



ExtremeWare™ Software User Guide


Software Version 6.2.1


Extreme Networks, Inc.
3585 Monroe Street
Santa Clara, California 95051
(888) 257-3000
<http://www.extremenetworks.com>

Published: February 2002
Part number: 100049-00 Rev. 04

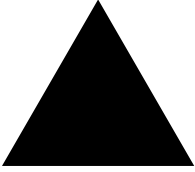
©2001 Extreme Networks, Inc. All rights reserved. Extreme Networks and BlackDiamond are registered trademarks of Extreme Networks, Inc. in the United States and certain other jurisdictions. ExtremeWare, ExtremeWare Vista, ExtremeWorks, ExtremeAssist, ExtremeAssist1, ExtremeAssist2, PartnerAssist, Extreme Standby Router Protocol, ESRP, SmartTraps, Alpine, Summit, Summit1, Summit 1i, Summit4, Summit4/FX, Summit 5i, Summit7i, Summit24, Summit48, Summit 48i, Summit Virtual Chassis, SummitLink, SummitGbX, SummitRPS, and the Extreme Networks logo are trademarks of Extreme Networks, Inc., which may be registered or pending registration in certain jurisdictions. The Extreme Turbodriven logo is a service mark of Extreme Networks, which may be registered or pending registration in certain jurisdictions. Specifications are subject to change without notice.

NetWare and Novell are registered trademarks of Novell, Inc. Merit is a registered trademark of Merit Network, Inc. Solaris is a trademark of Sun Microsystems, Inc. F5, BIG/ip, and 3DNS are registered trademarks of F5 Networks, Inc. see/IT is a trademark of F5 Networks, Inc.

 "Data Fellows", the triangle symbol, and Data Fellows product names and symbols/logos are trademarks of Data Fellows.

 F-Secure SSH is a registered trademark of Data Fellows.

All other registered trademarks, trademarks and service marks are property of their respective owners.



Contents

Preface	
Introduction	xvii
Terminology	xviii
Conventions	xviii
Related Publications	xix
Part 1: Using ExtremeWare	
1 ExtremeWare Overview	
Summary of Features	1-1
Virtual LANs (VLANs)	1-3
Spanning Tree Protocol	1-3
Quality of Service	1-3
Unicast Routing	1-4
IP Multicast Routing	1-4
Load Sharing	1-4
“I” Chipset Products	1-5
“I” Chipset Feature Differences	1-5
Software Licensing	1-5
Router Licensing	1-6
Security Licensing	1-7
Software Factory Defaults	1-8

2	Accessing the Switch	
	Understanding the Command Syntax	2-1
	Syntax Helper	2-2
	Command Shortcuts	2-3
	Modular Switch Numerical Ranges	2-3
	Stand-alone Switch Numerical Ranges	2-4
	Names	2-4
	Symbols	2-4
	Line-Editing Keys	2-5
	Command History	2-7
	Common Commands	2-7
	Configuring Management Access	2-10
	User Account	2-11
	Administrator Account	2-11
	Default Accounts	2-12
	Creating a Management Account	2-13
	Domain Name Service Client Services	2-14
	Checking Basic Connectivity	2-14
	Ping	2-14
	Traceroute	2-15
3	Managing the Switch	
	Overview	3-1
	Using the Console Interface	3-2
	Using the 10/100 UTP Management Port	3-2
	Using Telnet	3-3
	Connecting to Another Host Using Telnet	3-3
	Configuring Switch IP Parameters	3-4
	Disconnecting a Telnet Session	3-6
	Controlling Telnet Access	3-6

Using Secure Shell 2 (SSH2)	3-7
Enabling SSH2 for Inbound Switch Access	3-8
Using SCP2 from an External SSH2 Client	3-9
SSH2 Client Functions on the Switch	3-10
Using ExtremeWare Vista	3-11
Controlling Web Access	3-12
Setting Up Your Browser	3-12
Accessing ExtremeWare Vista	3-13
Navigating ExtremeWare Vista	3-14
Saving Changes	3-15
Filtering Information	3-16
Do a GET When Configuring a VLAN	3-16
Sending Screen Output to Extreme Networks	3-17
Using SNMP	3-17
Accessing Switch Agents	3-17
Supported MIBs	3-18
Configuring SNMP Settings	3-18
Displaying SNMP Settings	3-19
Authenticating Users	3-20
RADIUS Client	3-20
Configuring TACACS+	3-27
Using Network Login	3-28
Using Network Login in Campus Mode	3-29
Using Network Login in ISP Mode	3-32
DHCP Server on the Switch	3-33
Displaying Network Login Settings	3-33
Disabling Network Login	3-34
Using the Simple Network Time Protocol	3-34
Configuring and Using SNTP	3-34
SNTP Example	3-39

4	Configuring Slots and Ports on a Switch	
	Configuring a Slot on a Modular Switch	4-1
	Configuring Ports on a Switch	4-2
	Enabling and Disabling Switch Ports	4-3
	Configuring Switch Port Speed and Duplex Setting	4-4
	Jumbo Frames	4-5
	Enabling Jumbo Frames	4-5
	Path MTU Discovery	4-6
	IP Fragmentation with Jumbo Frames	4-6
	IP Fragmentation within a VLAN	4-7
	Load Sharing on the Switch	4-8
	Load-Sharing Algorithms	4-8
	Configuring Switch Load Sharing	4-10
	Load-Sharing Examples	4-13
	Verifying the Load-Sharing Configuration	4-14
	Switch Port-Mirroring	4-15
	Modular Switch Port-Mirroring Example	4-15
	Stand-alone Switch Port-Mirroring Example	4-16
	Extreme Discovery Protocol	4-16
	Software-Controlled Redundant Port	4-17
	Theory of Operation	4-18
	Configuring Software-Controlled Redundant Port	4-20
	Multicast Performance Enhancements (BlackDiamond)	4-20
	Performance Enhancements for Load Sharing	4-21
5	Virtual LANs (VLANs)	
	Overview of Virtual LANs	5-1
	Benefits	5-2
	Types of VLANs	5-2
	Port-Based VLANs	5-2
	Tagged VLANs	5-6
	Protocol-Based VLANs	5-9
	Precedence of Tagged Packets Over Protocol Filters	5-12

VLAN Names	5-12
Default VLAN	5-13
Renaming a VLAN	5-13
Configuring VLANs on the Switch	5-13
VLAN Configuration Examples	5-14
Displaying VLAN Settings	5-15
Displaying VLAN Statistics	5-16
Displaying VLAN Statistics Per Port	5-16
Displaying Protocol Information	5-17
VLAN Tunneling (VMANs)	5-17
MAC-Based VLANs	5-19
MAC-Based VLAN Guidelines	5-19
MAC-Based VLAN Limitations	5-20
MAC-Based VLAN Example	5-20
Timed Configuration Download for MAC-Based VLANs	5-21
6 Forwarding Database (FDB)	
Overview of the FDB	6-1
FDB Contents	6-1
How FDB Entries Get Added	6-2
FDB Entry Types	6-2
Disabling MAC Address Learning	6-4
Associating QoS Profiles with an FDB Entry	6-5
FDB Configuration Examples	6-6
MAC-Based Security	6-7
Limiting Dynamic MAC Addresses	6-7
MAC Address Lock Down	6-9
Displaying FDB Entries	6-10
7 Quality of Service (QoS)	
Overview of Policy-Based Quality of Service	7-2

Applications and Types of QoS	7-3
Voice Applications	7-3
Video Applications	7-3
Critical Database Applications	7-4
Web Browsing Applications	7-4
File Server Applications	7-4
Configuring QoS	7-5
QoS Profiles	7-6
Traffic Groupings	7-8
IP-Based Traffic Groupings	7-9
MAC-Based Traffic Groupings	7-9
Explicit Class of Service (802.1p and DiffServ) Traffic Groupings	7-11
Configuring DiffServ	7-14
Physical and Logical Groupings	7-18
Configuring QoS Traffic Grouping Priorities	7-19
Verifying and Resetting QoS Traffic Grouping Priorities	7-20
Verifying Configuration and Performance	7-20
QoS Monitor	7-20
Displaying QoS Profile Information	7-21
Modifying a QoS Configuration	7-22
Bi-Directional Rate Shaping	7-22
Configuring Bi-Directional Rate Shaping	7-23
Bandwidth Settings	7-23
Bi-Directional Rate Shaping Limitations	7-25
Dynamic Link Context System	7-26
DLCS Guidelines	7-26
DLCS Limitations	7-27
8 Access Policies	
Overview of Access Policies	8-1
IP Access Lists	8-2
Routing Access Policies	8-2
Route Maps	8-2

Using IP Access Lists	8-2
How IP Access Lists Work	8-3
Precedence Numbers	8-3
Specifying a Default Rule	8-3
The permit-established Keyword	8-4
Adding and Deleting Access List Entries	8-4
Access Lists for ICMP	8-5
Verifying Access List Configurations	8-6
IP Access List Examples	8-6
Using Routing Access Policies	8-11
Creating an Access Profile	8-12
Configuring an Access Profile Mode	8-12
Adding an Access Profile Entry	8-13
Deleting an Access Profile Entry	8-16
Applying Access Profiles	8-16
Routing Access Policies for RIP	8-16
Routing Access Policies for IPX	8-18
Routing Access Policies for OSPF	8-18
Routing Access Policies for DVMRP	8-20
Routing Access Policies for PIM	8-22
Routing Access Policies for BGP	8-23
Making Changes to a Routing Access Policy	8-23
Removing a Routing Access Policy	8-24
Using Route Maps	8-24
Creating a Route Map	8-24
Add Entries to the Route Map	8-25
Add Statements to the Route Map Entries	8-25
Route Map Operation	8-28
Changes to Route Maps	8-29
Route Maps in BGP	8-30
9 Network Address Translation (NAT)	
Overview	9-1
Internet IP Addressing	9-3

Configuring VLANs for NAT	9-3
NAT Modes	9-4
Configuring NAT	9-5
Creating NAT Rules	9-6
Creating Static and Dynamic NAT Rules	9-6
Creating Portmap NAT Rules	9-7
Creating Auto-Constrain NAT Rules	9-7
Advanced Rule Matching	9-8
Configuring Time-outs	9-8
Displaying NAT Settings	9-8
Disabling NAT	9-9
10 Server Load Balancing (SLB)	
Overview	10-1
SLB Components	10-2
Nodes	10-3
Pools	10-3
Virtual Servers	10-3
Node, Pool, and Virtual Server Relationships	10-5
SLB Traffic Types	10-7
Forwarding Modes	10-7
Transparent Mode	10-8
Translation Mode	10-11
Port Translation Mode	10-12
GoGo Mode	10-13
Load-Balancing Methods	10-15
Round-Robin	10-15
Ratio	10-15
Least Connections	10-16
Priority	10-16
Advanced SLB Application Example	10-17

Using Persistence	10-21
Persistence Methods	10-21
Persistence Levels	10-22
Persistence Types	10-23
Using High Availability System Features	10-24
Server Load Balancing with ESRP	10-24
Active-Active Operation	10-28
Health Checking	10-32
Ping-Check	10-32
TCP-Port-Check	10-33
Service-Check	10-33
3DNS Health Checking	10-34
Maintenance Mode	10-34
Health Checking in GoGo Mode	10-34
Flow Redirection	10-35
Web Cache Redirection	10-35
Policy-Based Routing	10-37
11 Ethernet Automatic Protection Switching	
Overview of the EAPS Protocol	11-1
Fault Detection and Recovery	11-3
Polling	11-4
Trap Message Sent by a Transit Node	11-5
Restoration Operations	11-5
Multiple EAPS Domains Per Switch	11-6
Creating and Deleting an EAPS Domain	11-7
Defining the EAPS Mode of the Switch	11-8
Configuring EAPS Polling Timers	11-8
Configuring the Primary and Secondary Ports	11-9
Configuring the EAPS Control VLAN	11-10
Configuring the EAPS Protected VLANs	11-11
Enabling and Disabling an EAPS Domain	11-11
Enabling and Disabling EAPS	11-11
Unconfiguring an EAPS Ring Port	11-12
Displaying EAPS Status Information	11-12

12 Status Monitoring and Statistics

Status Monitoring	12-2
Slot Diagnostics	12-2
Runtime Diagnostics (BlackDiamond)	12-3
Port Statistics	12-4
Port Errors	12-5
Port Monitoring Display Keys	12-6
System Health Checking (BlackDiamond)	12-7
Setting the System Recovery Level	12-9
Logging	12-9
Local Logging	12-10
Remote Logging	12-11
Logging Configuration Changes	12-12
Configuring and Monitoring Flow Statistics	12-12
Flow Statistics Background Information	12-13
Collection Port and Filtering Options	12-16
Collection Architecture Scalability and Reliability	12-16
Export Criteria	12-17
RMON	12-23
About RMON	12-24
RMON Features of the Switch	12-24
Configuring RMON	12-25
Event Actions	12-26

Part 2: Using Routing Protocols

13 Spanning Tree Protocol (STP)

Overview of the Spanning Tree Protocol	13-2
Spanning Tree Domains	13-2
Defaults	13-2
Port Modes	13-3

STPD BPDU Tunneling	13-3
Rapid Root Failover	13-4
STP Configurations	13-4
Basic STP Configuration	13-4
Multiple STPDs on a Port	13-7
VLAN Spanning Multiple STPDs	13-8
EMISTP Deployment Constraints	13-9
Per-VLAN Spanning Tree	13-11
STPD VLAN Mapping	13-12
Native VLAN	13-12
STP Rules and Restrictions	13-12
Configuring Basic STP on the Switch	13-13
STP Configuration Examples	13-14
Displaying STP Settings	13-16
14 Extreme Standby Router Protocol	
Overview	14-1
ESRP-Aware Switches	14-2
ESRP Basics	14-2
Determining the ESRP Master	14-3
ESRP Tracking	14-4
ESRP Election Algorithms	14-8
Master Switch Behavior	14-9
Standby Switch Behavior	14-9
Electing the Master Switch	14-9
Failover Time	14-10
Grouping Blocks of 10/100 Ports	14-10
ESRP Options	14-13
ESRP Host Attach	14-13
ESRP Domains	14-14
ESRP Groups	14-16
Linking ESRP Switches	14-17
Configuring ESRP and Multinetting	14-17

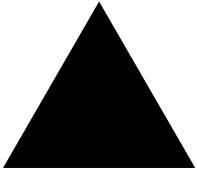
ESRP and Spanning Tree	14-17
ESRP Port Restart	14-18
ESRP and VLAN Aggregation	14-18
ESRP Examples	14-19
Displaying ESRP Information	14-23
15 Virtual Router Redundancy Protocol	
Overview	15-1
Determining the VRRP Master	15-3
VRRP Tracking	15-3
Electing the Master Router	15-6
Additional VRRP Highlights	15-7
VRRP Port Restart	15-7
VRRP Operation	15-8
Simple VRRP Network Configuration	15-8
Fully-Redundant VRRP Network	15-9
VRRP Configuration Parameters	15-11
VRRP Examples	15-12
Configuring the Simple VRRP Network	15-12
Configuring the Fully-Redundant VRRP Network	15-13
16 IP Unicast Routing	
Overview of IP Unicast Routing	16-2
Router Interfaces	16-2
Populating the Routing Table	16-3
Subnet-Directed Broadcast Forwarding	16-6
Proxy ARP	16-6
ARP-Incapable Devices	16-7
Proxy ARP Between Subnets	16-7
Relative Route Priorities	16-8
Configuring IP Unicast Routing	16-8
Verifying the IP Unicast Routing Configuration	16-9
Routing Configuration Example	16-9

IP Multinetting	16-11
IP Multinetting Operation	16-12
IP Multinetting Examples	16-13
Configuring DHCP/BOOTP Relay	16-14
Verifying the DHCP/BOOTP Relay Configuration	16-15
UDP-Forwarding	16-15
Configuring UDP-Forwarding	16-16
UDP-Forwarding Example	16-16
ICMP Packet Processing	16-16
VLAN Aggregation	16-17
VLAN Aggregation Properties	16-18
VLAN Aggregation Limitations	16-19
VLAN Aggregation SubVLAN Address Range Checking	16-19
Isolation Option for Communication Between Sub-VLANs	16-20
VLAN Aggregation Example	16-20
Verifying the VLAN Aggregation Configuration	16-21
17 Interior Gateway Routing Protocols	
Overview	17-2
RIP Versus OSPF	17-2
Overview of RIP	17-3
Routing Table	17-3
Split Horizon	17-4
Poison Reverse	17-4
Triggered Updates	17-4
Route Advertisement of VLANs	17-4
RIP Version 1 Versus RIP Version 2	17-4
Overview of OSPF	17-5
Link-State Database	17-5
Areas	17-7
Point-to-Point Support	17-11
Route Re-Distribution	17-12
Configuring Route Re-Distribution	17-13
OSPF Timers and Authentication	17-15
RIP Configuration Example	17-15

Configuring OSPF	17-17
Configuring OSPF Wait Interval	17-17
OSPF Configuration Example	17-18
Configuration for ABR1	17-20
Configuration for IR1	17-21
Displaying OSPF Settings	17-21
OSPF LSD Display	17-21
18 Exterior Gateway Routing Protocols	
Overview	18-2
BGP Attributes	18-2
BGP Communities	18-3
BGP Features	18-3
Route Reflectors	18-3
Route Confederations	18-4
Route Aggregation	18-8
IGP Synchronization	18-9
Using the Loopback Interface	18-9
BGP Peer Groups	18-9
BGP Route Selection	18-10
Stripping Out Private AS Numbers from Route Updates	18-11
Route Re-Distribution	18-11
Configuring Route Re-Distribution	18-12
19 IP Multicast Routing	
Overview	19-2
DVMRP Overview	19-2
PIM Overview	19-2
IGMP Overview	19-4
Performance Enhancements for the BlackDiamond Switch	19-5
Configuring IP Multicasting Routing	19-5
Configuration Examples	19-6
PIM-DM Configuration Example	19-7

	Configuration for IR1	19-8
	Configuration for ABR1	19-10
20	IPX Routing	
	Overview of IPX	20-1
	Router Interfaces	20-1
	IPX Routing Performance	20-3
	IPX Load Sharing	20-3
	IPX Encapsulation Types	20-3
	Tagged IPX VLANs	20-4
	Populating the Routing Table	20-5
	IPX/RIP Routing	20-5
	Routing SAP Advertisements	20-6
	Configuring IPX	20-7
	Verifying IPX Router Configuration	20-7
	Protocol-Based VLANs for IPX	20-8
	IPX Configuration Example	20-8
Part 3:	Appendixes	
A	Supported Protocols and Standards	
B	Software Upgrade and Boot Options	
	Downloading a New Image	B-1
	Rebooting the Switch	B-2
	Saving Configuration Changes	B-3
	Returning to Factory Defaults	B-3
	Using TFTP to Upload the Configuration	B-4
	Using TFTP to Download the Configuration	B-5
	Downloading a Complete Configuration	B-5
	Downloading an Incremental Configuration	B-5
	Scheduled Incremental Configuration Download	B-6
	Remember to Save	B-6

Synchronizing MSMs	B-7
Upgrading and Accessing BootROM	B-7
Upgrading BootROM	B-7
Accessing the BootROM menu	B-7
C Troubleshooting	
LEDs	C-1
Using the Command-Line Interface	C-3
Port Configuration	C-5
VLANs	C-6
STP	C-7
Debug Tracing	C-8
TOP Command	C-9
System Health Check	C-9
Contacting Extreme Technical Support	C-10
Index	
Index of Commands	



Preface

This Preface provides an overview of this guide, describes guide conventions, and lists other publications that may be useful.

Introduction

This guide provides the required information to configure ExtremeWare™ software running on either modular or stand-alone switches from Extreme Networks.

This guide is intended for use by network administrators who are responsible for installing and setting up network equipment. It assumes a basic working knowledge of:

- Local area networks (LANs).
- Ethernet concepts.
- Ethernet switching and bridging concepts.
- Routing concepts.
- Internet Protocol (IP) concepts.
- Routing Information Protocol (RIP) and Open Shortest Path First (OSPF).
- Border Gateway Protocol (BGP-4) concepts.
- IP Multicast concepts.
- Distance Vector Multicast Routing Protocol (DVMRP) concepts.
- Protocol Independent Multicast (PIM) concepts.

- Internet Packet Exchange (IPX) concepts.
- Server Load Balancing (SLB) concepts.
- Simple Network Management Protocol (SNMP).



Note: If the information in the release notes shipped with your switch differs from the information in this guide, follow the release notes.

Terminology

When features, functionality, or operation is specific to a modular or stand-alone switch family, the family name is used. Explanations about features and operations that are the same across all product families simply refer to the product as the “switch.”

Conventions

Table 1 and Table 2 list conventions that are used throughout this guide.

Table 1: Notice Icons




Icon	Notice Type	Alerts you to...
	Note	Important features or instructions.
	Caution	Risk of personal injury, system damage, or loss of data.
	Warning	Risk of severe personal injury.

Table 2: Text Conventions

Convention	Description
Screen displays	This typeface indicates command syntax, or represents information as it appears on the screen.
The words “enter” and “type”	When you see the word “enter” in this guide, you must type something, and then press the Return or Enter key. Do not press the Return or Enter key when an instruction simply says “type.”
[Key] names	Key names are written with brackets, such as [Return] or [Esc]. If you must press two or more keys simultaneously, the key names are linked with a plus sign (+). Example: Press [Ctrl]+[Alt]+[Del].
Letter in bold type	Letters within a command that appear in bold type indicate the keyboard shortcut for a command. When entering the command, you can use just the bolded letters instead of the entire word.
Words in <i>italicized</i> type	Italics emphasize a point or denote new terms at the place where they are defined in the text.

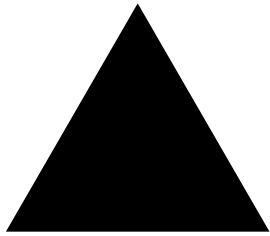
Related Publications

The publications related to this one are:

- ExtremeWare release notes.
- *ExtremeWare Software Command Reference Guide.*
- *ExtremeWare 6.2.1 Software Quick Reference Guide*
- *Extreme Networks Consolidated Hardware Guide.*

Documentation for Extreme Networks products is available on the World Wide Web at the following location:

- <http://www.extremenetworks.com/>



Part 1:

Using ExtremeWare



ExtremeWare Overview

This chapter covers the following topics:

- Summary of Features on page 1-1
- “i” Chipset Products on page 1-5
- Software Licensing on page 1-5
- Software Factory Defaults on page 1-8

ExtremeWare is the full-featured software operating system that is designed to run on the Extreme Networks families of modular and stand-alone Gigabit Ethernet switches.

Summary of Features

The features of ExtremeWare include:

- Virtual local area networks (VLANs) including support for IEEE 802.1Q and IEEE 802.1p.
- VLAN aggregation.
- Spanning Tree Protocol (STP) (IEEE 802.1D) with multiple STP domains.
- Policy-Based Quality of Service (PB-QoS).
- Wire-speed Internet Protocol (IP) routing.
- IP Multinetting.
- DHCP/BOOTP Relay.

- Extreme Standby Router Protocol (ESRP).
- Virtual Router Redundancy Protocol (VRRP).
- Routing Information Protocol (RIP) version 1 and RIP version 2.
- Open Shortest Path First (OSPF) routing protocol.
- Border Gateway Protocol (BGP) version 4.
- Wire-speed IP multicast routing support.
- Diffserv support.
- Access-policy support for routing protocols.
- Access list support for packet filtering.
- IGMP snooping to control IP multicast traffic.
- Distance Vector Multicast Routing Protocol (DVMRP).
- Protocol Independent Multicast-Dense Mode (PIM-DM).
- Protocol Independent Multicast-Sparse Mode (PIM-SM).
- Wire-speed IPX, IPX/RIP, and IPX/SAP support.
- Server Load Balancing (SLB) support.
- Load sharing on multiple ports, across all blades (modular switches only).
- RADIUS client and per-command authentication support.
- TACACS+ support.
- Console command-line interface (CLI) connection.
- Telnet CLI connection.
- SSH2 connection.
- ExtremeWare Vista Web-based management interface.
- Simple Network Management Protocol (SNMP) support.
- Remote Monitoring (RMON).
- System Monitoring (SMON).
- Traffic mirroring for all ports, across all blades (modular switches only).



Note: For more information on Extreme Networks switch components (the BlackDiamond 6800 family, the Alpine 3800 family, or the Summit switch family, refer to the Extreme Networks Consolidated Hardware Guide.

Virtual LANs (VLANs)

ExtremeWare has a VLAN feature that enables you to construct your broadcast domains without being restricted by physical connections. A VLAN is a group of location- and topology-independent devices that communicate as if they were on the same physical local area network (LAN).

Implementing VLANs on your network has the following three advantages:

- They help to control broadcast traffic. If a device in VLAN *Marketing* transmits a broadcast frame, only VLAN *Marketing* devices receive the frame.
- They provide extra security. Devices in VLAN *Marketing* can only communicate with devices on VLAN *Sales* using routing services.
- They ease the change and movement of devices on networks.



Note: For more information on VLANs, refer to Chapter 5.

Spanning Tree Protocol

The switch supports the IEEE 802.1D Spanning Tree Protocol (STP), which is a bridge-based mechanism for providing fault tolerance on networks. STP enables you to implement parallel paths for network traffic, and ensure that:

- Redundant paths are disabled when the main paths are operational.
- Redundant paths are enabled if the main traffic paths fail.

A single spanning tree can span multiple VLANs.



Note: For more information on STP, refer to Chapter 13.

Quality of Service

ExtremeWare has Policy-Based Quality of Service (QoS) features that enable you to specify service levels for different traffic groups. By default, all traffic is assigned the *normal* QoS policy profile. If needed, you can create other QoS policies and apply them to different traffic types so that they have different guaranteed minimum bandwidth, maximum bandwidth, and priority.



Note: For more information on Quality of Service, refer to Chapter 7.

Unicast Routing

The switch can route IP or IPX traffic between the VLANs that are configured as virtual router interfaces. Both dynamic and static IP routes are maintained in the routing table. The following routing protocols are supported:

- RIP version 1
- RIP version 2
- OSPF
- IPX/RIP
- BGP version 4



Note: For more information on IP unicast routing, refer to Chapter 16. For more information on IPX/RIP, refer to Chapter 20.

IP Multicast Routing

The switch can use IP multicasting to allow a single IP host to transmit a packet to a group of IP hosts. ExtremeWare supports multicast routes that are learned by way of the Distance Vector Multicast Routing Protocol (DVMRP) or the Protocol Independent Multicast (dense mode or sparse mode).



Note: For more information on IP multicast routing, refer to Chapter 19.

Load Sharing

Load sharing allows you to increase bandwidth and resiliency by using a group of ports to carry traffic in parallel between systems. The sharing algorithm allows the switch to use multiple ports as a single logical port. For example, VLANs see the load-sharing group as a single virtual port. The algorithm also guarantees packet sequencing between clients.



Note: For information on load sharing, refer to Chapter 4.

“i” Chipset Products

Switches and switch modules that use naming conventions ending with an “i” have additional capabilities that are documented throughout this user guide. For the most current list of products supporting the “i” chipset, consult your release notes.

Unless otherwise specified, a feature requiring the “i” chipset requires the use of both an “i” chipset-based management module, such as the MSM64i, and an “i” chipset-based I/O module, such as the G8Xi.

“i” Chipset Feature Differences

The following list summarizes the feature areas specific to the “i” chipset products:

- QoS and access policies – Complete use of IP access lists (products without the “i” chipset are capable of a subset of this functionality); support for IP DiffServ; and support for eight QoS queues per port, instead of four.
- Bridging/Switching – Support for jumbo frames; support for address- and round-robin-based load-sharing algorithms; ports belonging to a load-sharing group do not need to be contiguous.
- Routing – Wire-speed IPX routing
- BGP-4 – Requires the use of the “i” chipset, but requires only the MSM64i on the BlackDiamond.
- Server load balancing – Requires the use of the “i” chipset.
- Web cache redirection – Requires the use of the “i” chipset.
- ESRP – No port blocking restrictions, use of the additional tracking ESRP feature.
- Load sharing – No contiguous port or speed difference restrictions.

Software Licensing

Some Extreme Networks products have capabilities that are enabled by using a license key. Keys are typically unique to the switch, and are not transferable. Keys are stored in NVRAM and, once entered, persist through reboots, software upgrades, and reconfigurations. The following sections describe the features that are associated with license keys.

Router Licensing

Some switches support software licensing for different levels of router functionality. In ExtremeWare version 6.0 and above, routing protocol support is separated into two sets: Basic and Full L3. Basic is a subset of Full L3.

Basic Functionality

Basic functionality requires *no license key*. All Extreme switches have Basic layer 3 functionality, without the requirement of a license key. Basic functionality includes all switching functions, and also includes all available layer 3 QoS, access list, and ESRP functions. Layer 3 routing functions include support for:

- IP routing using RIP version 1 and/or RIP version 2.
- IP routing between directly attached VLANs.
- IP routing using static routes.
- IPX routing (direct, static, and dynamic using IPX/RIP and IPX/SAP).

Full L3 Functionality

On switches that support router licensing, the Full L3 license enables support of additional routing protocols and functions, including:

- IP routing using OSPF.
- IP multicast routing using DVMRP.
- IP multicast routing using PIM (Dense Mode or Sparse Mode).
- IP routing using BGP.
- Server load balancing.
- Web cache redirection.

Product Support

The Summit1i switch and all BlackDiamond 6800 series switches ship with Full L3 functionality. All other Summit models and the Alpine 3800 series switches are available with either Basic or Full L3 functionality.

Verifying the Router License

To verify the router license, use the `show switch` command.

Obtaining a Router License

You can order the desired functionality from the factory, using the appropriate model of the desired product. If you order licensing from the factory, the switch arrives packaged with a certificate that contains the unique license key(s), and instructions for enabling the correct functionality on the switch. The certificate is typically packaged with the switch documentation. Once the license key is entered, it should not be necessary to enter the information again. However, we recommend keeping the certificate for your records.

You can upgrade the router licensing of an existing product by purchasing a voucher for the desired product and functionality. Please contact your supplier to purchase a voucher.

The voucher contains information and instructions on obtaining a license key for the switch using the Extreme Networks Support website at:

<http://www.extremenetworks.com/support/techsupport.asp>

or by phoning Extreme Networks Technical Support at:

- (800) 998-2408
- (408) 579-2826

Security Licensing

Certain additional ExtremeWare security features, such as the use of Secure Shell (SSH2) encryption, may be under United States export restriction control. Extreme Networks ships these security features in a disabled state. You can obtain information on enabling these features at no charge from Extreme Networks.

Obtaining a Security License

To obtain information on enabling features that require export restriction, access the Extreme Networks Support website at:

<http://www.extremenetworks.com/go/security.htm>

Fill out a contact form to indicate compliance or noncompliance with the export restrictions. If you are in compliance, you will be given information that will allow you to enable security features.

Security Features Under License Control

ExtremeWare version 6.0 and above supports the SSH2 protocol. SSH2 allows the encryption of Telnet session data between an SSH2 client and an Extreme Networks switch. ExtremeWare version 6.2.1 and later also enables the switch to function as an SSH2 client, sending encrypted data to an SSH2 server on a remote system. ExtremeWare 6.2.1 also supports the Secure Copy Protocol (SCP). The encryption methods used are under U.S. export restriction control.

Software Factory Defaults

Table 1-1 shows factory defaults for global ExtremeWare features.

Table 1-1: ExtremeWare Global Factory Defaults

Item	Default Setting
Serial or Telnet user account	<i>admin</i> with no password and <i>user</i> with no password
Web network management	Enabled
Telnet	Enabled
SSH2	Disabled
SNMP	Enabled
SNMP read community string	<i>public</i>
SNMP write community string	<i>private</i>
RMON	Disabled
BOOTP	Enabled on the default VLAN (<i>default</i>)
QoS	All traffic is part of the default queue
QoS monitoring	Automatic roving

Table 1-1: ExtremeWare Global Factory Defaults (continued)

Item	Default Setting
802.1p priority	Recognition enabled
802.3x flow control	Enabled on Gigabit Ethernet ports
Virtual LANs	Three VLANs predefined. VLAN named <i>default</i> contains all ports and belongs to the STPD named <i>s0</i> . VLAN <i>mgmt</i> exists only on switches that have an Ethernet management port, and contains only that port. The Ethernet management port is DTE only, and is not capable of switching or routing. VLAN <i>MacVlanDiscover</i> is used only when using the MAC VLAN feature.
802.1Q tagging	All packets are untagged on the default VLAN (<i>default</i>).
Spanning Tree Protocol	Disabled for the switch; enabled for each port in the STPD.
Forwarding database aging period	300 seconds (5 minutes)
IP Routing	Disabled
RIP	Disabled
OSPF	Disabled
IP multicast routing	Disabled
IGMP	Enabled
IGMP snooping	Enabled
DVMRP	Disabled
GVRP	Disabled
PIM-DM	Disabled
IPX routing	Disabled
NTP	Disabled
DNS	Disabled
Port mirroring	Disabled



Note: For default settings of individual ExtremeWare features, refer to individual chapters in this guide.

2

Accessing the Switch

This chapter covers the following topics:

- Understanding the Command Syntax on page 2-1
- Line-Editing Keys on page 2-5
- Command History on page 2-7
- Common Commands on page 2-7
- Configuring Management Access on page 2-10
- Domain Name Service Client Services on page 2-14
- Checking Basic Connectivity on page 2-14

Understanding the Command Syntax

This section describes the steps to take when entering a command. Refer to the sections that follow for detailed information on using the command-line interface.

ExtremeWare command syntax is described in detail in the *ExtremeWare Software Command Reference Guide*. Some commands are also described in this User Guide, in order to describe how to use the features of the ExtremeWare software. However, only a subset of commands are described here, and in some cases only a subset of the options that a command supports. The ExtremeWare Software Command Reference Guide should be considered the definitive source for information on ExtremeWare commands.

When entering a command at the prompt, ensure that you have the appropriate privilege level. Most configuration commands require you to have the administrator privilege level. To use the command-line interface (CLI), follow these steps:

1 Enter the command name.

If the command does not include a parameter or values, skip to step 3. If the command requires more information, continue to step 2.

2 If the command includes a parameter, enter the parameter name and values.

3 The value part of the command specifies how you want the parameter to be set. Values include numerics, strings, or addresses, depending on the parameter.

4 After entering the complete command, press [Return].



Note: If an asterisk () appears in front of the command-line prompt, it indicates that you have outstanding configuration changes that have not been saved. For more information on saving configuration changes, refer to Appendix B.*

Syntax Helper

The CLI has a built-in syntax helper. If you are unsure of the complete syntax for a particular command, enter as much of the command as possible and press [Tab]. The syntax helper provides a list of options for the remainder of the command, and places the cursor at the end of the command you have entered so far, ready for the next option.

If the command is one where the next option is a named component, such as a VLAN, access profile, or route map, the syntax helper will also list any currently configured names that might be used as the next option. In situations where this list might be very long, the syntax helper will list only one line of names, followed by an ellipses to indicate that there are more names than can be displayed.

The syntax helper also provides assistance if you have entered an incorrect command.

Abbreviated Syntax

Abbreviated syntax is the shortest unambiguous allowable abbreviation of a command or parameter. Typically, this is the first three letters of the command. If you do not enter enough letters to allow the switch to determine which command you mean, the syntax helper will provide a list of the options based on the portion of the command you have entered.



Note: When using abbreviated syntax, you must enter enough characters to make the command unambiguous and distinguishable to the switch.

Command Shortcuts

All named components of the switch configuration must have a unique name. Components are typically named using the `create` command. When you enter a command to configure a named component, you do not need to use the keyword of the component. For example, to create a VLAN, you must enter a unique VLAN name:

```
create vlan engineering
```

Once you have created the VLAN with a unique name, you can then eliminate the keyword `vlan` from all other commands that require the name to be entered. For example, instead of entering the modular switch command

```
configure vlan engineering delete port 1:3,4:6
```

you could enter the following shortcut:

```
configure engineering delete port 1:3,4:6
```

Similarly, on the stand-alone switch, instead of entering the command

```
configure vlan engineering delete port 1-3,6
```

you could enter the following shortcut:

```
configure engineering delete port 1-3,6
```

Modular Switch Numerical Ranges

Commands that require you to enter one or more port numbers on a modular switch use the parameter `<portlist>` in the syntax. A `<portlist>` can be one port on a particular slot. For example,

```
port 3:1
```

A `<portlist>` can be a range of numbers. For example,

```
port 3:1-3:3
```

You can add additional slot and port numbers to the list, separated by a comma:

```
port 3:1,4:8,6:10
```

You can specify all ports on a particular slot. For example,

```
port 3:*
```

indicates all ports on slot 3.

You can specify a range of slots and ports. For example,

```
port 2:3-4:5
```

indicates slot 2, port 3 through slot 4, port 5.

Stand-alone Switch Numerical Ranges

Commands that require you to enter one or more port numbers on a stand-alone switch use the parameter `<portlist>` in the syntax. A portlist can be a range of numbers, for example:

```
port 1-3
```

You can add additional port numbers to the list, separated by a comma:

```
port 1-3,6,8
```

Names

All named components of the switch configuration must have a unique name. Names must begin with an alphabetical character and are delimited by whitespace, unless enclosed in quotation marks.

Symbols

You may see a variety of symbols shown as part of the command syntax. These symbols explain how to enter the command, and you do not type them as part of the command itself. Table 2-1 summarizes command syntax symbols.

Table 2-1: Command Syntax Symbols

Symbol	Description
angle brackets < >	Enclose a variable or value. You must specify the variable or value. For example, in the syntax <pre>config vlan <name> ipaddress <ip_address></pre> you must supply a VLAN name for <name> and an address for <ip_address> when entering the command. Do not type the angle brackets.
square brackets []	Enclose a required value or list of required arguments. One or more values or arguments can be specified. For example, in the syntax <pre>use image [primary secondary]</pre> you must specify either the primary or secondary image when entering the command. Do not type the square brackets.
vertical bar	Separates mutually exclusive items in a list, one of which must be entered. For example, in the syntax <pre>config snmp community [read-only read-write] <string></pre> you must specify either the read or write community string in the command. Do not type the vertical bar.
braces { }	Enclose an optional value or a list of optional arguments. One or more values or arguments can be specified. For example, in the syntax <pre>reboot {<date> <time> cancel}</pre> you can specify either a particular date and time combination, or the keyword <code>cancel</code> to cancel a previously scheduled reboot. If you do not specify an argument, the command will prompt, asking if you want to reboot the switch now. Do not type the braces.

Line-Editing Keys

Table 2-2 describes the line-editing keys available using the CLI.

Table 2-2: Line-Editing Keys

Key(s)	Description
Backspace	Deletes character to left of cursor and shifts remainder of line to left.
Delete or [Ctrl] + D	Deletes character under cursor and shifts remainder of line to left.
[Ctrl] + K	Deletes characters from under cursor to end of line.

Table 2-2: Line-Editing Keys (continued)

Key(s)	Description
Insert	Toggles on and off. When toggled on, inserts text and shifts previous text to right.
Left Arrow	Moves cursor to left.
Right Arrow	Moves cursor to right.
Home or [Ctrl] + A	Moves cursor to first character in line.
End or [Ctrl] + E	Moves cursor to last character in line.
[Ctrl] + L	Clears screen and moves cursor to beginning of line.
[Ctrl] + P or Up Arrow	Displays previous command in command history buffer and places cursor at end of command.
[Ctrl] + N or Down Arrow	Displays next command in command history buffer and places cursor at end of command.

Table 2-2: Line-Editing Keys (continued)

Key(s)	Description
[Ctrl] + U	Clears all characters typed from cursor to beginning of line.
[Ctrl] + W	Deletes previous word.

Command History

ExtremeWare “remembers” the last 49 commands you entered. You can display a list of these commands by using the following command:

```
history
```

Common Commands

Table 2-3 describes some of the common commands used to manage the switch. Commands specific to a particular feature may also be described in other chapters of this guide. For a detailed description of the commands and their options, see the *ExtremeWare Software Command Reference Guide*.

Table 2-3: Common Commands

Command	Description
<code>clear session <number></code>	Terminates a Telnet session from the switch.
<code>config account <username></code>	Configures a user account password. The switch will interactively prompt for a new password, and for reentry of the password to verify it. Passwords must have a minimum of 1 character and can have a maximum of 32 characters. User names and passwords are case-sensitive.

Table 2-3: Common Commands (continued)

Command	Description
config banner	Configures the banner string. You can enter up to 24 rows of 79-column text that is displayed before the login prompt of each session. Press [Return] at the beginning of a line to terminate the command and apply the banner. To clear the banner, press [Return] at the beginning of the first line.
config ports <portlist> auto off {speed [10 100 1000]} duplex [half full]	Manually configures the port speed and duplex setting of one or more ports on a switch.
config slot <slot number> module <module name>	Configures a slot for a particular I/O module card.
config ssh2 key {pregenerated}	Generates the SSH2 host key.
config sys-recovery-level [none [critical all] [shutdown reboot]]	Configures a recovery option for instances where an exception occurs in ExtremeWare.
config time <date> <time>	Configures the system date and time. The format is as follows: mm/dd/yyyy hh:mm:ss The time uses a 24-hour clock format. You cannot set the year past 2036.
config timezone <gmt_offset> {autodst noautodst}	Configures the time zone information to the configured offset from GMT time. The format of <code>gmt_offset</code> is +/- minutes from GMT time. The <code>autodst</code> and <code>noautodst</code> options enable and disable automatic Daylight Saving Time change based on the North American standard. Additional options are described in the <i>ExtremeWare Software Command Reference Guide</i> .
config vlan <name> ipaddress <ip_address> {<mask>}	Configures an IP address and subnet mask for a VLAN.
create account [admin user] <username> {<password>}	Creates a user account. This command is available to admin-level users and to users with RADIUS command authorization. The username is between 1 and 32 characters, the password is between 0 and 16 characters.

Table 2-3: Common Commands (continued)

Command	Description
create vlan <name>	Creates a VLAN.
delete account <username>	Deletes a user account.
delete vlan <name>	Deletes a VLAN.
disable bootp vlan [<name> all]	Disables BOOTP for one or more VLANs.
disable cli-config-logging	Disables logging of CLI commands to the Syslog.
disable clipaging	Disables pausing of the screen display when a show command output reaches the end of the page.
disable idletimeout	Disables the timer that disconnects all sessions. Once disabled, console sessions remain open until the switch is rebooted or you logoff. Telnet sessions remain open until you close the Telnet client.
disable ports [<portlist> all]	Disables a port on the switch.
disable ssh2	Disables SSH2 Telnet access to the switch.
disable telnet	Disables Telnet access to the switch.
disable web	Disables Web access to the switch.
enable bootp vlan [<name> all]	Enables BOOTP for one or more VLANs.
enable cli-config-logging	Enables the logging of CLI configuration commands to the Syslog for auditing purposes. The default setting is enabled.
enable clipaging	Enables pausing of the screen display when show command output reaches the end of the page. The default setting is enabled.
enable idletimeout	Enables a timer that disconnects all sessions (both Telnet and console) after 20 minutes of inactivity. The default setting is disabled.
enable license fullL3 <license_key>	Enables a particular software feature license. Specify <license_key> as an integer. The command <code>unconfig switch all</code> does not clear licensing information. This license cannot be disabled once it is enabled on the switch.

Table 2-3: Common Commands (continued)

Command	Description
enable ssh2 {access-profile [<access_profile> none]} {port <tcp_port_number>}	Enables SSH2 Telnet sessions. By default, SSH2 is enabled with no access profile, and uses TCP port number 22. To cancel a previously configured access-profile, use the <code>none</code> option.
enable telnet {access-profile [<access_profile> none]} {port <tcp_port_number>}	Enables Telnet access to the switch. By default, Telnet is enabled with no access profile, and uses TCP port number 23. To cancel a previously configured access-profile, use the <code>none</code> option.
enable web {access-profile [<access_profile> none]} {port <tcp_port_number>}	Enables ExtremeWare Vista Web access to the switch. By default, Web access is enabled with no access profile, using TCP port number 80. Use the <code>none</code> option to cancel a previously configured access-profile. You must reboot the switch for this command to take effect.
history	Displays the previous 49 commands entered on the switch.
show banner	Displays the user-configured banner.
unconfig switch {all}	Resets all switch parameters (with the exception of defined user accounts, and date and time information) to the factory defaults. If you specify the keyword <code>all</code> , the switch erases the currently selected configuration image in flash memory and reboots. As a result, all parameters are reset to default settings.

Configuring Management Access

ExtremeWare supports the following two levels of management:

- User
- Administrator

In addition to the management levels, you can optionally use an external RADIUS server to provide CLI command authorization checking for each command. For more information on RADIUS, refer to “RADIUS Client” in Chapter 3.

User Account

A user-level account has viewing access to all manageable parameters, with the exception of:

- User account database.
- SNMP community strings.

A user-level account can use the `ping` command to test device reachability, and change the password assigned to the account name. If you have logged on with user capabilities, the command-line prompt ends with a (>) sign. For example:

```
Summit1:2>
```

Administrator Account

An administrator-level account can view and change all switch parameters. It can also add and delete users, and change the password associated with any account name. The administrator can disconnect a management session that has been established by way of a Telnet connection. If this happens, the user logged on by way of the Telnet connection is notified that the session has been terminated.

If you have logged on with administrator capabilities, the command-line prompt ends with a (#) sign. For example:

```
Summit1:18#
```

Prompt Text

The prompt text is taken from the SNMP `sysname` setting. The number that follows the colon indicates the sequential line/command number.

If an asterisk (*) appears in front of the command-line prompt, it indicates that you have outstanding configuration changes that have not been saved. For example:

```
*Summit1:19#
```

Default Accounts

By default, the switch is configured with two accounts, as shown in Table 2-4.

Table 2-4: Default Accounts

Account Name	Access Level
admin	This user can access and change all manageable parameters. The admin account cannot be deleted.
user	This user can view (but not change) all manageable parameters, with the following exceptions: <ul style="list-style-type: none"> ■ This user cannot view the user account database. ■ This user cannot view the SNMP community strings.

Changing the Default Password

Default accounts do not have passwords assigned to them. Passwords must have a minimum of four characters and can have a maximum of 12 characters.



Note: User names and passwords are case-sensitive.

To add a password to the default admin account, follow these steps:

- 1 Log in to the switch using the name *admin*.
- 2 At the password prompt, press [Return].
- 3 Add a default admin password by entering the following command:

```
config account admin
```
- 4 Enter the new password at the prompt.
- 5 Re-enter the new password at the prompt.

To add a password to the default user account, follow these steps:

- 1 Log in to the switch using the name *admin*.
- 2 At the password prompt, press [Return], or enter the password that you have configured for the *admin* account.
- 3 Add a default user password by entering the following command:

```
config account user
```
- 4 Enter the new password at the prompt.

- 5 Re-enter the new password at the prompt.



Note: If you forget your password while logged out of the command-line interface, contact your local technical support representative, who will advise on your next course of action.

Creating a Management Account

The switch can have a total of 16 management accounts. You can use the default names (*admin* and *user*), or you can create new names and passwords for the accounts. Passwords can have a minimum of 0 characters and can have a maximum of 31 characters.

To create a new account, follow these steps:

- 1 Log in to the switch as *admin*.
- 2 At the password prompt, press [Return], or enter the password that you have configured for the *admin* account.
- 3 Add a new user by using the following command:


```
create account [admin | user] <username>
```
- 4 Enter the password at the prompt.
- 5 Re-enter the password at the prompt.

Viewing Accounts

To view the accounts that have been created, you must have administrator privileges. Use the following command to see the accounts:

```
show accounts
```

Deleting an Account

To delete a account, you must have administrator privileges. To delete an account, use the following command:

```
delete account <username>
```



Note: The account name admin cannot be deleted.

Domain Name Service Client Services

The Domain Name Service (DNS) client in ExtremeWare augments the following commands to allow them to accept either IP addresses or host names:

- telnet
- download [bootrom | configuration | image]
- upload configuration
- ping
- traceroute

In addition, the `nslookup` utility can be used to return the IP address of a hostname.

You can specify up to eight DNS servers for use by the DNS client using the following command:

```
config dns-client add <ipaddress>
```

You can specify a default domain for use when a host name is used without a domain. Use the following command:

```
config dns-client default-domain <domain name>
```

For example, if you specify the domain “xyz-inc.com” as the default domain, then a command such as `ping accounting1` will be taken as if it had been entered `ping accounting1.xyz-inc.com`.

Checking Basic Connectivity

The switch offers the following commands for checking basic connectivity:

- ping
- traceroute

Ping

The `ping` command enables you to send Internet Control Message Protocol (ICMP) echo messages to a remote IP device. The `ping` command is available for both the user and administrator privilege level.

The ping command syntax is:

```
ping {udp} {continuous} {size <start_size> {- <end_size>}} [<ip_address>
| <hostname>] {from <src_address> | with record-route | from
<src_ipaddress> with record-route}
```

Options for the ping command are described in Table 2-5.

Table 2-5: Ping Command Parameters

Parameter	Description
udp	Specifies that UDP messages should be sent instead of ICMP echo messages. When specified, <code>from</code> and <code>with record-route</code> options are not supported.
continuous	Specifies ICMP echo messages to be sent continuously. This option can be interrupted by pressing any key.
size	Specifies the size of the ICMP request. If both the <code>start_size</code> and <code>end_size</code> are specified, transmits ICMP requests using 1 byte increments, per packet. If no <code>end_size</code> is specified, packets of <code>start_size</code> are sent.
<ipaddress>	Specifies the IP address of the host.
<hostname>	Specifies the name of the host. To use the <code>hostname</code> , you must first configure DNS.
from	Uses the specified source address in the ICMP packet. If not specified, the address of the transmitting interface is used.
with record-route	Decodes the list of recorded routes and displays them when the ICMP echo reply is received.

If a ping request fails, the switch continues to send ping messages until interrupted. Press any key to interrupt a ping request.

Traceroute

The `traceroute` command enables you to trace the routed path between the switch and a destination endstation. The `traceroute` command syntax is:

```
traceroute [<ip_address> | <hostname>] {from <src_ipaddress>} {ttl
<TTL>} {port <port>}
```

where:

- `ip_address` is the IP address of the destination endstation.

- `hostname` is the hostname of the destination endstation. To use the hostname, you must first configure DNS.
- `from` uses the specified source address in the ICMP packet. If not specified, the address of the transmitting interface is used.
- `t11` configures the switch to trace up to the time-to-live number of the switch.
- `port` uses the specified UDP port number.



3 Managing the Switch

This chapter covers the following topics:

- Overview on page 3-1
- Using the Console Interface on page 3-2
- Using Telnet on page 3-3
- Using Secure Shell 2 (SSH2) on page 3-7
- Using ExtremeWare Vista on page 3-11
- Using SNMP on page 3-17
- Authenticating Users on page 3-20
- Using Network Login on page 3-28
- Using the Simple Network Time Protocol on page 3-34

Overview

Using ExtremeWare, you can manage the switch using the following methods:

- Access the CLI by connecting a terminal (or workstation with terminal-emulation software) to the console port.
- Access the switch remotely using TCP/IP through one of the switch ports or through the dedicated 10/100 unshielded twisted pair (UTP) Ethernet management port (on switches that are so equipped). Remote access includes:
 - Telnet using the CLI interface.

- SSH2 using the CLI interface.
- ExtremeWare Vista Web access using a standard Web browser.
- SNMP access using EPICenter or another SNMP manager.

The switch supports up to the following number of concurrent user sessions:

- One console session
 - Two console sessions are available on a modular switch that has two management modules installed.
- Eight Telnet sessions
- Eight SSH2 sessions
- One Web session

Using the Console Interface

The CLI built into the switch is accessible by way of the 9-pin, RS-232 port labeled *console*, located on the back of the stand-alone switch, or on the front of the modular switch management module.



Note: For more information on the console port pinouts, refer to the hardware installation guide that shipped with your switch.

Once the connection is established, you will see the switch prompt and you can log in.

Using the 10/100 UTP Management Port

Some Extreme switch models provide a dedicated 10/100 UTP management port. This port provides dedicated remote access to the switch using TCP/IP. It supports the following management methods:

- Telnet using the CLI interface
- ExtremeWare Vista Web access using a standard Web browser
- SNMP access using EPICenter or another SNMP manager

The management port is a DTE port, and is not capable of supporting switching or routing functions. The TCP/IP configuration for the management port is done using the

same syntax as used for VLAN configuration. The VLAN *mgmt* comes preconfigured with only the 10/100 UTP management port as a member.

You can configure the IP address, subnet mask, and default router for the VLAN *mgmt*, using the following commands:

```
config vlan mgmt ipaddress <ip_address>/<subnet_mask>
config iproute add default <gateway>
```

Using Telnet

Any workstation with a Telnet facility should be able to communicate with the switch over a TCP/IP network.

Up to eight active Telnet sessions can access the switch concurrently. If `idletimeouts` are enabled, the Telnet connection will time out after 20 minutes of inactivity. If a connection to a Telnet session is lost inadvertently, the switch terminates the session within two hours.

Before you can start a Telnet session, you must set up the IP parameters described in “Configuring Switch IP Parameters” later in this chapter. Telnet is enabled by default.

To open the Telnet session, you must specify the IP address of the device that you want to manage. Check the user manual supplied with the Telnet facility if you are unsure of how to do this.

Once the connection is established, you will see the switch prompt and you may log in.

Connecting to Another Host Using Telnet

You can Telnet from the current CLI session to another host using the following command:

```
telnet [<ipaddress> | <hostname>] {<port_number>}
```

If the TCP port number is not specified, the Telnet session defaults to port 23. Only VT100 emulation is supported.

