          (Add port EXT5 to trunk group 1)
>> Trunk group 1# add EXT6          (Add port EXT6 to trunk group 1)
>> Trunk group 1# ena              (Enable trunk group 1)
```

3. **Configure Failover parameters.**

```
>> # /cfg/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 your new configuration changes.**

```
>> Auto Monitor# 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 206](#) 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 pre-emption, 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.

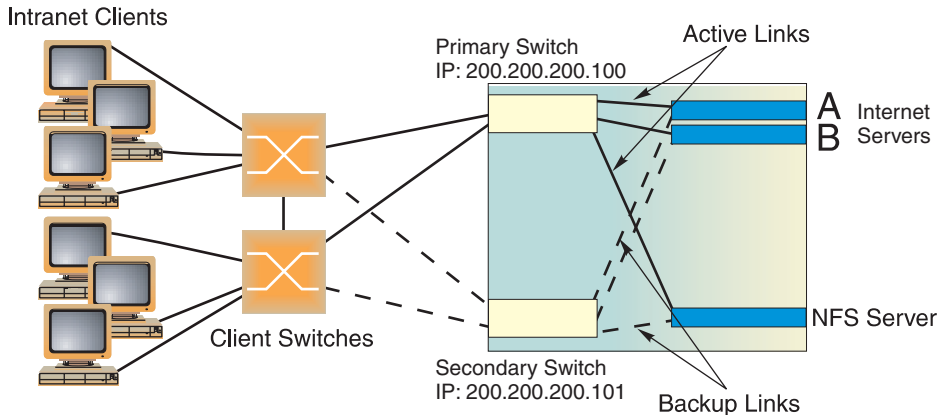
---

**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 13-4** 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.

Alteon OS high availability configurations are based on VRRP. The Alteon OS implementation of VRRP includes proprietary extensions.

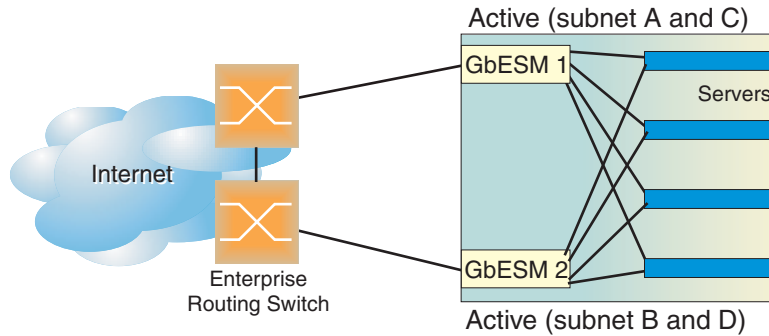
The Alteon OS implementation of VRRP supports the following modes of high availability:

- **Active-Active**—based on proprietary Alteon OS extensions to VRRP
- **Hot-Standby**—supports Network Adapter Teaming on your server blades

## Active-Active Redundancy

In an active-active configuration, shown in [Figure 13-5](#), two switches provide redundancy for each other, with both active at the same time. Each switch processes traffic on a different subnet. When a failure occurs, the remaining switch can process traffic on all subnets.

For a configuration example, see [“Active-Active Configuration”](#) on [page 213](#).



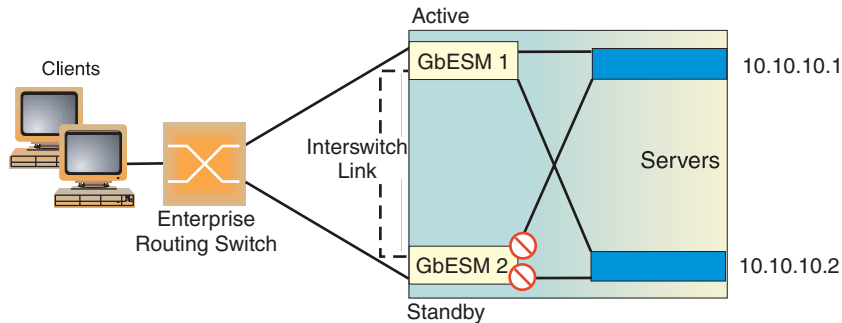
**Figure 13-5** Active-Active Redundancy



## Hot-Standby Redundancy

The primary application for VRRP-based hot-standby is to support Server Load Balancing when you have configured 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 “Configuring Teaming” in the Broadcom NetXtreme™ Gigabit Ethernet Adapter User Guide (<http://www.ibm.com/pc/support/site.wss/MIGR-43152.html>).

The hot-standby model is shown in Figure 13-6.



**Figure 13-6** Hot-Standby Redundancy

## 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.

## Alteon OS extensions to VRRP

This section describes the following VRRP enhancements that are implemented in Alteon OS:

- [Tracking VRRP Router Priority](#)

### Tracking VRRP Router Priority

Alteon OS 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.

Alteon OS can track the attributes listed in [Table 13-1](#):

**Table 13-1** 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. <b>Note:</b> 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/vr</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 211](#) for an example on how to configure the switch for tracking VRRP priority.

# Virtual Router Deployment Considerations

---

Review the following issues described in this section to prevent network problems when deploying virtual routers:

- [Assigning VRRP Virtual Router ID](#)
- [Configuring the Switch for Tracking](#)

## 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 #/vrid`) must be assigned. The virtual router ID may be configured as any number between 1 and 255.

## Configuring the Switch for Tracking

Tracking configuration largely depends on user preferences and network environment. Consider the configuration shown in [Figure 13-5 on page 208](#). 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.

---

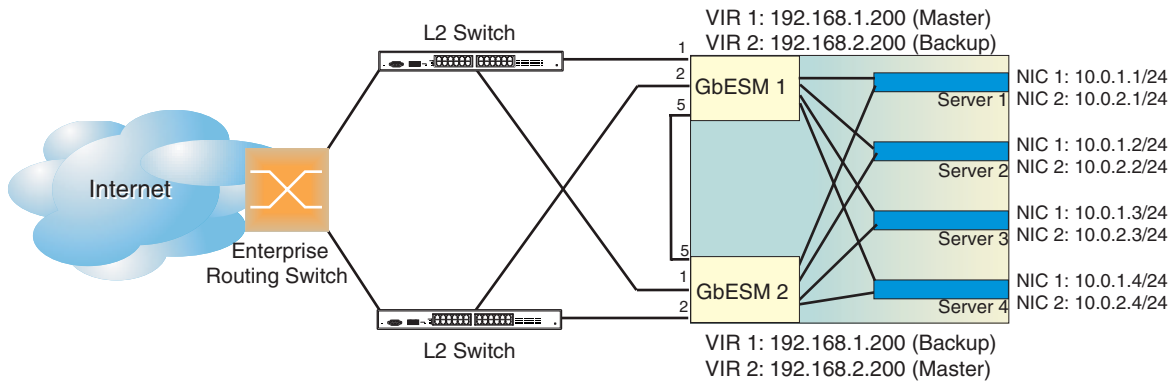
# High Availability Configurations

GbE Switch Modules offer flexibility in implementing redundant configurations. This section discusses the more useful and easily deployed configurations:

- “Active-Active Configuration” on page 213
- “Hot-Standby Configuration” on page 218

## Active-Active Configuration

Figure 13-7 shows an example configuration where two GbE Switch Modules are used as VRRP routers in an active-active configuration. In this configuration, both switches respond to packets.



**Figure 13-7** 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 13-7, 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.

<code>/cfg/l3/if 1</code>	<i>(Select interface 1)</i>
<code>&gt;&gt; IP Interface 1# <b>addr</b> 192.168.1.100</code>	<i>(Define IP address for interface 1)</i>
<code>&gt;&gt; IP Interface 1# <b>vlan</b> 10</code>	<i>(Assign VLAN 10 to interface 1)</i>
<code>&gt;&gt; IP Interface 1# <b>ena</b></code>	<i>(Enable interface 1)</i>
<code>&gt;&gt; IP Interface 1# ..</code>	
<code>&gt;&gt; Layer 3# <b>if</b> 2</code>	<i>(Select interface 2)</i>
<code>&gt;&gt; IP Interface 2# <b>addr</b> 192.168.2.101</code>	<i>(Define IP address for interface 2)</i>
<code>&gt;&gt; IP Interface 1# <b>vlan</b> 20</code>	<i>(Assign VLAN 20 to interface 2)</i>
<code>&gt;&gt; IP Interface 2# <b>ena</b></code>	<i>(Enable interface 2)</i>
<code>&gt;&gt; IP Interface 2# ..</code>	
<code>&gt;&gt; Layer 3# <b>if</b> 3</code>	<i>(Select interface 3)</i>
<code>&gt;&gt; IP Interface 3# <b>addr</b> 10.0.1.100</code>	<i>(Define IP address for interface 3)</i>
<code>&gt;&gt; IP Interface 3# <b>mask</b> 255.255.255.0</code>	<i>(Define subnet mask for interface 3)</i>
<code>&gt;&gt; IP Interface 3# <b>ena</b></code>	<i>(Enable interface 3)</i>
<code>&gt;&gt; IP Interface 2# ..</code>	
<code>&gt;&gt; Layer 3# <b>if</b> 4</code>	<i>(Select interface 4)</i>
<code>&gt;&gt; IP Interface 4# <b>addr</b> 10.0.2.101</code>	<i>(Define IP address for interface 4)</i>
<code>&gt;&gt; IP Interface 4# <b>mask</b> 255.255.255.0</code>	<i>(Define subnet mask for interface 4)</i>
<code>&gt;&gt; IP Interface 4# <b>ena</b></code>	<i>(Enable interface 4)</i>

### 2. Configure the default gateways. Each default gateway points to one of the Layer 2 routers.

<code>/cfg/l3/gw 1</code>	<i>(Select default gateway 1)</i>
<code>&gt;&gt; Default gateway 1# <b>addr</b> 192.168.1.1</code>	<i>(Point gateway to the first L2 router)</i>
<code>&gt;&gt; Default gateway 1# <b>ena</b></code>	<i>(Enable the default gateway)</i>
<code>&gt;&gt; Default gateway 1# ..</code>	
<code>&gt;&gt; Layer 3# <b>gw</b> 2</code>	<i>(Select default gateway 2)</i>
<code>&gt;&gt; Default gateway 1# <b>addr</b> 192.168.2.1</code>	<i>(Point gateway to the second router)</i>
<code>&gt;&gt; Default gateway 1# <b>ena</b></code>	<i>(Enable the default gateway)</i>

### 3. Turn on VRRP and configure two Virtual Interface Routers.