Configuring Switch IP Parameters

To manage the switch by way of a Telnet connection or by using an SNMP Network Manager, you must first configure the switch IP parameters.

Using a BOOTP Server

If you are using IP and you have a Bootstrap Protocol (BOOTP) server set up correctly on your network, you must add the following information to the BOOTP server:

- Switch Media Access Control (MAC) address, found on the rear label of the switch
- IP address
- Subnet address mask (optional)

Once this is done, the IP address and subnet mask for the switch will be downloaded automatically. You can then start managing the switch without further configuration.

You can enable BOOTP on a per-VLAN basis by using the following command:

```
enable bootp vlan [<name> | all]
```

By default, BOOTP is enabled on the *default* VLAN.

If you configure the switch to use BOOTP, the switch IP address is not retained through a power cycle, even if the configuration has been saved. To retain the IP address through a power cycle, you must configure the IP address of the VLAN using the command-line interface, Telnet, or Web interface.

All VLANs within a switch that are configured to use BOOTP to get their IP address use the same MAC address. Therefore, if you are using BOOTP relay through a router, the BOOTP server must be capable of differentiating its relay based on the gateway portion of the BOOTP packet.



Note: For more information on DHCP/BOOTP relay, refer to Chapter 16.

Manually Configuring the IP Settings

If you are using IP without a BOOTP server, you must enter the IP parameters for the switch in order for the SNMP Network Manager, Telnet software, or Web interface to communicate with the device. To assign IP parameters to the switch, you must perform the following tasks:

- Log in to the switch with administrator privileges.

- Assign an IP address and subnet mask to a VLAN.

The switch comes configured with a default VLAN named *default*. To use Telnet or an SNMP Network Manager, you must have at least one VLAN on the switch, and it must be assigned an IP address and subnet mask. IP addresses are always assigned to a VLAN. The switch can be assigned multiple IP addresses.



Note: For information on creating and configuring VLANs, refer to Chapter 5.

To manually configure the IP settings, follow these steps:

- 1 Connect a terminal or workstation running terminal-emulation software to the console port.
- 2 At your terminal, press [Return] one or more times until you see the login prompt.
- 3 At the login prompt, enter your user name and password. Note that they are both case-sensitive. Ensure that you have entered a user name and password with administrator privileges.

- If you are logging in for the first time, use the default user name *admin* to log in with administrator privileges. For example:

```
login: admin
```

Administrator capabilities enable you to access all switch functions. The default user names have no passwords assigned.

- If you have been assigned a user name and password with administrator privileges, enter them at the login prompt.

- 4 At the password prompt, enter the password and press [Return].

When you have successfully logged in to the switch, the command-line prompt displays the name of the switch in its prompt.

- 5 Assign an IP address and subnetwork mask for the default VLAN by using the following command:

```
config vlan <name> ipaddress <ipaddress> {<subnet_mask>}
```

For example:

```
config vlan default ipaddress 123.45.67.8 255.255.255.0
```

Your changes take effect immediately.



Note: As a general rule, when configuring any IP addresses for the switch, you can express a subnet mask by using dotted decimal notation, or by using classless inter-domain routing notation (CIDR). CIDR uses a forward slash plus the number of bits in the subnet mask. Using CIDR notation, the command identical to the one above would be:

```
config vlan default ipaddress 123.45.67.8 / 24
```

6 Configure the default route for the switch using the following command:

```
config iproute add default <gateway> {<metric>}
```

For example:

```
config iproute add default 123.45.67.1
```

7 Save your configuration changes so that they will be in effect after the next switch reboot, by typing:

```
save
```

8 When you are finished using the facility, log out of the switch by typing:

```
logout or quit
```

Disconnecting a Telnet Session

An administrator-level account can disconnect a Telnet management session. If this happens, the user logged in by way of the Telnet connection is notified that the session has been terminated.

To terminate a Telnet session, follow these steps:

- 1** Log in to the switch with administrator privileges.
- 2** Determine the session number of the session you want to terminate by using the following command:

```
show session
```

- 3** Terminate the session by using the following command:

```
clear session <session_number>
```

Controlling Telnet Access

By default, Telnet services are enabled on the switch. Telnet access can be restricted by the use of an access profile. An access profile permits or denies a named list of IP

addresses and subnet masks. To configure Telnet to use an access profile, use the following command:

```
enable telnet {access-profile [<access_profile> | none]} {port
<tcp_port_number>}
```

Use the `none` option to remove a previously configured access profile.

To display the status of Telnet, use the following command:

```
show management
```

You can choose to disable Telnet by using the following command:

```
disable telnet
```

To re-enable Telnet on the switch, at the console port use the following:

```
enable telnet
```

You must be logged in as an administrator to enable or disable Telnet.



Note: For more information on Access Profiles, refer to Chapter 8.

Using Secure Shell 2 (SSH2)

Secure Shell 2 (SSH2) is a feature of ExtremeWare that allows you to encrypt Telnet session data between a network administrator using SSH2 client software and the switch, or to send encrypted data from the switch to an SSH2 client on a remote system. Image and configuration files may also be transferred to the switch using the Secure Copy Protocol 2 (SCP2). The ExtremeWare CLI provides a command that enable the switch to function as an SSH2 client, sending commands to a remote system via an SSH2 session. It also provides commands to copy image and configuration files to the switch using the SCP2.

The ExtremeWare SSH2 switch application is based on the Data Fellows™ SSH2 server implementation. It is highly recommended that you use the F-Secure® SSH client products from Data Fellows corporation. These applications are available for most operating systems. For more information, refer to the Data Fellows website at:

<http://www.datafellows.com>.



Note: SSH2 is compatible with the Data Fellows SSH2 client version 2.0.12 or above. SSH2 is not compatible with SSH1.

The ExtremeWare SSH2 switch application also works with SSH2 client and server (version 2.x or later) from SSH Communication Security, and the free SSH2 and SCP2 implementation (version 2.5 or later) from OpenSSH. The SFTP file transfer protocol is required for file transfer using SCP2.

Enabling SSH2 for Inbound Switch Access

Because SSH2 is currently under U.S. export restrictions, you must first obtain a security-enabled version of the ExtremeWare software from Extreme Networks before you can enable SSH2. The procedure for obtaining a security-enabled version of the ExtremeWare software is described in Chapter 1.

You must enable SSH2 on the switch before you can connect to it using an external SSH2 client. Enabling SSH2 involves two steps:

- Enabling SSH2 access, which may include specifying a list of clients that can access the switch, and specifying a TCP port to be used for communication.
By default, if you have a security license, SSH2 is enabled using TCP port 22, with no restrictions on client access.
- Generating or specifying an authentication key for the SSH2 session.

To enable SSH2, use the following command:

```
enable ssh2 {access-profile [<access_profile> | none] {port  
<tcp_port_number>}}
```

You can specify a list of predefined clients that are allowed SSH2 access to the switch. To do this, you must create an access profile that contains a list of allowed IP addresses. For more information on creating access profiles, refer to Chapter 8.

You can also specify a TCP port number to be used for SSH2 communication. By default the TCP port number is 22.

The supported cipher is 3DES-CBC. The supported key exchange is DSA.

An authentication key must be generated before the switch can accept incoming SSH2 sessions. This can be done automatically by the switch, or you can enter a previously generated key. To have the key generated by the switch, use the following command:

```
config ssh2 key
```

You are prompted to enter information to be used in generating the key. The key generation process takes approximately ten minutes. Once the key has been generated, you should save your configuration to preserve the key.

To use a key that has been previously created, use the following command:

```
config ssh2 key pregenerated
```

You are prompted to enter the pregenerated key.

The key generation process generates the SSH2 private host key. The SSH2 public host key is derived from the private host key, and is automatically transmitted to the SSH2 client at the beginning of an SSH2 session.

Before you initiate a session from an SSH2 client, ensure that the client is configured for any nondefault access list or TCP port information that you have configured on the switch. Once these tasks are accomplished, you may establish an SSH2-encrypted session with the switch. Clients must have a valid user name and password on the switch in order to log into the switch after the SSH2 session has been established.

For additional information on the SSH protocol refer to [FIPS-186] Federal Information Processing Standards Publication (FIPSPUB) 186, Digital Signature Standard, 18 May 1994. This can be download from: <ftp://ftp.cs.hut.fi/pub/ssh>. General technical information is also available from <http://www.ssh.fi>.

Using SCP2 from an External SSH2 Client

In ExtremeWare version 6.2.1 or later, the SCP2 protocol is supported for transferring image and configuration files to the switch from the SSH2 client, and for copying the switch configuration from the switch to an SSH2 client.

The user must have administrator-level access to the switch. The switch can be specified by its switch name or IP address.

Configuration or image files stored on the system running the SSH2 client may be named as desired by the user. However, files on the switch have predefined names, as follows:

- `configuration.cfg`—The current configuration
- `incremental.cfg`—The current incremental configuration
- `primary.img`—The primary ExtremeWare image
- `secondary.img`—The secondary ExtremeWare image
- `bootrom.img`—The BootROM image

For example, to copy an image file saved as *image1.xtr* to switch with IP address 10.10.0.5 as the primary image using SCP2, you would enter the following command within your SSH2 session:

```
scp image1.xtr admin@10.20.0.5:primary.img
```

To copy the configuration from the switch and save it in file *config1.save* using SCP, you would enter the following command within your SSH2 session:

```
scp admin@10.10.0.5:configuration.cfg config1.save
```

SSH2 Client Functions on the Switch

In ExtremeWare version 6.2.1 or later, an Extreme Networks switch can function as an SSH2 client. This means you can connect from the switch to a remote device running an SSH2 server, and send commands to that device. You can also use SCP2 to transfer files to and from the remote device.

You do not need to enable SSH2 or generate an authentication key to use the SSH2 and SCP2 commands from the ExtremeWare CLI.

To send commands to a remote system using SSH2, use the following command:

```
ssh2 {cipher [3des | blowfish]} {port <portnum>} {compression [on | off]} {user <username>} {debug <debug_level>} {<username>}@<host> | <ipaddress> <remote commands>
```

The remote commands can be any commands acceptable by the remote system. You can specify the login user name as a separate argument, or as part of the `user@host` specification. If the login user name for the remote system is the same as your user name on the switch, you can omit the username parameter entirely.

To initiate a file copy from a remote system to the switch using SCP2, use the following command:

```
scp2 {cipher [3des | blowfish]} {port <portnum>} {debug <debug_level>}
<user>@[<hostname> | <ipaddress>]:<remote_file> [configuration
{incremental} | image [primary | secondary] | bootrom]
```

To initiate a file copy to a remote system from the switch using SCP2, use the following command:

```
scp2 {cipher [3des | blowfish]} {port <portnum>} {debug <debug_level>}
configuration <user>@[<hostname> | <ipaddress>]:<remote_file>
```

Using ExtremeWare Vista

ExtremeWare Vista is device-management software running in the switch that allows you to access the switch over a TCP/IP network using a standard Web browser. Any properly configured standard Web browser that supports frames and JavaScript (such as Netscape Navigator 3.0 or above, or Microsoft Internet Explorer 3.0 or above) can be used to manage the switch.

ExtremeWare Vista provides a subset of the command-line interface (CLI) commands available for configuring and monitoring the switch. If a particular command is not available using ExtremeWare Vista, you must use the CLI to access the desired functionality.

To use ExtremeWare Vista, at least one VLAN must be assigned an IP address.



Note: For more information on assigning an IP address, refer to “Configuring Switch IP Parameters” on page 3-4.

The default home page of the switch can be accessed using the following command:

```
http://<ipaddress>
```

When you access the home page of the switch, you are presented with the Logon screen.

Controlling Web Access

By default, Web access is enabled on the switch. Use of ExtremeWare Vista Web access can be restricted through the use of an access profile. An access profile permits or denies a named list of IP addresses and subnet masks. To configure Vista Web access to use an access profile, use the following command:

```
enable web {access-profile <access-profile> | none} {port  
<tcp_port_number>}
```

Use the `none` option to remove a previously configured access profile. Apply an access profile only when ExtremeWare Vista is enabled.

To display the status of Web access, use the following command:

```
show management
```

To disable ExtremeWare Vista, use the following command:

```
disable web
```

To re-enable Web access, use the `enable web` command.

By default, web access uses TCP port 80. To specify a different port, use the `port` option in the `enable web` command.

You will need to reboot the system in order for these changes to take effect.



Note: For more information on rebooting, refer to Appendix B.

Setting Up Your Browser

In general, the default settings that come configured on your browser work well with ExtremeWare Vista. The following are recommended settings that you can use to improve the display features and functionality of ExtremeWare Vista:

- After downloading a newer version of the switch image, clear the browser disk and memory cache to see the updated menu screens. You must clear the cache while at the main ExtremeWare Vista Logon screen, so that all underlying .GIF files are updated.
- Check for newer versions of stored pages. Every visit to the page should be selected as a cache setting.

If you are using Netscape Navigator, configure the cache option to check for changes “Every Time” you request a page.

If you are using Microsoft Internet Explorer, configure the Temporary Internet Files setting to check for newer versions of stored pages by selecting “Every visit to the page.”

- Images must be auto-loaded.
- Use a high-resolution monitor to maximize the amount of information displayed in the content frame. The recommended resolution is 1024 x 768 pixels. You can also use 800 x 600 pixels.
- Turn off one or more of the browser toolbars to maximize the viewing space of the ExtremeWare Vista content screen.
- If you will be using ExtremeWare Vista to send an e-mail to the Extreme Networks Technical Support department, configure the e-mail settings in your browser.
- Configure the browser to use the following recommended fonts:
 - Proportional font—Times New Roman
 - Fixed-width font—Courier New

Accessing ExtremeWare Vista

To access the default home page of the switch, enter the following URL in your browser:

```
http://<ip_address>
```

When you access the home page of the system, you are presented with the Login screen. Enter your user name and password in the appropriate fields, and click OK.

If you have entered the name and password of an administrator-level account, you have access to all ExtremeWare Vista pages. If you have used a user-level account name and password, you only have access to the Statistics and Support information.

If multiple people access the same switch using ExtremeWare Vista, you might see the following error message:

```
Web:server busy
```

To correct this situation, log out of the switch and log in again.

Navigating ExtremeWare Vista

After logging in to the switch, the ExtremeWare Vista home page is displayed.

ExtremeWare Vista divides the browser screen into the following sections:

- Task frame
- Content frame
- Standalone buttons

Task Frame

The task frame has two sections: menu buttons and submenu links. The four task menu buttons are:

- Configuration
- Statistics
- Support
- Logout

Below the task buttons are options. Options are specific to the task button that you select. When you select an option, the information displayed in the content frame changes. However, when you select a new task button, the content frame does not change until you select a new option.



Note: Submitting a configuration page with no change will result in an asterisk () appearing at the CLI prompt, even though actual configuration values have not changed.*

Content Frame

The content frame contains the main body of information in ExtremeWare Vista. For example, if you select an option from the Configuration task button, enter configuration parameters in the content frame. If you select the Statistics task button, statistics are displayed in the content frame.

Browser Controls. Browser controls include drop-down list boxes, check boxes, and multiselect list boxes. A multiselect list box has a scrollbar on the right side of the box. Using a multiselect list box, you can select a single item, all items, a set of contiguous items, or multiple noncontiguous items. Table 3-1 describes how to make selections from a multiselect list box.

Table 3-1: Multiselect List Box Key Definitions

Selection Type	Key Sequence
Single item	Click the item using the mouse.
All items	Click the first item, and drag to the last item.
Contiguous items	Click the first desired item, and drag to the last desired item.
Selected noncontiguous items	Hold down [Ctrl], click the first desired item, click the next desired item, and so on.

Status Messages

Status messages are displayed at the top of the content frame. The four types of status messages are:

- **Information** — Displays information that is useful to know prior to, or as a result of, changing configuration options.
- **Warning** — Displays warnings about the switch configuration.
- **Error** — Displays errors caused by incorrectly configured settings.
- **Success** — Displays informational messages after you click Submit. The message displayed reads, “Request was submitted successfully.”

Standalone Buttons

At the bottom of some of the content frames is a section that contains standalone buttons. Standalone buttons are used to perform tasks that are not associated with a particular configuration option. An example of this is the Reboot Switch button.

Saving Changes

You can save your changes to nonvolatile storage in either of two ways using ExtremeWare Vista:

- Select Save Configuration from the Configuration task button, Switch option.

This field contains a drop-down list box that allows you to select either the primary or secondary configuration area. After you select the configuration area, click Submit to save the changes.

- Click the Logout button.

If you attempt to log out without saving your changes, ExtremeWare Vista prompts you to save your changes.

If you select Yes, the changes are saved to the selected configuration area. To change the selected configuration area, you must go to the Configuration task button, Switch option.

Filtering Information

Some pages have a Filter button. The Filter button is used to display a subset of information on a given page. For example, on the OSPF configuration page, you can configure authentication based on the VLAN, area identifier, or virtual link. Once you select a filtering option and click the Filter button, the form that provides the configuration options displays the available interfaces in the drop-down menu, based on your filtering selection.

Similarly, in certain Configuration and Statistics pages, information is shown based on a particular slot.

Because modular switches allow you to preconfigure modules without having them physically available in the chassis, the configuration pages offer a drop-down menu to select any module card that has been configured on the system, whether or not the module is physically available. By default, information for the first configured module that is found in the chassis is displayed on the page. You can configure available slots and ports by filtering on a selected module from the Sort by Slot drop-down menu.

On the Statistics pages, you can only view information for cards that are configured and physically inserted into the chassis. On these pages, the Sort by Slot drop-down menu displays only these modules.

Do a GET When Configuring a VLAN

When configuring a VLAN using ExtremeWare Vista, prior to editing the VLAN configuration, you must first click the `get` button to ensure that subsequent edits are applied to the correct VLAN. If you do not click the `get` button and you submit the changes, the changes will be made to the VLAN that was previously displayed.

If you configure a VLAN and then delete it, the *default* VLAN is shown in the VLAN name window, but the VLAN information contained in the lower portion of the page is not updated. Click the `get` button to update the display.

Sending Screen Output to Extreme Networks

If Extreme Networks requests that you e-mail the output of a particular ExtremeWare Vista screen, follow these steps:

- 1 Click on the content frame of the screen that you must send.
- 2 From Netscape Navigator, select Save Frame As from the File menu, and enter a name for the file.

From Microsoft Internet Explorer 3.0, select Save As File from the File menu, and enter a name for the file.

From Microsoft Internet Explorer 4.0, right-click in the content frame, select View Source, and save the HTML text by copying it and pasting it into a text editor.
- 3 Attach the file to the e-mail message that you are sending to Extreme Networks.

Using SNMP

Any Network Manager running the Simple Network Management Protocol (SNMP) can manage the switch, provided the Management Information Base (MIB) is installed correctly on the management station. Each Network Manager provides its own user interface to the management facilities.

The following sections describe how to get started if you want to use an SNMP manager. It assumes you are already familiar with SNMP management. If not, refer to the following publication:

The Simple Book
by Marshall T. Rose
ISBN 0-13-8121611-9
Published by Prentice Hall.

Accessing Switch Agents

To have access to the SNMP agent residing in the switch, at least one VLAN must have an IP address assigned to it.

By default, SNMP access and SNMP traps are enabled. SNMP access and SNMP traps can be disabled and enabled independently—you can disable SNMP access but still allow SNMP traps to be sent, or vice versa.

Supported MIBs

In addition to private MIBs, the switch supports the standard MIBs listed in Appendix A.

Configuring SNMP Settings

The following SNMP parameters can be configured on the switch:

- **Authorized trap receivers** — An authorized trap receiver can be one or more network management stations on your network. The switch sends SNMP traps to all trap receivers. You can have a maximum of 16 trap receivers configured for each switch, and you can specify a community string and UDP port for individually for each trap receiver. All community strings must also be added to the switch using the `config snmp add community` command.

To configure a trap receiver on a switch, use the following command:

```
config snmp add trapreceiver <ip address> {port <udp_port>} community  
<community string> {from <source ip address>}
```

You can delete a trap receiver using the `config snmp delete trapreceiver` command.

Entries in the trap receiver list can also be created, modified, and deleted using the `RMON2 trapDestTable` MIB variable, as described in RFC 2021.

- **SNMP read access** — The ability to read SNMP information can be restricted through the use of an access profile. An access profile permits or denies a named list of IP addresses and subnet masks.

To configure SNMP read access to use an access profile, use the following command:

```
config snmp access-profile readonly [<access_profile> | none]
```

Use the `none` option to remove a previously configured access profile.

- **SNMP read/write access** — The ability to read and write SNMP information can be restricted through the use of an access profile. An access profile permits or denies a named list of IP addresses and subnet masks.

To configure SNMP read/write access to use an access profile, use the following command:

```
config snmp access-profile readwrite [<access_profile> | none]
```

Use the `none` option to remove a previously configured access-profile.

- **Community strings** — The community strings allow a simple method of authentication between the switch and the remote Network Manager. There are two

types of community strings on the switch. Read community strings provide read-only access to the switch. The default read-only community string is *public*. Read-write community strings provide read and write access to the switch. The default read-write community string is *private*. Up to 15 read and 15 read-write community strings can be configured on the switch. The community string for all authorized trap receivers must be configured on the switch for the trap receiver to receive switch-generated traps. SNMP community strings can contain up to 127 characters. To add a community string, use the command:

```
config snmp add community [readonly | readwrite] <string>
```

You can also change the value of the default community strings using the command:

```
config snmp community [readonly | readwrite] <string>
```

- **System contact** (optional) — The system contact is a text field that enables you to enter the name of the person(s) responsible for managing the switch.
- **System name** — The system name is the name that you have assigned to this switch. The default name is the model name of the switch (for example, Summit1 switch).
- **System location** (optional) — Using the system location field, you can enter an optional location for this switch.

Displaying SNMP Settings

To display the SNMP settings configured on the switch, use the following command:

```
show management
```

This command displays the following information:

- Enable/disable state for Telnet, SSH2, SNMP, and Web access, along with access profile information
- SNMP community strings
- Authorized SNMP station list
- SNMP trap receiver list
- RMON polling configuration
- Login statistics

Authenticating Users

ExtremeWare provides two methods to authenticate users who login to the switch:

- RADIUS client
- TACACS+

RADIUS Client

Remote Authentication Dial In User Service (RADIUS, RFC 2138) is a mechanism for authenticating and centrally administrating access to network nodes. The ExtremeWare RADIUS client implementation allows authentication for Telnet, Vista, or console access to the switch.



Note: You cannot configure RADIUS and TACACS+ at the same time.

You can define a primary and secondary RADIUS server for the switch to contact. When a user attempts to login using Telnet, http, or the console, the request is relayed to the primary RADIUS server, and then to the secondary RADIUS server, if the primary does not respond. If the RADIUS client is enabled, but access to the RADIUS primary and secondary server fails, the switch uses its local database for authentication.

The privileges assigned to the user (admin versus nonadmin) at the RADIUS server take precedence over the configuration in the local switch database.

To configure the RADIUS servers, use the following command:

```
config radius server [primary | secondary] <ipaddress> <UDP_port>  
client-ip <ipaddress>
```

Configuring the Shared Secret Password

In addition to specifying the RADIUS server IP information, RADIUS also contains a means to verify communication between network devices and the server. The *shared secret* is a password configured on the network device and RADIUS server, used by each to verify communication.

To configure the shared secret for RADIUS servers, use the following command:

```
config radius [primary | secondary] shared-secret {encrypted} <string>
```

Enabling and Disabling RADIUS

After server information is entered, you can start and stop RADIUS authentication as many times as necessary without needing to reconfigure server information.

To enable RADIUS authentication, use the following command:

```
enable radius
```

To disable RADIUS authentication, use the following command:

```
disable radius
```

Configuring RADIUS Accounting

Extreme switches are capable of sending RADIUS accounting information. As with RADIUS authentication, you can specify two servers for receipt of accounting information. You can configure RADIUS accounting servers to be the same as the authentication servers, but this is not required.

To specify RADIUS accounting servers, use the following command:

```
config radius-accounting [primary | secondary] server [<ipaddress> |  
<hostname>] {<udp_port>} client-ip <ipaddress>
```

RADIUS accounting also makes use of the shared secret password mechanism to validate communication between network access devices and RADIUS accounting servers.

To specify shared secret passwords for RADIUS accounting servers, use the following command:

```
config radius [primary | secondary] shared-secret {encrypted} <string>
```

After you configure RADIUS accounting server information, you must enable accounting before the switch begins transmitting the information. You must enable RADIUS authentication for accounting information to be generated. You can enable and disable accounting without affecting the current state of RADIUS authentication.

To enable RADIUS accounting, use the following command:

```
enable radius-accounting
```

To disable RADIUS accounting, use the following command:

```
disable radius-accounting
```

Per-Command Authentication Using RADIUS

The RADIUS implementation can be used to perform per-command authentication. Per-command authentication allows you to define several levels of user capabilities by controlling the permitted command sets based on the RADIUS username and password. You do not need to configure any additional switch parameters to take advantage of this capability. The RADIUS server implementation automatically negotiates the per-command authentication capability with the switch. For examples on per-command RADIUS configurations, refer to the next section.

Configuring RADIUS Client

You can define primary and secondary server communication information, and for each RADIUS server, the RADIUS port number to use when talking to the RADIUS server. The default port value is 1645. The client IP address is the IP address used by the RADIUS server for communicating back to the switch.

RADIUS RFC 2138 Attributes

The RADIUS RFC 2138 optional attributes supported are as follows:

- User-Name
- User-Password
- Service-Type
- Login-IP-Host

Using RADIUS Servers with Extreme Switches

Extreme Networks switches have two levels of user privilege:

- Read-only
- Read-write

Because there are no CLI command available to modify the privilege level, access rights are determined when you log in. For a RADIUS server to identify the administrative privileges of a user, Extreme switches expect a RADIUS server to transmit the Service-Type attribute in the Access-Accept packet, after successfully authenticating the user.

Extreme switches grant a RADIUS-authenticated user read-write privilege if a Service-Type value of 6 is transmitted as part of the Access-Accept message from the Radius server. Other Service-Type values, or no value, results in the switch granting

read-only access to the user. Different implementations of RADIUS handle attribute transmission differently. You should consult the documentation for your specific implement of RADIUS when you configure users for read-write access.

Cistron RADIUS

Cistron RADIUS is a popular server, distributed under GPL. Cistron RADIUS can be found at <http://www.miquels.cistron.nl/radius/>. When you configure the Cistron server for use with Extreme switches, you must pay close attention to the users file setup. The Cistron RADIUS dictionary associates the word Administrative-User with Service-Type value 6, and expects the Service-Type entry to appear alone on one line with a leading tab character.

The following is a user file example for read-write access:

```
adminuser  Auth-Type = System
           Service-Type = Administrative-User,
           Filter-Id = "unlim"
```

Livingston (Lucent) RADIUS

Livingston RADIUS is produced by Lucent Technologies primarily for use with their portmaster products. Version 2.1 is released under a BSD license agreement and can be found at <ftp://ftp.livingston.com/pub/le/radius/radius21.tar.Z>. As with Cistron RADIUS, the Livingston server default dictionary associates Administrative-User with Service-Type value 6. The administrative users file entry example for Cistron RADIUS also works with Livingston RADIUS.

RSA Ace

For users of their SecureID product, RSA offers RADIUS capability as part of their ACE server software. With some versions of ACE, the RADIUS shared-secret is incorrectly sent to the switch resulting in an inability to authenticate. As a work around, do *not* configure a shared-secret for RADIUS accounting and authentication servers on the switch.

Extreme RADIUS

Extreme Networks provides its users, free of charge, a radius server based on Merit RADIUS. Extreme RADIUS provides per-command authentication capabilities in addition to the standard set of radius features. Source code for Extreme RADIUS can be

obtained from the Extreme Networks Technical Assistance Center and has been tested on Red Hat Linux and Solaris.

When Extreme RADIUS is up and running, the two most commonly changed files will be users and profiles. The users file contains entries specifying login names and the profiles used for per-command authentication after they have logged in. Sending a HUP signal to the RADIUS process is sufficient to get changes in the users file to take place. Extreme RADIUS uses the file named profiles to specify command lists that are either permitted or denied to a user based on their login identity. Changes to the profiles file require the RADIUS server to be shutdown and restarted. Sending a HUP signal to the RADIUS process is not enough to force changes to the profiles file to take effect.

When you create command profiles, you can use an asterisk to indicate any possible ending to any particular command. The asterisk cannot be used as the beginning of a command. Reserved words for commands are matched exactly to those in the profiles file. Due to the exact match, it is not enough to simply enter “sh” for “show” in the profiles file, the complete word must be used. Commands can still be entered in the switch in partial format.

When you use per-command authentication, you must ensure that communication between the switch(es) and radius server(s) is not lost. Should the RADIUS server crash while users are logged in, they will have full administrative access to the switch until they log out. Using two RADIUS servers and enabling idle timeouts on all switches will greatly reduce the chance of a user gaining elevated access due to RADIUS server problems.

RADIUS Server Configuration Example (Merit)

Many implementations of RADIUS server use the publicly available Merit® AAA server application, available on the World Wide Web at:

<http://www.merit.edu/aaa>

Included below are excerpts from relevant portions of a sample Merit RADIUS server implementation. The example shows excerpts from the client and user configuration files. The client configuration file (`ClientCfg.txt`) defines the authorized source machine, source name, and access level. The user configuration file (`users`) defines username, password, and service type information.

`ClientCfg.txt`

#Client Name	Key	[type]	[version]	[prefix]
#-----	-----	-----	-----	-----
#10.1.2.3:256	test	type = nas	v2	px
#pml	%^\$%#*(&! (*&)+	type=nas		pml.

```
#pm2                :-):( ;^):-}!      type nas                pm2.
#merit.edu/homeless hmoemreilte.ses
#homeless           testing           type proxy             v1
#xyz.merit.edu      moretesting       type=Ascend:NAS       v1
#anyoldthing:1234  whoknows?           type=NAS+RAD_RFC+ACCT_RFC
10.202.1.3          andrew-linux         type=nas
10.203.1.41         eric                 type=nas
10.203.1.42         eric                 type=nas
10.0.52.14          samf                 type=nas
```

users

```
user      Password = ""
          Filter-Id = "unlim"
admin     Password = "", Service-Type = Administrative
          Filter-Id = "unlim"

eric      Password = "", Service-Type = Administrative
          Filter-Id = "unlim"

albert    Password = "password", Service-Type = Administrative
          Filter-Id = "unlim"

samuel    Password = "password", Service-Type = Administrative
          Filter-Id = "unlim"
```

RADIUS Per-Command Configuration Example

Building on this example configuration, you can use RADIUS to perform per-command authentication to differentiate user capabilities. To do so, use the Extreme-modified RADIUS Merit software that is available from the Extreme Networks web server at <http://www.extremenetworks.com/extreme/support/otherapps.htm> or by contacting Extreme Networks technical support. The software is available in compiled format for Solaris™ or Linux™ operating systems, as well as in source code format. For all clients that use RADIUS per-command authentication, you must add the following type to the client file:

```
type:extreme:nas + RAD_RFC + ACCT_RFC
```

Within the `users` configuration file, additional keywords are available for `Profile-Name` and `Extreme-CLI-Authorization`. To use per-command authentication, enable the CLI authorization function and indicate a profile name for that user. If

authorization is enabled without specifying a valid profile, the user is unable to perform any commands.

Next, define the desired profiles in an ASCII configuration file called `profiles`. This file contains named profiles of exact or partial strings of CLI commands. A named profile is linked with a user through the `users` file. A profile with the `permit` on keywords allows use of only the listed commands. A profile with the `deny` keyword allows use of all commands *except* the listed commands.

CLI commands can be defined easily in a hierarchal manner by using an asterisk (*) to indicate any possible subsequent entry. The parser performs exact string matches on other text to validate commands. Commands are separated by a comma (,) or newline.

Looking at the following example content in profiles for the profile named `PROFILE1`, which uses the `deny` keyword, the following attributes are associated with the user of this profile:

- Cannot use any command starting with `enable`.
- Cannot issue the `disable ipforwarding` command.
- Cannot issue a `show switch` command.
- Can perform all other commands.

We know from the `users` file that this applies to the users `albert` and `lulu`. We also know that `eric` is able to log in, but is unable to perform any commands, because he has no valid profile assigned.

In `PROFILE2`, a user associated with this profile can use any `enable` command, the `clear counter` command and the `show management` command, but can perform no other functions on the switch. We also know from the `users` file that `gerald` has these capabilities.

The following lists the contents of the file `users` with support for per-command authentication:

```

user      Password = ""
          Filter-Id = "unlim"

admin     Password = "", Service-Type = Administrative
          Filter-Id = "unlim"

eric      Password = "", Service-Type = Administrative, Profile-Name = ""
          Filter-Id = "unlim"

```

```

Extreme:Extreme-CLI-Authorization = Enabled

albert Password = "", Service-Type = Administrative, Profile-Name =
"Profile1"
    Filter-Id = "unlim"
    Extreme:Extreme-CLI-Authorization = Enabled

lulu Password = "", Service-Type = Administrative, Profile-Name =
"Profile1"
    Filter-Id = "unlim"
    Extreme:Extreme-CLI-Authorization = Enabled

gerald Password = "", Service-Type = Administrative, Profile-Name
"Profile2"
    Filter-Id = "unlim"
    Extreme:Extreme-CLI-Authorization = Enabled

```

Contents of the file "profiles":

```

PROFILE1 deny
{
enable *, disable ipforwarding
show switch
}

PROFILE2
{
enable *, clear counters
show management
}

PROFILE3 deny
{
create vlan *, configure iproute *, disable *, show fdb
delete *, configure rip add
}

```

Configuring TACACS+

Terminal Access Controller Access Control System Plus (TACACS+) is a mechanism for providing authentication, authorization, and accounting on a centralized server, similar in function to the RADIUS client. The ExtremeWare version of TACACS+ is used to

authenticate prospective users who are attempting to administer the switch. TACACS+ is used to communicate between the switch and an authentication database.



Note: You cannot use RADIUS and TACACS+ at the same time.

You can configure two TACACS+ servers, specifying the primary server address, secondary server address, and UDP port number to be used for TACACS+ sessions.

Using Network Login

Network login is a feature designed to control the admission of user packets into a network by giving addresses only to users that have been properly authenticated. Network login is controlled by an administrator on a per port, per VLAN basis and uses an integration of DHCP, user authentication over the web interface, and, sometimes, a RADIUS server to provide a user database or specific configuration details.

When network login is enabled on a port in a VLAN, that port will not forward any packets until authentication takes place.



Note: Windows authentication is not supported via network login.

Network login has two modes of operation:

- **Campus mode**
Campus mode is used when a port in a VLAN will move to another VLAN when authentication has been completed successfully. This mode is for the roaming user who will not always be using the same port for authentication.
- **ISP mode**
ISP mode is used when the port and VLAN used will remain constant. All network settings are configured for that VLAN.

These two network login modes have the following functional similarities:

- Until authentication takes place, ports on the VLAN are kept in a non-forwarding state.
- each mode requires the user to open a web browser with the IP address of the switch. This is the only address that the client can reach in a non-authenticated state.

- The web server on the switch provides user authentication.
- After authentication takes place, ports are moved into a forwarding state and moved to the VLAN configuration on the RADIUS server.

Using Network Login in Campus Mode

Campus mode requires:

- A DHCP server
- A RADIUS server configuration

The RADIUS server must have the following options configured in its dictionary file for network login:

Extreme.attr Extreme-Netlogin-Vlan 203 string (1, 0, ENCAPS)

The following optional configuration parameters can also be specified:

Extreme.attr Extreme-Netlogin-Url 204 string (1, 0, ENCAPS)

Extreme.attr Extreme-Netlogin-Url-Desc 205 string (1, 0, ENCAPS)



Note: These settings are for the Merit 3.6 version of RADIUS. The syntax of these settings will vary based on the type of RADIUS server that you are using.

The RADIUS server must also contain entries in the user file for a permanent VLAN, the URL to be redirected to after authentication has taken place, and the description of that URL. For example:

```
auto Authentication-Type = Unix-PW, Service-Type = login
   Filter-Id = "unlim"
   Extreme:Extreme-Netlogin-Vlan = "corp"
   Extreme:Extreme-Netlogin-Url = "http://192.207.37.16"
   Extreme:Extreme-Netlogin-Url-Desc = "Extreme Networks Home"
```

In this example, the username is *auto*, the permanent VLAN is *corp*, and the URL to be redirected to is the Extreme Networks home page *http://192.207.37.16*.

Configuring Campus Mode

To configure the switch to use network login in campus mode, follow these steps:

- 1 Configure the switch as a RADIUS client. See “RADIUS Client” on page 3-22.

- 2 Configure a DHCP range for the port or ports in the VLAN on which you want to enable network login, using this command:

```
config vlan <name> dhcp-address-range <ipaddress1> - <ipaddress2>
```

The switch will assign a temporary DHCP address within the DHCP range to the client.

- 3 Enable network login on the port, using the command:

```
enable netlogin ports <portlist> vlan <name>
```



Note: Network login is used on a per port, per VLAN basis. A port that is tagged can belong to more than one VLAN. In this case, network login can be enabled on one port for each VLAN.

Example Configuration Using Campus Mode

This example creates a permanent VLAN named *corp* on the switch. This VLAN will be used for authentication through a RADIUS server. The RADIUS server is 10.201.26.243 and the IP address of the switch is 10.201.26.11. The secret is “secret”. A temporary VLAN named *temporary* is created and port 9 is added. Network login is enabled on the port.

```
create vlan corp
config corp ipaddress 10.201.26.11/24
config radius primary server 10.201.26.243 client-ip 10.201.26.11
config radius primary shared-secret secret
enable radius
create vlan temporary
config temporary add port 9
config temporary ipaddress 192.168.0.1/24
config temporary dhcp-address-range 192.168.0.20 - 192.168.0.100
enable netlogin ports 9 vlan temporary
```

User Login Using Campus Mode

To log in as a user from the client, the user will follow these steps:

- 1 Set up the Windows IP configuration for DHCP.
- 2 Plug into the port that has network login enabled.
In this example, the user will plug into port 9.
- 3 Log in to Windows.
- 4 Release any old IP settings and renew the DHCP lease.

This is done differently depending on the version of Windows the user is running:

- Windows 9x—use the `winipcfg` tool. Choose the Ethernet adapter that is connected to the port on which network login is enabled. Use the buttons to release the IP configuration and renew the DHCP lease.
- Windows NT/2000—use the `ipconfig` command line utility. Use the command `ipconfig/release` to release the IP configuration and `ipconfig/renew` to get the temporary IP address from the switch. If you have more than one Ethernet adapter, specify the adapter by using a number for the adapter following the `ipconfig` command. You can find the adapter number using the command `ipconfig/all`.

At this point, the client will have its temporary IP address. In this example, the client should have obtained the IP address 192.168.0.20.

- 5 Bring up the web browser and enter the IP address of the switch.



Note: It is important to use the IP address of a VLAN that is reachable from anywhere on the network

A page will open with a link for network login.

- 6 Click the network login link.

A dialog box opens requesting a username and password.

- 7 Enter the username and password configured on the RADIUS server.

After the user has successfully logged in, the user will be redirected to the URL configured on the RADIUS server.

During the user login process, the following takes place:

- Authentication is done through the RADIUS server.
- After successful authentication, the connection information configured on the RADIUS server is returned to the switch:
 - the permanent VLAN
 - the URL to be redirected to (optional)
 - the URL description (optional)
- The port is moved to the permanent VLAN.

You can verify this using the `show vlan` command. For more information on the `show vlan` command, see [Displaying VLAN Settings](#) on page 5-15.

After a successful login has been achieved, there are several ways that a port can return to a non-authenticated, non-forwarding state:

- The user successfully logs out using the logout web browser window.
- The link from the user to the switch's port is lost.
- There is no activity on the port for 20 minutes.
- An administrator changes the port state.



Note: Because network login is sensitive to state changes during the authentication process, Extreme Networks recommends that you do not log out until the login process is completed. The login process is completed when you receive a permanent address.

Using Network Login in ISP Mode

In ISP mode, a RADIUS server might be used to provide user authentication. No Extreme-specific lines are required for the dictionary or the user file.

Configuring ISP Mode

Configure the switch to use network login in ISP mode, using this command:

```
enable netlogin ports <portlist> vlan <name>
```



Note: Network login is used on a per port, per VLAN basis. A port that is tagged can belong to more than one VLAN. In this case, network login can be enabled on one port for each VLAN.

Example Configuration Using ISP Mode

This example creates a permanent VLAN named *corp* on the switch. This VLAN will be used for authentication through RADIUS. The radius server is 10.201.26.243 and the IP address of the switch is 10.201.26.11. The secret is “secret”. Port 9 is added to the VLAN *corp*. Network login is enabled on the port.

```
create vlan corp
config corp ipaddress 10.201.26.11/24
config radius primary server 10.201.26.243 client-ip 10.201.26.11
config radius primary shared-secret secret
```

```
enable radius
config corp add port 9
enable netlogin ports 9 vlan corp
```

DHCP Server on the Switch

A DHCP server with limited configuration capabilities is included in the switch to provide IP addresses to clients.

DHCP is enabled on a per port, per VLAN basis. To enable or disable DHCP on a port in a VLAN, use one of the following commands:

```
enable dhcp ports <portlist> vlan <name>
disable dhcp ports <portlist> vlan <name>
```

Displaying Network Login Settings

To display the network login settings, use the following command:

```
show netlogin info {ports <portlist> vlan <name>}
```

Example

```
#show netlogin info ports 9 vlan temporary
Port 9:                VLAN: temporary
Port State:           Not Authenticated
Temp IP:              Unknown
DHCP:                 Not Enabled
User: Unknown        MAC: Unknown
```

In this example, the user is using campus mode and no authentication has taken place. Therefore, the port state displays as not authenticated. No packets sent by the user on port 9 will get past the port until authentication takes place. After authentication has taken place and the permanent IP address is obtained, the show command displays the port state as authenticated.

```
#show netlogin info ports 9 vlan corp
Port 9:                VLAN: corp
Port State:           Authenticated
Temp IP:              Unknown
DHCP:                 Not Enabled
User: auto            MAC: 00:10:A4:A9:11:3B
```

Disabling Network Login

Network login must be disabled on a port before you can delete a VLAN that contains that port. To disable network login, use the following command:

```
disable netlogin ports <portlist> vlan <name>
```

Using the Simple Network Time Protocol

ExtremeWare supports the client portion of the Simple Network Time Protocol (SNTP) Version 3 based on RFC1769. SNTP can be used by the switch to update and synchronize its internal clock from a Network Time Protocol (NTP) server. When enabled, the switch sends out a periodic query to the indicated NTP server, or the switch listens to broadcast NTP updates. In addition, the switch supports the configured setting for Greenwich Mean time (GMT) offset and the use of Daylight Saving Time. These features have been tested for year 2000 compliance.

Configuring and Using SNTP

To use SNTP, follow these steps:

- 1 Identify the host(s) that are configured as NTP server(s). Additionally, identify the preferred method for obtaining NTP updates. The options are for the NTP server to send out broadcasts, or for switches using NTP to query the NTP server(s) directly. A combination of both methods is possible. You must identify the method that should be used for the switch being configured.
- 2 Configure the Greenwich Mean Time (GMT) offset and Daylight Saving Time preference. The command syntax to configure GMT offset and usage of Daylight Saving Time is as follows:

```
config timezone {name <std_timezone_ID>} <GMT_offset>
{autodst {name <dst_timezone_ID>} {<dst_offset>}}
{begins [every <floatingday> | on <absoluteday>] {at <time_of_day>}}
{ends [every <floatingday> | on <absoluteday>] {at <time_of_day>}}}
```

By default, Daylight Saving Time is assumed to begin on the first Sunday in April at 2:00 AM, and end the last Sunday in October at 2:00 AM, and be offset from standard time by one hour. If this is the case in your timezone, you can set up automatic daylight savings adjustment with the command:

```
config timezone <GMT_offset> autodst
```

If your timezone uses starting and ending dates and times that differ from the default, you can specify the starting and ending date and time in terms of a floating day, as follows:

```
config timezone name MET 60 autodst name MDT begins every last
sunday march at 1 ends every last sunday october at 1
```

You can also specify a specific date and time, as shown in the following command.

```
config timezone name NZST 720 autodst name NZDT 60 begins every
first sunday october at 2 ends on 3/16/2002 at 2
```

The optional timezone IDs are used to identify the timezone in display commands such as `show switch`.

Table 3-2: describes the command options in detail:

Table 3-2: Time Zone Configuration Command Options

GMT_offset	Specifies a Greenwich Mean Time (GMT) offset, in + or - minutes.
std-timezone-ID	Specifies an optional name for this timezone specification. May be up to six characters in length. The default is an empty string.
autodst	Enables automatic Daylight Savings Time.
dst-timezone-ID	Specifies an optional name for this DST specification. May be up to six characters in length. The default is an empty string.
dst_offset	Specifies an offset from standard time, in minutes. Value is in the range of 1 to 60. Default is 60 minutes.
floating_day	<p>Specifies the day, week, and month of the year to begin or end DST each year. Format is:</p> <p><week><day><month> where:</p> <ul style="list-style-type: none"> ■ <week> is specified as [first second third fourth last] or 1-5 ■ <day> is specified as [sunday monday tuesday wednesday thursday friday saturday] or 1-7 (where 1 is Sunday) ■ <month> is specified as [january february march april may june july august september october november december] or 1-12 <p>Default for beginning is first sunday april; default for ending is last sunday october.</p>

Table 3-2: Time Zone Configuration Command Options

absolute_day	Specifies a specific day of a specific year on which to begin or end DST. Format is: <month>/<day>/<year> where: <ul style="list-style-type: none"> ■ <month> is specified as 1-12 ■ <day> is specified as 1-31 ■ <year> is specified as 1970 - 2035 The year must be the same for the begin and end dates.
time_of_day	Specifies the time of day to begin or end Daylight Savings Time. May be specified as an hour (0-23) or as hour:minutes. Default is 2:00.
noautodst	Disables automatic Daylight Savings Time.

Automatic Daylight Savings Time (DST) changes can be enabled or disabled. The default setting is enabled. To disable automatic DST, use the command:

```
config timezone {name <std_timezone_ID>} <GMT_offset> noautodst
```

3 Enable the SNTP client using the following command:

```
enable sntp-client
```

Once enabled, the switch sends out a periodic query to the NTP servers defined later (if configured) or listens to broadcast NTP updates from the network. The network time information is automatically saved into the on-board real-time clock.

4 If you would like this switch to use a directed query to the NTP server, configure the switch to use the NTP server(s). If the switch listens to NTP broadcasts, skip this step. To configure the switch to use a directed query, use the following command:

```
config sntp-client [primary | secondary] server [<ip_address> | <hostname>]
```

NTP queries are first sent to the primary server. If the primary server does not respond within 1 second, or if it is not synchronized, the switch queries the secondary server (if one is configured). If the switch cannot obtain the time, it restarts the query process. Otherwise, the switch waits for the `sntp-client update interval` before querying again.

5 Optionally, the interval for which the SNTP client updates the real-time clock of the switch can be changed using the following command:

```
config sntp-client update-interval <seconds>
```

The default `sntp-client update-interval` value is 64 seconds.

6 You can verify the configuration using the following commands:

- `show sntp-client`

This command provides configuration and statistics associated with SNTP and its connectivity to the NTP server.

- `show switch`

This command indicates the GMT offset, the Daylight Savings Time configuration and status, and the current local time.

NTP updates are distributed using GMT time. To properly display the local time in logs and other timestamp information, the switch should be configured with the appropriate offset to GMT based on geographical location. Table 3-3 describes GMT offsets.

Table 3-3: Greenwich Mean Time Offsets

GMT Offset in Hours	GMT Offset in Minutes	Common Time Zone References	Cities
+0:00	+0	GMT - Greenwich Mean UT or UTC - Universal (Coordinated) WET - Western European	London, England; Dublin, Ireland; Edinburgh, Scotland; Lisbon, Portugal; Reykjavik, Iceland; Casablanca, Morocco
-1:00	-60	WAT - West Africa	Azores, Cape Verde Islands
-2:00	-120	AT - Azores	
-3:00	-180		Brasilia, Brazil; Buenos Aires, Argentina; Georgetown, Guyana;
-4:00	-240	AST - Atlantic Standard	Caracas; La Paz
-5:00	-300	EST - Eastern Standard	Bogota, Columbia; Lima, Peru; New York, NY, Trevor City, MI USA
-6:00	-360	CST - Central Standard	Mexico City, Mexico
-7:00	-420	MST - Mountain Standard	Saskatchewan, Canada
-8:00	-480	PST - Pacific Standard	Los Angeles, CA, Cupertino, CA, Seattle, WA USA
-9:00	-540	YST - Yukon Standard	
-10:00	-600	AHST - Alaska-Hawaii Standard CAT - Central Alaska HST - Hawaii Standard	

Table 3-3: Greenwich Mean Time Offsets (continued)

GMT Offset in Hours	GMT Offset in Minutes	Common Time Zone References	Cities
-11:00	-660	NT - Nome	
-12:00	-720	IDLW - International Date Line West	
+1:00	+60	CET - Central European FWT - French Winter MET - Middle European MEWT - Middle European Winter SWT - Swedish Winter	Paris, France; Berlin, Germany; Amsterdam, The Netherlands; Brussels, Belgium; Vienna, Austria; Madrid, Spain; Rome, Italy; Bern, Switzerland; Stockholm, Sweden; Oslo, Norway
+2:00	+120	EET - Eastern European, Russia Zone 1	Athens, Greece; Helsinki, Finland; Istanbul, Turkey; Jerusalem, Israel; Harare, Zimbabwe
+3:00	+180	BT - Baghdad, Russia Zone 2	Kuwait; Nairobi, Kenya; Riyadh, Saudi Arabia; Moscow, Russia; Tehran, Iran
+4:00	+240	ZP4 - Russia Zone 3	Abu Dhabi, UAE; Muscat; Tblisi; Volgograd; Kabul
+5:00	+300	ZP5 - Russia Zone 4	
+5:30	+330	IST – India Standard Time	New Delhi, Pune, Allahabad, India
+6:00	+360	ZP6 - Russia Zone 5	
+7:00	+420	WAST - West Australian Standard	
+8:00	+480	CCT - China Coast, Russia Zone 7	
+9:00	+540	JST - Japan Standard, Russia Zone 8	
+10:00	+600	EAST - East Australian Standard GST - Guam Standard Russia Zone 9	
+11:00	+660		

Table 3-3: Greenwich Mean Time Offsets (continued)

GMT Offset in Hours	GMT Offset in Minutes	Common Time Zone References	Cities
+12:00	+720	IDLE - International Date Line East NZST - New Zealand Standard NZT - New Zealand	Wellington, New Zealand; Fiji, Marshall Islands

SNTP Example

In this example, the switch queries a specific NTP server and a backup NTP server. The switch is located in Cupertino, CA, and an update occurs every 20 minutes. The commands to configure the switch are as follows:

```
config timezone -480 autodst
config sntp-client update interval 1200
enable sntp-client
config sntp-client primary server 10.0.1.1
config sntp-client secondary server 10.0.1.2
```


4

Configuring Slots and Ports on a Switch

This chapter covers the following topics:

- Configuring a Slot on a Modular Switch on page 4-1
- Configuring Ports on a Switch on page 4-2
- Jumbo Frames on page 4-5
- Load Sharing on the Switch on page 4-8
- Switch Port-Mirroring on page 4-15
- Software-Controlled Redundant Port on page 4-17
- Multicast Performance Enhancements (BlackDiamond) on page 4-20

Configuring a Slot on a Modular Switch

If a slot has not been configured for a particular type of I/O module, then any type of module is accepted in that slot, and a default port and VLAN configuration is automatically generated.

Once any port on the module is configured (for example, a VLAN association, a VLAN tag configuration, or port parameters), all the port information and the module type for that slot must be saved to non-volatile storage. Otherwise, if the modular switch is rebooted or the module is removed from the slot, the port, VLAN, and module configuration information is not saved.



Note: For information on saving the configuration, refer to Appendix B.

You can configure the modular switch with the type of I/O module that is installed in each I/O slot. To do this on the BlackDiamond switch, use the following command:

```
config slot <slot number> module [f32t | f32fi | f32f | f48t | f96t |  
g4x | g6x | g8x | g8t | g12x | g12ts | wdmi]
```

To configure the I/O module on the Alpine switch, use the following command:

```
config slot <slot number> module [fm24f | fm32t | gm4sx | gm4t | gm4x  
| wdmi | fm24sf | fm24te]
```

You can also preconfigure the slot before inserting the module card. This allows you to begin configuring the module and ports before installing the card in the chassis.

If a slot is configured for one type of module, and a different type of module is inserted, the inserted module is put into a mismatch state, and is not brought online. To use the new module type in a slot, the slot configuration must be cleared or configured for the new module type. To clear the slot of a previously assigned I/O module type, use the following command:

```
clear slot <slot number>
```

All configuration information related to the slot and the ports on the module is erased. If a module is present when you issue this command, the module is reset to default settings.

To display information about a particular slot, use the following command:

```
show slot {<slot number>}
```

Information displayed includes:

- Card type, serial number, part number.
- Current state (power down, operational, diagnostic, mismatch).
- Port information.

If no slot is specified, information for all slots is displayed.

Configuring Ports on a Switch

On a modular switch, the port number is a combination of the slot number and the port number. The nomenclature for the port number is as follows:

```
slot:port
```

For example, if a G4X I/O module (having a total of four ports) is installed in slot 2 of the BlackDiamond 6808 chassis, the following ports are valid:

- 2:1
- 2:2
- 2:3
- 2:4

You can also use wildcard combinations (*) to specify multiple modular slot and port combinations. The following wildcard combinations are allowed:

- `slot:*` — Specifies all ports on a particular I/O module.
- `slot:x-slot:y` — Specifies a contiguous series of ports on a particular I/O module.
- `slota:x-slotb:y` — Specifies a contiguous series of ports that begin on one I/O module and end on another I/O module.