```

/cfg/13/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/13/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/12/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/12/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.

<code>/cfg/l3/if 1</code>	<i>(Select interface 1)</i>
<code>&gt;&gt; IP Interface 1# addr 192.168.1.101</code>	<i>(Define IP address for interface 1)</i>
<code>&gt;&gt; IP Interface 1# vlan 10</code>	<i>(Assign VLAN 10 to interface 1)</i>
<code>&gt;&gt; IP Interface 1# ena</code>	<i>(Enable interface 1)</i>
<code>&gt;&gt; IP Interface 1# ..</code>	
<code>&gt;&gt; Layer 3# if 2</code>	<i>(Select interface 2)</i>
<code>&gt;&gt; IP Interface 2# addr 192.168.2.100</code>	<i>(Define IP address for interface 2)</i>
<code>&gt;&gt; IP Interface 1# vlan 20</code>	<i>(Assign VLAN 20 to interface 2)</i>
<code>&gt;&gt; IP Interface 2# ena</code>	<i>(Enable interface 2)</i>
<code>&gt;&gt; IP Interface 2# ..</code>	
<code>&gt;&gt; Layer 3# if 3</code>	<i>(Select interface 3)</i>
<code>&gt;&gt; IP Interface 3# addr 10.0.1.101</code>	<i>(Define IP address for interface 3)</i>
<code>&gt;&gt; IP Interface 3# mask 255.255.255.0</code>	<i>(Define subnet mask for interface 3)</i>
<code>&gt;&gt; IP Interface 3# ena</code>	<i>(Enable interface 3)</i>
<code>&gt;&gt; IP Interface 2# ..</code>	
<code>&gt;&gt; Layer 3# if 4</code>	<i>(Select interface 4)</i>
<code>&gt;&gt; IP Interface 4# addr 10.0.2.100</code>	<i>(Define IP address for interface 4)</i>
<code>&gt;&gt; IP Interface 4# mask 255.255.255.0</code>	<i>(Define subnet mask for interface 4)</i>
<code>&gt;&gt; IP Interface 4# ena</code>	<i>(Enable interface 4)</i>

### 2. Configure the default gateways. Each default gateway points to one of the Layer 2 routers.

<code>/cfg/l3/gw 1</code>	<i>(Select default gateway 1)</i>
<code>&gt;&gt; Default gateway 1# addr 192.168.2.1</code>	<i>(Point gateway to the first L2 router)</i>
<code>&gt;&gt; Default gateway 1# ena</code>	<i>(Enable the default gateway)</i>
<code>&gt;&gt; Default gateway 1# ..</code>	
<code>&gt;&gt; Layer 3# gw 2</code>	<i>(Select default gateway 2)</i>
<code>&gt;&gt; Default gateway 1# addr 192.168.1.1</code>	<i>(Point gateway to the second router)</i>
<code>&gt;&gt; Default gateway 1# ena</code>	<i>(Enable the default gateway)</i>



### 3. Turn on VRRP and configure two Virtual Interface Routers.

```

/cfg/13/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/13/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/12/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/12/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 “Configuring Teaming” in the Broadcom NetX-treme™ Gigabit Ethernet Adapter User Guide (<http://www-306.ibm.com/pc/support/site.wss/MIGR-43152.html>).

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 13-8 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).

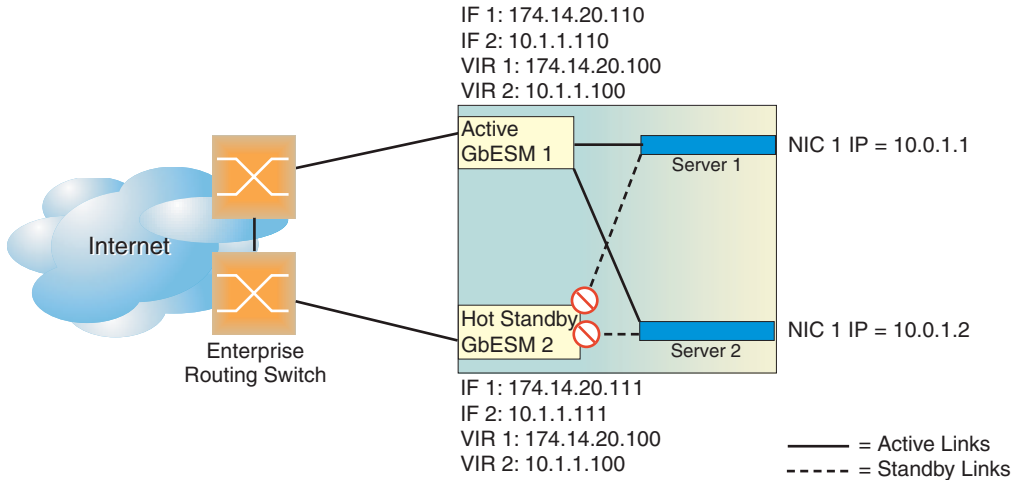


Figure 13-8 Hot-Standby Configuration

## 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/l3/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. Configure an external port Trunk Group.

```

/cfg/12/trunk 1
>> Trunk group 1# ena                               (Enable Trunk Group)
>> Trunk group 1# add EXT1                          (Add port EXT1 to Trunk Group)
>> Trunk group 1# add EXT2                          (Add port EXT2 to Trunk Group)
>> Trunk group 1# add EXT3                          (Add port EXT3 to Trunk Group)

```

## 6. 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 1, configure the interfaces for clients (174.14.20.111) and servers (10.1.1.111).

```

/cfg/l3/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/l3/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/l3/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/l3/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. Configure an external port Trunk Group.

```
/cfg/l2/trunk 1
>> Trunk group 1# ena                               (Enable Trunk Group)
>> Trunk group 1# add EXT1                           (Add port EXT1 to Trunk Group)
>> Trunk group 1# add EXT2                           (Add port EXT2 to Trunk Group)
>> Trunk group 1# add EXT3                           (Add port EXT3 to Trunk Group)
```