Enabling and Disabling Switch Ports

By default, all ports are enabled. To enable or disable one or more ports on a modular switch, use the following command:

```
[enable | disable] ports [all | mgmt | <portlist>]
```

To enable or disable one or more ports on a stand-alone switch, use the following command:

```
[enable | disable] ports [ all | <portlist>]
```

For example, to disable slot 7, ports 3, 5, and 12 through 15 on a modular switch, use the following command:

```
disable ports 7:3,7:5,7:12-7:15
```

For example, to disable ports 3, 5, and 12 through 15 on a stand-alone switch, use the following command:

```
disable ports 3,5,12-15
```

Even though a port is disabled, the link remains enabled for diagnostic purposes.

Configuring Switch Port Speed and Duplex Setting

By default, the switch is configured to use autonegotiation to determine the port speed and duplex setting for each port. You can manually configure the duplex setting and the speed of 10/100 Mbps ports, and you can manually configure the duplex setting of Gigabit Ethernet ports.

10BASE-T and 100BASE-TX ports can connect to either 10BASE-T or 100BASE-T networks. By default, the ports autonegotiate port speed. You can also configure each port for a particular speed (either 10 Mbps or 100 Mbps).

Gigabit Ethernet ports are statically set to 1 Gbps, and their speed cannot be modified.

All ports on a stand-alone switch can be configured for half-duplex or full-duplex operation. By default, the ports autonegotiate the duplex setting.

To configure port speed and duplex setting, use the following command:

```
config ports <portlist> auto off {speed [10 | 100 | 1000]} duplex [half  
| full]
```

To configure the system to autonegotiate, use the following command:

```
config ports <portlist> auto on
```

Flow control is fully supported only on Gigabit Ethernet ports. Gigabit ports both advertise support and respond to pause frames. 10/100 Mbps Ethernet ports also respond to pause frames, but do not advertise support. Neither 10/100 Mbps or Gigabit Ethernet ports initiate pause frames.

Flow Control is enabled or disabled as part of autonegotiation. If autonegotiation is set to off, flow control is disabled. When autonegotiation is turned on, flow control is enabled.

Turning Off Autonegotiation for a Gigabit Ethernet Port

In certain interoperability situations, you may need to turn autonegotiation off on a Gigabit Ethernet port. Even though a Gigabit Ethernet port runs only at full duplex, you must specify the duplex setting.

The following example turns autonegotiation off for port 1 on a G4X or G6X module located in slot 1 of a modular switch:

```
config ports 1:1 auto off duplex full
```

The following example turns autonegotiation off for port 4 (a Gigabit Ethernet port) on a stand-alone switch:

```
config ports 4 auto off duplex full
```

Jumbo Frames

Jumbo frames are Ethernet frames that are larger than 1523 bytes, including four bytes used for the cyclic redundancy check (CRC). Extreme products that use the “i” chipset support switching and routing of jumbo frames at wire-speed on all ports.

Jumbo frames are used between endstations that support larger frame sizes for more efficient transfers of bulk data. Both endstations involved in the transfer must be capable of supporting jumbo frames. The switch does not perform IP fragmentation, or participate in MTU negotiation on behalf of devices that do not support jumbo frames.

Enabling Jumbo Frames

To enable jumbo frame support, you must configure the maximum MTU size of a jumbo frame that will be allowed by the switch. To set the maximum MTU size, use the following command:

```
config jumbo-frame size <jumbo_frame_mtu>
```

The `jumbo_frame_mtu` range is 1523 to 9216. The value describes the maximum size “on the wire,” and includes 4 bytes of CRC plus another 4 bytes if 802.1Q tagging is being used.

Set the MTU size for the VLAN, using the following command:

```
config ip-mtu <size> vlan <vlan name>
```

The `ip-mtu` value can be 1500 or 9216, with 1500 the default. If you enter a value other than 1500, the switch will recognize that value as 9216.

Next, enable support on the physical ports that will carry jumbo frames using the following command:

```
enable jumbo-frame ports [<portlist> | all]
```



Note: Some network interface cards (NICs) have a configured maximum MTU size that does not include the additional 4 bytes of CRC. Ensure that the NIC maximum MTU size is at or below the maximum MTU size configured on the switch. Frames that are larger than the MTU size configured on the switch are dropped at the ingress port.

Path MTU Discovery

Using path MTU discovery, a source host assumes that the path MTU is the MTU of the first hop (which is known). The host sends all datagrams on that path with the “don’t fragment” (DF) bit set, which restricts fragmentation. If any of the datagrams must be fragmented by an Extreme switch along the path, the Extreme switch discards the datagrams and returns an ICMP Destination Unreachable message to the sending host, with a code meaning "fragmentation needed and DF set". When the source host receives the message (sometimes called a "Datagram Too Big" message), the source host reduces its assumed path MTU and retransmits the datagrams.

The path MTU discovery process ends when one of the following is true:

- The source host sets the path MTU low enough that its datagrams can be delivered without fragmentation.
- The source host does not set the DF bit in the datagram headers.

If it is willing to have datagrams fragmented, a source host can choose not to set the DF bit in datagram headers. Normally, the host continues to set DF in all datagrams, so that if the route changes and the new path MTU is lower, the host can perform path MTU discovery again.

IP Fragmentation with Jumbo Frames

ExtremeWare supports the fragmenting of IP packets. If an IP packet originates in a local network that allows large packets and those packets traverse a network that limits packets to a smaller size, the packets are fragmented instead of discarded.

This feature is designed to be used in conjunction with jumbo frames. Frames that are fragmented are not processed at wire-speed within the switch fabric.



Note: Jumbo frame-to-jumbo frame fragmentation is not supported. Only jumbo frame-to-normal frame fragmentation is supported.

To configure VLANs for IP fragmentation, follow these steps:

- 1 Enable jumbo frames on the incoming port.
- 2 Add the port to a VLAN.
- 3 Assign an IP address to the VLAN.
- 4 Enable ipforwarding on the VLAN.
- 5 Set the MTU size for the VLAN, using the following command:

```
config ip-mtu <size> vlan <vlan name>
```

The ip-mtu value can be 1500 or 9216, with 1500 the default. If you enter a value other than 1500, the switch will recognize that value as 9216.



Note: To set the MTU size greater than 1500, all ports in the VLAN must have jumbo frames enabled.

IP Fragmentation within a VLAN

ExtremeWare supports IP fragmentation within a VLAN. This feature does not require you to configure the MTU size. To use IP fragmentation within a VLAN, follow these steps:

- 1 Enable jumbo frames on the incoming port.
- 2 Add the port to a VLAN.
- 3 Assign an IP address to the VLAN.
- 4 Enable ipforwarding on the VLAN.

If you leave the MTU size configured to the default value, when you enable jumbo frame support on a port on the VLAN you will receive a warning that the ip-mtu size for the VLAN is not set at maximum jumbo frame size. You can ignore this warning if you want IP fragmentation within the VLAN, only. However, if you do not use jumbo frames, IP fragmentation can only be used for traffic that stays within the same VLAN. For traffic that is set to other VLANs, to use IP fragmentation, all ports in the VLAN must be configured for jumbo frame support.

Load Sharing on the Switch

Load sharing with switches allows you to increase bandwidth and resiliency by using a group of ports to carry traffic in parallel between switches. The sharing algorithm allows the switch to use multiple ports as a single logical port. For example, VLANs see the load-sharing group as a single logical port. Most load-sharing algorithms guarantee packet sequencing between clients.

If a port in a load-sharing group fails, traffic is redistributed to the remaining ports in the load-sharing group. If the failed port becomes active again, traffic is redistributed to include that port.



Note: Load sharing must be enabled on both ends of the link or a network loop may result. The load-sharing algorithms do not need to be the same on both ends.

This feature is supported between Extreme Networks switches only, but may be compatible with third-party trunking or link-aggregation algorithms. Check with an Extreme Networks technical representative for more information.

Load-Sharing Algorithms

Load-sharing algorithms allow you to select the distribution technique used by the load-sharing group to determine the output port selection. Algorithm selection is not intended for use in predictive traffic engineering.

You can configure one of three load-sharing algorithms on the switch, as follows:

- **Port-based** — Uses the ingress port to determine which physical port in the load-sharing group is used to forward traffic out of the switch.
- **Address-based** — Uses addressing information to determine which physical port in the load-sharing group to use for forwarding traffic out of the switch. Addressing information is based on the packet protocol, as follows:
 - **IP packets** — Uses the source and destination MAC and IP addresses, and the TCP port number.
 - **IPX packets** — Uses the source and destination MAC address, and IPX network identifiers.
 - **All other packets** — Uses the source and destination MAC address.

- Round-robin — When the switch receives a stream of packets, it forwards one packet out of each physical port in the load-sharing group using a round-robin scheme.



Note: Using the round-robin algorithm, packet sequencing between clients is not guaranteed.

If you do not explicitly select an algorithm, the port-based scheme is used. However, the address-based algorithm has a more even distribution and is the recommended choice.

The address-based and round-robin load-sharing algorithms are supported by all Alpine 3800 switch modules and by BlackDiamond switch modules and Summit switches that use the “i” chipset. The model designation of these modules and switches end with an “i” (for example, G12SXi and Summit7i). Modules designated with an “i” require the use of a management module also designated with an “i”. For more information on “i” chipset products, refer to Chapter 1.

Configured IP Address-Based Load Sharing

When you configure load sharing, the switch examines a specific place in the packet to determine which egress port to use for forwarding traffic:

- For layer 2 load sharing, the switch uses the MAC source address and destination address.
- For layer 3 load sharing, the switch uses the IP source address and destination address.
- For layer 4 load sharing, the switch using the UDP or TCP well-known port number.

You can control the field examined by the switch for IP address-based load sharing, using the following command:

```
config sharing address-based <12 | 12_13 | 12_13_14>
```

where:

- 12 — Indicates that the switch should examine the MAC source and destination address.
- 13 — Indicates that the switch should examine the IP source and destination address.
- 14 — Indicates that the switch should examine the UDP or TCP well-known port number.

This feature is available for the address-based load-sharing algorithm, only.

To verify your configuration, use the following command:

```
show sharing address-based
```

Configuring Switch Load Sharing

To set up a switch to load share among ports, you must create a load-sharing group of ports. The first port in the load-sharing group is configured to be the “master” logical port. This is the reference port used in configuration commands. It can be thought of as the logical port representing the entire port group.

On I/O modules or switches that do not have the “i” chipset, the following rules apply:

- Ports in a load-sharing group must be contiguous.
- Ports on the I/O module or switch are divided into groups of two or four.
- Address-based and round-robin load sharing algorithms do not apply.

Follow the outlined boxes in Table 4-1 through Table 4-9 to determine the valid port combinations for specific modules and switches.

Table 4-1, Table 4-2, and Table 4-3 show the possible load-sharing port group combinations for the G4X module, the G6X module, and the F32T and F32F modules, respectively.

Table 4-1: Port Combinations for the G4X Module

Load-Sharing Group	1	2	3	4
4-port groups	x	x	x	x
2-port groups	x	x	x	x

Table 4-2: Port Combinations for the G6X Module

Load-Sharing Group	1	2	3	4	5	6
4-port groups			x	x	x	x
2-port groups	x	x	x	x	x	x

Table 4-3: Port Combinations for the F32T and F32F Modules

Load-Sharing Group									1	1	1	1	1	1	1	
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Load-Sharing Group	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	
	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

BlackDiamond switch modules that use the “i” chipset do not require contiguous ports in the load-sharing group. On “r” chipset BlackDiamond switch modules, the following rules apply:

- One group can contain up to 8 ports.
- The ports in the group must be on the same I/O module.
- The ports in the group do not need to be contiguous.

Table 4-4, Table 4-5, Table 4-6, Table 4-7, Table 4-8, and Table 4-3 show the possible load-sharing port group combinations for the Summit1, Summit2, Summit3, Summit24, Summit4 and Summit4/FX, and Summit48 switches, respectively.

Table 4-4: Port Combinations for the Summit1 Switch

Load-Sharing Group								
	1	2	3	4	5	6	7	8
4-port groups				x	x	x	x	
2-port groups	*	x	x	x	x	x	x	*

* In addition, ports 1 and 8 can be combined into a two-port load-sharing group on the Summit1.

Table 4-5: Port Combinations for the Summit2 Switch

Load-Sharing Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 4-6: Port Combinations for the Summit3 Switch

Load-Sharing Group	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	

Table 4-7: Port Combinations for the Summit4 Switch and Summit4/FX Switch

Load-Sharing Group	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	2	2	2
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x							x	x	x	x	
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	

Table 4-8: Port Combinations for the Summit24 Switch

Load-Sharing Group	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	2	2	2	2
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		

Table 4-9: Port Combinations for the Summit48 Switch

Load-Sharing Group	1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2																							
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Load-Sharing Group	2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4																							
	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
4-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2-port groups	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Load-Sharing Group	4	5
	9	0
4-port groups		
2-port groups	x	x

On all other Summit switch models, the following rules apply:

- A group can contain up to 8 ports.
- The ports in a group do not need to be contiguous.

To define a load-sharing group, you assign a group of ports to a single, logical port number. To enable or disable a load-sharing group, use the following commands:

```
enable sharing <port> grouping <portlist> {port-based | address-based |
round-robin}
disable sharing <port>
```

Load-Sharing Examples

This section provides examples of how to define load-sharing on modular and stand-alone switches.

Load Sharing on a Modular Switch

The following example defines a load-sharing group on slot 3 that contains ports 9 through 12, and uses the first port in the group as the master logical port 9:

```
enable sharing 3:9 grouping 3:9-3:12
```

In this example, logical port 3:9 represents physical ports 3:9 through 3:12.

When using load sharing, you should always reference the master logical port of the load-sharing group (port 3:9 in the previous example) when configuring or viewing VLANs. VLANs configured to use other ports in the load-sharing group will have those ports deleted from the VLAN when load sharing becomes enabled.



Note: Do not disable a port that is part of a load-sharing group. Disabling the port prevents it from forwarding traffic, but still allows the link to initialize. As a result, a partner switch does not receive a valid indication that the port is not in a forwarding state, and the partner switch will continue to forward packets.

Load Sharing on a Stand-Alone Switch

The following example defines a load-sharing group that contains ports 9 through 12, and uses the first port in the group as the master logical port 9:

```
enable sharing 9 grouping 9-12
```

In this example, logical port 9 represents physical ports 9 through 12.

When using load sharing, you should always reference the master logical port of the load-sharing group (port 9 in the previous example) when configuring or viewing VLANs. VLANs configured to use other ports in the load-sharing group will have those ports deleted from the VLAN when load sharing becomes enabled.



Note: Do not disable a port that is part of a load-sharing group. Disabling the port prevents it from forwarding traffic, but still allows the link to initialize. As a result, a partner switch does not receive a valid indication that the port is not in a forwarding state, and the partner switch will continue to forward packets.

Verifying the Load-Sharing Configuration

The screen output resulting from the `show ports configuration` command lists the ports that are involved in load sharing and the master logical port identity.

Switch Port-Mirroring

Port-mirroring configures the switch to copy all traffic associated with one or more ports. The monitor port can be connected to a network analyzer or RMON probe for packet analysis. The system uses a traffic filter that copies a group of traffic to the monitor port.

The traffic filter can be defined based on one of the following criteria:

- **Physical port** — All data that traverses the port, regardless of VLAN configuration, is copied to the monitor port.
- **VLAN** — All data to and from a particular VLAN, regardless of the physical port configuration, is copied to the monitor port.
- **Virtual port** — All data specific to a VLAN on a specific port is copied to the monitor port.

Up to eight mirroring filters and one monitor port can be configured. Once a port is specified as a monitor port, it cannot be used for any other function.



Note: Frames that contain errors are not mirrored.

The mirrored port transmits tagged or untagged frames. This allows you to mirror multiple ports or VLANs to a mirror port, while preserving the ability of a single protocol analyzer to track and differentiate traffic within a broadcast domain (VLAN) and across broadcast domains (for example, across VLANs when routing).



Note: On switches that do not support the “i” chipset, mirrored frames that are transmitted from the switch do not contain 802.1Q VLAN tagging information.

Modular Switch Port-Mirroring Example

The following example selects slot 7, port 3 as the mirror port, and sends all traffic coming into or out of a modular switch on slot 7, port 1 to the mirror port:

```
config mirroring add port 7:3
config mirroring add port 7:1
enable mirroring to port 7:3
```

The following example sends all traffic coming into or out of the system on slot 8, port 1 and the VLAN *default* to the mirror port:

```
enable mirroring to port 8:4
```

```
config mirroring add port 8:1 vlan default
```

Stand-alone Switch Port-Mirroring Example

The following example selects port 3 as the mirror port and sends all traffic coming into or out of the switch on port 1 to the mirror port:

```
enable mirroring to port 3  
config mirroring add port 1
```

The following example sends all traffic coming into or out of the switch on port 1 and the VLAN *default* to the mirror port:

```
config mirroring add port 1 vlan default
```

Extreme Discovery Protocol

The Extreme Discovery Protocol (EDP) is used to gather information about neighbor Extreme Networks switches. EDP is used to by the switches to exchange topology information. EDP is also used by the Extreme Standby Router Protocol (ESRP), described in Chapter 14. Information communicated using EDP includes:

- Switch MAC address (switch ID).
- Switch software version information.
- Switch IP address.
- Switch VLAN-IP information.
- Switch port number.

EDP is enabled on all ports by default.

To disable EDP on one or more ports, use the following command:

```
disable edp ports <portlist>
```

To enable EDP on specified ports, use the following command:

```
enable edp ports <portlist>
```

To view EDP port information on the switch, use the following command:

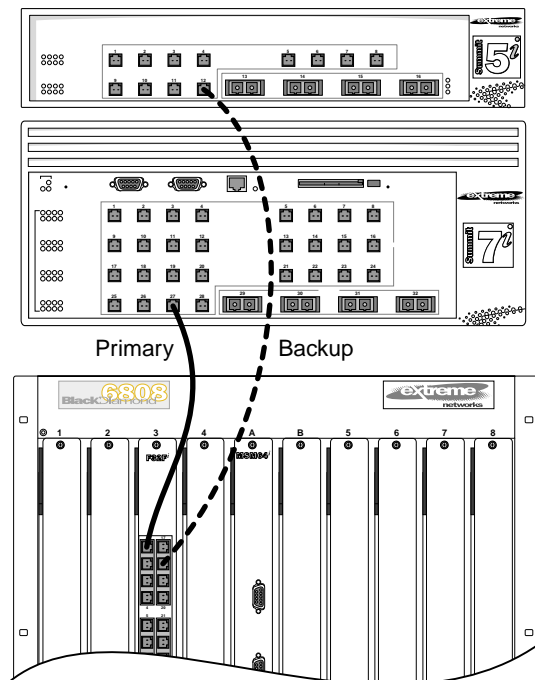
```
show edp
```

Software-Controlled Redundant Port

Using software-controlled redundant port you can back up a specified Ethernet port with a redundant, dedicated Ethernet port. If the active port fails, the backup port establishes a link and takes over for the failed port.

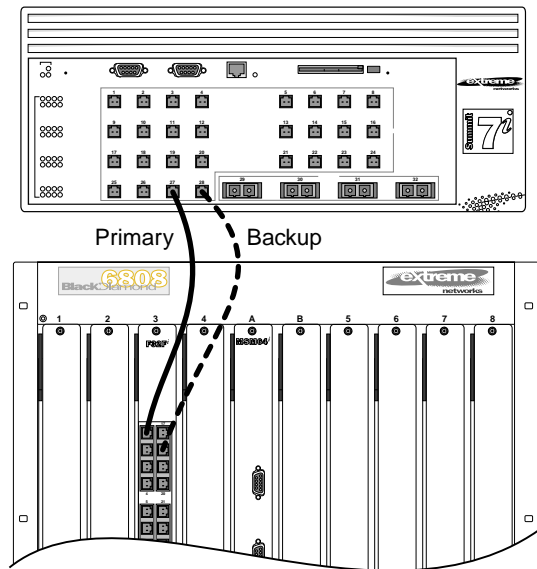
A load-shared group of Ethernet ports can be backed up with a set of load-shared redundant Ethernet ports. If a link in the active load-shared group fails, the entire group fails over to the redundant group.

Typical configurations of software-controlled redundant ports include dual-homing from a single switch to two different switches (shown in Figure 4-1) and redundant links between two switches (shown in Figure 4-2).



EW_076

Figure 4-1: Dual-homed redundant link



EW_075

Figure 4-2: Redundant link between two switches

Only one side of the link needs to be configured as redundant, since the redundant port link is held in standby state on both sides of the link.

Theory of Operation

A software-controlled redundant port is configured to backup a specified primary port. The redundant port tracks the link state of the associated primary port, and if the link on the primary port fails, the redundant port establishes a link and becomes active. You must manually configure the primary and redundant ports identically in terms of VLANs, QoS settings, access lists, and so on.

While the redundant port is inactive, the link between the primary port and the redundant port is held down by forcing an auto-negotiation error. Because of this, auto-negotiation must be enabled on both the primary and redundant port.

When the redundant port is activated, the primary port is ready to accept a link. After a link is established on the primary port, any link present on the redundant port is taken down.

Smart Redundancy and Software Redundant Port

The primary port will only take over for an active redundant port if the smart redundancy feature is enabled, which is the default setting. When smart redundancy is enabled, the primary port (or primary group of load-shared ports) becomes the “preferred” path, provided that the primary port link is active (or in the case of load sharing, the number of active links in the primary group is greater than or equal to the number of links active on the redundant group).

If smart redundancy is disabled, the active path remains the first path to establish a link (or in the case of load sharing, the first group to establish the highest number of active links). For example, if the primary port is active and fails, the redundant port takes over. If the primary is then reconnected and smart redundancy is disabled, the active path remains on the redundant port. If smart redundancy is enabled, the active path switches back to the primary (original) port.

Software Redundant Port and Load Sharing

A load-shared group of ports can be backed up with a redundant group of ports. Each port in the primary group is configured with a unique redundant port. The redundant ports must be grouped together as a separate load-shared group (independent from the primary load-shared group).

If the primary group is active and one or more links in that group fail, each port in the group fails over to its redundant port in the backup group, provided that enough links can be established on the redundant ports to be greater than the number of active links remaining in the primary group.

Limitations

The following lists the limitations of the software redundant port feature:

- Software redundant port only cover failures where both the TX and RX paths fail. If a single strand of a fiber is pulled, the software redundant port cannot correctly recover from the failure.
- Auto-negotiation must be enabled on the primary and redundant ports.
- You cannot configure hardware redundant ports (such as ports 49 and 50 on a Summit48i) as software controlled redundant ports.
- Software redundant port is supported only on products that use the “i” chipset.
- You must configure the software redundant port. VLANs, QoS, access control lists, and so on, must be configured separately on both the primary and redundant ports.

- Only one side of the link should be configured as redundant. For example if ports 1 and 2 are connected between switches A and B, only switch A should be configured with redundant ports.
- Software redundant port is not supported on 1000BASE-T ports.
- You can configure only one redundant port for each primary port.

Configuring Software-Controlled Redundant Port

To configure a software-controlled redundant port, use the following command:

```
config ports <portlist> redundant <portlist>
```

The first specified port is the primary port. The second specified port is the redundant port. For load-shared groups, this command must be entered for each port in the group. If a redundant port is not specified for a port in a load-shared group, it will not fail over, which could then result in a split group.

To unconfigure the port, use the following command:

```
unconfig ports <portlist> redundant
```

Any additional port-specific configuration (for example, VLANs, Spanning Tree, QoS, access lists, and so on) must be independently configured.

Multicast Performance Enhancements (BlackDiamond)

The BlackDiamond switch optimizes multicast data forwarding performance for modules that use the “i” chipset. To increase the performance of multicast applications, you can disable I/O modules that do not have the “i” series chipset. When you disable support for older modules (without the “i” chipset), the modules are not powered up, and they do not pass traffic in a BlackDiamond system. You must save and reboot for these changes to take effect. The default setting is enabled.

To enable I/O modules that do not have the “i” series chipset, use the following command:

```
enable gl-module support
```

To disable I/O modules that do not have the “i” series chipset, use the following command:

```
disable g1-module support
```

Performance Enhancements for Load Sharing

You can modify the backplane load-sharing policy for more robust support of multicast streams. The round-robin algorithm is not supported on modules that do not have the “i” series chipset. The default backplane load-sharing policy is port-based. You must save for changes to be saved across reboots.

To configure the switch backplane load-sharing policy, use the following command:

```
config backplane-ls-policy <address-based | port-based | round-robin>
```


5

Virtual LANs (VLANs)

This chapter covers the following topics:

- Overview of Virtual LANs on page 5-1
- Types of VLANs on page 5-2
- VLAN Names on page 5-12
- Configuring VLANs on the Switch on page 5-13
- Displaying VLAN Settings on page 5-15
- VLAN Tunneling (VMANs) on page 5-17
- MAC-Based VLANs on page 5-19

Setting up Virtual Local Area Networks (VLANs) on the switch eases many time-consuming tasks of network administration while increasing efficiency in network operations.

Overview of Virtual LANs

The term “VLAN” is used to refer to a collection of devices that communicate as if they were on the same physical LAN. Any set of ports (including all ports on the switch) is considered a VLAN. LAN segments are not restricted by the hardware that physically connects them. The segments are defined by flexible user groups you create with the command-line interface.

Benefits

Implementing VLANs on your networks has the following advantages:

- **VLANs help to control traffic.**

With traditional networks, congestion can be caused by broadcast traffic that is directed to all network devices, regardless of whether they require it. VLANs increase the efficiency of your network because each VLAN can be set up to contain only those devices that must communicate with each other.

- **VLANs provide extra security.**

Devices within each VLAN can only communicate with member devices in the same VLAN. If a device in VLAN *Marketing* must communicate with devices in VLAN *Sales*, the traffic must cross a routing device.

- **VLANs ease the change and movement of devices.**

With traditional networks, network administrators spend much of their time dealing with moves and changes. If users move to a different subnetwork, the addresses of each endstation must be updated manually.

Types of VLANs

VLANs can be created according to the following criteria:

- Physical port
- 802.1Q tag
- Ethernet, LLC SAP, or LLC/SNAP Ethernet protocol type
- MAC address
- A combination of these criteria

Port-Based VLANs

In a port-based VLAN, a VLAN name is given to a group of one or more ports on the switch. All ports are members of the port-based VLAN *default*. Before you can add any port to another port-based VLAN, you must remove it from the default VLAN, unless the new VLAN uses a protocol other than the default protocol *any*. A port can be a member of only one port-based VLAN.

On the Summit7i switch in Figure 5-1, ports 9 through 14 are part of VLAN *Marketing*; ports 25 through 29 are part of VLAN *Sales*; and ports 21 through 24 and 30 through 32 are in VLAN *Finance*.

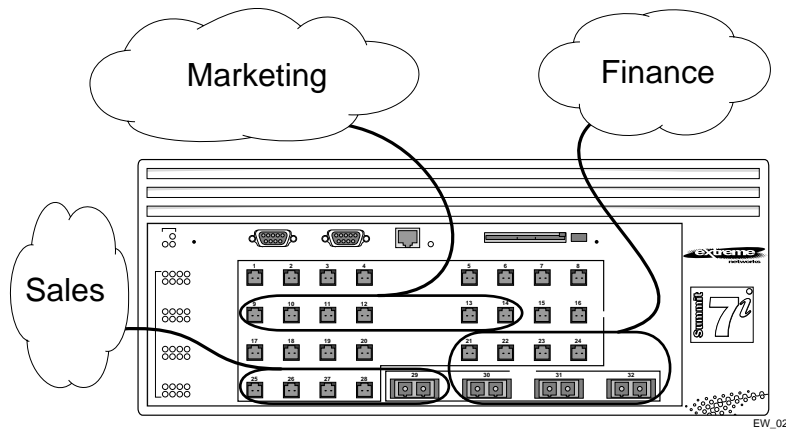


Figure 5-1: Example of a port-based VLAN on the Summit7i switch

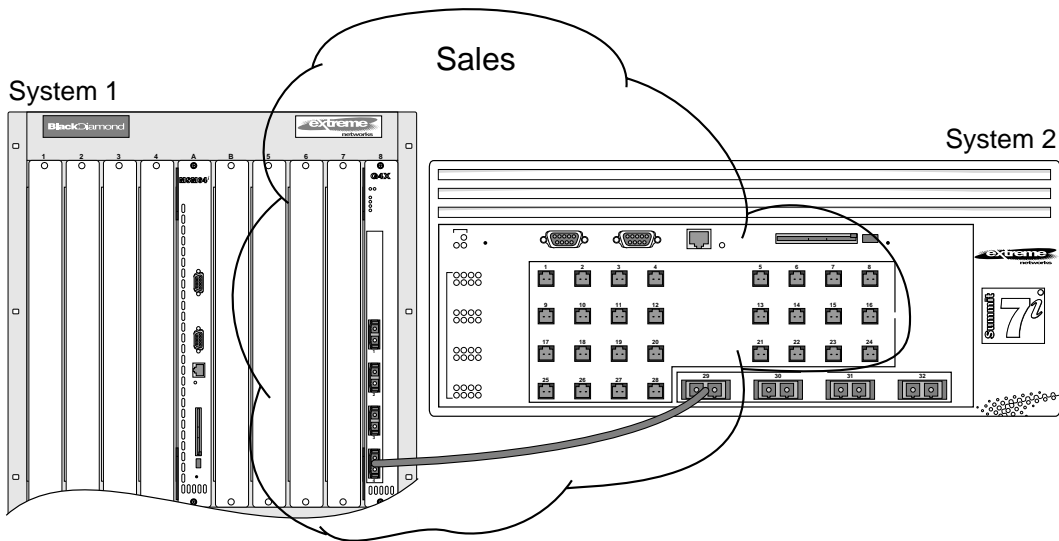
For the members of the different IP VLANs to communicate, the traffic must be routed by the switch, even if they are physically part of the same I/O module. This means that each VLAN must be configured as a router interface with a unique IP address.

Spanning Switches with Port-Based VLANs

To create a port-based VLAN that spans two switches, you must do two things:

- 1 Assign the port on each switch to the VLAN.
- 2 Cable the two switches together using one port on each switch per VLAN.

Figure 5-2 illustrates a single VLAN that spans a BlackDiamond switch and a Summit7i switch. All ports on the BlackDiamond switch belong to VLAN *Sales*. Ports 1 through 29 on the Summit 7i switch also belong to VLAN *Sales*. The two switches are connected using slot 8, port 4 on system 1 (the BlackDiamond switch), and port 29 on system 2 (the Summit7i switch).

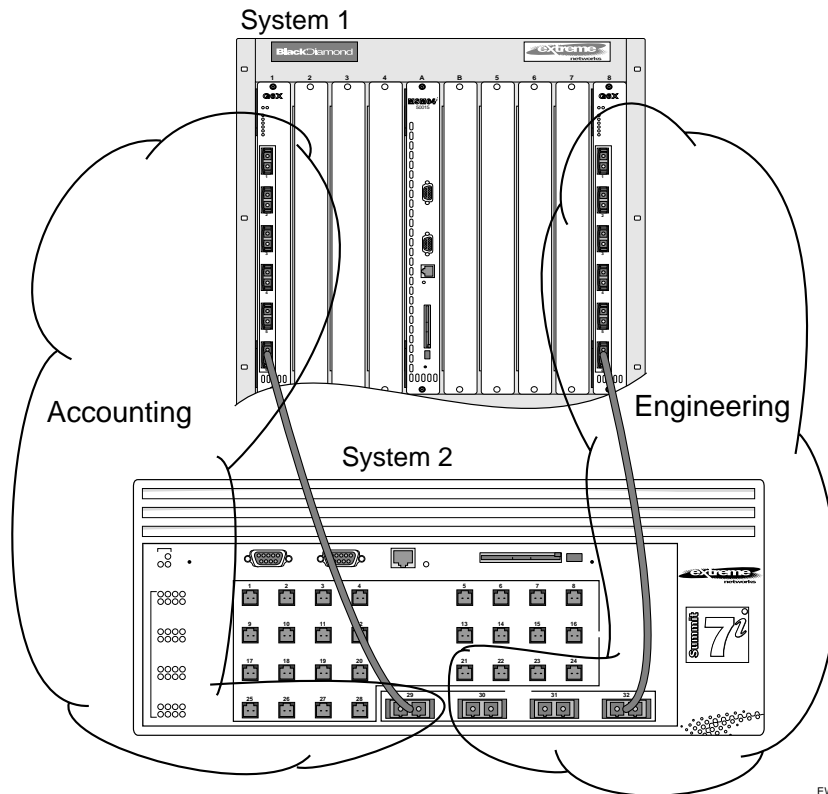


EW_028

Figure 5-2: Single port-based VLAN spanning two switches

To create multiple VLANs that span two switches in a port-based VLAN, a port on system 1 must be cabled to a port on system 2 for each VLAN you want to have span across the switches. At least one port on each switch must be a member of the corresponding VLANs, as well.

Figure 5-3 illustrates two VLANs spanning two switches. On system 2, ports 25 through 29 are part of VLAN *Accounting*; ports 21 through 24 and ports 30 through 32 are part of VLAN *Engineering*. On system 1, all port on slot 1 are part of VLAN *Accounting*; all ports on slot 8 are part of VLAN *Engineering*.



EW_030

Figure 5-3: Two port-based VLANs spanning two switches

VLAN *Accounting* spans system 1 and system 2 by way of a connection between system 2, port 29 and system 1, slot 1, port 6. VLAN *Engineering* spans system 1 and system 2 by way of a connection between system 2, port 32, and system 1, slot 8, port 6.

Using this configuration, you can create multiple VLANs that span multiple switches, in a daisy-chained fashion. Each switch must have a dedicated port for each VLAN. Each dedicated port must be connected to a port that is a member of its VLAN on the next switch.

Tagged VLANs

Tagging is a process that inserts a marker (called a *tag*) into the Ethernet frame. The tag contains the identification number of a specific VLAN, called the *VLANid*.



Note: The use of 802.1Q tagged packets may lead to the appearance of packets slightly bigger than the current IEEE 802.3/Ethernet maximum of 1,518 bytes. This may affect packet error counters in other devices, and may also lead to connectivity problems if non-802.1Q bridges or routers are placed in the path.

Uses of Tagged VLANs

Tagging is most commonly used to create VLANs that span switches. The switch-to-switch connections are typically called *trunks*. Using tags, multiple VLANs can span multiple switches using one or more trunks. In a port-based VLAN, each VLAN requires its own pair of trunk ports, as shown in Figure 5-3. Using tags, multiple VLANs can span two switches with a single trunk.

Another benefit of tagged VLANs is the ability to have a port be a member of multiple VLANs. This is particularly useful if you have a device (such as a server) that must belong to multiple VLANs. The device must have a NIC that supports 802.1Q tagging.

A single port can be a member of only one port-based VLAN. All additional VLAN membership for the port must be accompanied by tags. In addition to configuring the VLAN tag for the port, the server must have a *Network Interface Card (NIC)* that supports 802.1Q tagging.

Assigning a VLAN Tag

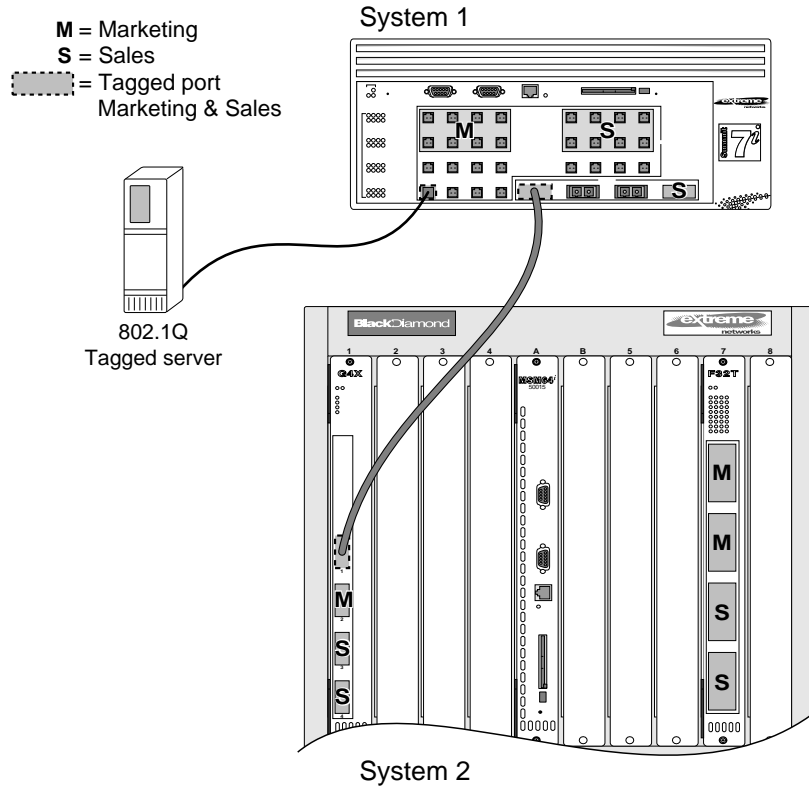
Each VLAN may be assigned an 802.1Q VLAN tag. As ports are added to a VLAN with an 802.1Q tag defined, you decide whether each port will use tagging for that VLAN. The default mode of the switch is to have all ports assigned to the VLAN named *default* with an 802.1Q VLAN tag (VLANid) of 1 assigned.

Not all ports in the VLAN must be tagged. As traffic from a port is forwarded out of the switch, the switch determines (in real time) if each destination port should use tagged or untagged packet formats for that VLAN. The switch adds and strips tags, as required, by the port configuration for that VLAN.



Note: Packets arriving tagged with a VLANid that is not configured on a port will be discarded.

Figure 5-4 illustrates the physical view of a network that uses tagged and untagged traffic.



EW_029

Figure 5-4: Physical diagram of tagged and untagged traffic

Figure 5-5 is a logical diagram of the same network.

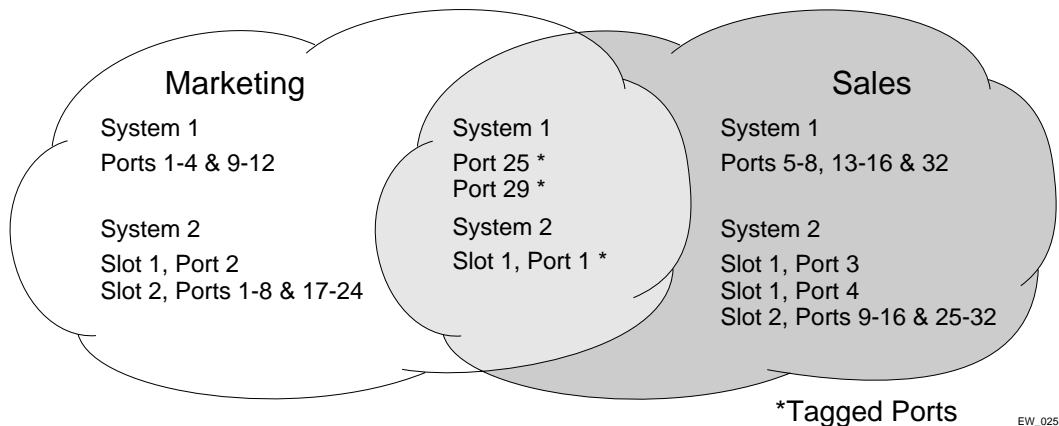


Figure 5-5: Logical diagram of tagged and untagged traffic

In Figure 5-4 and Figure 5-5:

- The trunk port on each switch carries traffic for both VLAN *Marketing* and VLAN *Sales*.
- The trunk port on each switch is tagged.
- The server connected to port 25 on system 1 has a NIC that supports 802.1Q tagging.
- The server connected to port 25 on system 1 is a member of both VLAN *Marketing* and VLAN *Sales*.
- All other stations use untagged traffic.

As data passes out of the switch, the switch determines if the destination port requires the frames to be tagged or untagged. All traffic coming from and going to the server is tagged. Traffic coming from and going to the trunk ports is tagged. The traffic that comes from and goes to the other stations on this network is not tagged.

Mixing Port-Based and Tagged VLANs

You can configure the switch using a combination of port-based and tagged VLANs. A given port can be a member of multiple VLANs, with the stipulation that only one of its VLANs uses untagged traffic. In other words, a port can simultaneously be a member of one port-based VLAN and multiple tag-based VLANs.



Note: For the purposes of VLAN classification, packets arriving on a port with an 802.1Q tag containing a VLANid of zero are treated as untagged.

Protocol-Based VLANs

Protocol-based VLANs enable you to define a packet filter that the switch uses as the matching criteria to determine if a particular packet belongs to a particular VLAN.

Protocol-based VLANs are most often used in situations where network segments contain hosts running multiple protocols. For example, in Figure 5-6, the hosts are running both the IP and NetBIOS protocols.

The IP traffic has been divided into two IP subnets, 192.207.35.0 and 192.207.36.0. The subnets are internally routed by the switch. The subnets are assigned different VLAN names, *Finance* and *Personnel*, respectively. The remainder of the traffic belongs to the VLAN named *MyCompany*. All ports are members of the VLAN *MyCompany*.

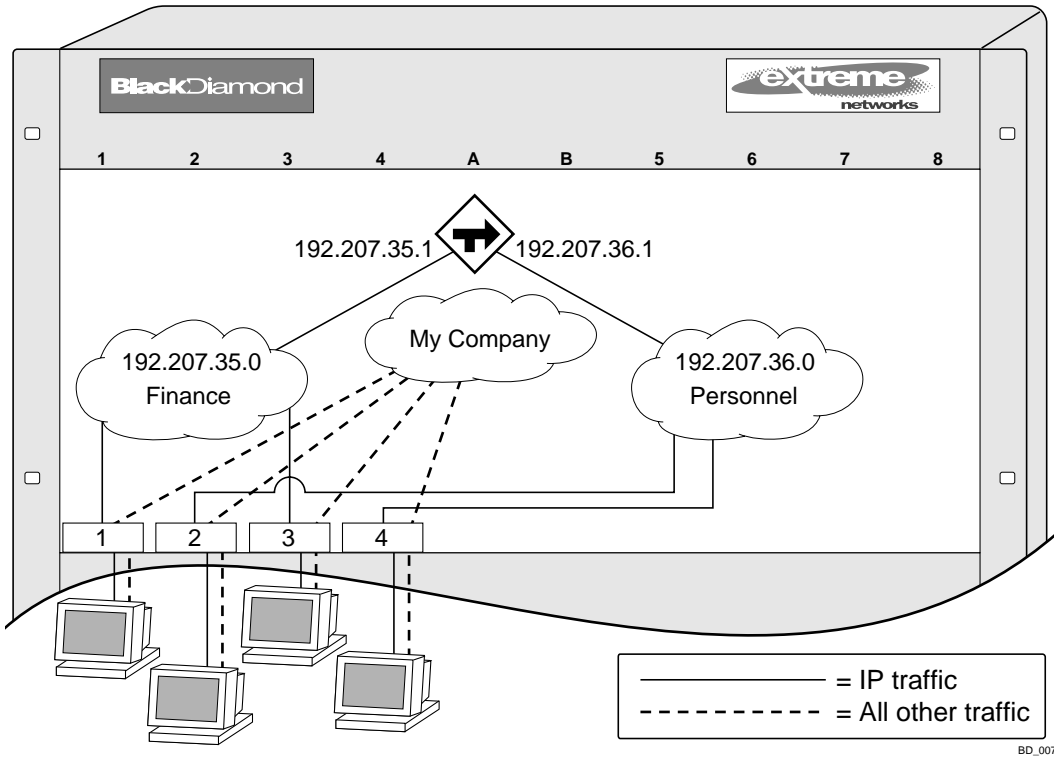


Figure 5-6: Protocol-based VLANs

Predefined Protocol Filters

The following protocol filters are predefined on the switch:

- IP
- IPX
- NetBIOS
- DECNet
- IPX_8022
- IPX_SNAP
- AppleTalk

Defining Protocol Filters

If necessary, you can define a customized protocol filter based on EtherType, Logical Link Control (LLC), and/or Subnetwork Access Protocol (SNAP). Up to six protocols may be part of a protocol filter. To define a protocol filter, follow these steps:

- 1 Create a protocol using the following command:

```
create protocol <protocol_name>
```

For example:

```
create protocol fred
```

The protocol name can have a maximum of 32 characters.

- 2 Configure the protocol using the following command:

```
config protocol <protocol_name> add <protocol_type> <hex_value>
```

Supported protocol types include:

- `etype` — EtherType.

The values for `etype` are four-digit hexadecimal numbers taken from a list maintained by the IEEE. This list can be found at the following URL:

<http://standards.ieee.org/regauth/ethertype/index.html>

- `llc` — LLC Service Advertising Protocol (SAP).

The values for `llc` are four-digit hexadecimal numbers that are created by concatenating a two-digit LLC Destination SAP (DSAP) and a two-digit LLC Source SAP (SSAP).

- `snap` — Ethertype inside an IEEE SNAP packet encapsulation.

The values for `snap` are the same as the values for `etype`, described previously.

For example:

```
config protocol fred add llc feff
```

```
config protocol fred add snap 9999
```

A maximum of 15 protocol filters, each containing a maximum of six protocols, can be defined. On products that use the Inferno chip set, all 15 protocol filters can be active and configured for use. On all other platforms, no more than seven protocols can be active and configured for use.



Note: For more information on SNAP for Ethernet protocol types, see TR 11802-5:1997 (ISO/IEC) [ANSI/IEEE std. 802.1H, 1997 Edition].

Deleting a Protocol Filter

If a protocol filter is deleted from a VLAN, the VLAN is assigned a protocol filter of `none`. You can continue to configure the VLAN. However, no traffic is forwarded to the VLAN until a protocol is assigned to it.

Precedence of Tagged Packets Over Protocol Filters

If a VLAN is configured to accept tagged packets on a particular port, incoming packets that match the tag configuration take precedence over any protocol filters associated with the VLAN.

VLAN Names

Each VLAN is given a name that can be up to 32 characters. VLAN names can use standard alphanumeric characters. The following characters are not permitted in a VLAN name:

- Space
- Comma
- Quotation mark

VLAN names must begin with an alphabetical letter. Quotation marks can be used to enclose a VLAN name that includes special characters, including single quotation marks or commas. Spaces may not be included, even within quotation marks. For example, the names `test`, `test1`, and `test_15` are acceptable VLAN names. The names “`test&5`” and “`joe’s`” may be used if enclosed in quotation marks. Names such as “`5test`” or “`test 5`” are not permitted.

VLAN names are locally significant. That is, VLAN names used on one switch are only meaningful to that switch. If another switch is connected to it, the VLAN names have no significance to the other switch.



Note: You should use VLAN names consistently across your entire network.

Default VLAN

The switch ships with one default VLAN that has the following properties:

- The VLAN name is *default*.
- It contains all the ports on a new or initialized switch.
- The default VLAN is untagged on all ports. It has an internal VLANid of 1.

Renaming a VLAN

To rename an existing VLAN, use the following command:

```
config vlan <old_name> name <new_name>
```

The following rules apply to renaming VLANs:

- Once you change the name of the default VLAN, it cannot be changed back to *default*.
- You cannot create a new VLAN named *default*.
- You cannot change the VLAN name *MacVlanDiscover*. Although the switch accepts a name change, once it is rebooted, the original name is recreated.

Configuring VLANs on the Switch

This section describes the commands associated with setting up VLANs on the switch. Configuring a VLAN involves the following steps:

- 1 Create and name the VLAN.
- 2 Assign an IP address and mask (if applicable) to the VLAN, if needed.



Note: Each IP address and mask assigned to a VLAN must represent a unique IP subnet. You cannot configure the same IP subnet on different VLANs.



Note: If you plan to use this VLAN as a control VLAN for an EAPS domain, do NOT assign an IP address to the VLAN.

- 3 Assign a VLANid, if any ports in this VLAN will use a tag.
- 4 Assign one or more ports to the VLAN.

As you add each port to the VLAN, decide if the port will use an 802.1Q tag.

VLAN Configuration Examples

The following modular switch example creates a port-based VLAN named *accounting*, assigns the IP address 132.15.121.1, and assigns slot 2, ports 1, 2, 3, and 6, and slot 4, ports 1 and 2 to it:

```
create vlan accounting
config accounting ipaddress 132.15.121.1
config default delete port 2:1-2:3,2:6,4:1,4:2
config accounting add port 2:1-2:3,2:6,4:1,4:2
```



Note: Because VLAN names are unique, you do not need to enter the keyword `vlan` after you have created the unique VLAN name. You can use the VLAN name alone.

The following stand-alone switch example creates a tag-based VLAN named *video*. It assigns the VLANid 1000. Ports 4 through 8 are added as tagged ports to the VLAN.

```
create vlan video
config video tag 1000
config video add port 4-8 tagged
```

The following stand-alone switch example creates a VLAN named *sales*, with the VLANid 120. The VLAN uses both tagged and untagged ports. Ports 1 through 3 are tagged, and ports 4 and 7 are untagged. Note that when not explicitly specified, ports are added as untagged.

```
create vlan sales
config sales tag 120
config sales add port 1-3 tagged
config default delete port 4,7
config sales add port 4,7
```

The following modular switch example creates a protocol-based VLAN named *ipsales*. Slot 5, ports 6 through 8, and slot 6, ports 1, 3, and 4-6 are assigned to the VLAN. In this example, you can add untagged ports to a new VLAN without first deleting them from the default VLAN, because the new VLAN uses a protocol other than the default protocol.

```
create vlan ipsales
config ipsales protocol ip
config ipsales add port 5:6-5:8,6:1,6:3-6:6
```

The following modular switch example defines a protocol filter, *myprotocol* and applies it to the VLAN named *myvlan*. This is an example only, and has no real-world application.

```
create protocol myprotocol
config protocol myprotocol add etype 0xf0f0
config protocol myprotocol add etype 0xffff
create vlan myvlan
config myvlan protocol myprotocol
```

Displaying VLAN Settings

To display VLAN settings, use the following command:

```
show vlan {<name> | detail}
```

The `show` command displays summary information about each VLAN, which includes:

- Name.
- VLANid.
- How the VLAN was created.
- IP address.
- IPX address (if configured).
- STPD information.
- Protocol information.
- QoS profile information.
- Ports assigned.
- Tagged/untagged status for each port.
- How the ports were added to the VLAN.
- Number of VLANs configured on the switch.

Use the `detail` option to display the detailed format.

Displaying VLAN Statistics

To display VLAN statistics on switches that have the “I”-series chipset, use the following command:

```
show vlan stats vlan <name> ... <name>
```

The information displayed includes:

- Transmitted and received unicast packets.
- Transmitted and received multicast packets.
- Transmitted and received broadcast packets.
- Transmitted and received bytes.

You can display statistics for multiple VLANs by entering the name of each VLAN on the command line.

Displaying VLAN Statistics Per Port

In addition to displaying VLAN statistics on a per-VLAN basis, you can display VLAN statistics on a per-port basis, using the following command:

```
config ports <portlist> monitor vlan <name>
```

You can monitor up to four VLANs on the same port by issuing the command four times. For example, if you want to monitor VLAN dog1, dog2, dog3, and dog4 on port 1, use the following command configuration:

```
config ports 1:* monitor vlan dog1
config ports 1:* monitor vlan dog2
config ports 1:* monitor vlan dog3
config ports 1:* monitor vlan dog4
```

After you configure the port, you can use this command to display information for the configured port:

```
show ports <portlist> vlan statistics
```

After you have configured per-port monitoring, every time you issue the `show ports` command, the latest statistics are displayed directly from the hardware in real-time. This information is not logged.

To remove the port mask, use the following command:

```
unconfig ports <portlist> monitor vlan <name>
```

You must issue the unconfig command for each VLAN you have configured for the port. For example:

```
unconfig ports 1:* monitor vlan dog1
unconfig ports 1:* monitor vlan dog2
unconfig ports 1:* monitor vlan dog3
unconfig ports 1:* monitor vlan dog4
```

Displaying Protocol Information

To display protocol information, use the following command:

```
show protocol {<protocol>}
```

This `show` command displays protocol information, which includes:

- Protocol name.
- List of protocol fields.
- VLANs that use the protocol.

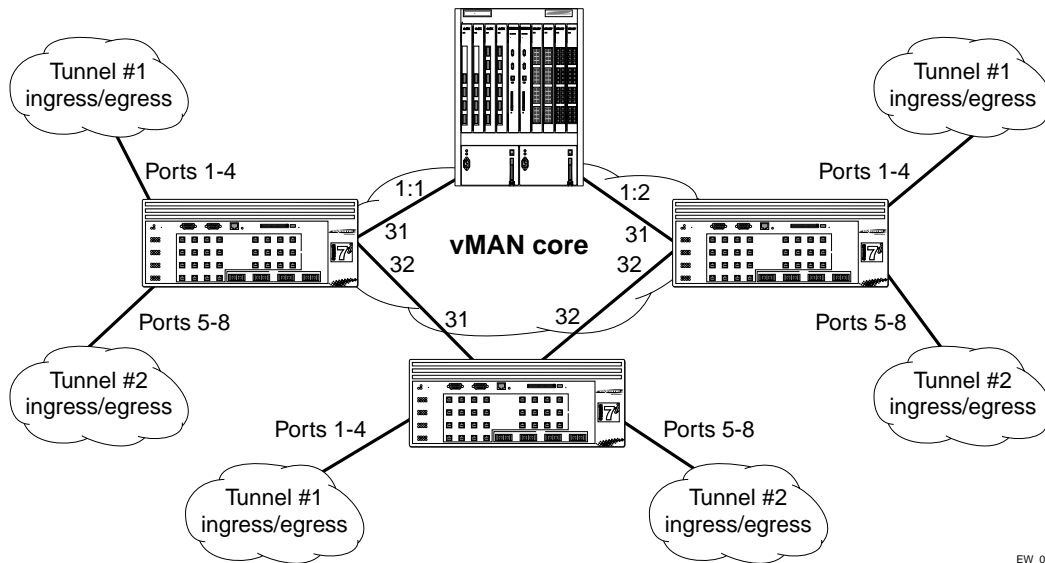
VLAN Tunneling (VMANs)

You can “tunnel” any number of 802.1Q and/or Cisco ISL VLANs into a single VLAN that can be switched through an Extreme Ethernet infrastructure. A given tunnel is completely isolated from other tunnels or VLANs. This feature is useful in building transparent private networks (VMANs) that need point-to-point or point-to-multipoint connectivity across an Ethernet infrastructure. The VLAN tagging methods used within the VMAN tunnel are transparent to the tunnel. For the MAN provider, the tagging numbers and methods used by the customer are transparent to the provider.