## 6. Turn off Spanning Tree Protocol globally. Apply and save changes.

```
/cfg/l2/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:

- Troubleshooting
- RADIUS Server Configuration Notes







## APPENDIX A

# Troubleshooting

---

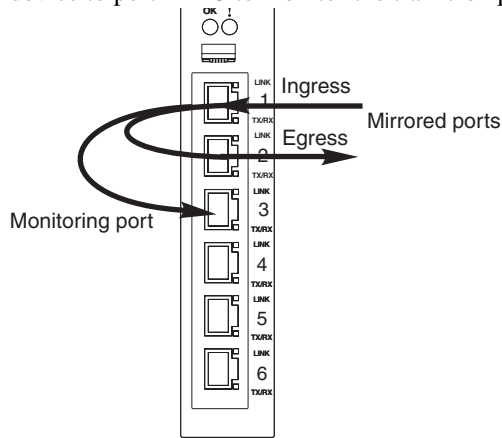
This section discusses some tools to help you troubleshoot common problems on the GbE Switch Module:

- [“Monitoring Ports” on page 226](#)

## Monitoring Ports

The port mirroring feature in the Alteon OS 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. Alteon OS 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 A-1](#), 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 A-1** Monitoring Ports

[Figure A-1](#) shows two mirrored ports monitored by a single port. Similarly, you can have a single or groups of:

- one mirrored port to one monitored port
- more than two mirrored ports to one monitored port

Alteon OS does not support a single port being monitored by multiple ports.

Ingress and egress traffic is duplicated and sent to the monitor port after processing.

---

**NOTE** – The GbESM cannot mirror LACPDU packets.

---



---

**NOTE** – Traffic on VLAN 4095 is not mirrored to the external ports.

---

## Port Mirroring behavior

This section describes the composition of monitored packets in the GbE Switch Module, based on the configuration of the ports.

If a tagged port's PVID is the same as its VLAN ID, then the egress traffic on that port is untagged. Therefore, the following port-mirroring cases are differentiated, as described below:

- Tagged or untagged frames from mirrored ports whose VLAN ID is the same as the monitoring port's PVID: mirrored frames viewed at the monitoring port are untagged
- Tagged or untagged frames from mirrored ports whose VLAN ID is different than the monitoring port's PVID: mirrored frames viewed at the monitoring port are tagged with the VLAN ID of ingress/egress port for ingress/egress mirroring respectively.

## Layer 2 Port Mirroring

In this scenario, you observe Layer 2 port mirroring.

- Same VLAN as Mirrored Port
  - Both ports tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port tagged, Monitoring port untagged: Untagged packet
  - Mirrored port untagged, Monitoring port tagged: Tagged packet with VID of the VLAN to which both ports belong
  - Both ports untagged: Untagged packet
- Different VLAN from Mirrored Port
  - Both ports tagged: Tagged packet with VID of mirrored port
  - Mirrored port tagged, Monitoring port untagged: Tagged packet with VID of mirrored port
  - Mirrored port untagged, Monitoring port tagged: Tagged packet with VID of mirrored port
  - Both ports untagged: Tagged packet with VID of mirrored port

## Layer 3 Port Mirroring (Monitoring Port and Egress Port in the same GEA)

In this scenario, you observe Layer 3 port mirroring on an egress port, and both the egress port and the monitoring port are in the same Gigabit Ethernet Aggregator (GEA) unit. To find out which GEA unit each port resides on, use the `/info/geaport` command.

The monitoring port always shows a tagged packet with a VLAN ID (VID) of the egress port.

---

**NOTE** – In general, the GbESM mirrors all Layer 3 ingress packets after the packets are routed, that is, all Layer 3 packets observed at the monitoring port have source MAC address, destination MAC address, VLAN, TTL, IP checksum changed as if they have gone through the routing process to the egress port.

---

## Layer 3 Port Mirroring (Both Ports in Different GEAs)

In this scenario, you observe Layer 3 port mirroring on an egress port, but the egress port and the monitoring port reside on different Gigabit Ethernet Aggregator (GEA) units. To find out which GEA unit each port resides on, use the `/info/geaport` command.

---

**NOTE** – In general, the GbESM mirrors all Layer 3 ingress packets after the packets are routed, that is, all Layer 3 packets observed at the monitoring port have source MAC address, destination MAC address, VLAN, TTL, IP checksum changed as if they have gone through the routing process to the egress port.

---

- Same VLAN as Mirrored Port
  - Both ports tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port tagged, Monitoring port untagged: Untagged packet
  - Mirrored port untagged, Monitoring port tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Both ports untagged: Untagged packet
- Different VLAN from Mirrored Port
  - Both ports tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port tagged, Monitoring port untagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port untagged, Monitoring port tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Both ports untagged: Tagged packet with VID of the VLAN to which the ports belong

---

**NOTE** – When the monitoring port belongs to a different VLAN, all mirrored packets have an 802.1q tag field with the VID of the VLAN to which the mirrored port belongs. The VID included in the tag field depends on which port is monitored.

For example, assume a packet traverses from port EXT1 (VLAN 1) to EXT2 (VLAN 2), and is monitored on port EXT3. When you monitor EXT1, the packet's tag includes the VID of VLAN 1. When you monitor EXT2, the packet's tag includes the VID of VLAN 2.

---

### Layer 3 Port Mirroring (MP Packets, Both Ports in the Same GEA)

MP packets are generated by the management processor, such as routing packets between direct interfaces. In this scenario, the mirrored port and the monitoring port reside on the same Gigabit Ethernet Aggregator (GEA) unit. To find out which GEA unit each port resides on, use the `/info/geaport` command.

The monitoring port always shows a tagged packet with a VLAN ID (VID) of the ingress port.

### Layer 3 Port Mirroring (MP Packets, Mirrored Port and Monitoring Port in Different GEAs)

MP packets are generated by the management processor, such as a ping to the switch interfaces. In this scenario, the mirrored port and the monitoring port reside on different Gigabit Ethernet Aggregator (GEA) units. To find out which GEA unit each port resides on, use the `/info/geaport` command.

---

**NOTE** – The switch doesn't monitor Layer 3 egress-routed packets from the MP when the mirrored port and the monitoring port are in different VLANs.

---

- Same VLAN as Mirrored Port
  - Both ports tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port tagged, Monitoring port untagged: Untagged packet
  - Mirrored port untagged, Monitoring port tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Both ports untagged: Untagged packet

- Different VLAN from Mirrored Port
  - Both ports tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port tagged, Monitoring port untagged: Tagged packet with VID of the VLAN to which the ports belong
  - Mirrored port untagged, Monitoring port tagged: Tagged packet with VID of the VLAN to which the ports belong
  - Both ports untagged: Tagged packet with VID of the VLAN to which the ports belong

---

**NOTE** – When the monitoring port belongs to a different VLAN, all mirrored packets have an 802.1q tag field with the VID of the VLAN to which the mirrored port belongs. The VID included in the tag field depends on which port is monitored.

For example, assume a packet traverses from port EXT1 (VLAN 1) to EXT2 (VLAN 2), and is monitored on port EXT3. When you monitor EXT1, the packet's tag includes the VID of VLAN 1. When you monitor EXT2, the packet's tag includes the VID of VLAN 2.

---

## Configuring Port Mirroring

To configure port mirroring for the example shown in [Figure A-1](#),

### 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
INT5                  none
-----
-----
-----
EXT1                  none
EXT2                  none
EXT3                  (EXT1, in) (EXT2, out)
EXT4                  none
EXT5                  none
EXT6                  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.

1. Create a dictionary file called `alteon.dct`, with the following content:

```
#####
# alteon.dct - RADLINX Alteon dictionary
#
# (See README.DCT for more details on the format of this file)
#####
#
# Use the Radius specification attributes in lieu of the
# RADLINX Alteon ones
#
@radius.dct

#
# Define additional RADLINX Alteon parameters
# (add RADLINX Alteon specific attributes below)

ATTRIBUTE   Radlinx-Vendor-Specific  26  [vid=648 data=string]  R

#####
# alteon.dct - RADLINX Alteon dictionary
#####
#Define Alteon GbESM Layer 2 & Layer 3 dictionary
#@radius.dct

@alteon.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):**

■ `@alteon.dct`

```
#####
# dictiona.dcm
#####
# Generic Radius

@radius.dct

#
# Specific Implementations (vendor specific)
#
@pprt1213.dct
@acc.dct
@accessbd.dct
@alteon.dct
.
.
.
#####
# dictiona.dcm
#####
```

**3. Open the vendor file (`vendor.ini`), and add the following data to the Vendor-Product identification list:**

```
vendor-product      = Alteon Blade-server module
dictionary          = alteon
ignore-ports        = no
help-id             = 0
```

# Glossary



<b>DIP (Destination IP Address)</b>	The destination IP address of a frame.
<b>Dport (Destination Port)</b>	The destination port (application socket: for example, http-80/https-443/DNS-53)
<b>NAT (Network Address Translation)</b>	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.
<b>Preemption</b>	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.
<b>Priority</b>	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.
<b>Proto (Protocol)</b>	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.)
<b>SIP (Source IP Address)</b>	The source IP address of a frame.
<b>SPort (Source Port)</b>	The source port (application socket: for example, HTTP-80/HTTPS-443/DNS-53).
<b>Tracking</b>	<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"><li>■ ifs: Active IP interfaces on the Web switch (increments priority by 2 for each)</li><li>■ ports: Active ports on the same VLAN (increments priority by 2 for each)</li><li>■ vrs: Number of virtual routers in master mode on the switch</li></ul>
<b>VIR (Virtual Interface Router)</b>	A VRRP address 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 GbE Switch Modules 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 value between 1 and 255 that 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}`. 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 whom to share with.