To configure a VMAN tunnel, follow these steps:

- 1 Modify the 802.1Q Ethertype the switch uses to recognize tagged frames.
- 2 Configure the switch to accept larger MTU size frames (jumbo frames).
- 3 Create tunnels by creating VLANs and configuring member ports as tagged on switch-to-switch ports and untagged on the ingress/egress ports of the tunnel.

Figure 5-7 illustrates a configuration with VMANs.



EW_059

Figure 5-7: VMAN example

Two tunnels are depicted that have ingress/egress ports on each Summit7i switch.

The configuration for the Summit7i switches shown in Figure 5-7 is:

```
config dot1q ethertype 9100
enable jumbo-frame ports 31,32
config jumbo-frame size 1530
create vlan Tunnel1
config vlan Tunnel1 tag 50
config vlan Tunnel1 add port 1-4 untag
config vlan Tunnel1 add port 31,32 tagged
create vlan Tunnel2
config vlan Tunnel2 tag 60
config vlan Tunnel2 add port 5-8 untag
create vlan Tunnel2 add port 31,32 tagged
```

On the BlackDiamond, the configuration is:

```
config dot1q ethertype 9100
enable jumbo-frame ports all
config jumbo-frame size 1530
```

```
create vlan tunnel1
config vlan tunnel1 tag 50
config vlan tunnel1 add port 1:1-1:2 tagged
create vlan tunnel2
config vlan tunnel2 tag 60
config vlan tunnel2 add port 1:1-1:2 tagged
```

Specific to this configuration, a Layer 1 or Layer 2 redundancy method would also be employed, such as Spanning Tree or other methods ExtremeWare offers.

MAC-Based VLANs

MAC-Based VLANs allow physical ports to be mapped to a VLAN based on the source MAC address learned in the FDB. This feature allows you to designate a set of ports that have their VLAN membership dynamically determined by the MAC address of the end station that plugs into the physical port. You can configure the source MAC address-to-VLAN mapping either offline or dynamically on the switch. For example, you could use this application for a roaming user who wants to connect to a network from a conference room. In each room, the user plugs into one of the designated ports on the switch and is mapped to the appropriate VLAN. Connectivity is maintained to the network with all of the benefits of the configured VLAN in terms of QoS, routing, and protocol support.

MAC-Based VLAN Guidelines

When using the MAC-to-VLAN mapping, consider the following guidelines:

- A port can only accept connections from an endstation/host and should not be connected to a layer-2 repeater device. Connecting to a layer-2 repeater device can cause certain addresses to not be mapped to their respective VLAN if they are not correctly configured in the MAC-VLAN configuration database. If a repeater device is connected to a MAC-Based VLAN port, and the configured MAC-to-VLAN mapped station enters on the repeater, any endstation that is attached to the repeater can be mapped to that VLAN while the configured endstation is active in that VLAN. Upon removal of the configured MAC-to-VLAN endstation, all other endstations lose connectivity.
- Groups are used as a security measure to allow a MAC address to enter into a VLAN only when the group mapping matches the port mapping.

As an example, the following configuration allows MAC 00:00:00:00:00:aa to enter into the VLAN only on ports 10 and 11 because of membership in group 100:

```
* Summit48:50 # show mac
Port      Vlan          Group      State
10        MacVlanDiscover 100        Discover
11        MacVlanDiscover 100        Discover
12        MacVlanDiscover any         Discover
13        MacVlanDiscover any         Discover
14        MacVlanDiscover any         Discover
Total Entries in Database:2
  Mac          Vlan      Group
00:00:00:00:00:aa  sales    100
00:00:00:00:00:01  sales    any
2 matching entries
```

- The group “any” is equivalent to the group “0”. Ports that are configured as “any” allow any MAC address to be assigned to a VLAN, regardless of group association.
- Partial configurations of the MAC to VLAN database can be downloaded to the switch using the timed download configuration feature.

MAC-Based VLAN Limitations

The following list contains the limitations of MAC-based VLANs:

- Ports participating in MAC VLANs must first be removed from any static VLANs.
- The MAC- to-VLAN mapping can only be associated with VLANs that exist on the switch.
- A MAC address cannot be configured to associate with more than 1 VLAN. If this is attempted, the MAC address is associated with the most recent VLAN entry in the MAC-to-VLAN database.
- The feature is intended to support one client per physical port. Once a client MAC address has successfully registered, the VLAN association remains until the port connection is dropped or the FDB entry ages out.

MAC-Based VLAN Example

In this following example, three VLANs are created: *engineering*, *marketing*, and *sales*. A single MAC address is associated with each VLAN. The MAC address 00:00:00:00:00:02 has a group number of “any” or “0” associated with it, allowing it to be plugged into any port that is in MacVlanDiscover mode (ports 10-15 in this case). The MAC address

00:00:00:00:00:01 has a group number of 10 associated with it, and can only be assigned to a VLAN if inserted into ports 16 or 17. The MAC address 00:00:00:00:00:03 has a group number of 200 associated with it and can only be inserted into ports 18 through 20.

```
enable mac-vlan mac-group any ports 10-15
enable mac-vlan mac-group 10 ports 16-17
enable mac-vlan mac-group 200 ports 18-20
config mac-vlan add mac-address 00:00:00:00:00:01 mac-group 10
engineering
config mac-vlan add mac-address 00:00:00:00:00:02 mac-group any marketing
config mac-vlan add mac-address 00:00:00:00:00:03 mac-group 200 sales
```

Timed Configuration Download for MAC-Based VLANs

To allow centralized control of MAC-based VLANs over multiple switches, a timed TFTP configuration download allows you to download incremental configuration files from a primary or secondary server at specified time intervals. The timed downloads are configurable in 24 hour intervals. When a switch reboots, the configuration is automatically downloaded immediately after booting, per the configured primary and secondary servers.

To configure the primary and/or secondary server and file name, use the following command:

```
config download server [primary | secondary] [<host_name> | <ip_address>]
<filename>
```

To enable timed interval downloads, use the following command:

```
download configuration every <hour:minute>
```

To display timed download information, use the following command:

```
show switch
```

Example

In relation to MAC-based VLANs, the downloaded file is an ASCII file that consists of CLI commands used to configure the most recent MAC-to-VLAN database. This feature is different from the normal download configuration command in that it allows incremental configuration without the automatic rebooting of the switch.

The following example shows an incremental configuration file for MAC-based VLAN information that updates the database and saves changes:

```
config mac-vlan add mac-address 00:00:00:00:00:01 mac-group any
engineering
config mac-vlan add mac-address 00:00:00:00:ab:02 mac-group any
engineering
config mac-vlan add mac-address 00:00:00:00:cd:04 mac-group any sales
.
.
config mac-vlan add mac-address 00:00:00:00:ab:50 mac-group any sales
config mac-vlan add mac-address 00:00:00:00:cd:60 mac-group any sales
save
```



6 Forwarding Database (FDB)

This chapter describes the following topics:

- Overview of the FDB on page 6-1
- Associating QoS Profiles with an FDB Entry on page 6-5
- MAC-Based Security on page 6-7
- Displaying FDB Entries on page 6-10

Overview of the FDB

The switch maintains a database of all media access control (MAC) addresses received on all of its ports. It uses the information in this database to decide whether a frame should be forwarded or filtered.

FDB Contents

Each FDB entry consists of the MAC address of the device, an identifier for the port and VLAN on which it was received, and the age of the entry. Frames destined for MAC addresses that are not in the FDB are flooded to all members of the VLAN.

How FDB Entries Get Added

Entries are added into the FDB in the following ways:

- The switch can learn entries by examining packets it receives. The system updates its FDB with the source MAC address from a packet, the VLAN, and the port identifier on which the source packet is received.

The ability to learn MAC addresses can be enabled or disabled on a port-by-port basis. You can also limit the number of addresses that can be learned, or you can “lock down” the current entries and prevent additional MAC address learning.

- You can enter and update entries using the command-line interface (CLI).
- Certain static entries are added by the system upon switch boot up.

FDB Entry Types

FDB entries may be dynamic or static, and may be permanent or non-permanent. The following describes the types of entries that can exist in the FDB:

- **Dynamic entries** — A dynamic entry is learned by the switch by examining packets to determine the source MAC address, VLAN, and port information. The switch then creates or updates an FDB entry for that MAC address. Initially, all entries in the database are dynamic, except for certain entries created by the switch at boot up.

Dynamic entries are flushed and relearned (updated) when any of the following take place:

- A VLAN is deleted.
- A VLAN identifier (VLANid) is changed.
- A port mode is changed (tagged/untagged).
- A port is deleted from a VLAN.
- A port is disabled.
- A port enters blocking state.
- A port QoS setting is changed.
- A port goes down (link down).

A *non-permanent dynamic entry* is initially created when the switch identifies a new source MAC address that does not yet have an entry in the FDB. The entry may then be updated as the switch continues to encounter the address in the packets it examines. These entries are identified by the “d” flag in `show fdb` output.

A *permanent dynamic entry* is created by command through the CLI, but may then be updated as the switch encounters the MAC address in the packets that it examines. A permanent dynamic entry is typically used to associate QoS profiles with the FDB entry. Permanent dynamic entries are identified by the “p” and “d” flags in `show fdb` output.

Both types of dynamic entries age—a dynamic entry will be removed from the FDB (aged-out) if the device does not transmit for a specified period of time (the aging time). This prevents the database from becoming full with obsolete entries by ensuring that when a device is removed from the network, its entry is deleted from the database. The aging time is configurable. For more information about setting the aging time, see “Configuring the FDB Aging Time” later in this chapter.

- **Static entries** —A static entry does not age, and does not get updated through the learning process. It is maintained exactly as it was created. Conditions that cause dynamic entries to be updated, such as VLAN or port configuration changes, do not affect static entries.

If the same MAC address is detected on another virtual port that is not defined in the static FDB entry for the MAC address, it is handled as a blackhole entry.

A *permanent static entry* is created through the command line interface, and can be used to associate QoS profiles with a non-aging FDB entry. Permanent static entries are identified by the “s” and “p” flags in `show fdb` output.

A *locked static entry* is an entry that was originally learned dynamically, but has been made static (locked) using the MAC address lock-down feature. It is identified by the “s” and “l” flags in `show fdb` output. See “MAC Address Lock Down” on page 6-9 for more information about MAC address lock-down.

Non-permanent static entries are created by the switch software for various reasons, typically upon switch boot up. They are identified by the “s” flag in `show fdb` output.

If the FDB entry aging time is set to zero, all entries in the database are considered static, non-aging entries. This means that they do not age, but they are still deleted if the switch is reset.

- **Permanent entries** — Permanent entries are retained in the database if the switch is reset or a power off/on cycle occurs. Permanent entries must be created by the system administrator through the command line interface. A permanent entry can either be a unicast or multicast MAC address.

Permanent entries may be static, meaning they do not age or get updated, or they may be dynamic, meaning that they do age and can be updated via learning.

Permanent entries can have QoS profiles associated with the MAC address. A different QoS profiles may be associated with the MAC address when it is a

destination address (an egress QoS profile) than when it is a source address (ingress QoS profile).

The stand-alone switches can support a maximum of 64 permanent entries, and the modular switches support a maximum of 254 permanent entries.

- **Blackhole entries** — A blackhole entry configures the switch to discard packets with a specified MAC address. Blackhole entries are useful as a security measure or in special circumstances where a specific source or destination address must be discarded. Blackhole entries may be created through the CLI, or they may be created by the switch when a port's learning limit has been exceeded.

Blackhole entries are treated like permanent entries in the event of a switch reset or power off/on cycle. Blackhole entries are never aged-out of the database.

Disabling MAC Address Learning

By default, MAC address learning is enabled on all ports. You can disable learning on specified ports using the following command:

```
disable learning {flood-list} ports <portlist>
```

If MAC address learning is disabled, only broadcast traffic, EDP traffic, and packets destined to a permanent MAC address matching that port number, are forwarded. Use this command in a secure environment where access is granted via permanent forwarding databases (FDBs) per port.

Using the `flood-list` option disables port learning and configures the ports to act like a hub. When flooding is enabled on a particular port, *all* frames and packets are passed on to other member ports that also have flooding enabled. This includes all broadcast, multicast, known and unknown unicast packets (including EPD). To make effective use of this feature you should have flooding enabled on more than one port.

Learning and flooding are mutually exclusive. To enable flooding, learning must be disabled. When ports are configured for flooding, the FDB will be flushed for the entire system, which means all the entries in the dynamic FDB must be relearned.

To disable flooding, enable port learning on the affected ports.

Associating QoS Profiles with an FDB Entry

You can associate QoS profiles with a MAC address (and VLAN) of a device by creating a permanent FDB entry and specifying QoS profiles for ingress or egress, or both. The permanent FDB entry can be either dynamic (it is learned and can be aged out) or static.

To associate a QoS profile with a dynamic FDB entry, use the following command:

```
create fdbentry <mac_address> vlan <name> dynamic
[ qosprofile <qosprofile> { ingress-qosprofile <iqosprofile> } |
ingress-qosprofile <qosprofile> { qosprofile <qosprofile> } ]
```

This command associates QoS profiles with packets received from or destined for the specified MAC address, while still allowing the FDB entry to be dynamically learned. If you specify only the ingress QoS profile, the egress QoS profile defaults to none, and vice-versa. If both profiles are specified, the source MAC address of an ingress packet and the destination MAC address of an egress packet are examined for QoS profile assignment.

The FDB entry is not actually created until the MAC address is encountered as the source MAC address in a packet. Thus, initially the entry may not appear in the `show fdb` output. Once the entry has been learned, it is created as a permanent dynamic entry, designated by “dpm” in the flags field of the `show fdb` output.

You can use the `show fdb permanent` command to display permanent FDB entries, including their QoS profile associations.

To associate a QoS profile with a permanent FDB entry, use the following command:

```
create fdbentry <mac_address> vlan <name> ports [<portlist | all>]
{ qosprofile <qosprofile> } { ingress-qosprofile <iqosprofile> }
```

This entry will not be aged out, and no learning will occur. If the same MAC address is encountered through a virtual port not specified in the portlist, it will be handled as a blackhole entry.



Note: For more information on QoS profiles, see Chapter 7.

FDB Configuration Examples

The following example adds a permanent static entry to the FDB:

```
create fdbentry 00:E0:2B:12:34:56 vlan marketing port 3:4
```

The permanent entry has the following characteristics:

- MAC address is 00:E0:2B:12:34:56.
- VLAN name is *marketing*.
- Slot number for this device is 3.
- Port number for this device is 4.

If the MAC address 00:E0:2B:12:34:56 is encountered on any port/VLAN other than VLAN *marketing*, port 3:4, it will be handled as a blackhole entry, and packets from that source will be dropped.

This example associates the QoS profile *qp2* with a dynamic entry for the device at MAC address 00:A0:23:12:34:56 on VLAN *net34* that will be learned by the FDB:

```
create fdbentry 00:A0:23:12:34:56 vlan net34 dynamic qosprofile qp2
```

This entry has the following characteristics:

- MAC address is 00:A0:23:12:34:56.
- VLAN name is *net34*.
- The entry will be learned dynamically.
- QoS profile *qp2* will be applied as an egress QoS profile when the entry is learned.

Configuring the FDB Aging Time

You can configure the aging time for dynamic FDB entries using the following command:

```
config fdb agingtime <seconds>
```

If the aging time is set to zero, all aging entries in the database are defined as static, nonaging entries. This means they will not age out, but non-permanent static entries can be deleted if the switch is reset.

MAC-Based Security

MAC-based security allows you to control the way the FDB is learned and populated. By managing entries in the FDB, you can block, assign priority (queues), and control packet flows on a per-address basis.

MAC-based security allows you to limit the number of dynamically-learned MAC addresses allowed per virtual port. You can also “lock” the FDB entries for a virtual port, so that the current entries will not change, and no additional addresses can be learned on the port.

You can also prioritize or stop packet flows based on the source MAC address of the ingress VLAN or the destination MAC address of the egress VLAN.

Limiting Dynamic MAC Addresses

You can set a predefined limit on the number of dynamic MAC addresses that can participate in the network. After the FDB reaches the MAC limit, all new source MAC addresses are blackholed at both the ingress and egress points. These dynamic blackhole entries prevent the MAC addresses from learning and responding to Internet control message protocol (ICMP) and address resolution protocol (ARP) packets.

To limit the number of dynamic MAC addresses that can participate in the network, use the following commands:

```
config ports [<portlist> | all] vlan <name> limit-learning <number>
```

This command specifies the number of dynamically-learned MAC entries allowed for these ports in this VLAN. The range is 0 to 500,000 addresses.

When the learned limit is reached, all new source MAC addresses are blackholed at both the ingress and egress points. This prevent these MAC addresses from learning and responding to Internet control message protocol (ICMP) and address resolution protocol (ARP) packets.

Dynamically learned entries still get aged, and can be cleared. If entries are cleared or aged out after the learning limit has been reached, new entries will then be able to be learned until the limit is reached again.

Permanent static and permanent dynamic entries can still be added and deleted using the `create fdbentry` and `delete fdbentry` commands. These override any dynamically learned entries.

For ports that have a learning limit in place, the following traffic will still flow to the port:

- Packets destined for permanent MAC addresses and other non-blackholed MAC addresses
- Broadcast traffic
- EDP traffic

Traffic from the permanent MAC and any other non-blackholed MAC addresses will still flow from the virtual port.

To remove the learning limit, use the following command:

```
config ports [<portlist> | all] vlan <name> unlimited-learning
```

To verify the configuration, use the following commands:

```
show vlan <name> security
```

This command displays the MAC security information for the specified VLAN.

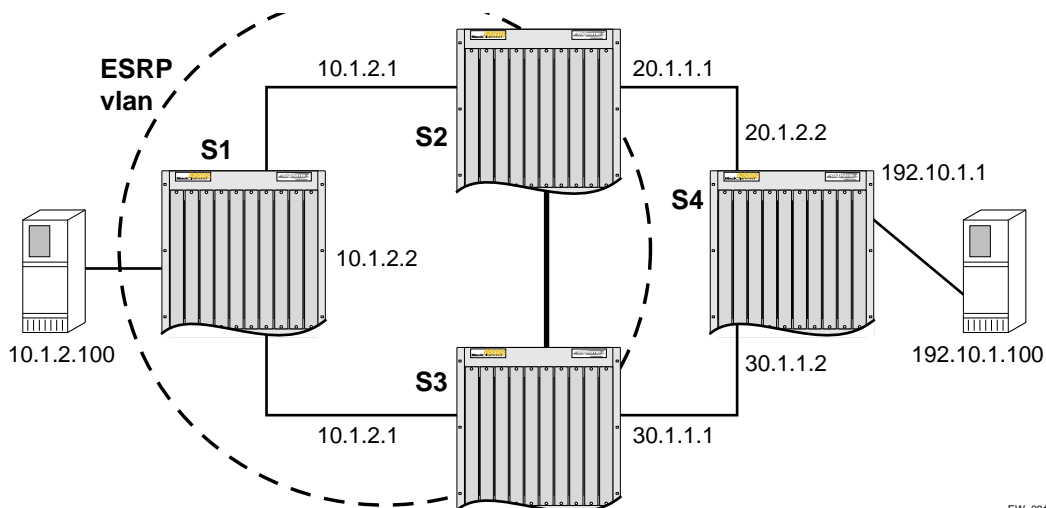
```
show ports <portlist> info detail
```

This command displays detailed information, including MAC security information, for the specified port.

Limiting MAC Addresses with ESRP Enabled

If you configure a MAC address limit on VLANs that have ESRP enabled, you should add an additional back-to-back link (that has no MAC address limit on these ports) between the ESRP-enabled switches. Doing so prevents ESRP PDU from being dropped due to MAC address limit settings.

Figure 6-1 is an example of configuring a MAC address limit on an ESRP-enabled VLAN.



EW_081

Figure 6-1: MAC address limits and ESRP-enabled VLANs

In Figure 6-1, S2 and S3 are ESRP-enabled switches, while S1 is an ESRP-aware (regular layer 2) switch. Configuring a MAC address limit on all S1 ports might prevent ESRP communication between S2 and S3. To resolve this, you should add a back-to-back link between S2 and S3. This link is not needed if MAC address limiting is configured only on S2 and S3, but not on S1.

MAC Address Lock Down

In addition to limiting learning on virtual ports, you can lock down the existing dynamic FDB entries and prevent any additional learning using the following command:

```
config ports [<portlist> | all] vlan <name> lock-learning
```

This command causes all dynamic FDB entries associated with the specified VLAN and ports to be converted to locked static entries. It also sets the learning limit to zero, so that no new entries can be learned. All new source MAC addresses are blackholed.

Locked entries do not get aged, but can be cleared. Permanent static entries can still be added and deleted. Permanent dynamic entries do not override locked static entries.

For ports that have lock-down in effect, the following traffic will still flow to the port:

- Packets destined for the permanent MAC and other non-blackholed MAC addresses
- Broadcast traffic
- EDP traffic

Traffic from the permanent MAC will still flow from the virtual port.

To remove MAC address lock down, use the following command:

```
config ports [<portlist> | all] vlan <name> unlock-learning
```

When you remove the lock down using the unlock-learning option, the learning-limit is reset to unlimited, and all associated entries in the FDB are flushed.

Displaying FDB Entries

To display FDB entries, use the following command:

```
show fdb {<mac_address> | vlan <name> | ports <portlist> | permanent}
```

where the following is true:

- `mac_address` — Displays the entry for a particular MAC address.
- `vlan <name>` — Displays the entries for a VLAN.
- `ports <portlist>` — Displays the entries for a set of ports or slots and ports.
- `permanent` — Displays all permanent entries, including the ingress and egress QoS profiles.

With no options, the command displays all FDB entries.

See the *ExtremeWare Command Reference Guide* for details of the commands related to the FDB.

7

Quality of Service (QoS)

This chapter covers the following topics:

- Overview of Policy-Based Quality of Service on page 7-2
- Applications and Types of QoS on page 7-3
- Configuring QoS on page 7-5
- QoS Profiles on page 7-6
- Traffic Groupings on page 7-8
 - IP-Based Traffic Groupings on page 7-9
 - MAC-Based Traffic Groupings on page 7-9
 - Explicit Class of Service (802.1p and DiffServ) Traffic Groupings on page 7-11
 - Physical and Logical Groupings on page 7-18
- Configuring QoS Traffic Grouping Priorities on page 7-19
- Verifying Configuration and Performance on page 7-20
- Modifying a QoS Configuration on page 7-22
- Bi-Directional Rate Shaping on page 7-22
- Dynamic Link Context System on page 7-26

Policy-based Quality of Service (QoS) is a feature of ExtremeWare and the Extreme switch architecture that allows you to specify different service levels for traffic traversing the switch. Policy-based QoS is an effective control mechanism for networks that have heterogeneous traffic patterns. Using Policy-based QoS, you can specify the service level that a particular traffic type receives.

Overview of Policy-Based Quality of Service

Policy-based QoS allows you to protect bandwidth for important categories of applications or specifically limit the bandwidth associated with less critical traffic. For example, if voice-over-IP traffic requires a reserved amount of bandwidth to function properly, using policy-based QoS, you can reserve sufficient bandwidth critical to this type of application. Other applications deemed less critical can be limited so as to not consume excessive bandwidth. The switch contains separate hardware queues on every physical port. Each hardware queue is programmed by ExtremeWare with bandwidth management and prioritization parameters. The bandwidth management and prioritization parameters that modify the forwarding behavior of the switch affect how the switch transmits traffic for a given hardware queue on a physical port.

The switch tracks and enforces the minimum and maximum percentage of bandwidth utilization transmitted on every hardware queue for every port. When two or more hardware queues on the same physical port are contending for transmission, the switch prioritizes bandwidth use so long as their respective bandwidth management parameters are satisfied. Switch products with the “i” chipset can be configured with up to eight physical queues per port, while other Extreme switches can be configured with up to four physical queues per port.



Note: Policy-based QoS has no impact on switch performance. Using even the most complex traffic groupings has no cost in terms of switch performance.

Policy-based QoS can be configured to perform per-port Random Early Detection (RED). Using this capability, the switch detects when traffic is filling up in any of the eight hardware queues, and performs a random discard on subsequent packets, based on the configured RED drop-probability.

Instead of dropping sessions during times when the queue depth is exceeded, RED causes the switch to lower session throughput. The destination node detects the dropped packet, and, using standard TCP windowing mechanisms, slows the transmission from the source node. RED drop-probability is configured on a system-wide basis, and has a valid range from 0% to 100%. Only switches and modules with the “i” chipset can use RED.

Applications and Types of QoS

Different applications have different QoS requirements. The following applications are ones that you will most commonly encounter and need to prioritize:

- Voice applications
- Video applications
- Critical database applications
- Web browsing applications
- File server applications

General guidelines for each traffic type are given below and summarized in Table 7-1. Consider them as general guidelines and not strict recommendations. Once QoS parameters are set, you can monitor the performance of the application to determine if the actual behavior of the applications matches your expectations. It is very important to understand the needs and behavior of the particular applications you wish to protect or limit. Behavioral aspects to consider include bandwidth needs, sensitivity to latency and jitter, and sensitivity and impact of packet loss.

Voice Applications

Voice applications typically demand small amounts of bandwidth. However, the bandwidth must be constant and predictable because voice applications are typically sensitive to latency (inter-packet delay) and jitter (variation in inter-packet delay). The most important QoS parameter to establish for voice applications is minimum bandwidth, followed by priority.

Video Applications

Video applications are similar in needs to voice applications, with the exception that bandwidth requirements are somewhat larger, depending on the encoding. It is important to understand the behavior of the video application being used. For example, in the playback of stored video streams, some applications can transmit large amounts of data for multiple streams in one “spike,” with the expectation that the end-stations will buffer significant amounts of video-stream data. This can present a problem to the network infrastructure, because it must be capable of buffering the transmitted spikes where there are speed differences (for example, going from Gigabit Ethernet to Fast Ethernet). Key QoS parameters for video applications include minimum bandwidth, priority, and possibly buffering (depending upon the behavior of the application).

Critical Database Applications

Database applications, such as those associated with ERP, typically do not demand significant bandwidth and are tolerant of delay. You can establish a minimum bandwidth using a priority less than that of delay-sensitive applications.

Web Browsing Applications

QoS needs for Web browsing applications cannot be generalized into a single category. For example, ERP applications that use a browser front-end may be more important than retrieving daily news information. Traffic groupings can typically be distinguished from each other by their server source and destinations. Most browser-based applications are distinguished by the dataflow being asymmetric (small dataflows from the browser client, large dataflows from the server to the browser client).

An exception to this may be created by some Java™ -based applications. In addition, Web-based applications are generally tolerant of latency, jitter, and some packet loss, however small packet-loss may have a large impact on perceived performance due to the nature of TCP. The relevant parameter for protecting browser applications is minimum bandwidth. The relevant parameter for preventing non-critical browser applications from overwhelming the network is maximum bandwidth. In addition, RED can be used to reduce session loss if the queue that floods Web traffic becomes over-subscribed.

File Server Applications

With some dependencies on the network operating system, file serving typically poses the greatest demand on bandwidth, although file server applications are very tolerant of latency, jitter, and some packet loss, depending on the network operating system and the use of TCP or UDP.



Note: Full-duplex links should be used when deploying policy-based QoS. Half-duplex operation on links can make delivery of guaranteed minimum bandwidth impossible.

Table 7-1 summarizes QoS guidelines for the different types of network traffic.

Table 7-1: Traffic Type and QoS Guidelines

Traffic Type	Key QoS Parameters
Voice	Minimum bandwidth, priority
Video	Minimum bandwidth, priority, buffering (varies)
Database	Minimum bandwidth
Web browsing	Minimum bandwidth for critical applications, maximum bandwidth for non-critical applications, RED
File server	Minimum bandwidth

Configuring QoS

To configure QoS, you define how your switch responds to different categories of traffic by creating and configuring QoS profiles. You then group traffic into categories (according to application, as previously discussed) and assign each category to a QoS profile. Configuring QoS is a three-step process:

1 Configure the QoS profile.

QoS profile — A class of service that is defined through minimum and maximum bandwidth parameters, configuration of buffering and RED, and prioritization settings. The bandwidth and level of service that a particular type of traffic or traffic grouping receives is determined by assigning it to a QoS profile.

2 Create traffic groupings.

Traffic grouping — A classification or traffic type that has one or more attributes in common. These can range from a physical port to a VLAN to IP Layer 4 port information. You assign traffic groupings to QoS profiles to modify switch forwarding behavior. Traffic groupings transmitting out the same port that are assigned to a particular QoS profile share the assigned bandwidth and prioritization characteristics, and hence share the class of service.

3 Monitor the performance of the application with the QoS monitor to determine whether the policies are meeting the desired results.

The next sections describe each of these QoS components in detail.

QoS Profiles

A QoS profile defines a class of service by specifying traffic behavior attributes, such as bandwidth. The parameters that make up a QoS profile include:

- **Minimum bandwidth** — The minimum percentage of total link bandwidth that is reserved for use by a hardware queue on a physical port. Bandwidth unused by the queue can be used by other queues. The minimum bandwidth for all queues should add up to less than 90%. The default value on all minimum bandwidth parameters is 0%.
- **Maximum bandwidth** — The maximum percentage of total link bandwidth that can be transmitted by a hardware queue on a physical port. The default value on all maximum bandwidth parameters is 100%.
- **Priority** — The level of priority assigned to a hardware queue on a physical port. Switch products that use the “i” chipset have eight different available priority settings. Other Extreme switches have four available priority settings. By default, each of the default QoS profiles is assigned a unique priority. You would use prioritization when two or more hardware queues on the same physical port are contending for transmission on the same physical port, only after their respective bandwidth management parameters have been satisfied. If two hardware queues on the same physical port have the same priority, a round-robin algorithm is used for transmission, depending on the available link bandwidth.
 - When configured to do so, the priority of a QoS profile can determine the 802.1p bits used in the priority field of a transmitted packet (described later).
 - On switch products using the “i” chipset, the priority of a QoS profile determines the DiffServ code point value used in an IP packet when the packet is transmitted (described later).
- **Buffer** — This parameter reserves buffer memory for use exclusively by a QoS profile across all affected ports. The default value for buffer settings is 0%. The sum of all QoS profile buffer parameters should not exceed 100%. The `maxbuf` parameter allows you to set a maximum buffer size (in Kbytes or Mbytes) for each queue, so that a single queue will not consume all of the un-allocated buffer space. The default buffer size is 256K. You should not modify the buffer parameter unless specific situations and application behavior indicate.

A QoS profile does not alter the behavior of the switch until it is assigned to a traffic grouping. Recall that QoS profiles are linked to hardware queues. There are multiple hardware queues per physical port. By default, a QoS profile links to the identical hardware queue across all the physical ports of the switch.

Four or eight default QoS profiles are provided, depending on the chipset used in the switch. The default QoS profiles cannot be deleted. Also by default, a QoS profile maps directly to a specific hardware queue across all physical ports. The settings for the default QoS profiles for Summit chipset products are summarized in Table 7-2. The settings for the default QoS parameters for “i” chipset products are summarized in Table 7-3.

Table 7-2: QoS Profiles for Summit Chipset Products

Profile Name	Hardware Queue	Priority	Buffer	Minimum Bandwidth	Maximum Bandwidth
Qp1	Q0	Low	0	0%	100%
Qp2	Q1	Normal	0	0%	100%
Qp3	Q2	Medium	0	0%	100%
Qp4	Q3	High	0	0%	100%

Table 7-3: QoS Parameters for “i” Chipset Products

Profile Name	Hardware Queue	Priority	Buffer	Minimum Bandwidth	Maximum Bandwidth
Qp1	Q0	Low	0	0%	100%
Qp2	Q1	Lowhi	0	0%	100%
Qp3	Q2	Normal	0	0%	100%
Qp4	Q3	Normalhi	0	0%	100%
Qp5	Q4	Medium	0	0%	100%
Qp6	Q5	Mediumhi	0	0%	100%
Qp7	Q6	High	0	0%	100%
Qp8	Q7	Highhi	0	0%	100%

When using Summit chipset modules in a BlackDiamond chassis, profiles map as shown in Table 7-4.

Table 7-4: First Generation BlackDiamond Module Profile Mapping

Summit Chipset Module QoS Profile	“i” series Module QoS Profile
Qp1	Qp1
Qp1	Qp2
Qp2	Qp3
Qp2	Qp4
Qp3	Qp5
Qp3	Qp6
Qp4	Qp7
Qp4	Qp8

Traffic Groupings

Once a QoS profile is modified for bandwidth and priority, you assign traffic a grouping to the profile. A *traffic grouping* is a classification of traffic that has one or more attributes in common. Traffic is typically grouped based on the applications discussed starting on page 7-3.

Traffic groupings are separated into the following categories for discussion:

- IP-based information, such as IP source/destination and TCP/UDP port information
- Destination MAC (MAC QoS groupings)
- Explicit packet class of service information, such as 802.1p or DiffServ (IP TOS)
- Physical/logical configuration (physical source port or VLAN association)

In the event that a given packet matches two or more grouping criteria, there is a predetermined precedence for which traffic grouping will apply. In general, the more specific traffic grouping takes precedence. By default, all traffic groupings are placed in the QoS profile Qp1. The supported traffic groupings are listed in Table 7-5. The groupings are listed in order of precedence (highest to lowest). The four types of traffic groupings are described in detail on the following pages.

Table 7-5: Traffic Groupings by Precedence

IP Information (Access Lists) Groupings

- Access list precedence determined by user configuration

Destination Address MAC-Based Groupings

- Permanent
- Dynamic
- Blackhole
- Broadcast/unknown rate limiting

Explicit Packet Class of Service Groupings

- DiffServ (IP TOS)
- 802.1P

Physical/Logical Groupings

- VLAN
 - Source port
-

IP-Based Traffic Groupings

IP-based traffic groupings are based on any combination of the following items:

- IP source or destination address
- TCP/UDP or other layer 4 protocol
- TCP/UDP port information

IP-based traffic groupings are defined using access lists. Access lists are discussed in detail in Chapter 8. By supplying a named QoS profile at the end of the access list command syntax, you can prescribe the bandwidth management and priority handling for that traffic grouping. This level of packet filtering has no impact on performance.

MAC-Based Traffic Groupings

QoS profiles can be assigned to destination MAC addresses. MAC-based traffic groupings are configured using the following command:

```
create fdbentry <mac_address> vlan <name> [blackhole | port <portlist>
| dynamic] qosprofile <qosprofile>
```

The MAC address options, defined below, are as follows:

- Permanent
- Dynamic
- Blackhole
- Broadcast/unknown rate limiting

Permanent MAC addresses

Permanent MAC addresses can be assigned a QoS profile whenever traffic is destined to the MAC address. This can be done when you create a permanent FDB entry. For example:

```
create fdbentry 00:11:22:33:44:55 vlan default port 4:1 qosprofile qp2
```

Dynamic MAC Addresses

Dynamic MAC addresses can be assigned a QoS profile whenever traffic is destined to the MAC address. For any port on which the specified MAC address is learned in the specified VLAN, the port is assigned the specified QoS profile. For example:

```
create fdbentry 00:11:22:33:44:55 vlan default dynamic qosprofile qp3
```

The QoS profile is assigned when the MAC address is learned. If a client's location moves, the assigned QoS profile moves with the device. If the MAC address entry already exists in the FDB, you can clear the forwarding database so that the QoS profile can be applied when the entry is added again. Use the following command to clear the FDB:

```
clear fdb
```

Blackhole MAC Address

Using the `blackhole` option configures the switch to not forward any packets to the destination MAC address on any ports for the VLAN specified. The `blackhole` option is configured using the following command:

```
create fdbentry 00:11:22:33:44:55 vlan default blackhole
```

Broadcast/Unknown Rate Limiting MAC Address

It is possible to assign broadcast and unknown destination packets to a QoS profile that has the desired priority and bandwidth parameters. Broadcast/unknown rate limiting is an extension of the QoS feature used for destination MAC addresses.

For example, if you want to limit broadcast and unknown traffic on the VLAN *default* to the bandwidth and priority defined in QoS profile *qp3*, the command is:

```
create fdbentry ff:ff:ff:ff:ff:ff vlan default dynamic qp3
```



Note: P multicast traffic is subject to broadcast and unknown rate limiting only when IGMP snooping is disabled.

Verifying MAC-Based QoS Settings

To verify any of the MAC-based QoS settings, use either the command

```
show fdb permanent
```

or the command

```
show qosprofile <qosprofile>
```

Explicit Class of Service (802.1p and DiffServ) Traffic Groupings

This category of traffic groupings describes what is sometimes referred to as *explicit packet marking*, and refers to information contained within a packet intended to explicitly determine a class of service. That information includes:

- IP DiffServ code points, formerly known as IP TOS bits
- Prioritization bits used in IEEE 802.1p packets

An advantage of explicit packet marking is that the class of service information can be carried throughout the network infrastructure, without repeating what can be complex traffic grouping policies at each switch location. Another advantage is that end stations can perform their own packet marking on an application-specific basis. Extreme switch products have the capability of observing and manipulating packet marking information with no performance penalty.

Extreme products that use the “i” chipset support DiffServ capabilities. Products that do not use the “i” chipset do not support DiffServ capabilities. The documented capabilities for 802.1p priority markings or DiffServ capabilities (if supported) are not

impacted by the switching or routing configuration of the switch. For example, 802.1p information can be preserved across a routed switch boundary and DiffServ code points can be observed or overwritten across a layer 2 switch boundary.

Configuring 802.1p Priority

Extreme switches support the standard 802.1p priority bits that are part of a tagged Ethernet packet. The 802.1p bits can be used to prioritize the packet, and assign it to a particular QoS profile.

When a packet arrives at the switch, the switch examines the 802.1p priority field maps it to a specific hardware queue when subsequently transmitting the packet. The 802.1p priority field is located directly following the 802.1Q type field, and preceding the 802.1Q VLAN ID, as shown in Figure 7-1.

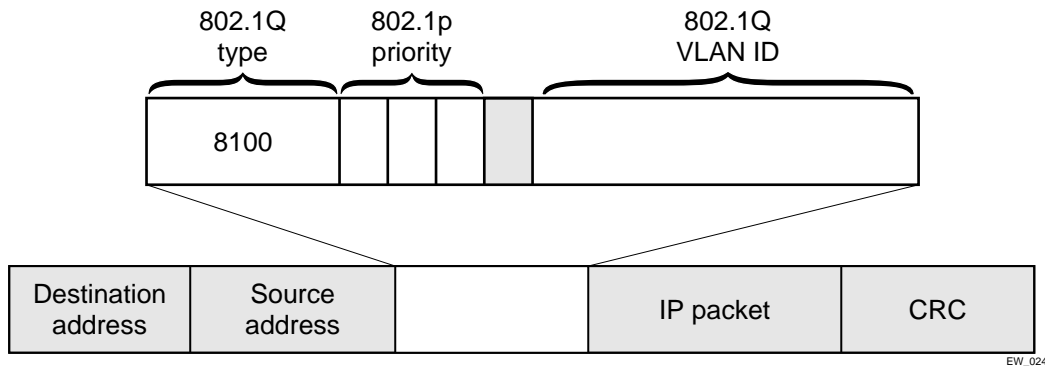


Figure 7-1: Ethernet packet encapsulation

Observing 802.1p Information

When ingress traffic that contains 802.1p prioritization information is detected by the switch, the traffic is mapped to various hardware queues on the egress port of the switch. Switches that use the “i” chipset support eight hardware queues, all other products support four hardware queues. The transmitting hardware queue determines the bandwidth management and priority characteristics used when transmitting packets.

To control the mapping of 802.1p prioritization values to hardware queues, 802.1p prioritization values can be mapped to a QoS profile. The default mapping of each 802.1p priority value to QoS profile is shown in Table 7-6.

Table 7-6: 802.1p Priority Value-to-QoS Profile Default Mapping

Priority Value	QoS Profile Summit Chipset	QoS Profile “i” Chipset
0	Qp1	Qp1
1	Qp1	Qp2
2	Qp2	Qp3
3	Qp2	Qp4
4	Qp3	Qp5
5	Qp3	Qp6
6	Qp4	Qp7
7	Qp4	Qp8

Changing the Default 802.1p Mapping

By default, a QoS profile is mapped to a hardware queue, and each QoS profile has configurable bandwidth parameters and priority. In this way, an 802.1p priority value seen on ingress can be mapped to a particular QoS profile and with specific bandwidth management and priority behavior.

To change the default mappings of QoS profiles to 802.1p priority values, use the following command:

```
config dot1p type <dot1p_priority> qosprofile <qosprofile>
```

Configuring 802.1p Priority

When a packet is transmitted by the switch, you can configure the 802.1p priority field that is placed in the 802.1Q tag. You can configure the priority to be a number between 0 and 7, using the following command:

```
config vlan <name> priority <number>
```

Replacing 802.1p Priority Information

By default, 802.1p priority information is not replaced or manipulated, and the information observed on ingress is preserved when transmitting the packet. This behavior is not affected by the switching or routing configuration of the switch.

However, the switch is capable of inserting and/or overwriting 802.1p priority information when it transmits an 802.1Q tagged frame. If 802.1p replacement is enabled,

the 802.1p priority information that is transmitted is determined by the hardware queue that is used when transmitting the packet. To replace 802.1p priority information, use the following command:

```
enable dot1p replacement ports [<portlist> | all]
```

802.1p priority information is replaced according to the hardware queue that is used when transmitting from the switch. The mapping is described in Table 7-7 for switches based on the “i” chipset and for other Extreme switches. This mapping cannot be changed.

Table 7-7: Queue to 802.1p Priority Replacement Value

Hardware Queue Summit Chipset	Hardware Queue “i” Chipset	802.1p Priority Replacement Value
Q0	Q0	0
	Q1	1
Q1	Q2	2
	Q3	3
Q2	Q4	4
	Q5	5
Q3	Q6	6
	Q7	7

Configuring DiffServ

Contained in the header of every IP packet is a field for IP Type of Service (TOS), now also called the DiffServ field. The TOS field is used by the switch to determine the type of service provided to the packet.

Observing DiffServ code points as a traffic grouping mechanism for defining QoS policies and overwriting the Diffserv code point fields are supported in switches using the “i” chipset.

Figure 7-2 shows the encapsulation of an IP packet header.

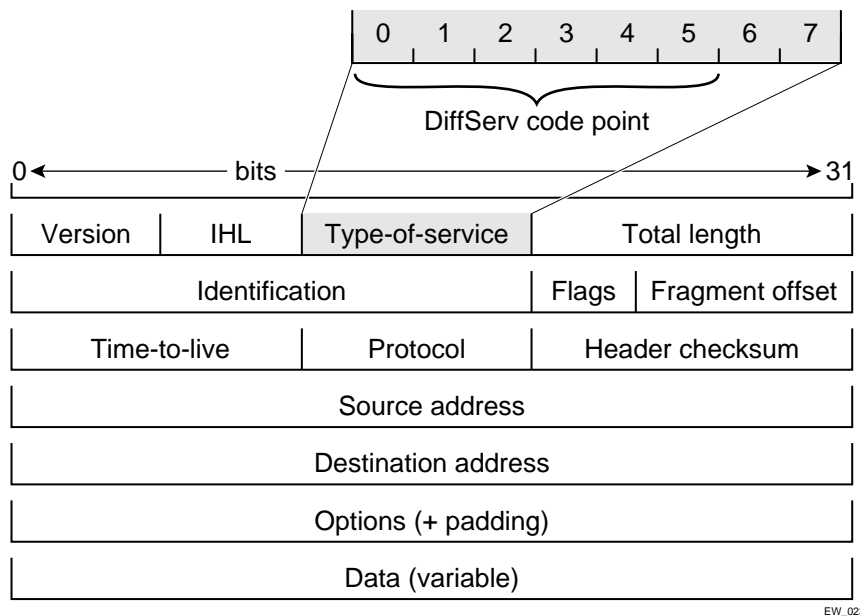


Figure 7-2: IP packet header encapsulation

Observing DiffServ Information

When a packet arrives at the switch on an ingress port, the switch examines the first six of eight TOS bits, called the *code point*. The switch can assign the QoS profile used to subsequently transmit the packet based on the code point. The QoS profile controls a hardware queue used when transmitting the packet out of the switch, and determines the forwarding characteristics of a particular code point. Viewing DiffServ information can be enabled or disabled; by default it is disabled. To view DiffServ information, use the following command:

```
enable diffserv examination ports [<portlist> | all]
```

Changing DiffServ Code point assignments in the QoS Profile

Because the code point uses six bits, it has 64 possible values ($2^6 = 64$). By default, the values are grouped and assigned to the default QoS profiles listed in Table 7-8.

Table 7-8: Default Code Point-to-QoS Profile Mapping

Code Point	QoS Profile
0-7	Qp1
8-15	Qp2
16-23	Qp3
24-31	Qp4
32-39	Qp5
40-47	Qp6
48-55	Qp7
56-63	Qp8

You can change the QoS profile assignment for all 64 code points using the following command:

```
config diffserv examination code-point <code_point> qosprofile
<qosprofile> ports [<portlist> | all]
```

Once assigned, the rest of the switches in the network prioritize the packet using the characteristics specified by the QoS profile.

Replacing DiffServ Code Points

The switch can be configured to change the DiffServ code point in the packet prior to the packet being transmitted by the switch. This is done with no impact on switch performance.

The DiffServ code point value used in overwriting a packet is determined by the 802.1p priority value. The 802.1p priority value is, in turn, determined by the hardware queue used when transmitting a packet, as described in “Replacing 802.1p Priority Information” on page 7-13.

It is not necessary to receive or transmit 802.1Q tagged frames, only to understand that the egress hardware queue, which also determines the 802.1p priority value, can also be configured to determine the DiffServ code point value if you want to replace the DiffServ code points.

To replace DiffServ code points you must enable both 802.1p replacement and DiffServ replacement using the following commands:

```
enable dot1p replacement ports [<portlist> | all]
enable diffserv replacement ports [<portlist> | all]
```

The default 802.1p priority value to code point mapping is described in Table 7-9.

Table 7-9: Default 802.1p Priority Value-to-Code Point Mapping

Hardware Queue "7" Chipset	802.1p Priority value	Code Point
Q0	0	0
Q1	1	8
Q2	2	16
Q3	3	24
Q4	4	32
Q5	5	40
Q6	6	48
Q7	7	56

You then change the 802.1p priority to DiffServ code point mapping to any code point value using the following command:

```
config diffserv replacement priority <vpri> code_point <code_point>
ports [<portlist> | all]
```

By doing so, the hardware queue used to transmit a packet determines the DiffServ value replaced in the IP packet.

To verify the DiffServ configuration, use the following command:

```
show ports <portlist> info {detail}
```

DiffServ Example

In this example, we use DiffServ to signal a class of service throughput and assign any traffic coming from network 10.1.2.x with a specific DiffServ code point. This allows all other network switches to send and observe the DiffServ code point instead of repeating the same QoS configuration on every network switch.

To configure the switch that handles incoming traffic from network 10.1.2.x, follow these steps:

1 Configure parameters of the QoS profile QP3:

```
config qp3 min 10 max 100
```

2 Assign a traffic grouping for traffic from network 10.1.2.x to qp3:

```
create access-list TenOneTwo
config TenOneTwo 10.1.2.0/24 permit qp3
```

3 To enable the switch to overwrite the DiffServ code point:

```
enable dot1p replacement
enable diffserv replacement
```

4 Configure the switch so that other switches can signal class of service that this switch should observe:

```
enable diffserv examination
```

Table 7-3 indicates that qp3 is tied to hardware queue Q2. We also know that when replacement is enabled all traffic sent out Q2 will contain code point value 16 (according to Table 7-9). If this is the desired code point to use, all traffic from 10.1.2.x will be sent out QP3 (at 10% minimum and 100% maximum) with a code point value of 16.

Physical and Logical Groupings

Two traffic groupings exist in this category:

- Source port
- VLAN

Source port

A source port traffic grouping implies that any traffic sourced from this physical port uses the indicated QoS profile when the traffic is transmitted out to any other port. To configure a source port traffic grouping, use the following command:

```
config ports <portlist> qosprofile <qosprofile>
```

In the following modular switch example, all traffic sourced from slot 5 port 7 uses the QoS profile named *qp3* when being transmitted.

```
config ports 5:7 qosprofile qp3
```

VLAN

A VLAN traffic grouping indicates that all intra-VLAN switched traffic and all routed traffic sourced from the named VLAN uses the indicated QoS profile. To configure a VLAN traffic grouping, use the following command:

```
config vlan <name> qosprofile <qosprofile>
```

For example, all devices on VLAN *servnet* require use of the QoS profile *qp4*. The command to configure this example is as follows:

```
config vlan servnet qosprofile qp4
```

Verifying Physical and Logical Groupings

To verify settings on ports or VLANs, use the following command:

```
show qosprofile <qosprofile>
```

The same information is also available for ports or VLANs using one of the following commands:

```
show ports <portlist> info {detail}
```

or

```
show vlan
```

Configuring QoS Traffic Grouping Priorities

Normally, there is a predetermined precedence for which traffic grouping applies to a given packet that matches two or more grouping criteria. In general, the more specific traffic grouping takes precedence. However, you can configure a new set of priorities using the following command:

```
config qostype priority [source-mac | dest-mac | access-list | vlan | diffserv | dot1p] <priority>
```

The valid priority values are 0 - 15. The default values are shown in Table 7-10.

Table 7-10: Traffic Grouping Priority Default Values

QoS Type	Default Value
source-mac	7
dest-mac	8

Table 7-10: Traffic Grouping Priority Default Values (continued)

QoS Type	Default Value
access-list	11
vlan	1
diffserv	3
dot1p	2

QoS types with a greater value take higher precedence. For example, to force FDB source-mac QoS to take a higher precedence over FDB dest-mac QoS, use the commands:

```
config qostype priority source-mac 9
```

where 9 is greater than the default value assigned to the dest-mac QoS type.

Traffic groupings based on the source port always have the lowest priority, and all other traffic groupings take priority. You cannot change the priority for source port-based traffic groupings.

Verifying and Resetting QoS Traffic Grouping Priorities

To verify QoS traffic grouping priority settings, use the command:

```
show qostype priority
```

To reset priority settings to their default values, use the command:

```
unconfig qostype priority
```

Verifying Configuration and Performance

Once you have created QoS policies that manage the traffic through the switch, you can use the QoS monitor to determine whether the application performance meets your expectations.

QoS Monitor

The QoS monitor is a utility that monitors the hardware queues associated with any port(s). The QoS monitor keeps track of the number of frames and the frames per second that a specific queue is responsible for transmitting on a physical port. Two

options are available: a real-time display, and a separate option for retrieving information in the background and writing it to the log.

Real-Time Performance Monitoring

The real-time display scrolls through the given portlist to provide statistics. You can choose screens for packet count and packets per second. The specific port being monitored is indicated by an asterisk (*) appearing after the port number in the display.

To view real-time switch per-port performance, use the following command:

```
show ports {<portlist>} qosmonitor
```

QoS monitor sampling is configured as follows:

- The port is monitored for 20 seconds before the switch moves on to the next port in the list.
- A port is sampled for five seconds before the packets per second (pps) value is displayed on the screen.

Background Performance Monitoring

Monitoring QoS in the background places transmit counter and any “overflow” information into the switch log. The log notification appears if one of the queues experiences an overflow condition since the last time it was sampled.

An overflow entry indicates that a queue was over-subscribed at least temporarily, and is useful for determining correct QoS settings and potential over-subscription issues.

Displaying QoS Profile Information

The QoS monitor can also be used to verify the QoS configuration and monitor the use of the QoS policies that are in place. To display QoS information on the switch, use the following command:

```
show qosprofile <qosprofile>
```

Displayed information includes:

- QoS profile name
- Minimum bandwidth
- Maximum bandwidth

- Priority
- A list of all traffic groups to which the QoS profile is applied

Additionally, QoS information can be displayed from the traffic grouping perspective by using one or more of the following commands:

- `show fdb permanent` — Displays destination MAC entries and their QoS profiles.
- `show switch` — Displays information including PACE enable/disable information.
- `show vlan` — Displays the QoS profile assignments to the VLAN.
- `show ports <portlist> info {detail}` — Displays information including QoS information for the port.

Modifying a QoS Configuration

If you make a change to the parameters of a QoS profile after implementing your configuration, the timing of the configuration change depends on the traffic grouping involved. The following rules apply:

- For destination MAC-based grouping (other than permanent), clear the MAC FDB using the command `clear fdb`. This command should also be issued after a configuration is implemented, as the configuration must be in place before an entry is made in the MAC FDB. For permanent destination MAC-based grouping, re-apply the QoS profile to the static FDB entry, as documented. You can also save and reboot the switch.
- For physical and logical groupings of a source port or VLAN, re-apply the QoS profile to the source port or VLAN, as documented. You can also save and reboot the switch.