**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.

# Index

## Symbols

[ ] ..... 15

## Numerics

802.1Q VLAN tagging ..... 68

## A

accessing the switch  
    defining source IP addresses ..... 40  
    RADIUS authentication ..... 41  
    security ..... 39  
    using the Browser-based Interface ..... 27  
active-active redundancy ..... 208  
administrator account ..... 44  
aggregating routes ..... 161  
    example ..... 167  
application ports ..... 112  
authenticating, in OSPF ..... 181  
autonomous systems (AS) ..... 174

## B

BBI  
    See Browser-Based Interface ..... 175  
Border Gateway Protocol (BGP) ..... 155  
    attributes ..... 162  
    failover configuration ..... 164  
    route aggregation ..... 161  
    route maps ..... 157  
    selecting route paths ..... 163  
Bridge Protocol Data Unit (BPDU) ..... 93  
broadcast domains ..... 65, 135  
Browser-Based Interface ..... 175

## C

Cisco EtherChannel ..... 84, 87  
CIST ..... 107  
Class of Service queue ..... 125  
command conventions ..... 15  
Command Line Interface ..... 175  
configuration rules  
    port mirroring ..... 84  
    spanning tree ..... 84  
    Trunking ..... 84  
    VLANs ..... 84  
configuring  
    BGP failover ..... 164  
    IP routing ..... 133  
    OSPF ..... 185  
    port trunking ..... 86  
    spanning tree groups ..... 99

## D

default gateway ..... 132  
    configuration example ..... 78, 134  
    VLAN-based ..... 75  
default password ..... 44  
default route  
    OSPF ..... 179  
Differentiated Services Code Point (DSCP) ..... 119

## E

EtherChannel ..... 82  
    as used with port trunking ..... 84, 87  
Extensible Authentication Protocol over LAN (EAPoL)  
    57  
external routing ..... 156, 174

**F**

Failover.....	198
failover	
overview .....	207
fault tolerance	
port trunking .....	83
fragmenting jumbo frames .....	130, 133
frame tagging. <i>See</i> VLANs tagging.	

**G**

gateway. *See* default gateway.

**H**

high-availability .....	197
Host routes	
OSPF .....	184
hot-standby redundancy .....	209
HP-OpenView .....	31

**I**

IBM Director .....	31
ICMP .....	112
IEEE standards	
802.1d.....	92
802.1s .....	103, 107
802.1x.....	57
IGMP .....	112
IGMP Snooping .....	147
incoming route maps .....	158
internal routing .....	156, 174
Internet Group Management Protocol (IGMP) .....	148
IP address	
local route cache ranges.....	137
routing example.....	77, 133
IP interfaces	
example configuration.....	77, 133, 136
IP routing	
cross-subnet example .....	130
default gateway configuration .....	78, 134
IP interface configuration .....	77, 133, 136
IP subnets .....	131
network diagram.....	131
subnet configuration example.....	133
switch-based topology .....	132

IP subnets .....	132
routing.....	131, 132
VLANs.....	65
ISL Trunking .....	82

**J**

jumbo frames	
fragmenting to normal size .....	130, 133
routing.....	130, 133

**L**

LACP.....	88
Link Aggregation Control Protocol .....	88
lmask (local route cache parameter) .....	137
lnet (local route cache parameter).....	137
local route cache address range .....	137
local route cache parameters	
lmask .....	137
lnet.....	137
logical segment. <i>See</i> IP subnets.	
LSAs.....	173

**M**

management module .....	20
Management Processor (MP)	
use in switch security.....	40
manual style conventions .....	15
meter .....	116
mirroring ports .....	226
monitoring ports .....	226
MSTP.....	107
multi-links between switches	
using port trunking .....	81
multiple spanning tree groups .....	97
Multiple Spanning Tree Protocol.....	107