Bi-Directional Rate Shaping

Bi-directional rate shaping allows you to manage bandwidth on layer 2 and layer 3 traffic flowing both to and from the switch. You can configure up to eight ingress queues per VLAN and up to eight egress queues per physical port. By defining minimum and maximum bandwidth for each queue, you define committed information rates for each queue. You can define different rates for ingress and egress queues.

You can then provide traffic groupings (such as physical port, VLAN, .1P, DiffServ, IP address, or layer 4 flow) for the predefined QoS Profiles, thereby directing specific types of traffic to the desired queue.

Configuring Bi-Directional Rate Shaping

Each VLAN requires a loopback port; all traffic from rate-shaped ports is directed through the loopback port for that VLAN. To rate-shape ingress traffic, configure QoS normally on the loopback port for the VLAN. The maximum bandwidth and traffic grouping defined in the QoS profile for the loopback port defines the rate limit for ingress traffic on rate-shaped ports in that VLAN.

Use the following guidelines for bi-directional rate shaping:

- You must configure a loopback port before adding rate-shaped ports.
- A loopback port cannot be used by an external device.
- You must configure the loopback port with a unique loopback VLAN tag ID.
- Ingress traffic on a port that is configured to use the loopback port will be rate-shaped.
- Ingress traffic on a port that is not configured to use the loopback port will not be rate-shaped.
- Unicast traffic from a non-rate-shaped port to a rate-shaped port within the VLAN will not be rate-shaped.
- The aggregate forwarding bandwidth of all rate-shaped ports in a VLAN is determined by the setting of the queue parameters of the loopback port.
- For 10/100 Mbps ports, you can configure the loopback port as a 10 Mbps port to achieve lower bandwidth values.

To remove the rate-shaping parameters of the loopback port, configure the QoS profile without specifying the buffer or portlist parameters.

Bandwidth Settings

You apply bandwidth settings to QoS profiles as a percentage of bandwidth. QoS profile bandwidth settings are in turn applied to queues on physical ports. The impact of the bandwidth setting is determined by the port speed (10, 100, or 1000 Mbps).

Maximum Bandwidth Settings

The maximum bandwidth settings determine the port bandwidth available to each queue. Use Table 7-11 to determine the bandwidth associated with each bandwidth setting at different port speeds.

Table 7-11: Maximum Bandwidth Settings

Bandwidth Setting (%)	Bandwidth at 10 Mbps	Bandwidth at 100 Mbps	Bandwidth at 1000 Mbps
2	200 Kbps	2 Mbps	20 Mbps
3	310 Kbps	31 Mbps	30 Mbps
5	490 Kbps	4.9 Mbps	50 Mbps
7	690 Kbps	6.9 Mbps	69 Mbps
8	790 Kbps	7.9 Mbps	79 Mbps
10	960 Kbps	9.6 Mbps	96 Mbps
11	1.12 Mbps	11.2 Mbps	112 Mbps
15	1.5 Mbps	15 Mbps	150 Mbps
20	1.9 Mbps	19 Mbps	190 Mbps
25	2.5 Mbps	25 Mbps	250 Mbps
30	3.3 Mbps	33 Mbps	330 Mbps
35	3.5 Mbps	35 Mbps	350 Mbps
40	4..2 Mbps	42 Mbps	420 Mbps
50	5 Mbps	50 Mbps	500 Mbps
60	5.7 Mbps	57 Mbps	570 Mbps
65	6.5 Mbps	65 Mbps	650 Mbps
70	7.3 Mbps	73 Mbps	730 Mbps
80	7.9 Mbps	79 Mbps	790 Mbps
95	9.5 Mbps	95 Mbps	950 Mbps
100	10 Mbps	100 Mbps	1000 Mbps

If you choose a setting not listed in Table 7-11, the setting is rounded up to the next value.

Minimum Bandwidth Settings

The minimum bandwidth settings determine the port bandwidth reserved for each queue. Use Table 7-12 to determine the bandwidth associated with each setting.

Table 7-12: Minimum Bandwidth Settings

Bandwidth Setting (%)	Bandwidth at 10 Mbps	Bandwidth at 100 Mbps	Bandwidth at 1000 Mbps
4	420 Kbps	4.2 Mbps	42 Mbps
6	570 Kbps	5.7 Mbps	57 Mbps
8	750 Kbps	7.5 Mbps	75 Mbps
9	930 Kbps	9.3 Mbps	93 Mbps
10	1 Mbps	10 Mbps	100 Mbps
20	1.87 Mbps	18.7 Mbps	187 Mbps
25	2.63 Mbps	26.3 Mbps	263 Mbps
35	3.4 Mbps	34 Mbps	340 Mbps
50	4.9 Mbps	49 Mbps	490 Mbps
60	6.3 Mbps	63 Mbps	630 Mbps
80	7.9 Mbps	79 Mbps	790 Mbps
89	9.4 Mbps	94 Mbps	940 Mbps



Note: Keep the sum of the minimum bandwidth values for the applied QoS profiles less than 90%. If the sum exceeds 90%, a lower priority queue might be unable to transmit in a sustained over-subscription situation.

If you choose a setting not listed in Table 7-12, the setting is rounded up to the next value. If the actual bandwidth used is below the minimum bandwidth, the additional bandwidth is available for other queues on that physical port.

Bi-Directional Rate Shaping Limitations

Consider the following limitations when configuring bi-directional rate shaping:

- You must delete all rate-shaped ports before deleting the loopback port.
- If rate-shaped ports within a VLAN use different bandwidth parameters, set the priority of the QoS profiles on the loopback port and rate-shaped ports to `low`.
- Layer 2 rate-shaping only affects a single VLAN.

- On a BlackDiamond switch, the loopback port must be on the same I/O module as the rate-shaped ports.
- You must enable IP forwarding on the VLAN prior to adding the loopback port to a VLAN for layer 2 rate shaping.
- You cannot use tagged ports for rate shaping.
- You cannot use load-shared ports for rate-shaping.
- You cannot run VRRP on a VLAN that is configured for ingress rate shaping.

Dynamic Link Context System

The Dynamic Link Context System (DLCS) is a feature that snoops WINS NetBIOS packets and creates a mapping between a user name, the IP address or MAC address, and the switch/port. Based on the information in the packet, DLCS can detect when an end station boots up or a user logs in or out, and dynamically maps the end station name to the current IP address and switch/port. This information is available for use by ExtremeWare Enterprise Manager (EEM) version 2.1 or later or ExtremeWare EPICenter in setting policies that can be applied to users and can dynamically follow a user's location. DLCS provides you with valuable information on a user's location and associated network attributes. For DLCS to operate within ExtremeWare, the user or end station must allow for automatic DLCS updates.

This feature is intended for use in conjunction with the EPICenter 3.1 Policy System. Refer to the EPICenter 3.1 documentation for more information.

DLCS Guidelines

Follow these guidelines when using DLCS:

- Only one user is allowed on one workstation at a given time.
- A user can be logged into many workstations simultaneously.
- An IP-address can be learned on only one port in the network at a given time.
- Multiple IP-addresses can be learned on the same port.
- DLCS mapping is flushed when a user logs in or logs out, or when an end-station is shutdown.

DLCS Limitations

Consider the following limitations concerning data received from WINS snooping:

- DLCS does not work for the WINS server. This is because the WINS server does not send NETBIOS packets on the network (these packets are address to itself).
- When the IP address of a host is changed, and the host is not immediately rebooted, the old host-to-IP address mapping is never deleted. You must delete the mapping of the host-to-IP address through the EEM Policy Manager or ExtremeWare EPICenter Policy Manager.
- When the host is moved from one port to another port on a switch, the old entry does not age out unless the host is rebooted or a user login operation is performed after the host is moved.
- DLCS information is dynamic, therefore, if the switch is rebooted, the information is lost. This information is still stored in the policy-server. To delete the information from the policy system, you must explicitly delete configuration parameters from the EEM or ExtremeWare EPICenter Policy Applet user interface. As a workaround, you can delete the switch that was rebooted from the list of managed devices in the EEM or EPICenter Inventory Applet, and re-add the switch to the Inventory Manager.
- DLCS is not supported on hosts that have multiple NIC cards.
- IPQoS is not supported to a WINS server that is serving more than one VLAN. If you attempt to add a WINS server to serve more than one VLAN, and there are IPQoS rules defined for that server, the command to add the WINS server is rejected.



Access Policies

This chapter describes the following topics:

- Overview of Access Policies on page 8-1
- Using IP Access Lists on page 8-2
- Using Routing Access Policies on page 8-11
- Making Changes to a Routing Access Policy on page 8-23
- Removing a Routing Access Policy on page 8-24
- Using Route Maps on page 8-24
- Using Route Maps on page 8-24

Overview of Access Policies

Access policies are a generalized category of features that impact forwarding and route forwarding decisions. Access policies are used primarily for security and quality of service (QoS) purposes.

The three categories of access policies are:

- Access lists.
- Routing access policies.
- Route maps.

IP Access Lists

IP access lists consist of IP access rules and are used to perform packet filtering and forwarding decisions on incoming traffic. Each packet arriving on an ingress port is compared to the access list in sequential order and is either forwarded to a specified QoS profile or dropped. Using access lists has no impact on switch performance.

Access lists are typically applied to traffic that crosses layer 3 router boundaries, but it is possible to use access lists within a layer 2 VLAN. Products that use the “i” chipset are capable of performing this function with no additional configuration. Products that do not use the “i” chipset require the enabling of Intra-subnet QoS (ISQ), to perform this function. For more information on ISQ, refer to Chapter 7.

Routing Access Policies

Routing access policies are used to control the advertisement or recognition of routing protocols, such as RIP, OSPF, or BGP. Routing access policies can be used to ‘hide’ entire networks, or to trust only specific sources for routes or ranges of routes. The capabilities of routing access policies are specific to the type of routing protocol involved, but are sometimes more efficient and easier to implement than access lists.

Route Maps

Route maps are used to modify or filter routes redistributed between two routing domains. They are also used to modify or filter the routing information exchanged between the domains.

Using IP Access Lists

Each entry that makes up an IP access list contains a unique name. It can also contain an optional, unique precedence number. The rules of an IP access list consist of a combination of the following six components:

- IP source address and mask
- IP destination address and mask
- TCP or UDP source port range
- TCP or UDP destination port range
- Physical source port
- Precedence number (optional)

How IP Access Lists Work

When a packet arrives on an ingress port, the packet is compared with the access list rules to determine a match. When a match is found, the packet is processed. If the access list is of type deny, the packet is dropped. If the list is of type permit, the packet is forwarded. A permit access list can also apply a QoS profile to the packet.

Precedence Numbers

The precedence number is optional, and determines the order in which each rule is examined by the switch. Access list entries that contain a precedence number are evaluated from highest to lowest. Precedence numbers range from 1 to 25,600, with the *number 1 having the highest precedence*.

You can specify overlapping rules; however, if you are using precedence numbers, overlapping rules without precedence numbers are ignored. Therefore, the precedence numbers must be specified among all overlapping rules. If a new rule without a precedence number is entered, and this rule overlaps with already existing rules, the switch rejects the new rule and resolves the precedences among all remaining overlapping rules.

Specifying a Default Rule

To begin constructing an access list, you should specify a default rule. A *default rule* is a rule that contains wildcards for destination and source IP address, with no Layer 4 information. A default rule determines if the behavior of the access list is an “implicit deny” or “implicit accept.” If no access list entry is satisfied, the default rule is used to determine whether the packet is forwarded or dropped. If no default rule is specified, the default implicit behavior is to forward the packet.

The following example shows a default entry that is used to specify an explicit deny:

```
create access-list denyall ip dest 0.0.0.0/0 source 0.0.0.0/0 deny
ports any
```

Once the default behavior of the access list is established, you can create additional entries using precedence numbers.

The following access-list example performs packet filtering in the following sequence, as determined by the precedence number:

- Deny UDP port 32 and TCP port 23 traffic to the 10.2.XX network.
- All other TCP port 23 traffic destined for other 10.X.X.X networks is permitted using QoS profile Qp4.
- All remaining traffic to 10.2.0.0 uses QoS profile Qp3.

With no default rule specified, all remaining traffic is allowed using the default QoS profile.

```
create access-list deny102_32 udp dest 10.2.0.0/16 ip-port 32 source
any ip-port any deny ports any precedence 10
```

```
create access-list deny102_23 tcp dest 10.2.0.0/16 ip-port 23 source
any ip-port any deny ports any precedence 20
```

```
create access-list allow10_23 tcp dest 10.0.0.0/8 ip-port 23 source any
ip-port any permit qosprofile qp4 ports any precedence 30
```

```
create access-list allow102 ip dest 10.2.0.0/16 source 0.0.0.0/0 permit
qosprofile qp3 ports any precedence 40
```

The permit-established Keyword

The `permit-established` keyword is used to directionally control attempts to open a TCP session. The `permit-established` keyword denies all traffic that matches the TCP source/destination, and has the SYN=1 and ACK=0 flags set. Thus, TCP session initiation can be explicitly blocked using this keyword. Traffic from TCP sessions that are already established continue to be permitted.



Note: For an example of using the permit-established keyword, refer to “Using the Permit-Established Keyword” on page 8-6.

Adding and Deleting Access List Entries

Entries can be added and deleted to the access list. To add an entry, you must supply a unique name and, optionally, a unique precedence number. To modify an existing entry, you must delete the entry and retype it, or create a new entry with a new unique name.

To delete an access list entry, use the following command:

```
delete access-list <name>
```


Maximum Entries

A maximum of 255 entries with an assigned precedence can be used. In addition to the 255 entries, entries that do not use precedence can also be created, with the following restrictions:

- A source IP address must use wildcards or a completely specified mask.
- The layer 4 source and destination ports must use wildcards or be completely specified (no ranges).
- No physical source port can be specified.
- Access list rules that apply to all physical ports are implemented on all BlackDiamond I/O modules.

On a BlackDiamond 6808 switch, the maximum number of access list entries is 255 entries per I/O module. One way to economize on the number of entries on a BlackDiamond switch is to provide a physical ingress port as a component of an access list rule. In this case, the rule is implemented only on the I/O modules that contain the specified ports. By restricting rules to specific I/O modules, you can extend the number of access list rules to 1024 (NVRAM limit).

Access Lists for ICMP

Access lists for ICMP traffic processing are handled in a slightly different manner. An access list for ICMP is only effective for traffic routed by the switch. ICMP traffic can either be forwarded (routed) by the switch or discarded, but cannot contain options for assigning a QoS profile. Other configuration options for filtering ICMP include:

- IP source and destination address and mask.
- ICMP type code.
- Physical source port (optional).
- Numbered precedence (optional).

Verifying Access List Configurations

To verify access list settings, you can view the access list configuration and see real-time statistics on which access list entries are being accessed when processing traffic.

To view the access list configuration and statistics screen, use the following command:

```
show access-list {<name> | port <portlist>}
```

To initiate and refresh a running display of access list statistics, use the following command:

```
show access-list-monitor
```

IP Access List Examples

This section presents two IP access list examples:

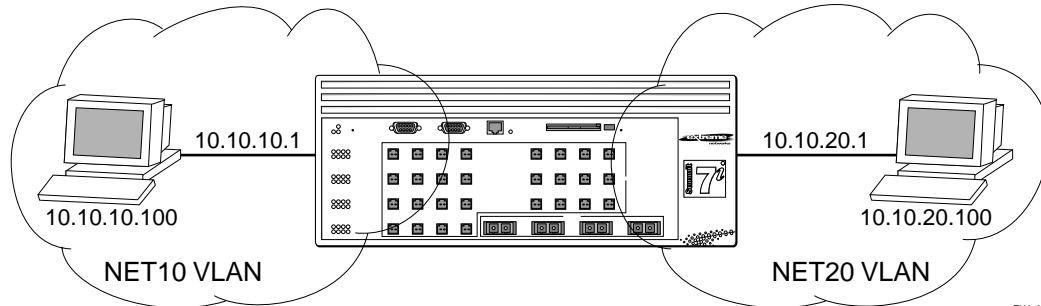
- Using the permit-establish keyword
- Filtering ICMP packets

Using the Permit-Established Keyword

This example uses an access list that permits TCP sessions (Telnet, FTP, and HTTP) to be established in one direction.

The Summit7i, shown in Figure 8-1, is configured as follows:

- Two VLANs, NET10 VLAN and NET20 VLAN, are defined.
- The IP addresses for NET10 VLAN is 10.10.10.1/24.
- The IP address for NET20 VLAN is 10.10.20.1/24.
- The workstations are configured using addresses 10.10.10.100 and 10.10.20.100.
- IP Forwarding is enabled.



EW_033

Figure 8-1: Permit-established access list example topology

The following sections describe the steps used to configure the example.

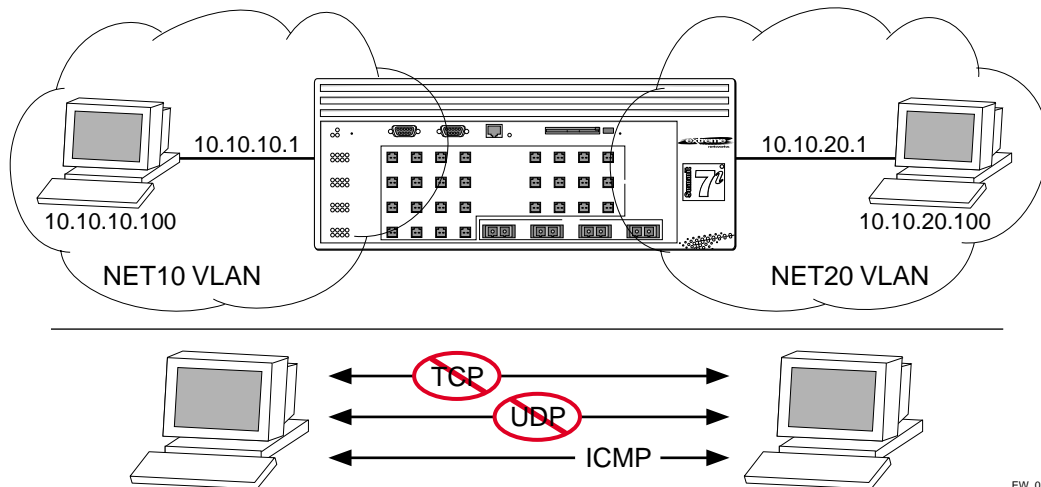
Step 1 – Deny IP Traffic.

First, create an access-list that blocks all IP-related traffic. This includes any TCP- and UDP-based traffic. Although ICMP is used in conjunction with IP, it is technically not an IP data packet. Thus, ICMP data traffic, such as ping traffic, is not affected.

The following command creates the access list:

```
create access-list denyall ip destination any source any deny ports any
```

Figure 8-2 illustrates the outcome of the access list.



EW_034

Figure 8-2: Access list denies all TCP and UDP traffic

Step 2 – Allow TCP traffic.

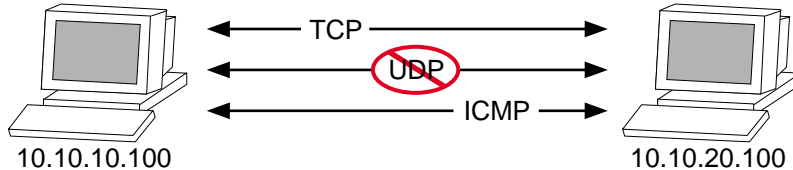
The next set of access list commands permits TCP-based traffic to flow. Because each session is bi-directional, an access list must be defined for each direction of the traffic flow. UDP traffic is still blocked.

The following commands create the access list:

```
create access-list tcp1 tcp destination 10.10.20.100/32 ip any source
10.10.10.100/32 ip any permit qpl ports any precedence 20
```

```
create access-list tcp2 tcp destination 10.10.10.100/32 ip any source
10.10.20.100/32 ip any permit qpl ports any precedence 21
```

Figure 8-3 illustrates the outcome of this access list.

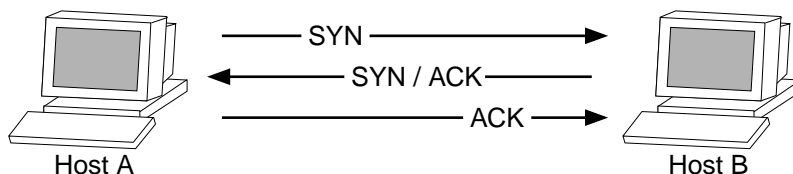


EW_035

Figure 8-3: Access list allows TCP traffic

Step 3 - Permit-Established Access List.

When a TCP session begins, there is a three-way handshake that includes a sequence of a SYN, SYN/ACK, and ACK packets. Figure 8-4 shows an illustration of the handshake that occurs when host A initiates a TCP session to host B. After this sequence, actual data can be passed.



EW_036

Figure 8-4: Host A initiates a TCP session to host B

An access list that uses the permit-established keyword filters the SYN packet in one direction.

Use the permit-established keyword to allow only host A to be able to establish a TCP session to host B and to prevent any TCP sessions from being initiated by host B, as illustrated in Figure 8-4. The syntax for this access list is as follows:

```
create access-list <name> tcp destination HostA ip-port 23 source HostB
ip-port any permit-established ports any pre 8
```



Note: This step may not be intuitive. Pay attention to the destination and source address, and the desired affect.

The exact command line entry for this example is as follows:

```
create access-list telnet-allow tcp destination 10.10.10.100/32 ip-port
23 source any ip-port any permit-established ports any pre 8
```



Note: This rule has a higher precedence than the rule “tcp2.”

Figure 8-5 shows the final outcome of this access list.

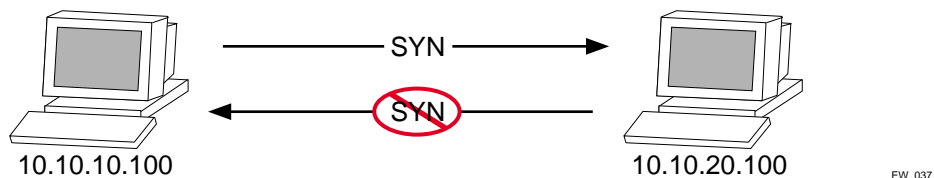


Figure 8-5: Permit-established access list filters out SYN packet to destination

Example 2: Filter ICMP Packets

This example creates an access list that filters out ping (ICMP echo) packets. ICMP echo packets are defined as type 8 code 0.

The command line syntax to create this access list is as follows:

```
create access-list denying icmp destination any source any type 8 code
0 deny ports any
```

The output for this access list is shown in Figure 8-6.

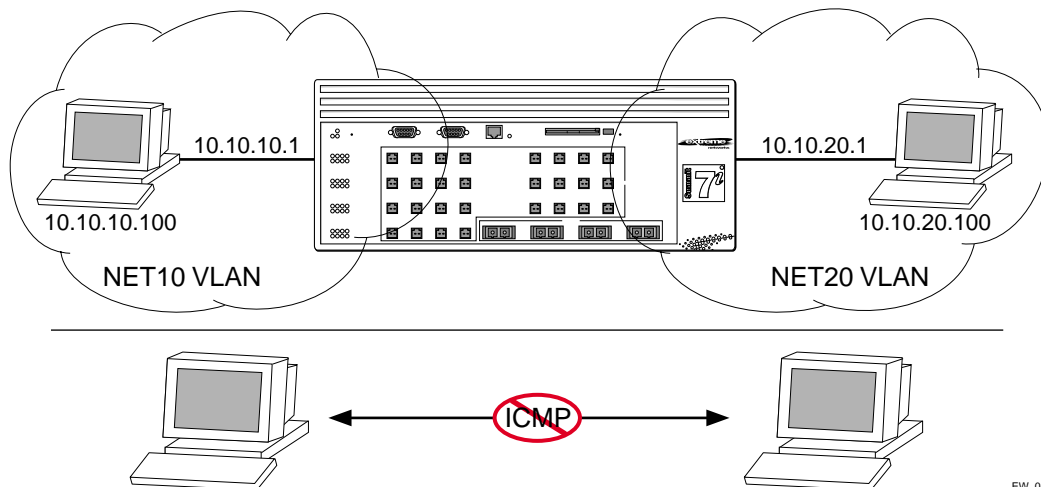


Figure 8-6: ICMP packets are filtered out

Using Routing Access Policies

To use routing access policies, you must perform the following steps:

- 1 Create an access profile.
- 2 Configure the access profile to be of type *permit*, *deny*, or *none*.
- 3 Add entries to the access profile. Entries can be one of the following types:
 - IP addresses and subnet masks
 - IPX node, IPX RIP, and IPX SAP
 - Autonomous system path expressions (as-paths) (BGP only)
 - BGP communities (BGP only)
- 4 Apply the access profile.

Creating an Access Profile

The first thing to do when using routing access policies is to create an *access profile*. An access profile has a unique name and contains one of the following entry types:

- A list of IP addresses and associated subnet masks
- A list of IPX NetIDs
- A list of IPX node addresses
- A list of IPX SAP advertisements
- One or more autonomous system path expressions (BGP only)
- One or more BGP community numbers (BGP only)

You must give the access profile a unique name (in the same manner as naming a VLAN, protocol filter, or Spanning Tree Domain). To create an access profile, use the following command:

```
create access-profile <access_profile> type [ipaddress | ipx-node |  
ipx-net | ipx-sap | as-path | bgp-community]
```

Configuring an Access Profile Mode

After the access profile is created, you must configure the access profile mode. The access profile mode determines whether the items in the list are to be permitted access or denied access.

Three modes are available:

- **Permit** — The permit access profile mode permits the operation, as long as it matches any entry in the access profile. If the operation does not match any entries in the list, the operation is denied.
- **Deny** — The deny access profile mode denies the operation, as long as it matches any entry in the access profile. If it does not match all specified entries in the list, the operation is permitted.
- **None** — Using the none mode, the access profile can contain a combination of permit and deny entries. Each entry must have a permit or deny attribute. The operation is compared with each entry in the list. Once a match is found, the operation is either permitted or denied, depending on the configuration of the matched entry. If no match is found, the operation is implicitly denied.

To configure the access profile mode, use the following command:

```
config access-profile <access_profile> mode [permit | deny | none]
```

Adding an Access Profile Entry

Next, configure the access profile, using the following command:

```
config access-profile <access_profile> add {<seq_number>} {permit | deny} [ipaddress <ipaddress> <mask> {exact} | ipx-node <net_id> <ipx_netid_mask> <ipx_nodeid> ipx-net <ipx_netid> <ipx_netid_mask> | ipx-sap <ipx_sap_type> <ipx_name>| as-path <path-expression> | bgp-community [internet | no-export | no-advertise | no-export-subconfed | <as_no:number> | number <community>]]
```

The following sections describe the `config access-profile add` command.

Specifying Subnet Masks

The subnet mask specified in the access profile command is interpreted as a *reverse mask*. A reverse mask indicates the bits that are significant in the IP address. In other words, a reverse mask specifies the part of the address that must match the IP address to which the profile is applied.

If you configure an IP address that is an exact match that is specifically denied or permitted, use a mask of /32 (for example, 141.251.24.28/32). If the IP address represents all addresses in a subnet address that you want to deny or permit, then configure the mask to cover only the subnet portion (for example, 141.251.10.0/24). The keyword `exact` can be used when you wish to match only against the subnet address, and ignore all addresses within the subnet.

If you are using off-byte boundary subnet masking, the same logic applies, but the configuration is more tricky. For example, the address 141.251.24.128/27 represents any host from subnet 141.251.24.128.

Sequence Numbering

You can specify the sequence number for each access profile entry. If you do not specify a sequence number, entries are sequenced in the order they are added. Each entry is assigned a value of 5 more than the sequence number of the last entry.

Permit and Deny Entries

If you have configured the access profile mode to be `none`, you must specify each entry type as either 'permit' or 'deny'. If you do not specify the entry type, it is added as a permit entry. If you have configured the access profile mode to be `permit` or `deny`, it is not necessary to specify a type for each entry.

IPX Routing Access Policies

IPX routing access policies consist of access rules and are used to perform packet filtering and forwarding decisions on incoming traffic. Each IPX RIP or SAP packet arriving on an ingress port is compared to each access profile rule in sequential order and is either forwarded or dropped.

Autonomous System Expressions

The `AS-path` keyword uses a regular expression string to match against the AS path. Regular expression notation can include any of the characters listed in Table 8-1.

Table 8-1: Regular Expression Notation

Character	Definition
N	As number
$N_1 - N_2$	Range of AS numbers, where N_1 and N_2 are AS numbers and $N_1 < N_2$
$[N_x \dots N_y]$	Group of AS numbers, where N_x and N_y are AS numbers or a range of AS numbers
$[\^N_x \dots N_y]$	Any AS numbers other than the ones in the group
.	Matches any number
^	Matches the beginning of the AS path
\$	Matches the end of the AS path
–	Matches the beginning or end, or a space
-	Separates the beginning and end of a range of numbers
*	Matches 0 or more instances
+	Matches 1 or more instances
?	Matches 0 or 1 instance
{	Start of AS SET segment in the AS path

Table 8-1: Regular Expression Notation (continued)

Character	Definition
}	End of AS SET segment in the AS path
(Start of a confederation segment in the AS path
)	End of a confederation segment in the AS path

Autonomous System Expression Example

The following example uses combinations of the autonomous system expressions to create a complicated access profile:

```
create access-profile AS1 type as-path
config access-profile AS1 mode none
```

These commands create the access profile.

```
config access-profile AS1 add 5 permit as-path "^65535$"
```

This command configures the access profile to permit AS paths that contain only (begin and end with) AS number 65535.

```
config access-profile AS1 add 10 permit as-path "^65535 14490$"
```

This command configures the access profile to permit AS paths beginning with AS number 65535, ending with AS number 14490, and containing no other AS paths.

```
config access-profile AS1 add 15 permit as-path "^1 2-8 [11 13 15]$"
```

This command configures the access profile to permit AS paths beginning with AS number 1, followed by any AS number from 2 - 8, and ending with either AS number 11, 13, or 15.

```
config access-profile AS1 add 20 deny as-path "111 [2-8]"
```

This command configures the access profile to deny AS paths beginning with AS number 111 and ending with any AS number from 2 - 8.

```
config access-profile AS1 add 25 permit as-path "111 .?"
```

This command configures the access profile to permit AS paths beginning with AS number 111 and ending with any additional AS number, or beginning and ending with AS number 111.

Deleting an Access Profile Entry

To delete an access profile entry, use the following command:

```
config access-profile <access_profile> delete <seq_number>
```

Applying Access Profiles

Once the access profile is defined, apply it to one or more routing protocols or VLANs. When an access profile is applied to a protocol function (for example, the export of RIP routes) or a VLAN, this forms an access policy. A profile can be used by multiple routing protocol functions or VLANs, but a protocol function or VLAN can use only one access profile.

Routing Access Policies for RIP

If you are using the RIP protocol, the switch can be configured to use an access profile to determine:

- **Trusted Neighbor** — Use an access profile to determine trusted RIP router neighbors for the VLAN on the switch running RIP. To configure a trusted neighbor policy, use the following command:

```
config rip vlan [<name> | all] trusted-gateway [<access_profile> | none]
```

- **Import Filter** — Use an access profile to determine which RIP routes are accepted as valid routes. This policy can be combined with the trusted neighbor policy to accept selected routes only from a set of trusted neighbors. To configure an import filter policy, use the following command:

```
config rip vlan [<name> | all] import-filter [<access_profile> | none]
```

- **Export Filter** — Use an access profile to determine which RIP routes are advertised into a particular VLAN, using the following command:

```
config rip vlan [<name> | all] export-filter [<access_profile> | none]
```

Examples

In the example shown in Figure 8-7, a switch is configured with two VLANs, *Engsvrs* and *Backbone*. The RIP protocol is used to communicate with other routers on the network. The administrator wants to allow all internal access to the VLANs on the switch, but no access to the router that connects to the Internet. The remote router that connects to the Internet has a local interface connected to the corporate backbone. The IP address of the local interface connected to the corporate backbone is 10.0.0.10/24.

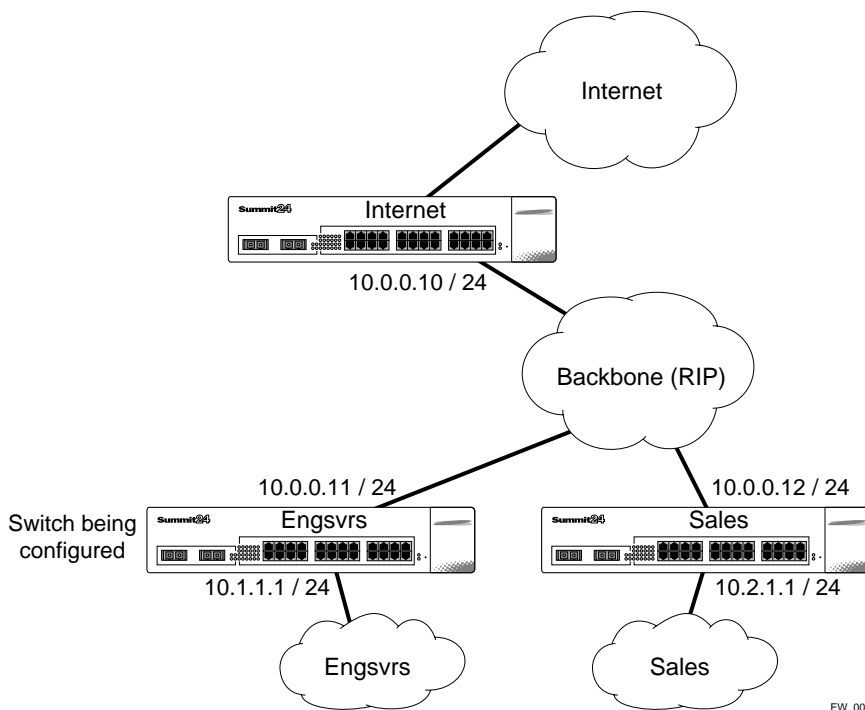


Figure 8-7: RIP access policy example

Assuming the backbone VLAN interconnects all the routers in the company (and, therefore, the Internet router does not have the best routes for other local subnets), the commands to build the access policy for the switch would be:

```
create access-profile nointernet ipaddress
config access-profile nointernet mode deny
config access-profile nointernet add 10.0.0.10/32
config rip vlan backbone trusted-gateway nointernet
```

In addition, if the administrator wants to restrict any user belonging to the VLAN *Engsvrs* from reaching the VLAN *Sales* (IP address 10.2.1.0/24), the additional access policy commands to build the access policy would be:

```
create access-profile nosales ipaddress
config access-profile nosales mode deny
config access-profile nosales add 10.2.1.0/24
config rip vlan backbone import-filter nosales
```

This configuration results in the switch having no route back to the VLAN *Sales*.

Routing Access Policies for IPX

If you are using the IPX protocol, the switch can be configured to use an access profile to determine:

- **Import Filter** — Use an access profile to determine which IPX/RIP or IPX/SAP routes are accepted as valid routes. To configure an import filter policy, use the following command:

```
config ipxrip vlan [<vlan name> | all] import-filter
[<access_profile> | none]
config ipxsap vlan [<vlan name> | all] import-filter
[<access_profile> | none]
```

- **Export Filter** — Use an access profile to determine which IPX/RIP and IPX/SAP routes are advertised into a particular VLAN, using the following command:

```
config ipxrip vlan [<vlan name> | all] export-filter
[<access_profile> | none]
config ipxsap vlan [<vlan name> | all] export-filter
[<access_profile> | none]
```

Routing Access Policies for OSPF

Because OSPF is a link-state protocol, the access policies associated with OSPF are different in nature than those associated with RIP. Access policies for OSPF are intended to extend the existing filtering and security capabilities of OSPF (for example, link authentication and the use of IP address ranges). If you are using the OSPF protocol, the switch can be configured to use an access profile to determine any of the following:

- **Inter-area Filter** — For switches configured to support multiple OSPF areas (an ABR function), an access profile can be applied to an OSPF area that filters a set of OSPF

inter-area routes from being sourced from any other areas. To configure an inter-area filter policy, use the following command:

```
config ospf area <area_id> interarea-filter [<access_profile> | none]
```

- **External Filter** — For switches configured to support multiple OSPF areas (an ABR function), an access profile can be applied to an OSPF area that filters a set of OSPF external routes from being advertised into that area. To configure an external filter policy, use the following command:

```
config ospf area <area_id> external-filter [<access_profile> | none]
```



Note: If any of the external routes specified in the filter have already been advertised, those routes will remain until the associated LSAs in that area time-out.

- **ASBR Filter** — For switches configured to support RIP and static route re-distribution into OSPF, an access profile can be used to limit the routes that are advertised into OSPF for the switch as a whole. To configure an ASBR filter policy, use the following command:

```
config ospf asbr-filter [<access_profile> | none]
```

- **Direct Filter** — For switches configured to support direct route re-distribution into OSPF, an access profile can be used to limit the routes that are advertised into OSPF for the switch as a whole. To configure a direct filter policy, use the following command:

```
config ospf direct-filter [<access_profile> | none]
```

Example

Figure 8-8 illustrates an OSPF network that is similar to the network used previously in the RIP example. In this example, access to the Internet is accomplished by using the ASBR function on the switch labeled Internet. As a result, all routes to the Internet will be done through external routes. Suppose the network administrator wishes to only allow access to certain internet addresses falling within the range 192.1.1.0/24 to the internal backbone.

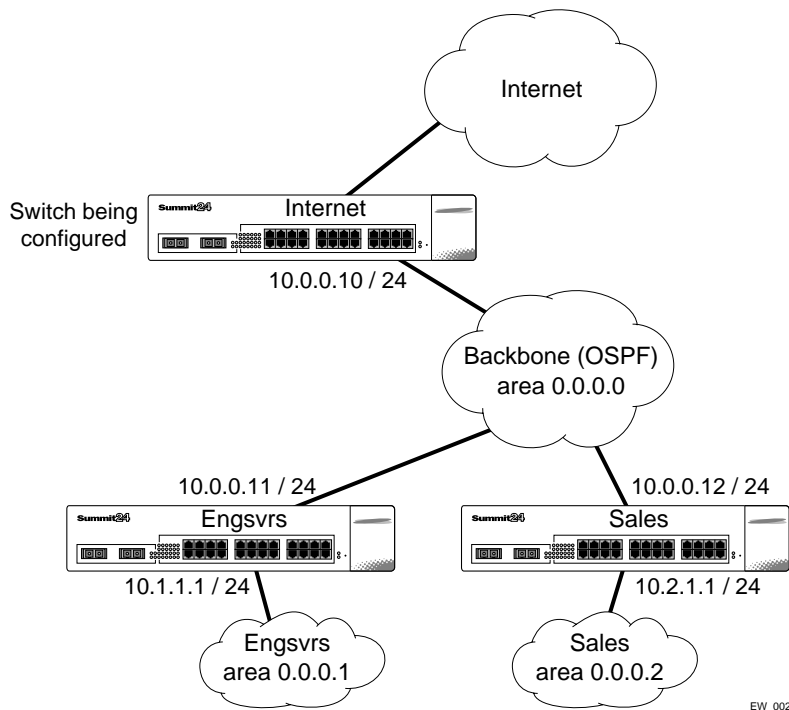


Figure 8-8: OSPF access policy example

To configure the switch labeled Internet, the commands would be as follows:

```
create access-profile okinternet ipaddress
config access-profile okinternet mode permit
config access-profile okinternet add 192.1.1.0/24
config ospf asbr-filter okinternet
```

Routing Access Policies for DVMRP

The access policy capabilities for DVMRP are very similar to those for RIP. If you are using the DVMRP protocol is used for routing IP multicast traffic, you can configure the switch to use an access profile to determine:

- **Trusted Neighbor** — Use an access profile to determine trusted DVMRP router neighbors for the VLAN on the switch running DVMRP. To configure a trusted neighbor policy, use the following command:


```
config dvmrp vlan [<name> | all] trusted-gateway [<access_profile> |
none]
```

- **Import Filter** — Use an access profile to determine which DVMRP routes are accepted as valid routes. To configure an import filter policy, use the following command:

```
config dvmrp vlan [<name> | all] import-filter [<access_profile> |
none]
```

- **Export-Filter** — Use an access profile to determine which DVMRP routes are advertised into a particular VLAN, using the following command:

```
config dvmrp vlan [<name> | all] export-filter [<access_profile> |
none]
```

Example

In this example, the network used in the previous RIP example is configured to run DVMRP. The network administrator wants to disallow Internet access for multicast traffic to users on the VLAN *Engsvrs*. This is accomplished by preventing the learning of routes that originate from the switch labeled Internet by way of DVMRP on the switch labeled *Engsvrs*.

To configure the switch labeled *Engsvrs*, use the following commands:

```
create access-profile nointernet ipaddress
config access-profile nointernet mode deny
config access-profile nointernet add 10.0.0.10/32
config dvmrp vlan backbone trusted-gateway nointernet
```

In addition, suppose the administrator wants to preclude users on the VLAN *Engsvrs* from seeing any multicast streams that are generated by the VLAN *Sales* across the backbone. The additional configuration of the switch labeled *Engsvrs* is as follows:

```
create access-profile nosales ipaddress
config access-profile nosales mode deny
config access-profile nosales add 10.2.1.0/24
config dvmrp vlan backbone import-filter nosales
```

Routing Access Policies for PIM

Because PIM leverages the unicast routing capability that is already present in the switch, the access policy capabilities are, by nature, different. If you are using the PIM protocol for routing IP multicast traffic, you can configure the switch to use an access profile to determine:

- **Trusted Neighbor** — Use an access profile to determine trusted PIM router neighbors for the VLAN on the switch running PIM. To configure a trusted neighbor policy, use the following command:

```
config pim vlan [<name> | all] trusted-gateway [<access_profile> | none]
```

Example

Using PIM, the unicast access policies can be used to restrict multicast traffic. In this example, a network similar to the example used in the previous RIP example is also running PIM. The network administrator wants to disallow Internet access for multicast traffic to users on the VLAN *Engsvrs*. This is accomplished by preventing the learning of routes that originate from the switch labeled Internet by way of PIM on the switch labeled *Engsvrs*.

To configure the switch labeled Engsvrs, the commands would be as follows:

```
create access-profile nointernet ipaddress
config access-profile nointernet mode deny
config access-profile nointernet add 10.0.0.10/32
config pim vlan backbone trusted-gateway nointernet
```

Routing Access Policies for BGP

If the BGP protocol is being used, the switch can be configured to use an access profile to determine:

- **NLRI filter** — Use an access profile to determine the NLRI information that must be exchanged with a neighbor. To configure an NLRI filter policy, use the following command:

```
config bgp neighbor [<ipaddress> | all] nlri-filter [in | out]
[<access_profile> | none]
```

The NLRI filter access policy can be applied to the ingress or egress updates, using the `in` and `out` keywords, respectively.

- **Autonomous system path filter** — Use an access profile to determine which NLRI information must be exchanged with a neighbor based on the AS path information present in the path attributes of the NLRI. To configure an autonomous system path filter policy, use the following command:

```
config bgp neighbor [<ipaddress> | all] as-path-filter [in | out]
[<access_profile> | none]
```

The autonomous system path filter can be applied to the ingress or egress updates, using the `in` and `out` keywords, respectively.

Making Changes to a Routing Access Policy

You can change the routing access policy by changing the associated access profile. However, the propagation of the change depends on the protocol and policy involved. Propagation of changes applied to RIP, DVMRP, and PIM access policies depend on the respective protocol timers to age-out entries.

In BGP, the change to the policy is immediately effective on the routing information exchanged after the policy changes. The changes can be applied on the routing information that had been exchanged before the policy changes by issuing a soft reset

on the ingress or egress side, depending on the change. For soft resets to be applied on the ingress side, the changes must have been previously enabled on the neighbor.



Note: Changes to profiles applied to OSPF typically require rebooting the switch, or disabling and re-enabling OSPF on the switch.

Removing a Routing Access Policy

To remove a routing access policy, you must remove the access profile from the routing protocol or VLAN. All the commands that apply an access profile to form an access policy also have the option of choosing `none` as the access profile. Using the `none` option removes any access profile of that particular type from the protocol or VLAN, and, therefore, removes the access policy.

Using Route Maps

Route maps are a mechanism that can be used to conditionally control the redistribution of routes between two routing domains, and to modify the routing information that is redistributed.

Route maps are used in conjunction with the match and set operations. A match operation specifies a criteria that must be matched. A set operation specifies a change that is made to the route when the match operation is successful.

To create a route map, follow these steps:

- 1 Create a route map.
- 2 Add entries to the route map.
- 3 Add statements to the route map entries.

Creating a Route Map

To create a route map, use the following command:

```
create route-map <route-map>
```

Add Entries to the Route Map

To add entries to the route map, use the following command:

```
config route-map <route-map> add <sequence number> [permit | deny]
{match-one | match-all}
```

where the following is true:

- The `sequence number` uniquely identifies the entry, and determines the position of the entry in the route map. Route maps are evaluated sequentially.
- The `permit` keyword permits the route; the `deny` keyword denies the route and is applied only if the entry is successful.
- The `match-one` keyword is a logical “or”. The route map is successful as long as at least one of the matching statements is true.
- The `match-all` keyword is a logical “and”. The route map is successful when all match statements are true. This is the default setting.

Add Statements to the Route Map Entries

To add statements to the route map entries, use one of the following three commands:

```
config route-map <route-map> <sequence number> add match [nlri-list
<access-profile> | as-path [access-profile <access-profile> | <as no>]
| community [access-profile <access-profile> | <as no> : <number> |
number <community> | no-advertise | no-export | no-export-subconfed] |
next-hop <ip address> | med <number> | origin [igp | egp | incomplete]
| tag <number> | origin [igp | egp | incomplete]]
```

```
config route-map <route-map> <sequence number> add set [as-path <as no>
| community [[access-profile <access-profile> | <as no> : <number> |
number <community> | no-advertise | no-export | no-export-subconfed] |
remove | [add | delete] [access-profile <access-profile> | <as no> :
<number> | number <community> | no-advertise | no-export |
no-export-subconfed]] | next-hop <ip address> | med [internal |
<number> | remove | [add | delete] <number> ] | local-preference
<number> | weight <number> | origin [igp | egp | incomplete] | tag
<number> | accounting index <number> value <number> | cost <number> |
cost-type [ase-type-1 | ase-type-2 ]]
```

```
config route-map <route-map> <sequence number> add goto <route-map>
```

where the following is true:

- The `route-map` is the name of the route map.
- The `sequence number` identifies the entry in the route map to which this statement is being added.
- The `match`, `set`, and `goto` keywords specify the operations to be performed. Within an entry, the statements are sequenced in the order of their operation. The `match` statements are first, followed by `set`, and then `goto`.
- The `nlri-list`, `as-path`, `community`, `next-hop`, `med`, `origin`, and `weight` keywords specify the type of values that must be applied using the specified operation against the corresponding attributes as described in Table 8-2 and Table 8-3.

Table 8-2: Match Operation Keywords

Keyword	Description
<code>nlri-list <access_profile></code>	Matches the NLRI against the specified access profile.
<code>as-path [<access_profile> <as-no>]</code>	Matches the AS path in the path attributes against the specified access profile or AS number.
<code>community [access-profile <access-profile> <as no> : <number> number <community> no-advertise no-export no-export-subconfed]</code>	Matches the communities in the path attribute against the specified BGP community access profile or the community number.
<code>next-hop <ipaddress></code>	Matches the next hop in the path attribute against the specified IP address.
<code>med <number></code>	Matches the MED in the path attribute against the specified MED number.
<code>origin [igp egp incomplete]</code>	Matches the origin in the path attribute against the specified origin.
<code>tag <number></code>	Matches the tag associated with the redistributed OSPF route.

Table 8-3: Set Operation Keywords

Keyword	Definition
<code>as-path <as no></code>	Prepends the specified AS number to the AS path in the path attribute.

Table 8-3: Set Operation Keywords (continued)

Keyword	Definition
community [[access-profile <access-profile> <as no> : <number> number <community> no-advertise no-export no-export-subconfed] remove [add delete] [access-profile <access-profile> <as no> : <number> number <community> no-advertise no-export no-export-subconfed]]	Adds the specified community to the existing community in the path attribute.
next-hop <ipaddress>	Sets the next hop in the path attribute to the specified IP address.
med [internal <number> remove [add delete] <number>]	<p>Modifies the MED in the path attribute as specified:</p> <ul style="list-style-type: none"> ■ <i>internal</i> — When used in the BGP neighbor output route map, the MED attribute is set to a value equal to the metric to reach the nexthop. ■ <i><number></i> — Sets the MED attribute to the specified value. ■ <i>remove</i> — Removes the MED attribute, if present. ■ <i>[add delete] <number></i> — Adds or deletes the specified value to or from the MED that is received. The final result is bound by 0 and 2147483647.
local-preference <number>	Sets the local preference in the path attribute to the specified local preference number.
weight <number>	Sets the weight associated with the NLRI to the specified number.
origin [igp egp incomplete]	Sets the origin in the path attributes to the specified origin.
tag <number>	Sets the tag in the route to the specified number.
cost <number>	Sets the cost of the route to the specified number
cost-type <number>	Sets the type of the cost associated with the route.
accounting index [ase-type-1 ase-type-2]	Sets the specified accounting index to the specified number.

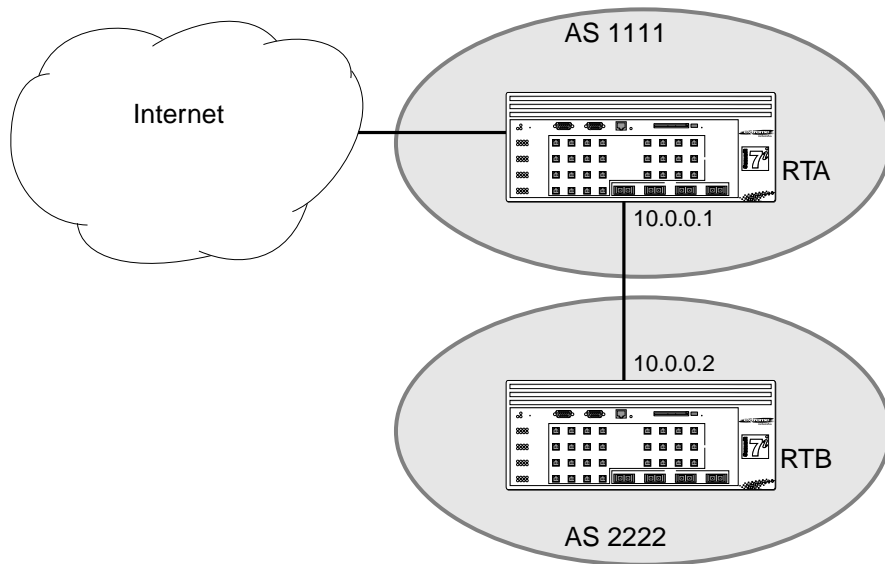
Route Map Operation

The entries in the route map are processed in the ascending order of the sequence number. Within the entry, the match statements are processed first. When the match operation is successful, the set and goto statements within the entry are processed, and the action associated with the entry is either applied, or else the next entry is processed. If the end of the route map is reached, it is implicitly denied.

When there are multiple match statement, the primitive match-one or match-all in the entry determines how many matches are required for success. When there are no match statements in an entry, the entry is considered a successful match.

Route Map Example

Figure 8-9 shows a topology in which route maps are used to filter or modify routing information that is exchanged between the neighbors RTA and RTB using BGP.



EW_048

Figure 8-9: Route maps

The following points apply to this example:

- RTA is a member of in AS 1111 and peers with a router in the Internet to receive the entire Internet routing table.
- RTB is a member of AS 2222, and has an EBGP connection with RTA through which it receives the Internet routing table.
- AS 1111 is acting as a transit AS for all traffic between AS 2222 and the Internet. If the administrator of AS 1111 wants to filter out route information about network 221.1.1.0 / 24 and its subnets from being passed on to AS 2222, s/he can configure a route-map on the egress side of RTA's EBGP connection with RTB and filter out the routes.

To configure RTA, use the following commands:

```
create access-profile iplist type ipaddress
config iplist add ipaddress 221.1.1.0 / 24

create route-map bgp-out
config bgp-out add 10 deny
config bgp-out 10 add match nlri-list iplist
config bgp-out add 20 permit

config bgp neighbor 10.0.0.2 route-map-filter out bgp-out
config bgp neighbor 10.0.0.2 soft-reset out
```

If you wish to modify the routing information originated from AS 300 to include a MED value of 200, the sequence of commands would be:

```
create access-profile aslist type as-path
config aslist add as-path "^300"

config bgp-out add 15 permit
config bgp-out 15 add match as-path access-profile aslist
config bgp-out 15 add set med 200

config bgp neighbor 10.0.0.2 soft-reset out
```

Changes to Route Maps

Changes to the route maps used to modify or filter NLRI information exchanged with neighbors is immediately effective on the routing information exchanged after the policy changes. The changes can be applied on the NLRI information that had been exchanged before the policy changes by issuing a soft reset on the ingress or egress side,

depending on the changes. For soft resets to be applied on the ingress side, the changes must be previously enabled on the neighbor.

Changes to the route maps associated with network aggregation or redistribution commands becomes effective after a maximum interval of 30 seconds. You can immediately apply them by using the soft reconfiguration command.

Route Maps in BGP

Route maps are used in BGP to modify/filter NLRI information exchanged with neighbors. They are also used NLRI information that originates by way of network command, aggregation, or redistribution.

9

Network Address Translation (NAT)

This chapter covers the following topics:

- Overview on page 9-1
- Internet IP Addressing on page 9-3
- Configuring VLANs for NAT on page 9-3
- Configuring NAT on page 9-5
- Displaying NAT Settings on page 9-8
- Disabling NAT on page 9-9

Overview

NAT is a feature that allows one set of IP addresses, typically private IP addresses, to be converted to another set of IP addresses, typically public Internet IP addresses. This conversion is done transparently by having a NAT device (any Extreme Networks switch using the “i” chipset) rewrite the source IP address and Layer 4 port of the packets.

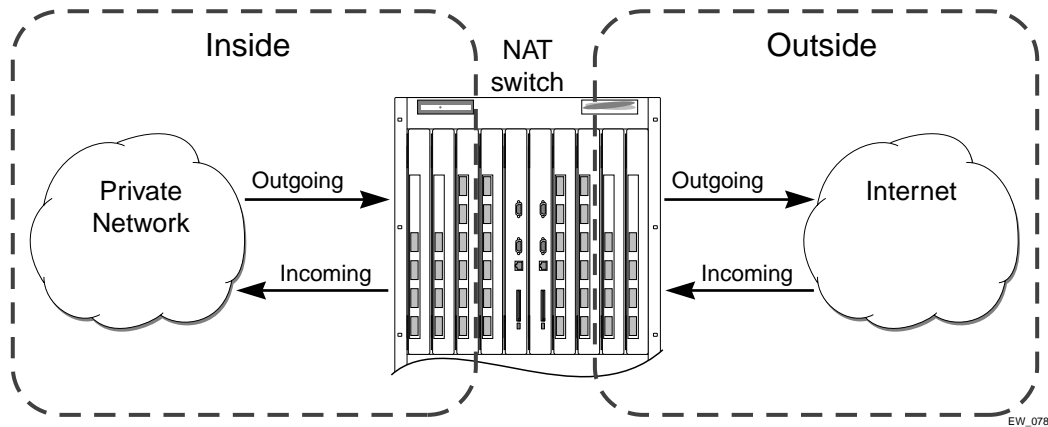


Figure 9-1: NAT Overview

You can configure NAT to conserve IP address space by mapping a large number of inside (private) addresses to a much smaller number of outside (public) addresses.

In implementing NAT, you must configure at least two separate VLANs involved. One VLAN is configured as inside, and corresponds to the private IP addresses you would like to translate into other IP addresses. The other type of VLAN is configured as outside, which corresponds to the public (probably Internet) IP addresses you want the inside addresses translated to. The mappings between inside and outside IP addresses are done via rules that specify the IP subnets involved and the algorithms used to translate the addresses.



The NAT modes in ExtremeWare 6.2 support translating traffic initiating only from inside addresses.

NAT rules are associated with a single outside VLAN. Multiple rules per outside VLAN are allowed. The rules take effect in the order they are displayed using the `show` command. Any number of inside VLANs can use a single outside VLAN, assuming that you have created proper rules. Similarly, a single inside VLAN can use any number of different outside VLANs, assuming that the rules and routing are set up properly.

Both TCP and UDP have Layer 4 port numbers ranging from 1 to 65535. These Layer 4 ports, in combination with the IP addresses, form a unique identifier which allows hosts (as well as the NAT switch) to distinguish between separate conversations. NAT operates by replacing the inside IP packet's source IP and Layer 4 port with an outside IP and Layer 4 port. The NAT switch maintains a connection table to map the return packets on the outside VLAN back into their corresponding inside sessions.

Internet IP Addressing

When implementing NAT in an Internet environment, it is strongly recommended that you use one of the reserved private IP address ranges for your inside IP addresses. These ranges have been reserved specifically for networks not directly attached to the Internet. Using IP addresses within these ranges prevents addressing conflicts with public Internet sites to which you want to connect. The ranges are as follows:

- 10.0.0.0/8—Reserved Class A private address space
- 172.16.0.0/12—Reserved Class B private address space
- 192.168.0.0/16—Reserved Class C private address space

Configuring VLANs for NAT

You must configure each VLAN participating in NAT as either an inside or outside VLAN. To configure a VLAN as an inside or outside VLAN, use the following command:

```
config nat vlan <name> [inside | outside | none]
```

When a VLAN is configured to be *inside*, traffic from that VLAN is translated only if it has a matching NAT rule. Any unmatched traffic will be routed normally and not be translated. Because all traffic runs through the central processing unit (CPU), it cannot run at line-rate.

When a VLAN is configured to be *outside*, it routes all traffic. Because all traffic runs through the CPU, it cannot run at line-rate. Normally, outside traffic will be able to initiate connections to the internal private IP addresses. If you want to prevent this, you can create IP and ICMP access-lists on the outside VLAN ports to deny traffic destined for the inside IP addresses. There is a NAT performance penalty when you do this.

When a VLAN is configured to be `none`, all NAT functions are disabled and the VLAN operates normally.

Below is a set of example ACL rules to deny outside traffic. These examples assume the inside network is `192.168.1.0/24` and the outside VLAN is on port 1.

```
create access-list deny_ip ip destination 192.168.1.0/24 source any
deny ports 1
create access-list deny_icmp icmp destination 192.168.1.0/24 source any
type any code any deny ports 1
```

NAT Modes

There are 4 different modes used to determine how the outside IP addresses and Layer 4 ports are assigned.

- Static mapping
- Dynamic mapping
- Port-mapping
- Auto-constraining

Static Mapping

When static mapping is used, each inside IP address uses a single outside IP address. The Layer 4 ports are not changed, only the IP address is rewritten. Because this mode requires a 1:1 mapping of internal to external addresses, it does not make efficient use of the external address space. However, it is useful when you have a small number of hosts that need to have their IP addresses rewritten without conflicting with other hosts. Because this mode does not rely on Layer 4 ports, ICMP traffic is translated and allowed to pass.

Dynamic Mapping

Dynamic mapping is similar to static mapping in that the Layer 4 ports are not rewritten during translation. Dynamic mapping is different in that the number of inside hosts can be greater than the number of outside hosts. The outside IP addresses are allocated on a first-come, first-serve basis to the inside IP addresses. When the last session for a specific inside IP address closes, that outside IP address can be used by other hosts. Since this mode does not rely on Layer 4 ports, ICMP traffic is translated and allowed to pass.

Port-mapping

Port-mapping gives you the most efficient use of the external address space. As each new connection is initiated from the inside, the NAT device picks the next available source Layer 4 port on the first available outside IP address. When all ports on a given IP address are in use, the NAT device uses ports off of the next outside IP address. Some systems reserve certain port ranges for specific types of traffic, so it is possible to map specific source Layer 4 port ranges on the inside to specific outside source ranges. However, this may cause a small performance penalty. In this case, you would need to make several rules using the same inside and outside IP addresses, one for each Layer 4 port range. ICMP traffic is not translated in this mode. You must add a dynamic NAT rule for the same IP address range to allow for ICMP traffic.

Auto-constraining

The auto-constraining algorithm for port-mapping limits the number of outside Layer 4 ports a single inside host can use simultaneously. The limitation is based on the ratio of inside to outside IP addresses. The outside IP address and Layer 4 port space is evenly distributed to all possible inside hosts. This guarantees that no single inside host can prevent other traffic from flowing through the NAT device. Because of the large number of simultaneous requests that can be made from a web browser, it is not recommended that this mode be used when a large number of inside hosts are being translated to a small number of outside IP addresses. ICMP traffic is not translated in this mode. You must add a dynamic NAT rule for the same IP address range to allow for ICMP traffic.

Configuring NAT

The behavior of NAT is determined by the rules you create to translate the IP addresses. You must attach each rule to a specific VLAN. All rules are processed in order. The options specified on the NAT rule determine the algorithm used to translate the inside IP addresses to the outside IP addresses. For outgoing (inside to outside) packets, the first rule to match is processed. All following rules are ignored. All return packets must arrive on the same outside VLAN on which the session went out. For most configurations, make sure that the outside IP addresses specified in the rule are part of the outside VLAN's subnet range, so that the switch can proxy the address resolution protocol (ARP) for those addresses.

To enable NAT functionality, use the following command:

```
enable nat
```

Creating NAT Rules

This section describes how to configure the various types of NAT (static, dynamic, portmap, and auto-constrain). In the examples in this section, advanced port and destination matching options have been removed. For information on how to use some of the more advanced rule matching features, refer to “Advanced Rule Matching” on page 9-8.

Creating Static and Dynamic NAT Rules

To create static or dynamic NAT rules, use this command:

```
config nat [add | delete] vlan <outside_vlan> map source [any |  
<ipaddress> [/<bits>| <netmask>]] to <ipaddress> [/<mask> | <netmask> |  
- <ipaddress>]
```

This is the simplest NAT rule. You specify the outside vlan name, and a subnet of inside IP addresses, which get translated to the outside IP address using the specified mode (static in this case). For the outside IP addresses, you can either specify an IP address and netmask or a starting and ending IP range to determine the IP addresses the switch will translate the inside IP addresses to. If the netmask for both the source and NAT addresses is /32, the switch will use static NAT translation. If the netmask for both the source and NAT addresses are not both /32, the switch will use dynamic NAT translation.

Static NAT Rule Example

```
config nat add out_vlan_1 map source 192.168.1.12/32 to 216.52.8.32/32
```

Dynamic NAT Rule Example

```
config nat add out_vlan_1 map source 192.168.1.0/24 to 216.52.8.1 -  
216.52.8.31
```


Creating Portmap NAT Rules

To configure portmap NAT rules, use this command:

```
config nat [add | delete] vlan <outside_vlan> map source [any |
<ipaddress> [/<bits>| <netmask>]] to <ip> [/<mask> | <netmask> | -
<ipaddress>] [{tcp |udp | both} portmap {<min> - <max>}]
```

The addition of an L4 protocol name and the `portmap` keyword tells the switch to use portmap mode. Optionally, you may specify the range of L4 ports the switch chooses on the translated IP addresses, but there is a performance penalty for doing this.

Remember that portmap mode will only translate TCP and/or UDP, so a dynamic NAT rule must be specified after the portmap rule in order to allow ICMP packets through without interfering with the portmapping.

Portmap NAT Rule Example

```
config nat add out_vlan_2 map source 192.168.2.0/25 to 216.52.8.32 /28
both portmap
```

Portmap Min-Max Example

```
config nat add out_vlan_2 map source 192.168.2.128/25 to 216.52.8.64/28
tcp portmap 1024 - 8192
```

Creating Auto-Constrain NAT Rules

To create auto-constrain NAT rules, use the following command:

```
config nat [add | delete] vlan <outside_vlan> map source [any |
<ipaddress> [/<bits>| <netmask>]] to <ip> [/<mask> | <netmask> | -
<ipaddress>] [{tcp |udp | both} auto-constrain}
```

This rule uses auto-constrain NAT. Remember that each inside IP address will be restricted in the number of simultaneous connections. Most installations should use portmap mode.

Auto-Constrain Example

```
config nat add out_vlan_3 map source 192.168.3.0/24 to 216.52.8.64/32
both auto-constrain
```

Advanced Rule Matching

By default, NAT rules only match connections based on the source IP address of the outgoing packets. Using the `L4-port` and `destination` keywords, you can further limit the scope of the NAT rule so that it only applied to specific TCP/UDP Layer 4 port numbers, or specific outside destination IP addresses.



Note: Once a single rule is matched, no other rules are processed.

Destination Specific NAT

```
config nat [add | delete] vlan <outside_vlan> map source [any |
<ipaddress> [ /<bits> | <netmask> ]] {destination <ipaddress/mask> } to
<ipaddress> [ /<mask> | <netmask> | - <ipaddress> ]
```

The addition of the `destination` optional keyword after the source IP address and mask allows the NAT rule to be applied to only packets with a specific destination IP address.

L4-Port Specific NAT

The addition of the `L4-port` optional keyword after the source IP address and mask allows the NAT rule to be applied to only packets with a specific L4 source or destination port. If you use the `L4-port` command after the source IP/mask, the rule will only match if the port(s) specified are the source L4-ports. If you use the `L4-port` command after the destination IP/mask, the rule will only match if the port(s) specified are the destination L4-ports. Both options may be used together to further limit the rule.

Configuring Time-outs

When an inside host initiates a session, a session table entry is created. Depending on the type of traffic or the current TCP state, the table entries time out after the configured time-out expires.

Displaying NAT Settings

To display NAT rules, use the following command:

```
show nat rules {vlan <outside_vlan>}
```

This command displays the NAT rules for a specific VLAN. Rules are displayed in the order they are processed, starting with the first one.

To display NAT traffic statistics, use the following command:

```
show nat stats
```

This command displays statistics for the NAT traffic, and includes:

- The number of rules
- The number of current connections
- The number of translated packets on the inside and outside VLANs
- Information on missed translations

To display NAT connection information, use the following command:

```
show nat connections
```

This command displays the current NAT connection table, including source IP/Layer 4 port mappings from inside to outside.

Disabling NAT

To disable NAT, use the following command:

```
disable nat
```


10

Server Load Balancing (SLB)

This chapter describes the following topics:

- Overview on page 10-1
- SLB Components on page 10-2
- SLB Traffic Types on page 10-7
- Forwarding Modes on page 10-7
- Load-Balancing Methods on page 10-15
- Advanced SLB Application Example on page 10-17
- Using Persistence on page 10-21
- Using High Availability System Features on page 10-24
- Health Checking on page 10-32
- Flow Redirection on page 10-35

Overview

Server load balancing (SLB) transparently distributes client requests among several servers. The main use for SLB is for web hosting (using redundant servers to increase the performance and reliability of busy websites).

You can use SLB to manage and balance traffic for client equipment such as web servers, cache servers, routers, and proxy servers. SLB is especially useful for e-commerce sites, Internet service providers, and managers of large intranets.

A basic SLB application is shown in Figure 10-1.

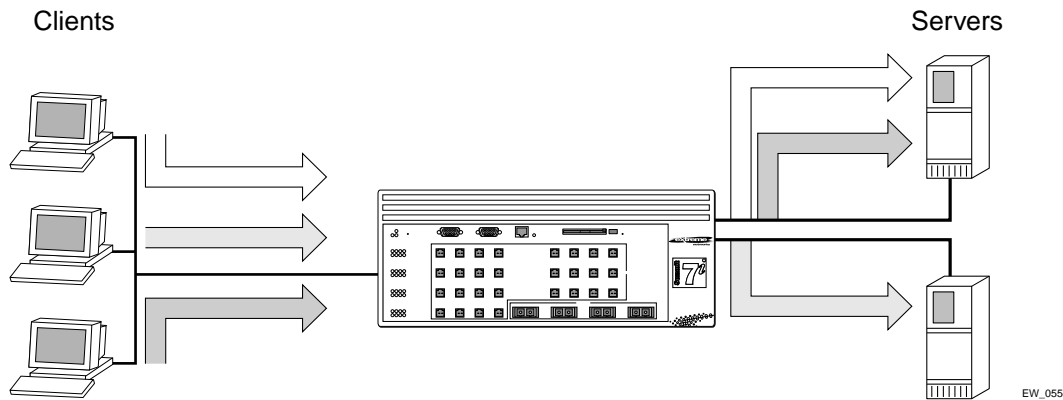


Figure 10-1: Basic SLB application

All content must be duplicated on all physical servers for server load balancing. To configure SLB, perform the following basic steps:

- 1 Create pools and configure the load balancing method for each pool.
- 2 Add nodes to the pools.
- 3 Create virtual servers and select the forwarding mode for each virtual server.
- 4 Assign an SLB traffic type to the server and client VLANs.
- 5 Enable IP forwarding on the server and client VLANs.
- 6 Enable SLB.

SLB Components

Three components comprise an SLB system:

- Nodes
- Pools
- Virtual servers

All three components are required for every SLB configuration.

Nodes

A *node* is an individual service on a physical server, and consists of an IP address and a port number. All nodes must have identical content. Nodes cannot belong to the same VLAN as the virtual servers they access.

Pools

A *pool* is a group of nodes that are mapped to a corresponding virtual server. You can use pools to easily scale large networks with many nodes.

Each pool contains its own load-balancing method. A pool must be associated with a virtual server to be used for load balancing. You must create pools before associating them with virtual servers, and must delete virtual servers before deleting their associated pools. You cannot delete pools that are still associated with a virtual server.

Virtual Servers

Virtual servers are the backbone of the SLB configuration. A virtual server is a virtual IP address that points to a group of servers. The switch then load balances those groups of servers (or other network equipment). Before you configure virtual servers, you need the following:

- The forwarding mode
- The name of the pool
- The virtual IP address
- The virtual port number

Virtual servers cannot belong to the same VLAN as the nodes in the pool they reference. Do not configure a virtual server with the same IP address as a VLAN.

Using Standard or Wildcard Virtual Servers

Each virtual server is associated with a single pool, which can be a group of content servers, routers, or cache servers.

You can configure two different types of virtual servers:

- Standard virtual servers

A standard virtual server represents a site (such as a web site or an FTP site), and provides load balancing for content. Configure the virtual server IP address to be the same IP address as that of the site that the virtual server represents.

- Wildcard virtual servers

A wildcard virtual server load balances transparent network devices such as routers or cache servers. Wildcard virtual servers use a special wildcard IP address (0.0.0.0), and require Transparent mode.



Note: For cache server applications, refer to Web Cache Redirection, on page 10-35.

Network Advertisement

Three modes are available for controlling network connectivity to virtual servers. The switch will automatically select a method based on the virtual server's subnet.

- Proxy ARP

If the virtual server is a member of an existing subnet to which the switch is directly attached, the switch will respond to ARP requests on behalf of the virtual server. This allows you to implement server load balancing on a layer 2 network. The VLAN containing the servers is in a different subnet than the client VLAN's subnet. The virtual server will appear to be a member of the client subnet.

- Host-Route

If the virtual server is not a member of an existing subnet to which the switch is directly attached, the switch will add a host-route entry to the routing table. In this situation, all clients require a routed path (to the virtual server) that points to the switch's IP address on the client VLAN.

- Subnet-Route

If the virtual server is separated from the switch by a router, the switch propagates the subnet containing the virtual server. You must create a loopback VLAN with the virtual server as a member of the loopback VLAN's subnet.

When you enable the routing protocol to advertise the entire subnet to directly connected routers, a single entry is created in the routing table for each subnet

advertised. For example, the following command enables RIP to advertise routes to all directly connected routers with a cost of 1:

```
enable rip export direct cost 1
```

When you enable the routing protocol to advertise specific virtual servers, an entry is created in the routing table for each virtual server you advertise. For example, the following command enables OSPF to advertise a specific virtual server with a cost of 1:

```
enable ospf export vip cost 1
```

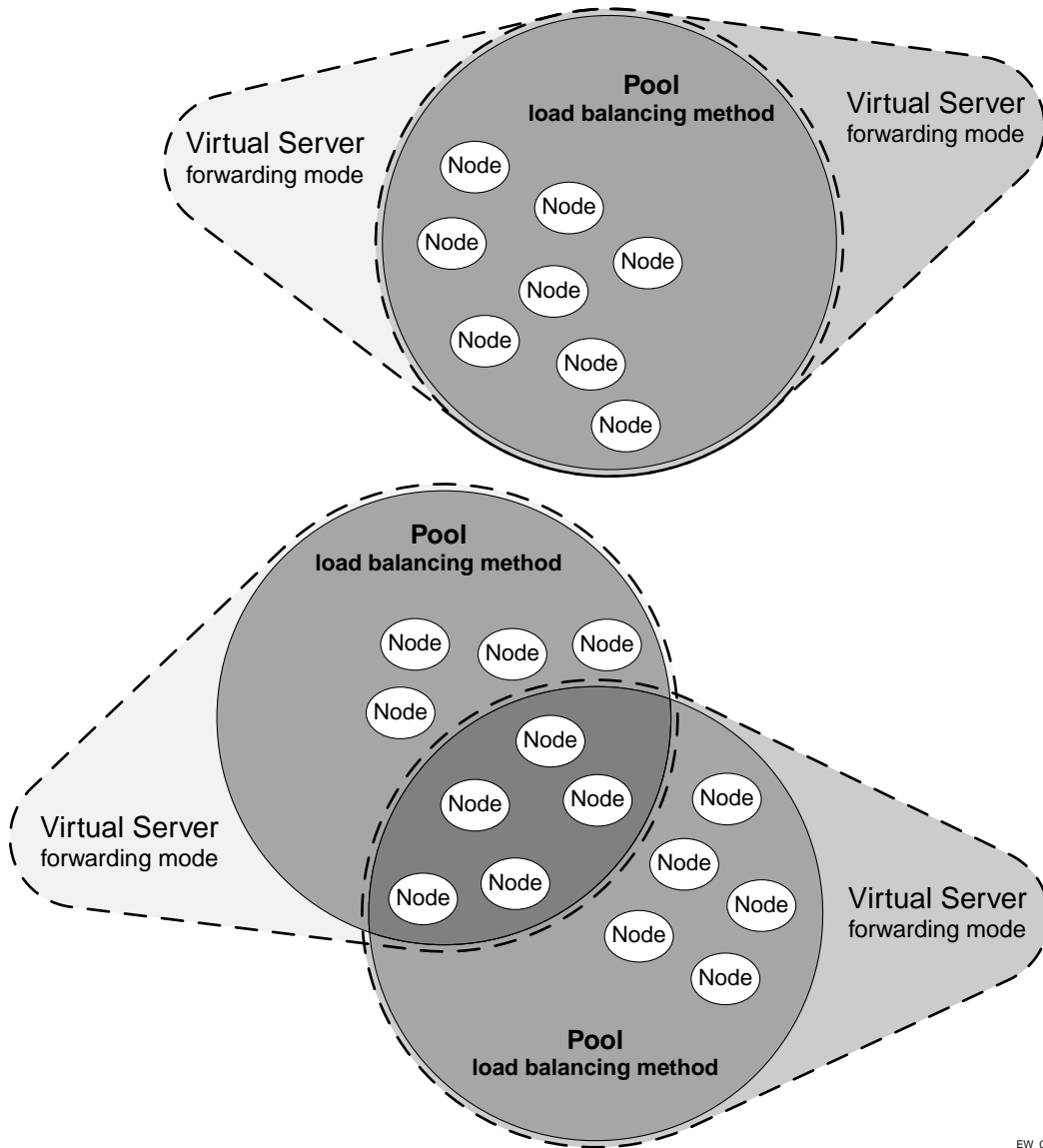
This command exports the virtual servers to the entire network. Extreme Networks recommends this method to advertise virtual servers.

Node, Pool, and Virtual Server Relationships

Nodes, pools, and virtual servers have the following relationships:

- Nodes can belong to multiple pools
- Pools can contain multiple nodes
- Pools can be associated with multiple virtual servers
- Virtual servers can be associated with only a single pool

Figure 10-2 illustrates the relationships of these basic components.



EW_069

Figure 10-2: Basic SLB components

SLB Traffic Types

SLB traffic must cross a routing boundary for SLB to work. To ensure that traffic crosses a routing boundary, assign clients and servers to separate VLANs. You must specify an SLB traffic type for each VLAN. The four SLB traffic types are:

- None—Disables SLB on the VLAN. This is the default setting.
- Client—Specifies that the VLAN contains clients, and originates requests for virtual servers.
- Server—Specifies that the VLAN contains nodes, and receives requests for virtual servers.
- Both—Specifies that the VLAN contains both clients and nodes. Clients in this VLAN can only access virtual servers whose nodes are outside this VLAN.



Note: You must enable IP forwarding on each VLAN involved in SLB.

You can assign the same SLB traffic type to as many different VLANs as necessary, up to the number of VLANs supported by the switch.

Forwarding Modes

The forwarding mode is the method the switch uses to forward traffic to the virtual servers. The forwarding mode determines what happens to the packets as they travel through the switch. The switch supports the following forwarding modes:

- Transparent
- Translation
- Port Translation
- GoGo

Table 10-1 summarizes the features supported by each forwarding mode.

Table 10-1: Forwarding Mode Feature Summary

	Transparent	Translation	Port Translation	GoGo
Performance	Hardware-based from server-to-client	CPU-based in both directions	CPU-based in both directions	Hardware-based in both directions
Load sharing algorithms	Round-robin, Ratio, Priority, Least Connections	Round-robin, Ratio, Priority, Least Connections	Round-robin, Ratio, Priority, Least Connections	Round-robin (hash)
Persistence	IPSA + Mask, IP list	IPSA + Mask, IP list	IPSA + Mask, IP list	IPSA
Health checking	Layer 3, 4, and 7	Layer 3, 4, and 7	Layer 3, 4, and 7	Layer 3, 4, and 7

Transparent Mode

In transparent mode, the switch does not modify IP addresses before forwarding traffic to the servers. You must configure all physical servers with the virtual IP address associated with the virtual server. This virtual IP address is the address seen by clients. You must configure the physical servers with this address as a loopback address. Configure the loopback address with the most specific subnet mask that your operating system supports.

In transparent mode, you can directly attach servers or have a layer 2 switch between the SLB switch and the servers. You cannot have a router between the SLB switch and the servers.

Use transparent mode when you want a balance between features and performance. Figure 10-3 shows transparent mode.

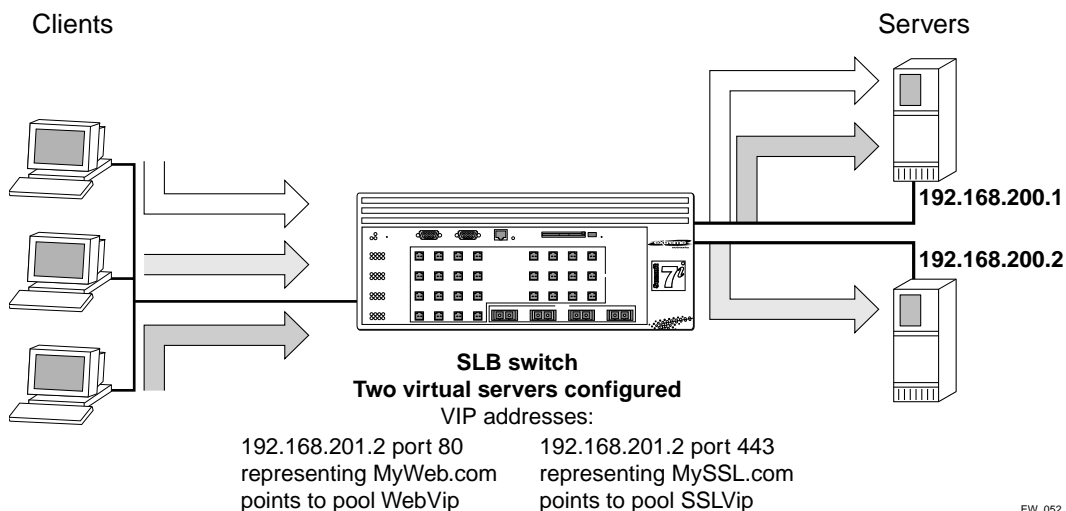


Figure 10-3: Transparent mode

In Figure 10-3, the switch is configured to respond to requests for the virtual server by forwarding them to the load-balanced servers.

The servers are configured as follows:

- The interface for server 1 is 192.168.200.1.
- The interface for server 2 is 192.168.200.2.
- The loopback address on the servers is 192.168.201.1 (virtual server).
- The service is configured to use the appropriate address and port, as specified in the switch configuration.

Use the following commands to configure the VLANs and the switch IP addresses and subnets:

```
create vlan srvr
create vlan clnt
create vlan vips
config srvr ipaddress 192.168.200.10 /24
config clnt ipaddress 10.1.1.1 /24
config vips ipaddress 192.168.201.1 /24
```

```
config srvr add port 29-32
config client add port 1-4
enable ipforwarding
```

Use the following commands to create a round-robin pool (MyWeb) and add nodes to the new pool.

```
create slb pool MyWeb lb-method round
config slb pool MyWeb add 192.168.200.1:80
config slb pool MyWeb add 192.168.200.2:80
```

Use the following command to create a transparent mode virtual server for the website and assign MyWeb to it:

```
create slb vip WebVip pool MyWeb mode transparent 192.168.201.2:80
```

Use the following commands to create a round-robin pool, MySSL and add nodes to the new pool.

```
create slb pool MySSL lb-method round-robin
config slb pool MySSL add 192.168.200.1:443
config slb pool MySSL add 192.168.200.2:443
```

Use the following command to create a transparent mode virtual server for the website and assign MySSL to it.

```
create slb vip SSLVip pool MySSL mode transparent 192.168.201.2:443
```

Use the following commands to enable SLB, configure the server VLAN to act as the server side, and configure the client VLAN to act as the client side:

```
enable slb
config vlan srvr slb-type server
config vlan clnt slb-type client
```

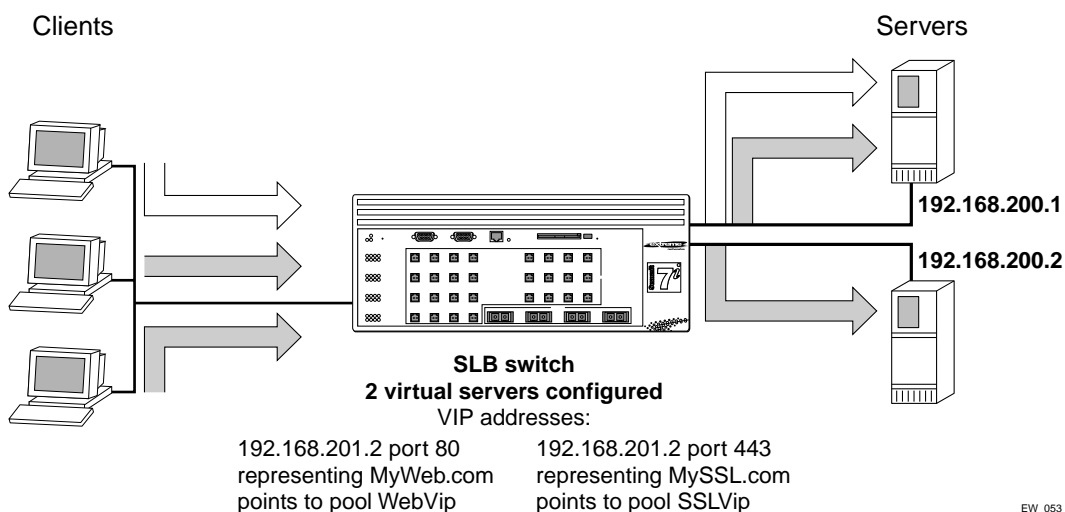
You must configure a loopback address for each IP address to which the server will respond.

Translation Mode

In translation mode, the switch translates the IP address to that of the server to be balanced. You do not need to configure a loopback address for translation mode.

In translation mode, you can directly attach servers or have a layer 2 switch between the SLB switch and the servers. You cannot have a router between the SLB switch and the servers.

Use translation mode when you cannot have a loopback address. Figure 10-4 shows translation mode.



EW_053

Figure 10-4: translation mode

In Figure 10-4, the switch is configured to respond to requests for the virtual server by translating them and forwarding them to the load balanced servers. No additional server configuration is needed.

Use the following commands to configure the VLANs and the switch IP addresses and subnets:

```
create vlan srvr
create vlan clnt
create vlan vips
config srvr ipaddress 192.168.200.10 /24
config clnt ipaddress 10.1.1.1 /24
```

```
config vips ipaddress 192.168.201.1 /24
config srvr add port 29-32
config client add port 1-4
enable ipforwarding
```

Use the following commands to create a round-robin pool, MyWeb, and add nodes to the new pool:

```
create slb pool MyWeb lb-method round
config slb pool MyWeb add 192.168.200.1:80
config slb pool MyWeb add 192.168.200.2:80
```

Use the following command to create a translation mode virtual server for the website and assign MyWeb to it:

```
create slb vip WebVip pool MyWeb mode translation 192.168.201.2:80
```

Use the following commands to create a round-robin pool, MySSL, and add nodes to the new pool:

```
create slb pool MySSL lb-method round
config slb pool MySSL add 192.168.200.1:443
config slb pool MySSL add 192.168.200.2:443
```

Use the following command to create a translation mode virtual server for the website and assign MySSL to it:

```
create slb vip SSLVip pool MySSL mode translation 192.168.201.2:443
```

Use the following commands to enable SLB, configure the server VLAN to act as the server side, and configure the client VLAN to act as the client side:

```
enable slb
config vlan srvr slb-type server
config vlan clnt slb-type client
```

Port Translation Mode

Port translation mode is similar to translation mode, except that the layer 4 port on the virtual server can be different from the layer 4 port on the nodes. The switch translates the IP address and port address to that of the servers to be balanced.

In port translation mode, you can directly attach servers or have a layer 2 switch between the SLB switch and the servers. You cannot have a router between the SLB switch and the servers.

Use port translation mode when you must translate layer 4 port numbers in addition to translating IP addresses.

GoGo Mode

GoGo mode is a line rate method of server load balancing that forwards traffic without changing packet content. You must directly attach servers to the SLB switch in GoGo mode.

The optimal configuration is groups of 2, 4, or 8 servers. Because you must directly attach servers, you do not need to configure nodes, pools, or virtual servers. Instead, you configure all servers with the same MAC and IP addresses. Clients then see the group of servers as a single server, much like port-based load sharing.

As in port-based load sharing, the first port in the GoGo mode group is designated the “master” logical port. Use this port to represent the entire GoGo mode group when configuring health checks or VLAN membership.

In GoGo mode, the load balancing method is fixed, based on a hashing of the client IP address. GoGo mode persistence is based on source IP information: a given source address will map to one, and only one, physical server.

Use GoGo mode when you require performance without any traffic management features. Figure 10-5 shows GoGo mode.

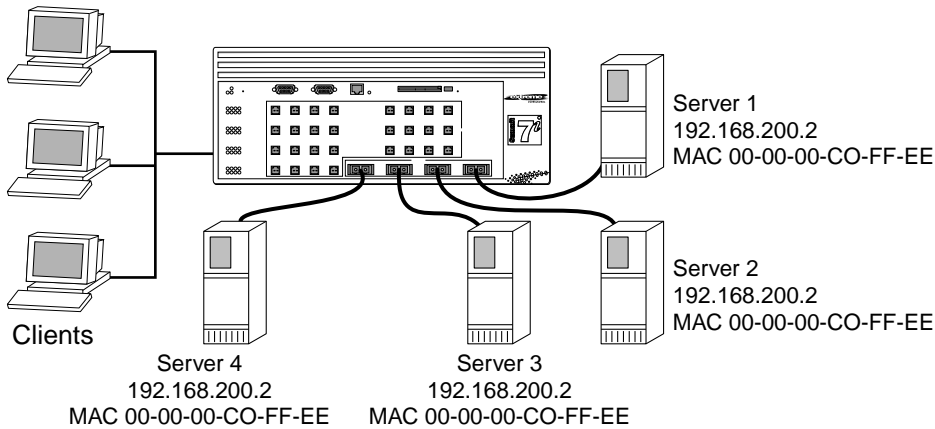


Figure 10-5: GoGo mode

In Figure 10-5, the switch is configured to balance all traffic sent to the virtual server based on the client IP address.

The servers are configured as follows:

- All servers have the same MAC address.
- All servers have the same IP address.
- All servers have the same content.

To configure the switch as indicated in the example, use the following commands:

```
create vlan server
create vlan client
config server ipaddress 192.168.200.2 /24
config client ipaddress 1.1.1.1 /24
config server add port 29-32
config client add port 1-4
enable slb gogo 29 grouping 29-32
enable ipforwarding
```

In this example, port 29 is designated the master port.

GoGo mode requires you to place clients and servers into separate VLANs.

Load-Balancing Methods

Load-balancing methods are algorithms that determine which node receives a connection hosted by a particular virtual server. The forwarding mode determines *how* the switch forwards traffic; the load-balancing method determines *where* the switch forwards traffic.

Individual load-balancing methods take into account dynamic factors such as current connection count. Because each application of SLB is unique, node performance depends on a number of different factors. We recommend that you experiment with different load-balancing methods and choose the one that offers the best performance in your particular environment.

The switch supports the following load balancing methods:

- Round-robin
- Ratio
- Least connections
- Priority

Round-Robin

Round robin passes each new connection request to the next server in line. Because round robin is simple and predictable, it is the default load-balancing method.

Use round-robin if the equipment that you are load balancing is roughly equal in processing speed and memory.

Ratio

Ratio distributes connections among servers according to ratio weights that you set. The number of connections that each server receives is proportionate to the ratio weight you defined for each server.

Use ratio if the equipment that you are load balancing varies significantly in processing speed and memory. For example, if you have one new, high-speed server and two older servers, you can set the ratio so that the high-speed server receives twice as many connections as either of the two older servers.

A ratio of 2 results in twice as much traffic as a ratio of 1. If all nodes use the same weight, connections are distributed equally among the nodes. The default ratio is 1.

Least Connections

Least connections method passes a new connection to the node having the least number of active sessions. The number of active sessions includes only those sessions occurring within the same virtual server.

Use least connections when the equipment that you are load balancing has similar capabilities. Because least connections requires more processing, it works best with small pools (under 25 nodes) when you require intelligent distribution.

Priority

Priority is a variant of round-robin designed to provide redundant “standby” nodes within a pool. When you add a node to a pool, you can assign a priority level ranging from 1 - 65535, with a higher number indicating a higher priority.

In priority, the switch uses round-robin to distribute traffic among the active nodes with the highest priority. If all nodes at that priority level become inactive or reach a session limit maximum, all new sessions are directed to the nodes at the next lowest priority. The switch monitors the status of the inactive nodes. As each node becomes active, the switch redistributes traffic according to the priorities.

For example, in a pool with six nodes divided evenly into two priority levels (2 and 1), all sessions are evenly distributed to the priority 2 nodes. If one of the priority 2 nodes becomes inactive, all traffic is assigned to the remaining priority 2 nodes. If all of the priority 2 nodes become inactive, all sessions are directed to the priority 1 nodes. If one of the level 2 nodes becomes active, all new sessions are assigned to it.

Use priority when you want a set of servers held in reserve, as a back-up pool.

Advanced SLB Application Example

The advanced features described in this section are:

- Persistence
- High availability
- 3DNS support
- Flow redirection
- Health checking

The advanced SLB application example builds upon the basic SLB application example. The advanced concepts included in this example are:

- Multiple pools.
- Multiple virtual servers.
- Multiple balancing algorithms.
- Multiple types of health checking.

Figure 10-6 is an example of an advanced SLB application.

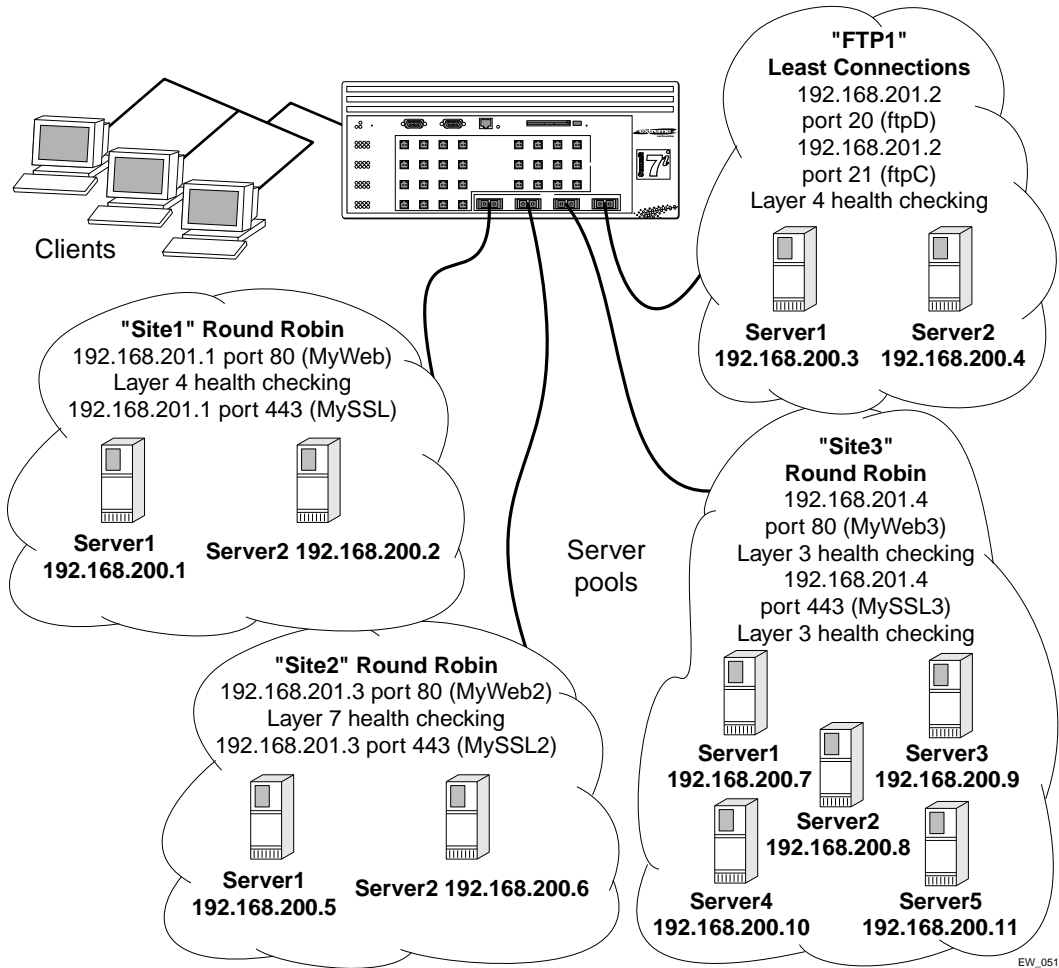


Figure 10-6: Advanced SLB configuration

To create the VLAN from which outside connections will come, use the following commands:

```
create vlan outside
config vlan outside ipaddress 172.16.0.1 /16
```

```
config vlan outside add ports 1-8
```

To create is the virtual IP VLAN, use the following commands:

```
create vlan sites
config vlan sites ipaddress 192.168.201.254 /24
```

All virtual servers will use this subnet. There are no ports associated with this VLAN.

Use the following commands to create the VLAN *servers* and enable IP forwarding:

```
create vlan servers
config vlan servers ipaddress 192.168.200.254 /24
config vlan servers add ports 9-16
enable ipforwarding
```

Use the following series of commands to create a web site. The site is defined as having two servers: 192.168.200.1 and 192.168.200.2, each with two services (HTTP and SSL). Two virtual servers point at the appropriate pools. The load-balancing method is round-robin. Both virtual servers use the same IP address; the difference is the port number. Finally, port checking is enabled to ensure fault tolerance on each of the servers.

```
create slb pool sitelweb
config slb sitel add 192.168.200.1:80
config slb sitel add 192.168.200.2:80
create slb pool sitelssl
config slb sitel add 192.168.200.1:443
config slb sitel add 192.168.200.2:443
create slb vip myweb pool sitelweb mode transparent 192.168.201.1:80
create slb vip myssl pool sitelssl mode transparent 192.168.201.1:443
enable slb node 192.168.200.1:80 tcp-port-check
enable slb node 192.168.200.2:80 tcp-port-check
enable slb node 192.168.200.1:443 tcp-port-check
enable slb node 192.168.200.2:443 tcp-port-check
```

Use the following series of commands to create a second web site. This site is similar to the first site, except that content checking is enabled on this site.

```
create slb pool site2web
config slb site2web add 192.168.200.5:80
config slb site2web add 192.168.200.6:80
create slb pool site2ssl
config slb site2ssl add 192.168.200.5:443
config slb site2ssl add 192.168.200.6:443
create slb vip myweb2 pool site2web mode transparent 192.168.201.3:80
create slb vip myssl2 pool site2ssl mode transparent 192.168.201.3:443
enable slb vip myweb2 service-check
config slb vip myweb2 service-check http url "/testpage.htm"
match-string "test successful"
```

Use the following series of commands to create a third web site. This example has one pool with a wildcard port. The wildcard port allows any port sent to it by the virtual server. All five servers respond to requests on both port 80 and port 443.

```
create slb pool site3web
config slb site3web add 192.168.200.7:0
config slb site3web add 192.168.200.8:0
config slb site3web add 192.168.200.9:0
config slb site3web add 192.168.200.10:0
config slb site3web add 192.168.200.11:0
create slb vip myweb3 pool site3web mode transparent 192.168.201.4:80
create slb vip myssl3 pool site3web mode transparent 192.168.201.4:443
```

Use the following series of commands to create an FTP site. The site has two servers: 192.168.200.3 and 192.168.200.4. The servers provide only FTP service. The two different virtual servers and port numbers refer to the control and data channels used by the FTP service. Two virtual servers point at the appropriate pools.

The load-balancing method is round-robin. Layer 7 health checking is enabled for the ftpc virtual server.

```
create slb pool ftp1c
config slb ftp1c add 192.168.200.3:21
config slb ftp1c add 192.168.200.4:21
create slb pool ftp1d
config slb ftp1d add 192.168.200.3:20
config slb ftp1d add 192.168.200.4:20
create slb vip ftpc pool ftp1c mode transparent 192.168.201.2:21
create slb vip ftpd pool ftp1d mode transparent 192.168.201.2:20
enable slb vip ftpc service-check
config slb vip ftpc service-check ftp user test password testpass
```


Finally, enable SLB and configure the VLANs to be either client or server, using the following commands.

```
enable slb
config vlan outside slb-type client
config vlan servers slb-type server
```

Using Persistence

Persistence ensures that subsequent connections from the same client attach to the same server. To configure persistence, you select:

- persistence method
- persistence level
- persistence type

Persistence is not affected by the load-balancing method unless you select GoGo mode, where the persistence is fixed, as described on page 10-13.

Persistence Methods

Persistence methods determine how the persistence table times-out entries. The persistence methods are:

- Per-session
- Per-packet

Per-Session Persistence

Per-session persistence creates a persistence entry when the first packet arrives from a client with the timeout value configured for the virtual server. The entry is removed after the timeout period. The entry is not refreshed. Per-session is the default persistence method.

Use per-session persistence when you want the smallest impact on performance and you can accurately gauge your total connection time.

Per-Packet Persistence

Per-packet persistence creates a persistence entry when the first packet arrives and refreshes that entry each time a new packet arrives. Thus, the entry remains as long as new packets continue to arrive.

Use per-packet persistence when you want to sacrifice a small amount of performance in return for a timeout period based on connection idle time instead of total connection time.

Persistence Levels

Persistence levels determine how the persistence entries affect multiple virtual servers. Use persistence levels if you have servers that provide multiple services to a single client (such as, HTTP and SSL). To use persistence levels, your virtual servers must contain the same physical servers.

The persistence levels are as follows:

- Same-VIP-same-port
- Same-VIP-any-port
- Any-VIP

Same-VIP-Same-Port Persistence

Same-VIP-same-port matches a new client request to a persistence entry only if the destination is the same virtual server and same port as the original client request. Same-VIP-same-port is the default persistence method.

Use same-VIP-same-port persistence to ensure that connections from a client are only persistent on the specific virtual server that is connecting to that client.

Same-VIP-Any-Port Persistence

Same-VIP-any-port persistence directs client requests to the same virtual server even if the port is different.

Use same-VIP-any-port persistence to ensure that connections from a client are persistent on the virtual server for all layer 4 services offered on that particular virtual server. For example, if you use HTTP (port 80) to build a shopping cart, then need to

use SSL (port 443) for the credit card transaction at the end, use same-VIP-any-port persistence to preserve the client session.

If you have virtual servers with the same IP address but a different port, you must configure associated pools with identical nodes that can service requests on either port.

Any-VIP Persistence

Any-VIP persistence directs all connections from a client to the same virtual server regardless of the destination.

Use any-VIP persistence to ensure that connections from a client always go to the same server no matter what layer 4 service they connect to. When you use any-VIP persistence, you must ensure that all servers have the same content for all services.

Persistence Types

The switch supports the following types of persistence:

- Client persistence
- Proxy client persistence
- Sticky persistence

Client Persistence

Client persistence provides a persist mask feature. You can define a range of IP addresses that map to a persistent connection. Any client whose source IP address falls within the range is considered a match for the entry.

Use client persistence to ensure that transactions, from your defined range of IP addresses, that span multiple TCP sessions are always connected to the same virtual servers. For example, if you assume that a single client uses multiple sessions to fill a shopping cart, you can force all traffic from the same range of IP addresses (which you assume to be the same client) to connect to the same virtual server.

Proxy Client Persistence

Some networks translate addresses internally with an array of proxy servers. Proxy client persistence allows the switch to maintain connections for clients in these types of networks. You can define ranges of IP addresses that map to a persistent connection.

Any client whose source IP address falls within one of the ranges is considered a match for the entry. You must add every range of possible source IP addresses.

Use proxy client persistence to provide persistence for users who are behind proxy servers that change the source IP address of client requests from session to session.

Sticky Persistence

Sticky persistence is available only on wildcard virtual servers and is especially useful for cache servers. Sticky persistence tracks destination IP addresses. When a client attempts to connect to a destination IP address, the switch directs the client to the same cache server previously used for that destination IP address. This helps you reduce the amount of duplicated content on cache servers in your network.

Use sticky persistence to provide persistence based on destination IP address. Sticky persistence is especially useful when you load balance caching proxy servers. A caching proxy server intercepts web requests and returns a cached web page (if that page is available). You can improve the efficiency of cache proxy servers by sending similar requests to the same server. Sticky persistence can cache a given web page on one proxy server instead of on every proxy server. This saves the other servers from duplicating the content.



Note: For additional cache server applications, refer to Web Cache Redirection, on page 10-35.

Using High Availability System Features

The switch supports several advanced redundant system features. Advanced redundant system features provide additional assurance that your content remains available if a switch experiences a problem. The advanced redundant system options include:

- SLB with ESRP
- Active-active operation

Server Load Balancing with ESRP

You can use ESRP to make SLB, along with the layer 2 and layer 3 services of the switch, redundant. SLB with ESRP allows single- or dual-attached servers to have redundant gateway services and very fast recovery from a fault. When you enable

ESRP, all servers can be online simultaneously, and recovery from a switch failure occurs in less than 8 seconds.

Figure 10-7 shows SLB enabled using ESRP and dual-attached servers.

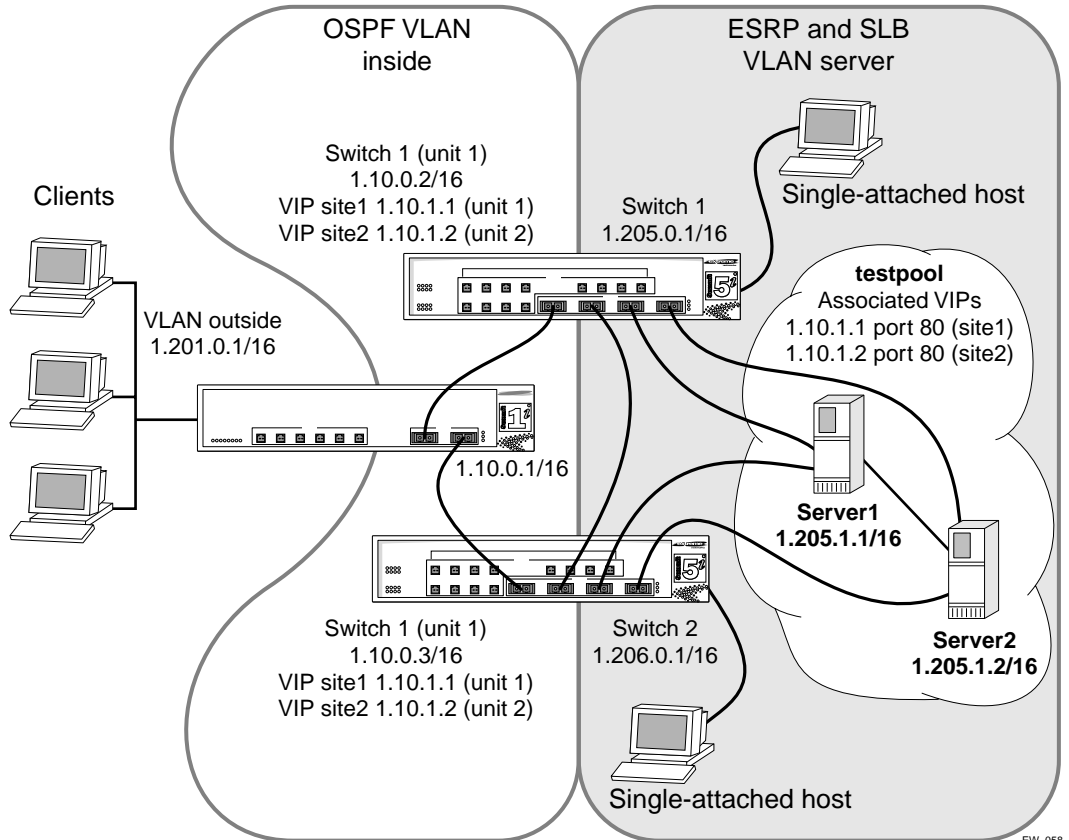


Figure 10-7: SLB using ESRP and dual-homed servers

EW_058

Configuring the Switches for SLB and ESRP

To create the VLANs, use the following commands:

```
create vlan inside
create vlan server
```

To connect the gateway to the VLAN *inside*, use the following commands:

```
config inside ipaddress 1.10.0.2 /16
config inside add port 31
```

To configure the servers to connect to the VLAN *server* on ports 1-4, and configure port 32 to connect to the other ESRP switch, use the following commands:

```
configure server ipaddress 1.205.0.1 /16
configure server add port 1-4, 32
```

To enable ipforwarding, create a server pool called *testpool*, and add four servers to it using TCP port 80, use the following commands:

```
enable ipforwarding
create slb pool testpool
config slb pool testpool add 1.205.1.1:80
config slb pool testpool add 1.205.1.2:80
config slb pool testpool add 1.205.1.3:80
config slb pool testpool add 1.205.1.4:80
```

To create SLB virtual server addresses for the two websites (*site1* and *site2*) and associate the server pool *testpool* with it, use the following commands:

```
create slb vip site1 pool testpool mode transparent 1.10.1.1:80
create slb vip site2 pool testpool mode transparent 1.10.1.2:80
```

To enable SLB and configure it for the appropriate VLANs (client connections enter from the VLAN *inside*), use the following commands:

```
enable slb
config inside slb client
config server slb server
```

To enable OSPF, use the following command:

```
enable ospf
```

To enable ESRP on the VLAN server and configure the ESRP direct-attached hosts mode to allow the proper failover of services, use the following commands:

```
enable esrp server
config esrp port-mode host ports 1-4, 32
```

the interconnection between the switches is also configured as a host port.