**N**

network management .....	31
--------------------------	----

**O****OSPF**

area types.....	170
authentication.....	181
configuration examples.....	186 to ??
default route.....	179
external routes.....	185
filtering criteria.....	112
host routes.....	184
link state database.....	173
neighbors.....	173
overview.....	170
redistributing routes.....	157, 161
route maps.....	157, 159
route summarization.....	178
router ID.....	181
virtual link.....	180
outgoing route maps.....	158

**P****password**

administrator account.....	44
default.....	44
user account.....	44

**Per Hop Behavior (PHB).....** 119**port mirroring.....** 226

configuration rules.....	84
--------------------------	----

**Port Trunking.....** 83**port trunking.....** 83

configuration example.....	85
description.....	87
EtherChannel.....	82
fault tolerance.....	83

**ports**

for services.....	112
monitoring.....	226
physical. <i>See</i> switch ports.	

**priority value (802.1p).....** 123**protocol types.....** 112**PVID (port VLAN ID).....** 67**R****RADIUS**

authentication.....	41
port 1812 and 1645.....	112
port 1813.....	112
SSH/SCP.....	54

**Rapid Spanning Tree Protocol.....** 104**Rapid Spanning Tree Protocol (RSTP).....** 104**redistributing routes.....** 157, 161, 167**redundancy**

active-active.....	208
hot-standby.....	209

**re-mark.....** 116**RIP (Routing Information Protocol)**

advertisements.....	142
distance vector protocol.....	141
hop count.....	141
metric.....	141
TCP/IP route information.....	14, 141
version 1.....	141

**route aggregation.....** 161, 167**route cache, sample.....** 76**route maps.....** 157

configuring.....	159
incoming and outgoing.....	158

**route paths in BGP.....** 163**Router ID**

OSPF.....	181
-----------	-----

**routers.....** 78, 131, 134

border.....	174
peer.....	174
port trunking.....	82
switch-based routing topology.....	132

**routes, advertising.....** 174**routing.....** 156

internal and external.....	174
----------------------------	-----

**Routing Information Protocol. *See* RIP****RSA keys.....** 53**RSTP.....** 104**S****SecurID.....** 55**security**

allowable SIP addresses.....	40
port mirroring.....	226
RADIUS authentication.....	41
switch management.....	40
VLANs.....	65

**segmentation. *See* IP subnets.****segments. *See* IP subnets.****service ports.....** 112**SNMP.....** 31, 175

HP-OpenView.....	31
------------------	----

spanning tree	
configuration rules .....	84
Spanning-Tree Protocol	
multiple instances .....	97
spoofing, prevention of .....	40
SSH	
RSA host and server keys .....	53
SSH/SCP	
configuring .....	50
statistical load distribution.....	83
summarizing routes .....	178
switch failover .....	207
switch management	
security .....	40
switch ports VLANs membership .....	68

## T

TACACS+ .....	45
tagging. <i>See</i> VLANs tagging.	
TCP .....	112
technical terms	
port VLAN identifier (PVID) .....	68
tagged frame .....	68
tagged member.....	69
untagged frame.....	68
untagged member .....	69
VLAN identifier (VID) .....	68
text conventions .....	15
thash .....	87
Trunk Hash algorithm.....	87
Trunking	
configuration rules .....	84
typographic conventions .....	15

## U

UDP .....	112
jumbo frame traffic fragmentation .....	133

user account .....	44
--------------------	----

## V

virtual interface router (VIR) .....	204
virtual link, OSPF.....	180
Virtual Local Area Networks. <i>See</i> VLANs.	
virtual router	
ID numbering .....	211
virtual router group .....	209
Virtual Router Redundancy Protocol	
tracking .....	210
VLANs	
broadcast domains.....	65, 135
configuration rules .....	84
default PVID .....	67
example showing multiple VLANs .....	73
gateway, default.....	75
ID numbers .....	66
IP interface configuration.....	136
multiple spanning trees .....	92
multiple VLANs .....	68
port members .....	68
PVID.....	67
routing.....	135
security.....	65
Spanning-Tree Protocol .....	92
tagging .....	68 to 74
topologies .....	72
VRRP (Virtual Router Redundancy Protocol)	
active-active redundancy.....	208
hot-standby redundancy .....	209
overview.....	204, 210
virtual interface router .....	204
virtual router ID numbering .....	211
vrid .....	204