To configure SLB to use ESRP, use the following command:

```
config slb esrp server add unit 1
```

Note the following about the configurations for the switches running SLB and ESRP:

- You must configure all switch ports connected directly to the servers as ESRP host ports.
- You must configure the link between the two switches as an ESRP host port.
- The configuration uses transparent mode and HTTP services, but can be configured to support any of the currently supported load-balancing protocols.
- Both switches are configured as unit 1.
- The SLB and ESRP port configurations are identical on both switches.

Web-Server Configuration

The services must match those configured on the switch; for example, HTTP services configured at TCP port 7080 on the switch require the servers to allow connections at port 7080. You must ensure that the SLB configuration is valid before trying to transfer the configuration to an ESRP/SLB configuration.

The two types of ESRP hosts that you can connect to the switches are single-attached hosts and dual-attached hosts. Single-attached hosts provide no server link redundancy, but allow hosts to be connected to the same VLAN as the web servers. Dual-attached hosts allow for redundant NICs in the servers, as well as connections to the switch. When configured as dual-attached hosts, the servers are supported fully by the ESRP redundant gateway services.



Note: For information on specific NIC card configurations, please contact your NIC vendor.

Active-Active Operation

Active-active operation is a redundant configuration of two switches. If one switch fails, the second switch takes over the SLB function. By preparing a redundant switch for the possibility of failover, you provide reliability and availability.

To create an active-active configuration, configure redundant switches with identical SLB configurations, except for the failover parameters.

Active-active operation uses a gateway ping-check to determine if the active SLB switch has network connectivity. If the specified IP address is unreachable for a specified duration, the gateway ping-check triggers a failover to the redundant switch.



Note: When you configure the gateway ping check, specify the IP address of a device other than the redundant SLB switch.

Configuring Active-Active Operation

Using active-active redundant SLB, you configure one switch as unit 1 and the other switch as unit 2. You then assign the virtual servers either to unit 1 or to unit 2 (unit 1 is the default). When both switches are active, each switch performs SLB only for the virtual servers assigned to it. If a switch fails, the remaining switch performs SLB for the virtual servers assigned to the failed switch.

Use the following command to assign the unit number:

```
config slb failover unit [1 | 2] remote-ip <ipaddress> local-ip
<ipaddress>:<L4Port> {alive-frequency <seconds> timeout <seconds>}
{dead-frequency <seconds>}
```

The `remote-ip` specifies the IP address of the redundant SLB switch. The `local-ip` specifies the IP address of the switch you are configuring.

You must assign virtual servers with the same virtual IP address to the same unit.

Sample Active-Active Configuration

Figure 10-8 is an example of an active-active failover configuration.

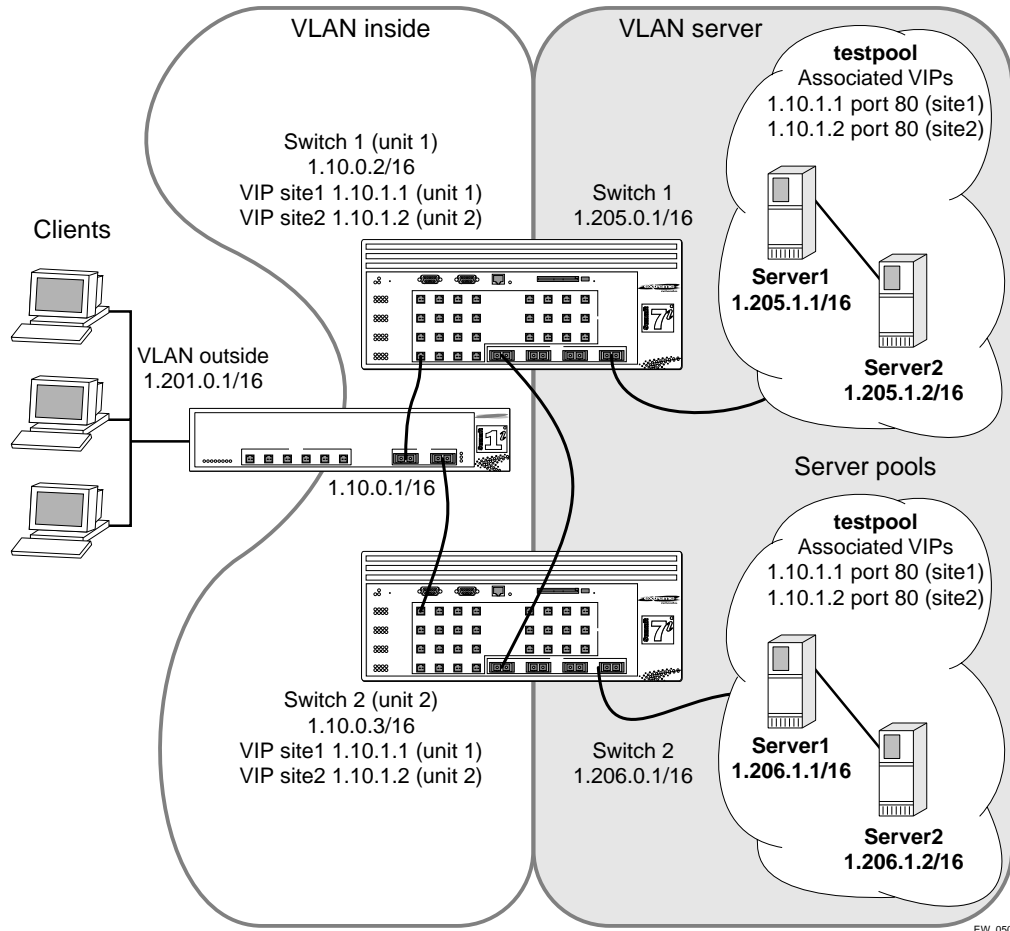


Figure 10-8: Active-active configuration

To configure this example on the first switch, use the following commands:

```
create vlan inside
create vlan server
config vlan inside ipaddress 1.10.0.2 /16
config vlan inside add port 31
config vlan server ipaddress 1.205.0.1 /16
```

```
config vlan server add port 29-30

enable ipforwarding

create slb pool testpool
config slb pool testpool add 1.205.1.1:80
config slb pool testpool add 1.205.1.2:80
create slb vip site1 pool testpool mode transparent 1.10.1.1:80
create slb vip site2 pool testpool mode transparent 1.10.1.2:80

enable slb
config vlan inside slb-type client
config vlan server slb-type server

config slb failover unit 1 remote 1.10.0.3 local 1.10.0.2:1028

enable slb failover

enable slb failover ping

config slb vip site1 unit 1
config slb vip site2 unit 2

config slb fail ping-check 1.10.0.1 freq 1
```

To configure this example on the second switch, use the following commands:

```
create vlan inside
create vlan server
config vlan inside ipaddress 1.10.0.3 /16
config vlan inside add port 31
config vlan server ipaddress 1.206.0.1 /16
config vlan server add port 29-30

enable ipforwarding

create slb pool testpool
config slb pool testpool add 1.206.1.1:80
config slb pool testpool add 1.206.1.2:80
create slb vip site1 pool testpool mode transparent 1.10.1.1:80
create slb vip site2 pool testpool mode transparent 1.10.1.2:80

enable slb
config vlan inside slb-type client
```

```

config vlan server slb-type server

config slb failover unit 2 remote 1.10.0.2 local 1.10.0.3:1028
enable slb failover
enable slb fail ping

config slb vip site1 unit 1
config slb vip site2 unit 2

config slb fail ping-check 1.10.0.1 freq 1

```

The differences between the configurations of these two switches are the IP addresses and the designation of the first switch as unit 1 of the active-active configuration.

If you use this configuration with only one virtual server, you have an active switch and a standby switch, because only one switch is actively performing SLB. This configuration is called “active-standby.”

Active-Active Configuration Notes

Note the following about active-active configurations:

- In the design shown in Figure 10-8, only the servers directly connected to the switch that is actively servicing the virtual server are used in the load-balancing scheme. Without ESRP, you must have another switch interconnecting all the servers.
- One switch is designated as unit 1 and the other is unit 2. This designation determines which virtual servers are active on each switch in the failover pair.
- In this configuration, *site1* is serviced by switch 1 and has two servers that respond to client requests. *Site2* is be serviced by the remote switch (switch 2) and has two other servers that respond to client requests.
- If you enable ping-check, do not direct it at the remote switch. The ping-check works best when directed at a gateway to ensure that a path out of the network is available to the switch.
- The configuration uses transparent mode and HTTP services, but can be configured to support any of the currently supported load balancing protocols.
- The configurations for the switches are identical, with the exception of the failover command.
- The remote switch is set to unit 2, and the remote/local IP addresses are reversed to accurately describe the network.

Using Manual Fail-Back

With manual fail-back, fail-back occurs only when you enter the fail-back command. In an active-active configuration, fail-back occurs automatically. If the minor disruption of fail-back makes automatic fail-back undesirable, you can enable manual fail-back.

Health Checking

Health checking allows you to actively poll nodes to determine their health. The switch makes new connections only if the virtual server and node are both enabled and passing health checks. The switch considers a virtual server or node active unless a health check fails. If a health check fails, the switch considers the virtual server or node inactive. A virtual server or node is also considered inactive if it is disabled and has zero active connections. If it is inactive for this reason, the switch stops ping-checks and port-checks on the virtual server or node to conserve system resources. The switch resumes ping checks and port checks when you enable the virtual server or node.

The switch does not establish new connections with an inactive node until that node passes all configured health checks. If a health check fails and you have enabled the `ign-reset` parameter on an associated virtual server, the switch closes all existing connections for the virtual server by sending a TCP reset to the client and node.

The switch supports three types of health checking:

- Ping-check
- Port-check
- Service-check

The switch also supports 3DNS health checking.

Ping-Check

Ping-check operates at layer 3 and is the default health check. The switch sends an ICMP ping to the configured server or next hop. The default ping frequency is 10 seconds and the default timeout is 30 seconds (three pings). If the node does not respond within 30 seconds, it is considered down. Ping check is the only health check that will accept a wildcard as the IP port.

TCP-Port-Check

TCP-port-check operates at layer 4 and tests the physical node. The default frequency is 30 seconds and the default timeout is 90 seconds. If the node does not respond within 90 seconds, it is considered down. You can use TCP-port-checking to determine if a TCP service, such as httpd, is available. If a TCP-port-check fails, the IP/port combination is considered unavailable.

Service-Check

Service-check operates at layer 7 and is an application-dependent check defined on a virtual server. The switch performs a service-check on each node in the pool. The default frequency is 60 seconds and the default timeout is 180 seconds. Table 10-2 describes the service-check parameters.

Table 10-2: Service-Check Parameters

Service	Attribute	Global Default Value
HTTP	URL	"/"
	Match-string	Any-content
FTP	Userid	"anonymous"
	Password	"anonymous"
Telnet	Userid	"anonymous"
	Password	"anonymous"
SMTP	Dns-domain	Same as the switch DNS domain. If no DNS domain is configured for the switch, the value is "".
NNTP	Newsgroup	"ebusiness"
POP3	Userid	"anonymous"
	Password	"anonymous"

If you do not specify the service-check parameters, the switch uses the global default values. You can configure the global default values.

For HTTP, you can specify both the URL to be retrieved, and a `match-string`, such as "Welcome." If the switch finds the `match-string` in the first 1000 bytes of the retrieved URL, the service-check passes. A `match-string` specified as `any-content` matches any retrieved text. Extreme Networks recommends that you create a text file that contains a single word, such as "ok."

The FTP, Telnet, and POP3 service-checks establish a connection between the switch and the next hop. Service-check logs on and off using the specified `userid` and `password`.

For SMTP, service-check identifies the switch by providing the DNS domain you configure. Extreme Networks recommends that you specify a DNS domain that is used only for identification.

The NNTP service-check connects to the node, logs in, and attaches to the newsgroup specified.

You configure service-checks for each virtual server, and nodes can be members of multiple virtual servers. Therefore, because each node can have multiple service-checks, some service-checks can fail while others pass. So to accept a new connection for a virtual server, a node must have passed the service-check configured for that virtual server. When showing detailed virtual server information, the status for individual nodes is shown with respect to that virtual server.

3DNS Health Checking

When you enable SLB, the switch reports health status to 3DNS using the iQuery™ protocol from F5 Networks®. The health status of the nodes within the server farm is based on layer 3, layer 4, layer 7, or external health check mechanisms.

Maintenance Mode

You can put a node or virtual server into maintenance mode by disabling the node or virtual server. In maintenance mode, existing connections remain active, but no new connections are permitted. The existing connections are either closed by the client and server or are aged out if idle for more than 600 seconds.

Health Checking in GoGo Mode

GoGo mode does not use nodes, pools, or virtual servers. Therefore, to configure health checking when using GoGo mode, you must use the GoGo mode health checking commands.

Flow Redirection

Flow redirection overrides routing decisions to transparently redirect client requests to a target device (or group of devices). Unlike SLB, you do not duplicate content on the target device(s).

The switch can only redirect traffic that crosses a VLAN boundary, because flow redirection operates at layer 3. Flow redirection examines traffic and redirects it based on the following criteria, in order of priority:

- 1 Destination IP address and mask
- 2 Layer 4 port
- 3 Source IP address and mask

You must prevent logical loops of redirected traffic. You can use flow redirection for the following:

- Web cache redirection
- Policy-based routing



Note: Extreme Networks recommends that you use flow redirection and SLB on separate switches; do not use flow redirection and SLB on the same switch.

Web Cache Redirection

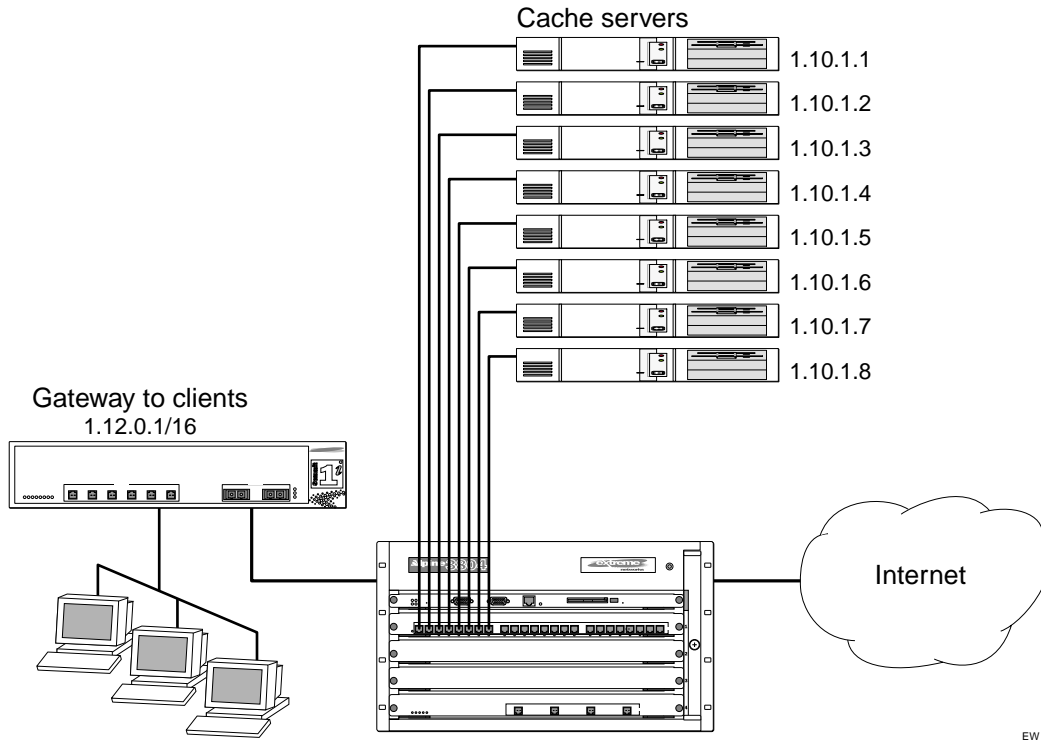
Web cache redirection operates at line rate to redirect traffic from the requested server to a web cache server. If the web cache server has a copy of the requested content, it sends the content to the client. If the web cache server does not have the requested content, it queries the server for the data, stores it locally, and sends a copy to the client.

When you have web cache redirection enabled, clients connect exclusively to your web cache servers; clients never connect to the requested server.

The switch automatically load-balances your cache servers based on the destination IP address of the requested content. Thus, subsequent requests for a destination IP address are redirected to the same web cache server, because that web cache server is most likely to contain the requested content. This load-balancing reduces the amount of content duplication on your web cache servers.

Web Cache Redirection Example

Figure 10-9 uses flow redirection to redirect Web traffic to cache servers. In this example, the clients and the cache devices are located on different networks. This is done by creating a different VLAN for the clients and cache devices.



EW_064

Figure 10-9: Web cache redirection example

To configure the switch in this example, use the following commands:

```
create vlan client
config vlan client add port 1
config vlan client ipaddress 1.12.0.1/16
```



```
create vlan cache
config vlan cache add port 2
config vlan cache ipaddress 1.10.1.1/24

create vlan internet
config vlan internet add port 3
config vlan internet ipaddress 1.11.1.1/16

enable ipforwarding

create flow-redirect flow1 tcp destination 1.11.1.0/24 ip-port 80
source any

config flow1 add next-hop 1.10.1.2
config flow1 add next-hop 1.10.1.3
config flow1 add next-hop 1.10.1.4
config flow1 add next-hop 1.10.1.5
config flow1 add next-hop 1.10.1.6
config flow1 add next-hop 1.10.1.7
config flow1 add next-hop 1.10.1.8
```

Policy-Based Routing

Policy based routing is an application of flow redirection that allows you to control routed traffic regardless of the routing protocol configured. For example, you can use policy-based routing to force SNMP traffic to follow a less efficient but more secure path.

As with web cache redirection, policy-based routing examines traffic and redirects it based on the following criteria (in order of priority):

- 1 Destination IP address and mask
- 2 Layer 4 port
- 3 Source IP address and mask

If the next-hop address is unavailable, the switch routes the traffic normally. You can define several rules; the precedence of rules is determined by the best match of the rule to the packet. If no rule is satisfied, no redirection occurs.

If you define multiple next-hop addresses, traffic satisfying the rule is load-shared across the next hop addresses based on destination IP address. If next hop addresses do not respond to ICMP pings, the switch resumes normal routing.

Policy-based routing has no impact on switch performance unless you use policy-based routing and SLB on the same switch.

11

Ethernet Automatic Protection Switching

This chapter describes the use of the Ethernet Automatic Protection Switching (EAPS™) protocol, and includes information on the following topics:

- Overview of the EAPS Protocol on page 11-1

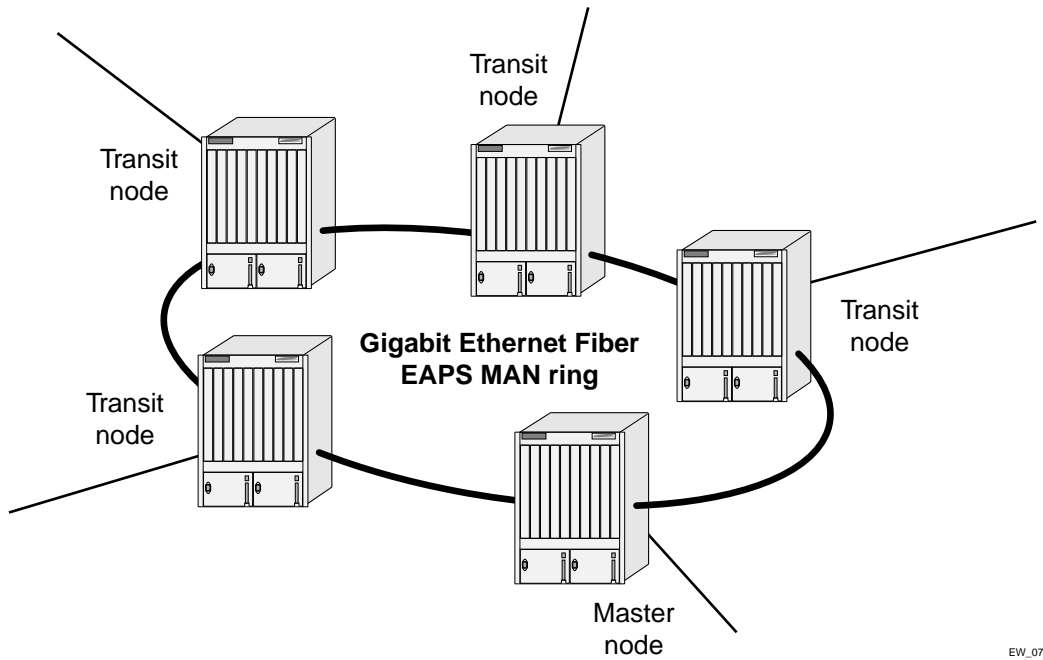
Overview of the EAPS Protocol

The EAPS protocol provides fast protection switching to Layer 2 switches interconnected in an Ethernet ring topology, such as a Metropolitan Area Network (MAN) or large campuses (see Figure 11-1).

EAPS protection switching is similar to what can be achieved with the Spanning Tree Protocol (STP), but offers the advantage of converging in less than a second when a link in the ring breaks.

In order to use EAPS, you must enable EDP on the switch. For more information on EDP, refer to Chapter 4.

EAPS operates by declaring an EAPS domain on a single ring. Any VLAN that warrants fault protection is configured on all ring ports in the ring, and is then assigned to an EAPS domain. On that ring domain, one switch, or node, is designated the *master* node (see Figure 11-2), while all other nodes are designated as *transit* nodes.



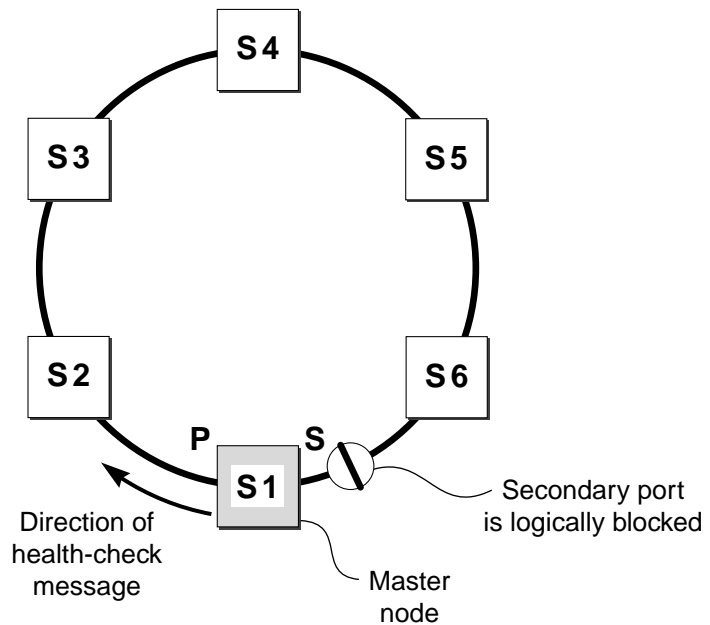
EW_070

Figure 11-1: Gigabit Ethernet fiber EAPS MAN ring

One port of the master node is designated the master node's *primary* port (P) to the ring; another port is designated as the master node's *secondary* port (S) to the ring. In normal operation, the master node blocks the secondary port for all non-control traffic belonging to this EAPS domain, thereby avoiding a loop in the ring, like STP. Layer 2 switching and learning mechanisms operate per existing standards on this ring.



Note: Like the master node, each transit node is also configured with a primary port and a secondary port on the ring, but the primary/secondary port distinction is ignored as long as the node is configured as a transit node.



EW_071

Figure 11-2: EAPS operation

If the ring is complete, the master node logically blocks all data traffic in the transmit and receive directions on the secondary port to prevent a loop. If the master node detects a break in the ring, it unblocks its secondary port and allows data traffic to be transmitted and received through it.

Fault Detection and Recovery

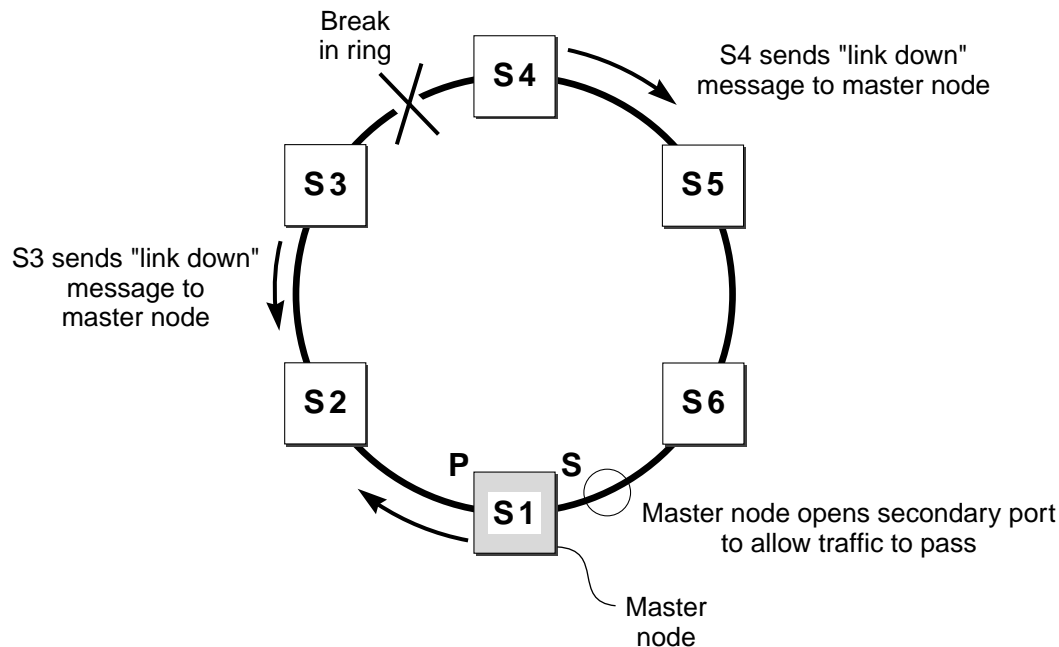
EAPS fault detection on a ring is based on a single *control* VLAN per EAPS domain. This EAPS domain provides protection to one or more data-carrying VLANs called *protected* VLANs.

The control VLAN is used only to send and receive EAPS messages; the protected VLANs carry the actual data traffic. As long as the ring is complete, the EAPS master node blocks the protected VLANs from accessing its secondary port.



Note: The control VLAN is not blocked. Messages sent on the control VLAN must be allowed into the switch for the master node to determine whether the ring is complete.

To avoid loops in the network, the control VLAN must be NOT be configured with an IP address, and ONLY ring ports may be added to the VLAN.



EW_072

Figure 11-3: EAPS fault detection and protection switching

A master node detects a ring fault in either of two ways:

- Polling response
- Trap message sent by a transit node

Polling

The master node transmits a health-check packet on the control VLAN at a user-configurable interval (see Figure 11-2). If the ring is complete, the master node will receive the health-check packet on its secondary port (the control VLAN is not blocked

on the secondary port). When the master node receives the health-check packet, it resets its fail-period timer and continues normal operation.

If the master node does not receive the health-check packet before the fail-period timer expires, it declares a “failed” state and opens its logically blocked secondary port on all the protected VLANs. Now, traffic can flow through the master’s secondary port. The master node also flushes its forwarding database (FDB) and sends a message on the control VLAN to all of its associated transit nodes to flush their forwarding databases as well, so that all of the switches can learn the new paths to Layer 2 end stations on the reconfigured ring topology.

Trap Message Sent by a Transit Node

When any transit node detects a loss of link connectivity on any of its ring ports, it immediately sends a “link down” message on the control VLAN using its good link to the master node.

When the master node receives the “link down” message (see Figure 11-3), it immediately declares a “failed” state and performs the same steps described above: it unblocks its secondary port for access by the protected VLANs, flushes its FDB, and sends a “flush FDB” message to its associated transit nodes.

Restoration Operations

The master node continues sending health-check packets out its primary port even when the master node is operating in the failed state. As long as there is a break in the ring, the fail-period timer of the master node will continue to expire and the master node will remain in the failed state.

When the broken link is restored, the master will receive its health-check packet back on its secondary port, and will once again declare the ring to be complete. It will logically block the protected VLANs on its secondary port, flush its FDB, and send a “flush FDB” message to its associated transit nodes.

During the time between when the transit node detects that the link is operable again and when the master node detects that the ring is complete, the secondary port on the master node is still open and data could start traversing the transit node port that just came up. To prevent the possibility of a such a temporary loop, when the transit node detects that its failed link is up again, it will perform these steps:

- 1 For the port that just came up, put all the protected VLANs traversing that port into a temporary blocked state.
- 2 Remember which port has been temporarily blocked.
- 3 Set the state to Preforwarding.

When the master node receives its health-check packet back on its secondary port, and detects that the ring is once again complete, it sends a message to all its associated transit nodes to flush their forwarding databases.

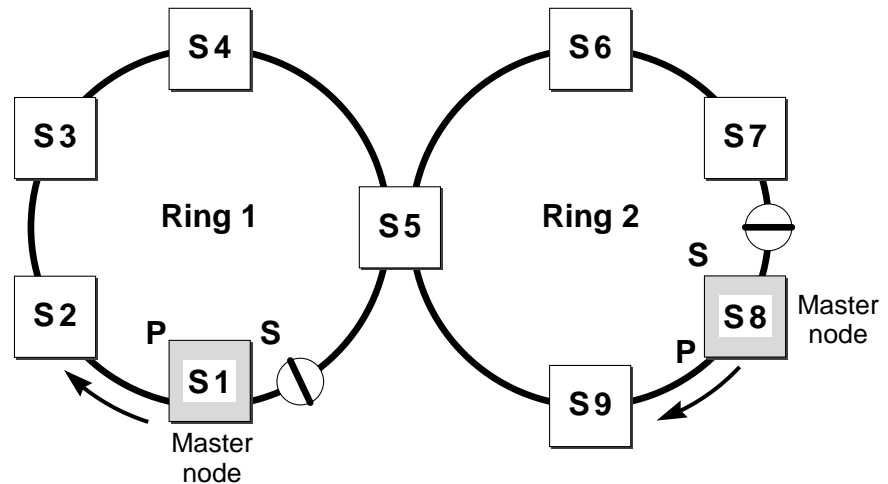
When the transit nodes receive the message to flush their forwarding databases, they perform these steps:

- 1 Flush their forwarding databases on the protected VLANs.
- 2 If the port state is set to Preforwarding, unblock all the previously blocked protected VLANs for the port.

Multiple EAPS Domains Per Switch

To take advantage of the Spatial Reuse technology and broaden the use of the ring's bandwidth, EAPS supports multiple EAPS domains running on the ring at the same time.

So, a single ring might have two EAPS domains running on it. Each EAPS domain would have a different EAPS master node. Each EAPS domain will protect its own set of protected VLANs.



EW_073

Figure 11-4: EAPS data VLAN spanning two rings interconnected by one switch

Figure 11-4 shows how a data VLAN could span two rings interconnected by a common switch—a “figure eight” topology. In this example, there is an EAPS domain with its own control VLAN running on ring 1 and another EAPS domain with its own control VLAN running on ring 2. A data VLAN that spans both rings will be added as a protected VLAN to both EAPS domains. In Figure 11-4, switch S5 will have two instances of EAPS domains running on it: one for each ring.

Creating and Deleting an EAPS Domain

Each EAPS domain is identified by a unique domain name.

To create an EAPS domain, use the following command:

```
create eaps <name>
```

The `name` parameter is a character string of up to 32 characters that identifies the EAPS domain to be created. EAPS domain names and VLAN names must be unique: Do not use the same name string to identify both an EAPS domain and a VLAN.

The following command example creates EAPS domain `eaps_1` on the BlackDiamond switch:

```
create eaps eaps_1
```

To delete an EAPS domain, use the following command:

```
delete eaps <name>
```

The following command example deletes the EAPS domain eaps_1:

```
delete eaps eaps_1
```

Defining the EAPS Mode of the Switch

To configure the EAPS node type of the switch, use the following command:

```
config eaps <name> mode [master | transit]
```

One node on the ring must be configured as the master node for the specified domain; all other nodes on the ring are configured as transit nodes for the same domain.

The following command example identifies this switch as the master node for the domain named eaps_1.

```
config eaps eaps_1 master
```

The following command example identifies this switch as a transit node for the domain named eaps_1.

```
config eaps eaps_1 transit
```

Configuring EAPS Polling Timers

To set the values of the polling timers the master node uses for the EAPS health-check packet that is circulated around the ring for an EAPS domain, use the following command:

```
config eaps <name> [hellotime <seconds> | failtime <seconds>]
```



Note: This command applies only to the master node. If you configure the polling timers for a transit node, they will be ignored. If you later reconfigure that transit node as the master node, the polling timer values will be used as the current values.

Use the `hellotime` keyword and its associated `seconds` parameter to specify the amount of time the master node waits between transmissions of health-check packets on the control VLAN. `seconds` must be greater than 0 when you are configuring a master node. The default value is one second.



Note: Increasing the `hellotime` value keeps the processor from sending and processing too many health-check packets. Increasing the `hellotime` value should not affect the network convergence time, because transit nodes are already sending “link down” notifications.

Use the `failtime` keyword and its associated `seconds` parameter to specify the amount of time the master node waits before declaring a failed state and opens the logically blocked VLANs on the secondary port. `seconds` must be greater than the configured value for `hellotime`. The default value is three seconds.



Note: Increasing the `failtime` value provides more protection against frequent “flapping” between the complete state and the failed state by waiting long enough to receive a health-check packet when the network is congested.



Note: When the master node declares a failed state, it also flushes its forwarding database (FDB) and sends a “flush FDB” message to all the transit switches on the ring by way of the control VLAN. The reason for flushing the FDB is so that the switches can relearn the new directions to reach Layer 2 end stations via the reconfigured topology.

The following command examples configure the `hellotime` value for the EAPS domain “`eaps_1`” to 2 seconds and the `failtime` value to 10 seconds.

```
config eaps eaps_1 hellotime 2
config eaps eaps_1 failtime 10
```

Configuring the Primary and Secondary Ports

Each node on the ring connects to the ring through two ring ports. As part of the protection switching scheme, one port must be configured as the *primary* port; the other must be configured as the *secondary* port.

If the ring is complete, the master node prevents a loop by logically blocking all data traffic in the transmit and receive directions on its secondary port. If the master node subsequently detects a break in the ring, it unblocks its secondary port and allows data traffic to be transmitted and received through it.

To configure a node port as primary or secondary, use the following command:

```
config eaps <name> [primary | secondary] port <port number>
```

The following command example adds port 1 of the module installed in slot 8 of the BlackDiamond switch to the EAPS domain *eaps_1* as the primary port.

```
config eaps eaps_1 primary port 8:1
```

Configuring the EAPS Control VLAN

You must configure one *control* VLAN for each EAPS domain. The control VLAN is used only to send and receive EAPS messages.



Note: A control VLAN cannot belong to more than one EAPS domain.

To configure the EAPS control VLAN for the domain, use the following command:

```
config eaps <name> add control vlan <name>
```



Note: The control VLAN must NOT be configured with an IP address. In addition, only ring ports may be added to this control VLAN. No other ports can be members of this VLAN. Failure to observe these restrictions can result in a loop in the network.



Note: When you configure the VLAN that will act as the control VLAN, that VLAN must be assigned a QoS profile of Qp8, and the ring ports of the control VLAN must be tagged.

By assigning the control VLAN a QoS profile of Qp8 (with the QoS profile `HighHi` priority setting), you ensure that EAPS control VLAN traffic is serviced before any other traffic and that control VLAN messages reach their intended destinations. For example, if the control VLAN is not assigned the highest priority and a broadcast storm occurs in the network, the control VLAN messages might be dropped at intermediate points. Assigning the control VLAN the highest priority prevents dropped control VLAN messages.

Because the QoS profile `HighHi` priority setting by itself should ensure that the control VLAN traffic gets through a congested port first, you should not need to set the QoS profile minimum bandwidth (`minbw`) or maximum bandwidth (`maxbw`) settings. However, if you plan to use QoS (profile priority and bandwidth settings) for other traffic, you might need to set a `minbw` value on Qp8 for control VLAN traffic. Whether you need to do this depends entirely on your configuration.

The following command example adds the control VLAN “keys” to the EAPS domain *eaps_1*.

```
config eaps eaps_1 add control vlan keys
```

Configuring the EAPS Protected VLANs

You must configure one or more *protected* VLANs for each EAPS domain. The protected VLANs are the data-carrying VLANs.



Note: When you configure the VLAN that will act as a protected VLAN, the ring ports of the protected VLAN must be tagged (except in the case of the default VLAN).

To configure an EAPS protected VLAN, use the following command:

```
config eaps <name> add protect vlan <name>
```



Note: As long as the ring is complete, the master node blocks the protected VLANs on its secondary port.

The following command example adds the protected VLAN “orchid” to the EAPS domain “eaps_1.”

```
config eaps eaps_1 add protect vlan orchid
```

Enabling and Disabling an EAPS Domain

To enable a specific EAPS domain, use the following command:

```
enable eaps <name>
```

To disable a specific EAPS domain, use the following command:

```
disable eaps <name>
```

Enabling and Disabling EAPS

To enable the EAPS function for the entire switch, use the following command:

```
enable eaps
```

To disable the EAPS function for the entire switch, use the following command:

```
disable eaps
```

Unconfiguring an EAPS Ring Port

Unconfiguring an EAPS port sets its internal configuration state to INVALID, which causes the port to appear in the Idle state with a port status of Unknown when you use the `show eaps {<name>} detail` command to display the status information about the port.

To unconfigure an EAPS primary or secondary ring port for an EAPS domain, use the following command:

```
unconfig eaps <name> [primary | secondary] port
```

The following command example unconfigures this node's EAPS primary ring port on the domain `eaps_1`:

```
unconfig eaps eaps_1 primary port
```

Displaying EAPS Status Information

To display EAPS status information, use the following command:

```
show eaps {<name>} [detail]
```

If you enter the `show eaps` command without an argument or keyword, the command displays a summary of status information for all configured EAPS domains. You can use the `detail` keyword to display more detailed status information.



Note: The output displayed by this command depends on whether the node is a transit node or a master node. The display for a transit node contains information fields that are not shown for a master node. Also, some state values are different on a transit node than on a master node.

The following example of the `show eaps {<name>} detail` command displays detailed EAPS information for a transit node. Table 11-1 describes the fields and values in the display.

```
* Summit5iTx:39 # show eaps detail
EAPS Enabled: Yes
Number of EAPS instances: 1
EAPSD-Bridge links: 2
```

```

Name: "eaps1" (instance=0)
State: Links-Up          [Running: Yes]
Enabled: Yes      Mode: Transit
Primary port: 13          Port status: Up          Tag status: Tagged
Secondary port: 14        Port status: Up          Tag status: Tagged
Hello Timer interval: 1 sec      Fail Timer interval: 3 sec
Preforwarding Timer interval: 3 sec
Last update: From Master Id 00:E0:2B:81:20:00, Sat Mar 17 17:03:37 2001
Eaps Domain has following Controller Vlan:
  Vlan Name          VID
  "rhsc"              0020
EAPS Domain has following Protected Vlan(s):
  Vlan Name          VID
  "traffic"          1001
Number of Protected Vlans: 1

```

The following example of the `show eaps {<name>}` detail command displays detailed EAPS information for a single EAPS domain named "eaps2" on the master node. Table 11-1 describes significant fields and values in the display.

```

* Baker15:4 # show eaps2 detail
Name: "eaps2" (instance=0)
State: Complete          [Running: Yes]
Enabled: Yes      Mode: Master
Primary port: 14          Port status: Up          Tag status: Tagged
Secondary port: 13        Port status: Blocked     Tag status: Tagged
Hello Timer interval: 1 sec      Fail Timer interval: 3 sec
Eaps Domain has following Controller Vlan:
  Vlan Name          VID
  "rhsc"              0020
EAPS Domain has following Protected Vlan(s):
  Vlan Name          VID
  "blue"              1003
  "traffic"           1001
Number of Protected Vlans: 2

```

Table 11-1: show eaps Display Fields

Field	Description
EAPS Enabled:	<p>Current state of EAPS on this switch:</p> <ul style="list-style-type: none"> ■ Yes—EAPS is enabled on the switch. ■ no—EAPS is not enabled.
Number of EAPS instances:	Number of EAPS domains created. The maximum number of EAPS domains per switch is 64.
EAPSD-Bridge links:	The total number of EAPS bridge links in the system. The maximum count is 4096. Each time a VLAN is added to EAPS, this count increments by 1.
Name:	The configured name for this EAPS domain.
(Instance=)	The instance number is created internally by the system.
State:	<p>On a transit node, the command displays one of the following states:</p> <ul style="list-style-type: none"> ■ Idle—The EAPS domain has been enabled, but the configuration is not complete. ■ Links-Up—This EAPS domain is running, and both its ports are up and in the FORWARDING state. ■ Links-Down—This EAPS domain is running, but one or both of its ports are down. ■ Preforwarding—This EAPS domain is running, and both of its ports are up, but one of them is in a temporary BLOCKED state. <p>On a master node, the command displays one of the following states:</p> <ul style="list-style-type: none"> ■ Idle—The EAPS domain has been enabled, but the configuration is not complete. ■ Complete—The ring is in the COMPLETE state for this EAPS domain. ■ Failed—There is a break in the ring for this EAPS domain.
[Running: ...]	<ul style="list-style-type: none"> ■ Yes—This EAPS domain is running. ■ No—This EAPS domain is not running.
Enabled:	<p>Indicates whether EAPS is enabled on this domain.</p> <ul style="list-style-type: none"> ■ Yes—EAPS is enabled on this domain. ■ no—EAPS is not enabled.

Table 11-1: show eaps Display Fields (continued)

Field	Description
Mode:	The configured EAPS mode for this switch: transit or master.
Primary/Secondary port:	The port numbers assigned as the EAPS primary and secondary ports. On the master node, the port distinction indicates which port is blocked to avoid a loop.
Port status:	<ul style="list-style-type: none"> ■ Unknown—This EAPS domain is not running, so the port status has not yet been determined. ■ Up—The port is up and is forwarding data. ■ Down—The port is down. ■ Blocked—The port is up, but data is blocked from being forwarded.
Tag status:	Tagged status of the control VLAN: <ul style="list-style-type: none"> ■ Tagged—The control VLAN has this port assigned to it, and the port is tagged in the VLAN. ■ Untagged—The control VLAN has this port assigned to it, but the port is untagged in the control VLAN. ■ Undetermined—Either a VLAN has not been added as the control VLAN to this EAPS domain or this port has not been added to the control VLAN.
Hello Timer interval:	The configured value of the timer.
Fail Timer interval:	The configured value of the timer.
Preforwarding Timer interval: ¹	The configured value of the timer. This value is set internally by the EAPS software.
Last update: ¹	Displayed only for transit nodes; indicates the last time the transit node received a hello packet from the master node (identified by its MAC address).
EAPS Domain has ... Controller Vlans:	Lists the assigned name and ID of the control VLAN.
EAPS Domain has ... Protected Vlans: ²	Lists the assigned names and VLAN IDs of all the protected VLANs configured on this EAPS domain.
Number of Protected Vlans:	The count of protected VLANs configured on this EAPS domain.

1. These fields apply only to transit nodes; they are not displayed for a master node.

2. This list is displayed when you use the `detail` keyword in the `show eaps` command.



Status Monitoring and Statistics

This chapter describes the following topics:

- Status Monitoring on page 12-2
- Slot Diagnostics on page 12-2
- Port Statistics on page 12-4
- Port Errors on page 12-5
- Port Monitoring Display Keys on page 12-6
- System Health Checking (BlackDiamond) on page 12-7
- Setting the System Recovery Level on page 12-9
- Logging on page 12-9
- Configuring and Monitoring Flow Statistics on page 12-12
- RMON on page 12-23

Viewing statistics on a regular basis allows you to see how well your network is performing. If you keep simple daily records, you will see trends emerging and notice problems arising before they cause major network faults. In this way, statistics can help you get the best out of your network.

Status Monitoring

The status monitoring facility provides information about the switch. This information may be useful for your technical support representative if you have a problem. ExtremeWare includes many show commands that display information about different switch functions and facilities.



Note: For more information about show commands for a specific ExtremeWare feature, refer to the appropriate chapter in this guide.

Slot Diagnostics

The BlackDiamond switch provides a facility for running normal or extended diagnostics on an I/O module or a Management Switch Fabric Module (MSM) without affecting the operation of the rest of the system.

If you select to run the diagnostic routine on an I/O module, that module is taken off-line while the diagnostic test is performed. Traffic to and from the ports on the module are temporarily unavailable. Once the diagnostic test is completed, the I/O module is reset and becomes operational again.

You can run normal or extended diagnostics on the slave MSM. The normal diagnostic routine is a short series of tests that do not test all the internal Application-Specific Integrated Circuit (ASIC) functions. The extended diagnostic routine tests coverage of all MSM components including the internal ASIC functions. The slave MSM is taken off-line while the diagnostic test is performed. It is reset and operational once the test is completed.

If you want the diagnostic routine to run on the master MSM every time the system boots, use the following command:

```
config diagnostics on
```

To turn off the diagnostic routine, use the following command:

```
config diagnostics off
```

If you want the diagnostic routine to run one time (rather than with each system boot), use the following command:

```
run diagnostics [normal | extended] [<slot> | msm-a | msm-b]
```

where the following is true:

- `normal` — Takes the switch fabric and ports offline, and performs a simple ASIC and packet loopback test on all ports. The test is completed in 30 seconds. CPU and out-of-band management ports are not tested in this mode. As a result, console and Telnet access from the management port is available during this routine.
- `extended` — Takes the switch fabric and ports offline, and performs extensive ASIC, ASIC-memory, and packet loopback tests. Extended diagnostic tests take a maximum of 15 minutes. The CPU is not tested. Console access is available during extended diagnostics.
- `<slot>` — Specifies the slot number of an I/O module. Once the diagnostics test is complete, the system attempts to bring the I/O module back online. This parameter is applicable to the BlackDiamond switch, only.
- `msm-a` | `msm-b` — Specifies the slot letter of an MSM64i. If the master MSM is specified, the diagnostic routine is performed when the system reboots. Both switch fabric and management ports are taken offline during diagnostics. This parameter is applicable to the BlackDiamond switch, only.

Runtime Diagnostics (BlackDiamond)

BlackDiamond runtime diagnostics perform a single test on a single I/O module. All error messages are logged. To perform diagnostics on an I/O module, use the following command:

```
run diagnostics [extended | normal] slot [<slot number> |
msm-a | msm-b]
```

Use the `normal` option when you want a fast (30 – 60 seconds) hardware status check. Use the `extended` option when you want a more thorough test. The `extended` option requires significantly more time to complete, depending on the number of ports on the blade.

You can also execute packet memory scanning for all packet memory associated with the specified I/O slot on a BlackDiamond 6808 or 6816, using the following command:

```
run diagnostics packet-memory slot <slot number>
```

The packet memory diagnostic scans the specified blade to detect single bit-related memory defects and their associated buffer locations. If packet memory defects are detected, their locations are recorded in the blade's EEPROM. Up to eight occurrences can be recorded. If a defect was found during the scan process, the card is reset, the

defective buffer is mapped out from further use, and the I/O card is returned to the operational state. If more than eight defects are detected, or if the defects cannot be mapped out, the card is treated as a failed card and left offline. The card should then be returned through the RMA process with Extreme Networks Technical Support.



Note: Only run extended or packet-memory diagnostics when the switch can be brought off-line. The tests conducted during these diagnostics are extensive and can affect traffic that must be processed by the system CPU.

To view results of the normal or extended diagnostics test, use the following command:

```
show diagnostics
```

To view the results of a packet memory scan, use the following command:

```
show diagnostics packet-memory slot <slot number>
```

Port Statistics

ExtremeWare provides a facility for viewing port statistic information. The summary information lists values for the current counter against each port on each operational module in the system, and it is refreshed approximately every 2 seconds. Values are displayed to nine digits of accuracy.

To view port statistics, use the following command:

```
show ports <portlist> stats
```

The following port statistic information is collected by the switch:

- **Link Status** — The current status of the link. Options are:
 - Ready (the port is ready to accept a link).
 - Active (the link is present at this port).
 - Chassis (the link is connected to a Summit Virtual Chassis).
- **Transmitted Packet Count (Tx Pkt Count)** — The number of packets that have been successfully transmitted by the port.
- **Transmitted Byte Count (Tx Byte Count)** — The total number of data bytes successfully transmitted by the port.
- **Received Packet Count (Rx Pkt Count)** — The total number of good packets that have been received by the port.

- **Received Byte Count (RX Byte Count)** — The total number of bytes that were received by the port, including bad or lost frames. This number includes bytes contained in the Frame Check Sequence (FCS), but excludes bytes in the preamble.
- **Received Broadcast (RX Bcast)** — The total number of frames received by the port that are addressed to a broadcast address.
- **Received Multicast (RX Mcast)** — The total number of frames received by the port that are addressed to a multicast address.

Port Errors

The switch keeps track of errors for each port.

To view port transmit errors, use the following command:

```
show ports <portlist> txerrors
```

The following port transmit error information is collected by the system:

- **Port Number**
- **Link Status** — The current status of the link. Options are:
 - Ready (the port is ready to accept a link).
 - Active (the link is present at this port).
- **Transmit Collisions (TX Coll)** — The total number of collisions seen by the port, regardless of whether a device connected to the port participated in any of the collisions.
- **Transmit Late Collisions (TX Late Coll)** — The total number of collisions that have occurred after the port's transmit window has expired.
- **Transmit Deferred Frames (TX Deferred)** — The total number of frames that were transmitted by the port after the first transmission attempt was deferred by other network traffic.
- **Transmit Errored Frames (TX Error)** — The total number of frames that were not completely transmitted by the port because of network errors (such as late collisions or excessive collisions).
- **Transmit Parity Frames (TX Parity)** — The bit summation has a parity mismatch.

To view port receive errors, use the following command:

```
show ports <portlist> rxerrors
```

The following port receive error information is collected by the switch:

- **Receive Bad CRC Frames (RX CRC)** — The total number of frames received by the port that were of the correct length, but contained a bad FCS value.
- **Receive Oversize Frames (RX Over)** — The total number of good frames received by the port greater than the supported maximum length of 1,522 bytes. For products that use the “i” chipset, ports with jumbo frames enabled do not increment this counter.
- **Receive Undersize Frames (RX Under)** — The total number of frames received by the port that were less than 64 bytes long.
- **Receive Fragmented Frames (RX Frag)** — The total number of frames received by the port were of incorrect length and contained a bad FCS value.
- **Receive Jabber Frames (RX Jab)** — The total number of frames received by the port that was of greater than the support maximum length and had a Cyclic Redundancy Check (CRC) error.
- **Receive Alignment Errors (RX Align)** — The total number of frames received by the port that occurs if a frame has a CRC error and does not contain an integral number of octets.
- **Receive Frames Lost (RX Lost)** — The total number of frames received by the port that were lost because of buffer overflow in the switch.

Port Monitoring Display Keys

Table 12-1 describes the keys used to control the displays that appear when you issue any of the `show port` commands.

Table 12-1: Port Monitoring Display Keys

Key(s)	Description
U	Displays the previous page of ports.
D	Displays the next page of ports.
[Esc] or [Return]	Exits from the screen.
0	Clears all counters.

Table 12-1: Port Monitoring Display Keys (continued)

Key(s)	Description
[Space]	<p>Cycles through the following screens:</p> <ul style="list-style-type: none"> ■ Packets per second ■ Bytes per second ■ Percentage of bandwidth <p>Available using the <code>show port utilization</code> command only.</p>

System Health Checking (BlackDiamond)

The system health checker tests the backplane, the CPU, and I/O modules on the BlackDiamond switch by periodically forwarding packets and checking for the validity of the forwarded packets.

The system health checker can be configured to handle a failure as an error condition, logging the problem to the syslog, or it can attempt auto-recovery of the module that generated the errors.

The alarm-level and auto-recovery options are mutually exclusive.

To enable the system health checker, use the following command:

```
enable sys-health-check
```

To disable the system health checker, use the following command:

```
disable sys-health-check
```

To configure the switch to respond to a failed health check based on an alarm-level, use the following command:

```
config sys-health-check alarm-level [card-down | default | log |
system-down | traps]
```

This command provides the following options:

- `card-down` — Posts a CRIT message to the log, sends a trap, and turns off the module.
- `log` — Posts a CRIT message to the log.

- `system-down` — Posts a CRIT message to the log, sends a trap, and turns off the system.
- `traps` — Posts a CRIT message to the log and sends a trap.

The default option is `log`.

To configure the switch for auto-recovery, use the following command:

```
config sys-health-check auto-recovery <number of tries>
```

The auto-recovery option is used to configure the number of times the system health checker attempts to automatically reset a faulty module and bring it online. If the module continues to fail more than the configured number of attempts, the system health checker sets the module to card-down.

When auto-recovery is configured, the occurrence of three consecutive checksum errors will cause a packet memory (PM) defect detection program to be run against the I/O module. Checksum errors can include internal and external MAC port parity errors, EDP checksum errors, and CPU packet or diagnostic packet checksum errors. If defects are detected, the card is taken offline, the memory defect information is recorded in the card EEPROM, the defective buffer is mapped out of further use, and the card is returned to operational state. A maximum of 8 defects can be stored in the EEPROM.

After the PM defect detection and mapping process has been run, a card is considered failed and is taken offline in the following circumstances:

- More than eight defects are detected.
- Three consecutive checksum errors were detected by the health checker, but no new PM defects were found by the PM defect detection process.
- After defects were detected and mapped out, the same checksum errors are again detected by the system health checker.

The auto-recovery repetition value is ignored in these cases. In any of these cases, please contact Extreme Technical Support.

To view the status of the system health checker, use the following command:

```
show diagnostics
```

Setting the System Recovery Level

You can configure the system to automatically reboot after a software task exception, using the following command:

```
config sys-recovery-level [none | critical | all]
```

Where the following is true:

- `none` — Configures the level to no recovery.
- `critical` — Configures ExtremeWare to log an error into the syslog and automatically reboot the system after a critical task exception.
- `all` — Configures ExtremeWare to log an error into the syslog and automatically reboot the system after any task exception.

The default setting is `none`.

Logging

The switch log tracks all configuration and fault information pertaining to the device. Each entry in the log contains the following information:

- **Timestamp** — The timestamp records the month and day of the event, along with the time (hours, minutes, and seconds) in the form HH:MM:SS. If the event was caused by a user, the user name is also provided.
- **Fault level** — Table 12-2 describes the three levels of importance that the system can assign to a fault.

Table 12-2: Fault Levels Assigned by the Switch

Level	Description
Critical	A desired switch function is inoperable. The switch may need to be reset.
Warning	A noncritical error that may lead to a function failure.
Informational	Actions and events that are consistent with expected behavior.
Debug	Information that is useful when performing detailed troubleshooting procedures.

By default, log entries that are assigned a critical or warning level remain in the log after a switch reboot. Issuing a clear log command does not remove these static entries. To remove log entries of all levels (including warning or critical), use the following command:

```
clear log static
```

- **Subsystem** — The subsystem refers to the specific functional area to which the error refers. Table 12-3 describes the subsystems.

Table 12-3: Fault Log Subsystems

Subsystem	Description
Syst	General system-related information. Examples include memory, power supply, security violations, fan failure, overheat condition, and configuration mode.
STP	STP information. Examples include an STP state change.
Brdg	Bridge-related functionality. Examples include low table space and queue overflow.
SNMP	SNMP information. Examples include community string violations.
Telnet	Information related to Telnet login and configuration performed by way of a Telnet session.
VLAN	VLAN-related configuration information.
Port	Port management-related configuration. Examples include port statistics and errors.

- **Message** — The message contains the log information with text that is specific to the problem.

Local Logging

The switch maintains 1,000 messages in its internal log. You can display a snapshot of the log at any time by using the following command:

```
show log {<severity>}
```

where the following is true:

- **severity** — Filters the log to display message with the selected severity or higher (more critical). Severities include (in order) emergency, critical, alert, error, warning, notice, info, and debug. If not specified, all messages are displayed.

Real-Time Display

In addition to viewing a snapshot of the log, you can configure the system to maintain a running real-time display of log messages on the console. To turn on the log display, use the following command:

```
enable log display
```

To configure the log display, use the following command:

```
config log display {<severity>}
```

If `severity` is not specified, only messages of critical priority are displayed.

If you enable the log display on a terminal connected to the console port, your settings will remain in effect even after your console session is ended (unless you explicitly disable the log display). The log display setting will be stored with the configuration upon a `save`.

You can enable a log display from a Telnet session using the `enable log display` command, but the log display will appear only on the console.

Remote Logging

In addition to maintaining an internal log, the switch supports remote logging by way of the UNIX syslog host facility. To enable remote logging, follow these steps:

- 1 Configure the syslog host to accept and log messages.
- 2 Enable remote logging by using the following command:

```
enable syslog
```

- 3 Configure remote logging by using the following command:

```
config syslog {add} <ipaddress> <facility> {<severity>}
```

Specify the following:

- `ipaddress` — The IP address of the syslog host.
- `facility` — The syslog facility level for local use. Options include `local0` through `local7`.
- `severity` — Filters the log to display message with the selected severity or higher (more critical). Severities include (in order) emergency, critical, alert, error, warning, notice, info, and debug. If not specified, all messages are sent to the syslog host.



Note: Refer to your UNIX documentation for more information about the syslog host facility.

Logging Configuration Changes

ExtremeWare allows you to record all configuration changes and their sources that are made using the CLI by way of Telnet or the local console. The changes are logged to the system log. Each log entry includes the user account name that performed the change and the source IP address of the client (if Telnet was used). Configuration logging applies only to commands that result in a configuration change. To enable configuration logging, use the following command:

```
enable cli-config-logging
```

To disable configuration logging, use the following command:

```
disable cli-config-logging
```

CLI configuration logging is enabled by default.

Configuring and Monitoring Flow Statistics



Note: This section describes the process of configuring and monitoring flow statistics on Ethernet links. If you plan to configure and monitor flow statistics on PoS links, see the Packet Over SONET Installation and User Guide for more information.

The broad growth in Internet and intranet usage has brought with it an increased demand for network bandwidth and performance that is based on predictable quality of service and security. This movement is paralleled by the related need for measurement technology that makes it possible to gather, analyze, and manipulate information about network and application use. NetFlow, originally developed by Cisco, provides a way for a switch to capture and export traffic classification or precedence information as data traverses, or flows, across portions of a network.

A network flow is defined as a unidirectional sequence of packets between a particular source device and destination device that share the same protocol and transport-layer information. Flows are defined by the following fields: source IP address, destination IP address, source port, destination port, and protocol type. Per-flow statistics are useful for many management purposes, including:

- Accounting and billing
- Network capacity planning and trend analysis
- Network monitoring
- Workload characterization
- User profiling
- Data warehousing and mining

Flow Statistics Background Information

Per-flow statistics are exported in the NetFlow Version 1 record format described in Table 12-4. NetFlow records are unidirectional in nature, which implies that two flow records are maintained for a typical TCP connection: one record for flow in the ingress direction; a second for the flow in the egress direction. Also, records are maintained only for TCP and UDP flows.

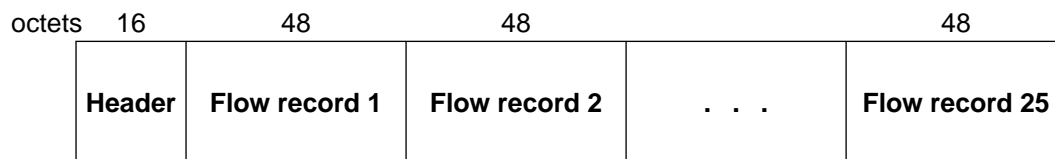
Table 12-4: NetFlow Version 1 Record Format

Field Name	Octets	Field Description
<i>srcaddr</i>	4	Source IP address.
<i>dstaddr</i>	4	Destination IP address.
<i>nexthop</i>	4	IP address of next-hop router; set to zero for per-flow statistics; set to xFFFF for filter-based aggregated statistics.
<i>input</i>	2	(Not supported.) Input interface index.
<i>output</i>	2	(Not supported.) Output interface index.
<i>dPkts</i>	4	Number of packets sent in this flow.
<i>dOctets</i>	4	(Not Supported.) Number of octets sent in this flow.
<i>First</i>	4	(Not supported.) SysUptime when flow record was created.
<i>Last</i>	4	(Not supported.) SysUptime at most-recent, or last packet of flow.
<i>srcport</i>	2	Source port number, valid only for TCP and UDP flows.
<i>dstport</i>	2	Destination port number, valid only for TCP and UDP flows.
<i>pad</i>	2	Unused field.
<i>prot</i>	1	Number identifying the IP protocol; for example, 6=TCP and 17=UDP.

Table 12-4: NetFlow Version 1 Record Format (continued)

Field Name	Octets	Field Description
<i>tos</i>	1	(Not supported.) IP Type-of-Service (TOS) field value from initial packet that caused this flow record to be created.
<i>tcp_flags</i>	1	(Not supported.) Cumulative OR of TCP flags field, valid only when <i>prot</i> =6.
<i>pad</i>	7	Unused field.

Flow records are grouped together into UDP datagrams for export to a flow-collector device. A NetFlow Version 1 export datagram can contain up to 25 flow records. Figure 12-1 shows the format of the export datagram; Table 12-5 describes the export datagram header.



EW_086

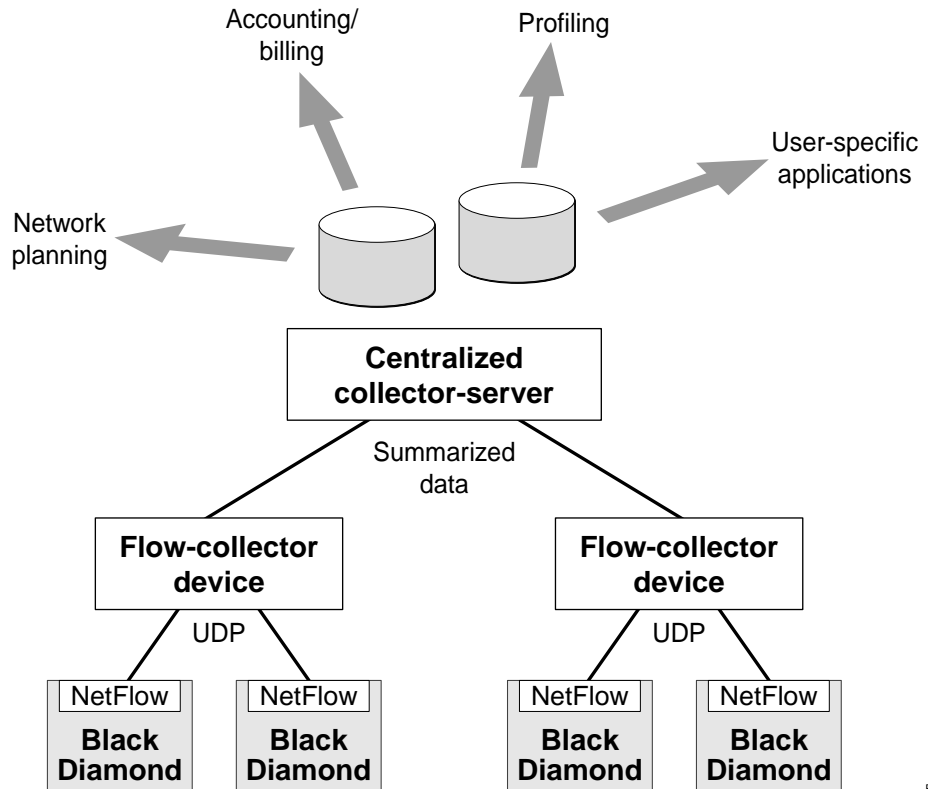
Figure 12-1: Format of NetFlow export datagram

Table 12-5: Format of NetFlow Version 1 Export Datagram Header

Field Name	Octets	Field Description
<i>version</i>	2	Header version=1.
<i>count</i>	2	Number of flow records in datagram.
<i>SysUptime</i>	4	Current time in milliseconds since the switch booted.
<i>unix_secs</i>	4	(Not Supported.) Current count of seconds since 0000 UTC 1970.
<i>unix_nsecs</i>	4	(Not Supported.) Current count of residual nanoseconds since 0000 UTV 1970.

The IP addresses (or host names) and UDP port numbers of the available flow collectors can be configured on a per-switch basis. The flow collection architecture example in Figure 12-2 illustrates how multiple BlackDiamond switches might export flow records to flow-collector devices that, in turn, feed records into a central collector-server. Other

flow-collector architectures are also possible. For example, each switch port configured for flow switching might export statistics to a dedicated flow-collector device.



PoS_024

Figure 12-2: NetFlow Collection Architecture Example

The ExtremeWare NetFlow implementation also enables a single port to distribute statistics across multiple groups of flow-collector devices. This NetFlow distribution feature enables a scalable collection architecture that is able to accommodate high volumes of exported data. The NetFlow distribution feature is enabled by configuring *export distribution groups* that contain the addresses of multiple flow-collector devices. The feature uses a distribution algorithm that ensures all of the records for a given flow are exported to the same collector. The algorithm also ensures that the flow records of the ingress direction of a TCP or UDP connection are exported to the same collector. (For Ethernet applications, only ingress traffic is monitored on Ethernet ports.) For example, multiple filters can be assigned to a set of ports for the same group. The flow

records that match the filters are then sent to one of the flow collector devices in that group. You can also establish redundancy by configuring multiple flow collector devices per group so that data is still collected as long as there is one working flow collector device in that group.

To implement flow-collector devices, you can choose from commercial software products and public-domain software packages.

Collection Port and Filtering Options

By default, each Ethernet port configured for flow switching maintains statistics for all the flows traversing the link in the ingress direction.

Generalized filtering options exist that enable you to configure an Ethernet port to maintain statistics selectively for only those flows that match a specified filter. For example, to monitor aggregated flow records on Ethernet ports, you could configure an aggregation filter that specifies a range of IP addresses or ports.

Up to eight filters are supported for each Ethernet port, with a total of 128 filters possible per each I/O module. The filters consist of a *{value, mask}* pair for each of the following flow components: destination IP address, source IP address, destination port, source port, and protocol. Conceptually, the filters work by logically ANDing the contents of each of these five components of a forwarded flow with the associated *masks* from the first filter. Statistics are maintained if the results of the AND operations match the configured filter values for all fields of the five flow components. If there is not a match on all fields of the five components, then the operation is repeated for the second filter, and so on. If there is no match for any of the filters, then statistics are not maintained for the flow.

Collection Architecture Scalability and Reliability

By supporting statistics distribution across groups of flow-collector devices, the NetFlow distribution function enables a scalable collection architecture that is able to accommodate high volumes of exported data. The function also includes a health-check feature that significantly improves the reliability of the collection architecture. The health-checker ensures that only responsive flow-collector devices are included in the effective export distribution lists.

Up to 32 export distribution groups can be configured on a Black Diamond 6800 series switch. Each of these groups can contain the addresses of up to eight flow-collector devices. A particular export group can then be specified for each filter, which provides a high-degree of flexibility.

A filter-based aggregation capability is also offered to further enhance scalability. Each filter can be configured to be either a *per-flow filter* or an *aggregation filter*. When a flow matches a filter that is configured as an aggregation, normal per-flow statistics are not maintained for the flow. Instead, a single set of statistics is maintained for all the flows that match the aggregation filter, which can substantially reduce the volume of exported data.

Aggregated flow statistics are also exported in the NetFlow Version 1 format. The *srcaddr*, *dstaddr*, *srcport*, *dstport*, and *prot* fields of an aggregated flow record are set to the corresponding value components of the associated filter specification.

Export Criteria

For Ethernet ports, flow records are exported on an age basis. If the age of the flow is greater than a configurable time, the record is exported.

An Ethernet port configured for flow switching transmits a NetFlow Export Datagram when 25 flow records are ready for export, or when at least one flow has been awaiting export for one second.

An Ethernet port configured for capturing flows transmits NetFlow export datagrams when the configured time-out expires and exports the data collected by the flow filters configured on that port. As the NetFlow on Ethernet links is modeled as port-based, individual ports maintain their configured time-outs and export the flows collected by the configured flow filters on the expiry of flow export time-out.

Enabling and Disabling the Flow Statistics Feature on a Switch

To enable the flow statistics feature on a switch, use the following command:

```
enable flowstats
```

The flow statistics feature is disabled by default.

To disable the flow statistics feature on a switch, use the following command:

```
disable flowstats
```

Enabling and Disabling Flow Statistics on a Port

To enable the flow statistics function on the specified port, use the following command:

```
enable flowstats ports <portlist>
```

The flow statistics function is disabled by default.

To disable the flow statistics function on the specified port, use the following command:

```
disable flowstats ports <portlist>
```

Configuring the Export Destination

A single port can distribute statistics across multiple groups of flow-collector devices. This NetFlow distribution capability makes it possible to create a collection architecture that scales to accommodate high volumes of exported data. It also offers a health-checking function that improves the reliability of the collection architecture by ensuring that only responsive flow-collector devices are included in active export distribution lists. The distribution algorithm also ensures that all the ingress flow records for a given flow are exported to the same collector.

NetFlow distribution is enabled by configuring export distribution groups that identify the addresses of multiple flow-collector devices. You can configure up to 32 export distribution groups on a BlackDiamond 6800 series switch, and each group can contain as many as eight flow-collector devices.

To configure the export groups and flow-collector devices to which NetFlow datagrams are exported, use the following command:

```
config flowstats export <group#> [add | delete] [<ipaddress> | <hostname>]  
    port <udp_port>
```

The `group#` parameter is an integer in the range from 1 through 32 that identifies the specific group for which the destination is being configured.

You can use the `add` and `delete` keywords to add or delete flow-collector destinations.

To export NetFlow datagrams to a group, you must configure at least one flow-collector destination. By default, no flow-collector destinations are configured. To configure a flow-collector destination, use either an IP address and UDP port number pair or a hostname and UDP port number pair to identify the flow-collector device to which NetFlow export datagrams are to be transmitted. You can configure up to eight flow-collector destinations for each group. When multiple flow-collectors are configured as members of the same group, the exported NetFlow datagrams are distributed across the available destinations.

Configuring the Source IP Address

To configure the IP address that is to be used as the source IP address for NetFlow datagrams to be exported, use the following command:

```
config flowstats source <ipaddress>
```

By default, flow records are exported with the VLAN interface address that has a route to the configured flow-collector device. Depending on how it is configured, a flow-collector device can use the source IP address of received NetFlow datagrams to identify the switch that sent the information.

The following command example specifies that the IP address 192.168.100.1 is to be used as the source IP address for exported NetFlow datagrams.

```
config flowstats source 192.168.100.1
```

Configuring Flow Record Time-out

Flow records are exported on an age basis. If the age of the flow record is greater than the configured time-out, the record is exported.

To configure the time-out value for flow records on the specified port, use the following command:

```
config flowstats timeout <minutes> ports [<portlist> | any]
```

The time-out value is the number of minutes to use in deciding when to export flow records. The default time-out is 5 minutes.

The following command example specifies a 10-minute time-out for exported NetFlow datagrams on port 1 of the Ethernet module installed in slot 8 of the BlackDiamond switch.

```
config flowstats timeout 10 ports 8:1
```

Configuring a Flow Record Filter

You can configure an Ethernet port to maintain statistics selectively for only those flows that match a specified filter. Each Ethernet port supports eight filters for ingress flows.

To configure a flow record filter for the specified Ethernet port, use the following command:

```
config flowstats filter-ingress <filter#> export <group#> ports <portlist>
  {aggregation} [<filterspec> | match-all-flows]
```

where:

filter#	The <code>filter#</code> parameter is an integer in the range from 1 to 8 that identifies the filter being defined.
<group#>	Specifies the group number that identifies the set of flow collector devices to which records for flows matching the filter are to be exported.
aggregation	To reduce the volume of exported data, use this optional keyword to maintain a single set of statistics for all the flows that match the specified filter.
filterspec	<p>Specifies a set of five parameters (four are value/mask pairs) that define the criteria by which a flow is evaluated to determine if it should be exported. The parameters are:</p> <pre>destination [<ipaddress/ipaddress_mask> any] ip-port [<portlist>/port_mask> any] source [<ipaddress/ipaddress_mask> any] ip-port [<portlist>/port_mask> any] [ip tcp udp]</pre> <p>All five specifications must be included in the order specified.</p> <p>The range for port/port_mask is calculated using the following formula: (minport = port, maxport = 2^(32-port_mask)-1).</p> <p>Conceptually, the filters work by ANDing the contents of each of the five components of a forwarded flow with the associated masks from the first defined filter (filter #1). Statistics are maintained if the results of the AND operations match the configured filter values for all fields of the sequence. If there is no match, then the operation is repeated for filter #2, and so on. If there is no match for any of the filters, then statistics are not maintained for the flow. Filters for any or all of the sequence components can be configured with a single command.</p>
match-all-flows	Specifies that the filter should match any flow.

The following command example configures a filter to collect aggregate statistics for all traffic flowing through ports 1-8 from the 192.170.0.0/16 subnet to the 192.171.132.0/24 subnet:

```
config flowstats filter-ingress 2 ports 1-8 export 1 aggregation
  destination 192.171.132.0/24 ip-port 0/0 source 192.170.0.0/16
  ip-port 0/0 ip
```

Likewise, the following command example configures a filter to collect aggregate statistics for all ingress traffic flowing from the 192.171.0.0/16 subnet to the 192.170.0.0/16 subnet and export the flows to group 3 for ports 6:1, 7:9, and 8:42

```
config flowstats filter-ingress 2 ports 6:1,7:9,8:42 export 3
  aggregation destination 192.170.0.0/16 ip-port 0/0
  source 192.171.0.0/16 ip-port 0/0 ip
```

Finally, the following command configures filter 3 to collect statistics on any flows for ports 4-32 that did not match the filters defined in the two previous commands:

```
config flowstats filter-ingress 3 ports 4-32 export 1 aggregation
  match-all-flows
```

Enabling and Disabling a Flow Record Filter

To enable a specified flow record filter for the specified Ethernet port, use the following command:

```
enable flowstats filter <filter#> ports <portlist>
```

By default, all filters are enabled after they are configured.

To disable a specified flow record filter for the specified Ethernet port, use the following command:

```
disable flowstats filter <filter#> ports <portlist>
```

where:

filter# The **filter#** parameter is an integer in the range from 1 to 8 that identifies the filter that is being enabled or disabled.

The following command example enables filter #2 on port 1 of the Ethernet module installed in slot 8 of the BlackDiamond switch.

```
enable flowstats filter 2 ports 8:1
```

The following command example disables filter #2 on port 1 of the Ethernet module installed in slot 8 of the BlackDiamond switch.

```
disable flowstats filter 2 ports 8:1
```

Enabling and Disabling Flow Statistics Ping-Check

To enable the flow statistics ping-check function for a specified group of collector devices, use the following command:

```
enable flowstats ping-check <group#>
```

The ping-check function is enabled by default.

When the ping-check function is enabled, each of the flow collector devices is pinged periodically to check its network connectivity. If a flow collector device is repetitively unresponsive, it is temporarily removed from the export distribution list for that group. The flow collector device will be returned to the export distribution list automatically when subsequent ping checks are consistently successful.

The following command example enables the ping-check function for export group 2.

```
enable flowstats ping-check 2
```

To disable the flow statistics ping-check function for a specified group of collector devices, use the following command:

```
disable flowstats ping-check <group#>
```

The following command example disables the ping-check function for export group 2.

```
disable flowstats ping-check 2
```

Unconfiguring Flow Statistics

To reset the flow statistics configuration parameters for a specified Ethernet port to their default values, use the following command:

```
unconfig flowstats ports <portlist>
```



Note: This command does not affect the enabled or disabled status of flow statistics collection, nor does it affect the configured export destinations.

The following command example resets the flow statistics configuration parameters for port 1 of the module installed in slot 8 of the BlackDiamond switch to their default values.

```
unconfig flowstats ports 8:1
```

Displaying Flow Statistics Status Information

To display status information for the flow statistics function, use the following command:

```
show flowstats {detail | group <group#> | ports <portlist>}
```


where:

detail	Use this optional keyword to display detailed NetFlow configuration information.
group#	Use this optional parameter with the <code>group</code> keyword to display status information for a specific export group.
portlist	Use this optional parameter to specify one or more ports or slots and ports for which status information is to be displayed.

If you enter the `show flowstats` command with none of the optional keywords or parameters, the command displays a summary of status information for all ports.

The summary status display for a port shows the values for all flow statistics configuration parameters for the port.

The summary status display for an export group includes the following information:

- Values for all configuration parameters
- Status of each export destination device

The detailed status display for an export group includes the summary information, plus the following management information:

- Counts of the number of times each flow collector destination has been taken out of service due to health-check (ping check) failures
- The source IP address configuration information

RMON

Using the Remote Monitoring (RMON) capabilities of the switch allows network administrators to improve system efficiency and reduce the load on the network.

The following sections explain more about the RMON concept and the RMON features supported by the switch.



Note: You can only use the RMON features of the system if you have an RMON management application, and have enabled RMON on the switch.

About RMON

RMON is the common abbreviation for the Remote Monitoring Management Information Base (MIB) system defined by the Internet Engineering Task Force (IETF) documents RFC 1271 and RFC 1757, which allows you to monitor LANs remotely.

A typical RMON setup consists of the following two components:

- **RMON probe** — An intelligent, remotely controlled device or software agent that continually collects statistics about a LAN segment or VLAN. The probe transfers the information to a management workstation on request, or when a predefined threshold is crossed.
- **Management workstation** — Communicates with the RMON probe and collects the statistics from it. The workstation does not have to be on the same network as the probe, and can manage the probe by in-band or out-of-band connections.

RMON Features of the Switch

The IETF defines nine groups of Ethernet RMON statistics. The switch supports the following four of these groups:

- Statistics
- History
- Alarms
- Events

This section describes these groups and discusses how they can be used.

Statistics

The RMON Ethernet Statistics group provides traffic and error statistics showing packets, bytes, broadcasts, multicasts, and errors on a LAN segment or VLAN.

Information from the Statistics group is used to detect changes in traffic and error patterns in critical areas of the network.

History

The History group provides historical views of network performance by taking periodic samples of the counters supplied by the Statistics group. The group features

user-defined sample intervals and bucket counters for complete customization of trend analysis.

The group is useful for analysis of traffic patterns and trends on a LAN segment or VLAN, and to establish baseline information indicating normal operating parameters.

Alarms

The Alarms group provides a versatile, general mechanism for setting threshold and sampling intervals to generate events on any RMON variable. Both rising and falling thresholds are supported, and thresholds can be on the absolute value of a variable or its delta value. In addition, alarm thresholds can be autocalibrated or set manually.

Alarms inform you of a network performance problem and can trigger automated action responses through the Events group.

Events

The Events group creates entries in an event log and/or sends SNMP traps to the management workstation. An event is triggered by an RMON alarm. The action taken can be configured to ignore it, to log the event, to send an SNMP trap to the receivers listed in the trap receiver table, or to both log and send a trap. The RMON traps are defined in RFC 1757 for rising and falling thresholds.

Effective use of the Events group saves you time. Rather than having to watch real-time graphs for important occurrences, you can depend on the Event group for notification. Through the SNMP traps, events can trigger other actions, which provides a mechanism for an automated response to certain occurrences.

Configuring RMON

RMON requires one probe per LAN segment, and standalone RMON probes traditionally have been expensive. Therefore, Extreme's approach has been to build an inexpensive RMON probe into the agent of each system. This allows RMON to be widely deployed around the network without costing more than traditional network management. The switch accurately maintains RMON statistics at the maximum line rate of all of its ports.

For example, statistics can be related to individual ports. Also, because a probe must be able to see all traffic, a stand-alone probe must be attached to a nonsecure port. Implementing RMON in the switch means that all ports can have security features enabled.

To enable or disable the collection of RMON statistics on the switch, use the following command:

```
[enable | disable] rmon
```

By default, RMON is disabled. However, even in the disabled state, the switch response to RMON queries and sets for alarms and events. By enabling RMON, the switch begins the processes necessary for collecting switch statistics.

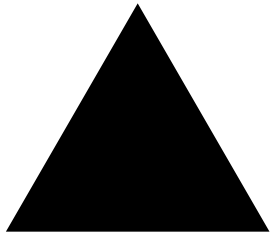
Event Actions

The actions that you can define for each alarm are shown in Table 12-6.

Table 12-6: Event Actions

Action	High Threshold
No action	
Notify only	Send trap to all trap receivers.
Notify and log	Send trap; place entry in RMON log.

To be notified of events using SNMP traps, you must configure one or more trap receivers, as described in Chapter 3.



Part 2:

Using Routing Protocols

13

Spanning Tree Protocol (STP)

This chapter covers the following topics:

- Overview of the Spanning Tree Protocol on page 13-2
- Spanning Tree Domains on page 13-2
- STP Configurations on page 13-4
- Per-VLAN Spanning Tree on page 13-11
- STP Rules and Restrictions on page 13-12
- Configuring Basic STP on the Switch on page 13-13
- Displaying STP Settings on page 13-16

Using the Spanning Tree Protocol (STP) functionality of the switch makes your network more fault tolerant. The following sections explain more about STP and the STP features supported by ExtremeWare.



Note: STP is a part of the 802.1D bridge specification defined by the IEEE Computer Society. To explain STP in terms used by the 802.1D specification, the switch will be referred to as a bridge.

Overview of the Spanning Tree Protocol

STP is a bridge-based mechanism for providing fault tolerance on networks. STP allows you to implement parallel paths for network traffic, and ensure that:

- Redundant paths are disabled when the main paths are operational.
- Redundant paths are enabled if the main path fails.

Spanning Tree Domains

The switch can be partitioned into multiple virtual bridges. Each virtual bridge can run an independent Spanning Tree instance. Each Spanning Tree instance is called a *Spanning Tree Domain* (STPD). Each STPD has its own root bridge and active path. After an STPD is created, one or more VLANs can be assigned to it.

A port can belong to multiple STPDs. In addition, a VLAN can span multiple STPDs.

The key points to remember when configuring VLANs and STP are:

- Each VLAN forms an independent broadcast domain.
- STP blocks paths to create a loop-free environment.
- When STP blocks a path, no data can be transmitted or received on the blocked port.
- Within any given STPD, all VLANs belonging to it use the same spanning tree.

If you delete a STPD, the VLANs that were members of that STPD are also deleted. You must remove all VLANs associated with the STP before deleting the STPD.

Defaults

The default device configuration contains a single STPD called *s0*. The default VLAN is a member of STPD *s0*.

All STP parameters default to the IEEE 802.1D values, as appropriate.

Port Modes

An STP port has three modes of operation:

- 802.1D mode

This mode is used for backward compatibility with previous STP versions and for compatibility with third-party switches using IEEE standard 802.1D. BPDUs are sent untagged in 1D mode. Because of this, on any given physical interface there can be only *one* STPD running in 1D mode.

- Extreme Multiple Instance Spanning Tree Protocol (EMISTP) mode

EMISTP mode is an extension of STP that allows a physical port to belong to multiple STPDs by assigning the port to multiple VLANs. EMISTP adds significant flexibility to STP network design. BPDUs are sent with an 802.1Q tag having an STPD instance Identifier (StpdID) in the VLANid field.

- PVST+ mode

This mode implements PVST+ in compatibility with third-party switches running this version of STP. The STPDs running in this mode have a one-to-one relationship with VLANs, and send and process packets in PVST+ format.

These port modes are for STP ports, not for physical ports. When a physical port belongs to multiple STPDs, it is associated with multiple STP ports. It is possible for the physical port to run in different modes for different domains to which it belongs.

STPD Identifier

An StpdID is used to identify each STP domain. You assign the StpdID when configuring the domain, and that VLAN cannot belong to another STPD.

An StpdID must be identical to the VLANid of one of the member VLANs in that STP domain.



Note: If an STPD contains at least one port not in 1D mode, the STPD must be configured with an StpdID.

STPD BPDU Tunneling

You can configure ExtremeWare to allow a BPDU to traverse a VLAN without being processed by STP, even if STP is enabled on the port. This is known as BPDU *tunneling*.

To enable and disable BPDU tunneling on a VLAN, use the following command:

```
[enable | disable] ignore-bpdu vlan <name>
```

If you have a known topology and have switches outside of your network within your STPD, use this feature to keep the root bridge within your network.

Rapid Root Failover

ExtremeWare supports rapid root failover for faster STP failover recovery times. The default setting is disabled.

To configure rapid root failover, use the following command:

```
[enable | disable] stpd <stpd_name> rapid-root-failover
```

To display the configuration, use the following command:

```
show stpd <stpd>
```

STP Configurations

When you assign VLANs to an STPD, pay careful attention to the STP configuration and its effect on the forwarding of VLAN traffic.

This section describes three types of STP configurations:

- Basic STP
- Multiple STPDs on a single port (EMISTP)
- A VLAN that spans multiple STPDs

Basic STP Configuration

This section describes a basic, 802.1d STP configuration. Figure 13-1 illustrates a network that uses VLAN tagging for trunk connections. The following four VLANs have been defined:

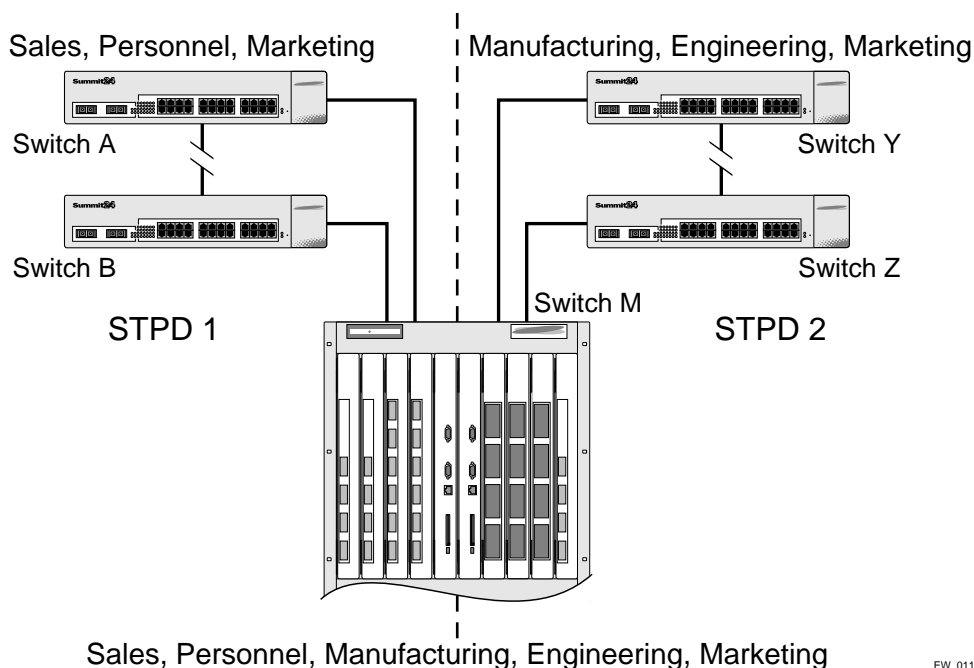
- *Sales* is defined on switch A, switch B, and switch M.
- *Personnel* is defined on switch A, switch B, and switch M.
- *Manufacturing* is defined on switch Y, switch Z, and switch M.

- *Engineering* is defined on switch Y, switch Z, and switch M.
- *Marketing* is defined on all switches (switch A, switch B, switch Y, switch Z, and switch M).

Two STPDs are defined:

- STPD1 contains VLANs *Sales* and *Personnel*.
- STPD2 contains VLANs *Manufacturing* and *Engineering*.

The VLAN *Marketing* is a member of both STPD1 and STPD2.



EW_011

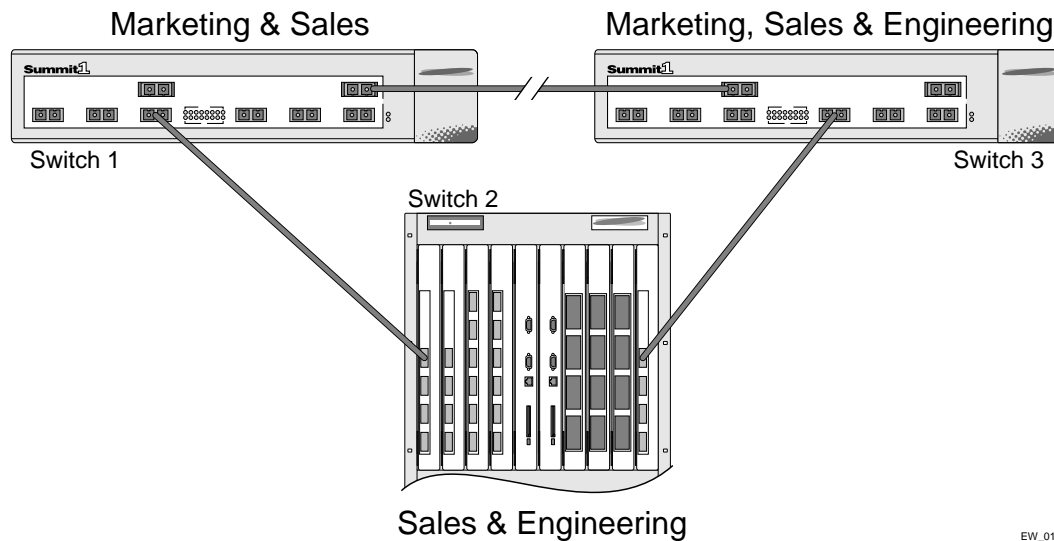
Figure 13-1: Multiple Spanning Tree Domains

When the switches in this configuration start up, STP configures each STPD such that there are no active loops in the topology. STP could configure the topology in a number of ways to make it loop-free.

In Figure 13-1, the connection between switch A and switch B is put into blocking state, and the connection between switch Y and switch Z is put into blocking state. After STP converges, all the VLANs can communicate, and all bridging loops are prevented.

The VLAN *Marketing*, which has been assigned to both STPD1 and STPD2, communicates using all five switches. The topology has no loops, because STP has already blocked the port connection between switch A and switch B, and between switch Y and switch Z.

Within a single STPD, you must be extra careful when configuring your VLANs. Figure 13-2 illustrates a network that has been incorrectly set up using a single STPD so that the STP configuration disables the ability of the switches to forward VLAN traffic.



EW_012

Figure 13-2: Tag-based STP configuration

The tag-based network in Figure 13-2 has the following configuration:

- Switch 1 contains VLAN *Marketing* and VLAN *Sales*.
- Switch 2 contains VLAN *Engineering* and VLAN *Sales*.
- Switch 3 contains VLAN *Marketing*, VLAN *Engineering*, and VLAN *Sales*.
- The tagged trunk connections for three switches form a triangular loop that is not permitted in an STP topology.
- All VLANs in each switch are members of the same STPD.

STP can block traffic between switch 1 and switch 3 by disabling the trunk ports for that connection on each switch.

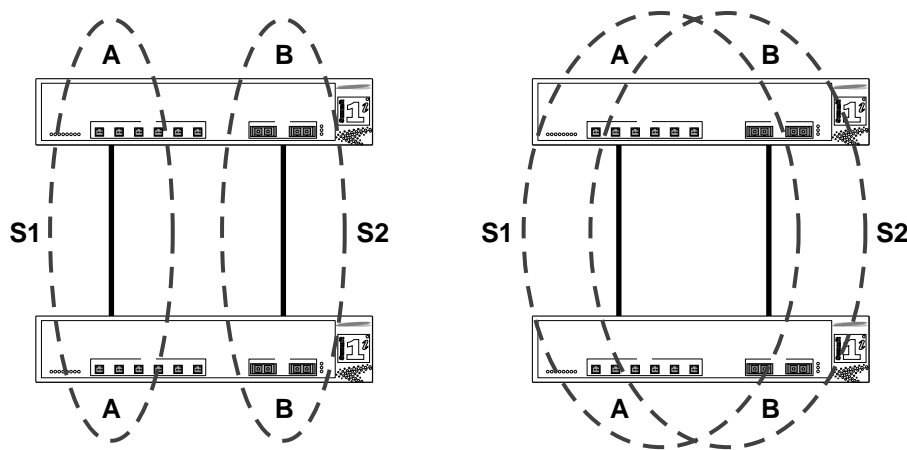
Switch 2 has no ports assigned to VLAN marketing. Therefore, if the trunk for VLAN marketing on switches 1 and 3 is blocked, the traffic for VLAN marketing will not be able to traverse the switches.



Note: If an STPD contains multiple VLANs, all VLANs must be configured on all ports in that domain, except for ports that connect to hosts (edge ports).

Multiple STPDs on a Port

Traditional 802.1d STP has some inherent limitations when addressing networks that have multiple VLANs and multiple STPDs. For example, consider the simple depicted in Figure 13-3.



EW_082

Figure 13-3: Limitations of Traditional STPD

The two switches are connected by a pair of parallel links. Both switches run two VLANs, A and B. To achieve load-balancing between the two links using the traditional approach, you would have to associate A and B with two different STPDs, called S1 and S2, respectively, and make the left link carry VLAN A traffic while the right link carry VLAN B traffic (or vice versa). If the right link fails, S2 is broken and VLAN B traffic is disrupted.

To optimize the solution, you can use the Extreme Multiple Instance Spanning (EMISTP) mode, which allows a port to belong to multiple STPDs. EMISTP adds

significant flexibility to STP network design. Referring to Figure 13-3, using EMISTP, you can configure all four ports to belong to both VLANs.

Assuming that S1 and S2 still correspond to VLANs A and B, respectively, you can fine-tune STP parameters to make the left link active in S1 and blocking in S2, while the right link is active in S2 and blocking in S1. Once again, if the right link fails, the left link is elected active by the STP algorithm for S2, without affecting normal switching of data traffic.

Using EMISTP, an STPD becomes more of an abstract concept. It does not necessarily correspond to a physical domain. It is better regarded as a vehicle to carry VLANs that have STP instances. Because VLANs can overlap, so do STPDs. However, even if the different STPDs share the entire topology or part of the redundant topology, the STPDs react to topology change events in an independent fashion.

VLAN Spanning Multiple STPDs

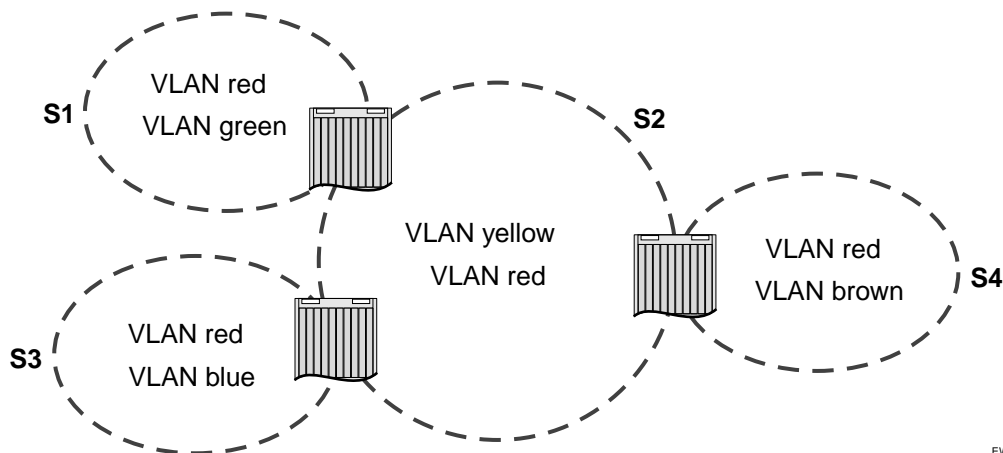
Traditionally, the mapping from VLANs to STP instances have been one-to-one, or many-to-one. In both cases, a VLAN is wholly contained in a single instance. In practical deployment there are cases in which a one-to-many mapping is desirable. In a typical large enterprise network, for example, VLANs span multiple sites and/or buildings. Each site represents a redundant looped area. However, between any two sites the topology is usually very simple.

Alternatively, the same VLAN may span multiple large geographical areas (because they belong to the same enterprise) and may traverse a great many nodes. In this case, it is desirable to have multiple STP domains operating in a single VLAN, one for each looped area. The justifications include the following:

- The complexity of the STP algorithm increases, and performance drops, with the size and complexity of the network. The 802.1d standard specifies a maximum network diameter of 7 hops. By segregating a big VLAN into multiple STPDs, you reduce complexity and enhance performance.
- Local to each site, there may be other smaller VLANs that share the same redundant looped area with the large VLAN. Some STPDs must be created to protect those VLAN. The ability to partition VLANs allows the large VLAN to be “piggybacked” in those STPDs in a site-specific fashion.

Figure 13-4 has five domains. VLANs green, blue, brown, and yellow are local to each domain. VLAN red spans all of the four domains. Using a VLAN that spans multiple

STPDS, you do not have to create a separate domain for VLAN red. Instead, VLAN red is “piggybacked” onto those domains local to other VLANs.



EW_083

Figure 13-4: VLAN Spanning Multiple STPDs

In addition, the configuration in Figure 13-4 has these features:

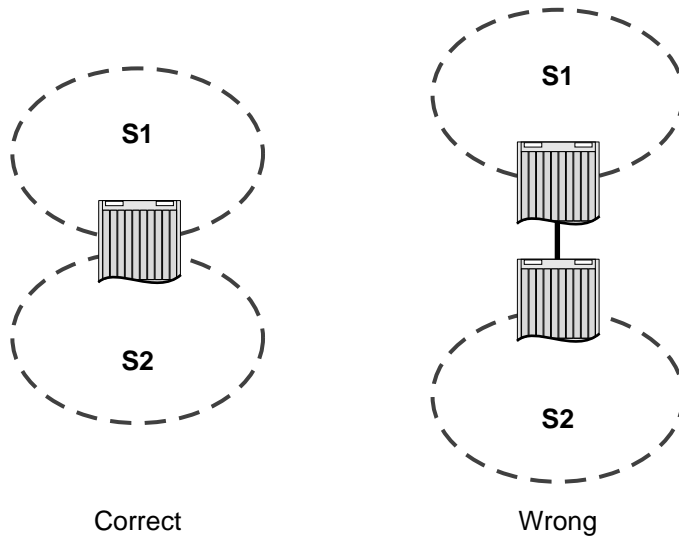
- Each site can be administered by a different organization or department within the enterprise. Having a site-specific STP implementation makes the administration more flexible and convenient.
- Between the sites the connection usually traverse distribution switches in ways that are known beforehand to be “safe” with STP. In other words, the looped areas are already well-defined.

EMISTP Deployment Constraints

While EMISTP greatly enhances STP capability, these features must be deployed with care. This section discusses configuration issues that, if not followed, could lead to an improper deployment of EMISTP. This section also provides the restrictive principles to abide by in network design.

- While a physical port can belong to multiple STPDs, any VLAN on that port can be in only one domain. Put another way, a VLAN can not belong to two domains on the same physical port.

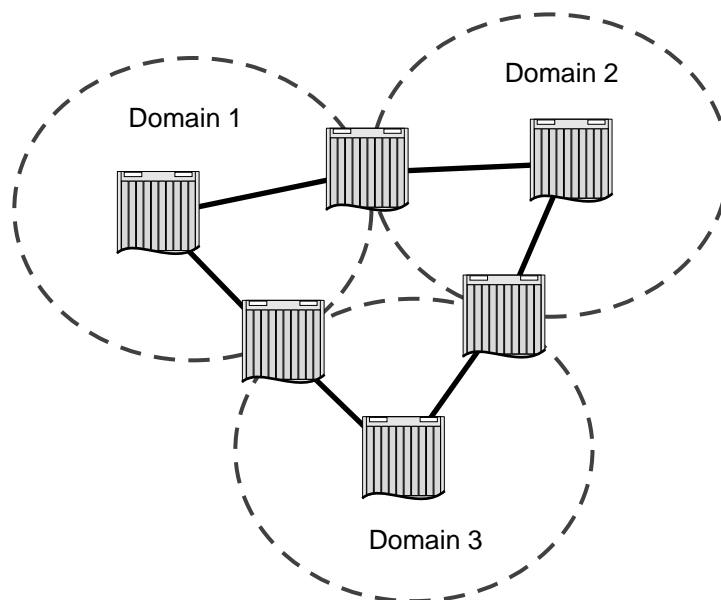
- While a VLAN can span multiple domains, any LAN segment in that VLAN must be in the same STPD. VLANs traverse domains only inside switches, not across links. On a single switch, however, bridge ports for the same VLAN can be assigned to different STPDs. This scenario is illustrated in Figure 13-5.



EW_084

Figure 13-5: VLANs traverse domains inside switches

- The VLAN partition feature is deployed under the premise that the overall inter-domain topology for that VLAN is loop-free. Consider the case in Figure 13-6, VLAN red (the only VLAN in the figure) spans domains 1, 2, and 3. Inside each domain, STP produces a loop-free topology. However, VLAN red is still looped, because the three domains form a ring among themselves.



EW_085

Figure 13-6: Looped VLAN topology

A necessary (but not sufficient) condition for a loop-free inter-domain topology is that every two domains only meet at a single crossing point.

Per-VLAN Spanning Tree

Switching products that implement Per-VLAN Spanning Tree (PVST) have been in existence for many years and are widely deployed. To support STP configurations that use PVST, ExtremeWare has an operational mode called PVST+.



Note: In this document, PVST and PVST+ are used interchangeably. PVST+ is an enhanced version of PVST that is interoperable with 802.1Q STP. The following discussions are in regard to PVST+, if not specifically mentioned.

STPD VLAN Mapping

Each VLAN participating in PVST+ must be in a separate STPD and the VLAN number must be the same as the STPD identifier (StpdID). As a result, PVST+ VLANs can not be partitioned.

This fact does not exclude other non-PVST+ VLANs from being grouped into the same STPD. A PVST+ VLAN can be joined by multiple non-PVST+ VLANs to be in the same STP domain.

Native VLAN

In PVST+, the native VLAN must be peered with default VLAN on Extreme devices, as both are the only VLAN allowed to send and receive untagged packets on the physical port.

Third-party PVST+ devices send VLAN 1 packets in a special manner. ExtremeWare does not support PVST+ for VLAN 1. Therefore, when the switch receives a packet for VLAN 1, the packet is dropped.

When a PVST+ instance is disabled, the fact that PVST+ uses a different packet format raises an issue. If the STPD also contains ports not in PVST+ mode, the flooded packet has an incompatible format with those ports. The packet is not recognized by the devices connected to those ports. Therefore, ExtremeWare has the following limitation:

- If an STPD contains both PVST+ and non-PVST+ ports, the STPD must not be disabled. Otherwise, the BPDUs are flooded in the format of the incoming STP port.

STP Rules and Restrictions

This section summarizes the rules and restrictions for configuring STP.

- The StpdID must be the VLANid of one of its member VLANs, and that VLAN can not be partitioned.
- A default VLAN can not be partitioned. If a VLAN traverses multiple STP domains, the VLAN must be tagged.
- An STPD can carry, at most, one VLAN running in PVST+ mode, and its StpdID must be identical with that VLANid. In addition, the PVST+ VLAN can not be partitioned.

- The default VLAN of a PVST+ port must be identical with the native VLAN on the PVST+ device connected to that port.
- If a port supports 802.1s-STPD, then the port must be configured with a default VLAN. If not, the BPDUs for that STPD are not flooded when the STPD is disabled.
- If an STPD contains both PVST+ and non-PVST+ ports, it must be enabled. If it is disabled, the BPDUs are flooded in the format of the incoming STP port, which may be incompatible with those of the connected devices.

Configuring Basic STP on the Switch

To configure basic STP, follow these steps:

- 1 Create one or more STP domains using the following command:

```
create stpd <stpd_name>
```



Note: STPD, VLAN, and QoS profile names must all be unique. For example, a name used to identify a VLAN cannot be used when you create an STPD or a QoS profile.

- 2 Add one or more VLANs to the STPD using the following command:

```
config stpd <stpd_name> add vlan <name>
```

- 3 Enable STP for one or more STP domains using the following command:

```
enable stpd {<stpd_name>}
```

After you have created the STPD, you can optionally configure STP parameters for the STPD.



Note: You should not configure any STP parameters unless you have considerable knowledge and experience with STP. The default STP parameters are adequate for most networks.

The following parameters can be configured on each STPD:

- Hello time
- Forward delay
- Max age

- Bridge priority
- StpdID

The following parameters can be configured on each port:

- Path cost
- Port priority
- Port mode



Note: The device supports the RFC 1493 Bridge MIB. Parameters of only the s0 default STPD are accessible through this MIB.



Note: If an STPD contains at least one port not in dot1D mode, the STPD must be configured with an StpdID.

STP Configuration Examples

This section provides two configuration examples:

- Basic 802.1d STP
- EMISTP

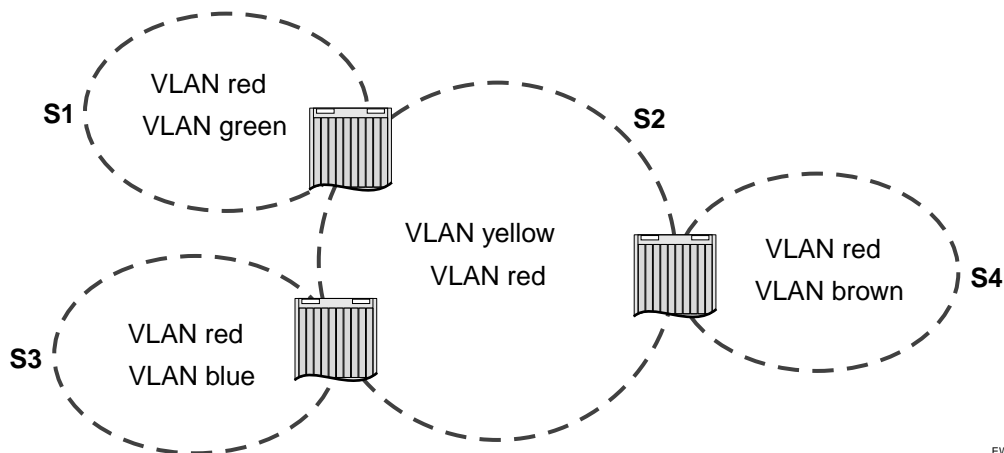
Basic 802.1d Configuration Example

The following modular switch example creates and enables an STPD named *Backbone_st*. It assigns the *Manufacturing* VLAN to the STPD. It disables STP on slot 2, ports 1 through 7, and slot 3 port 12.

```
create stpd backbone_st
config stpd backbone_st add vlan manufacturing
enable stpd backbone_st
disable stpd backbone_st port 2:1-2:7,3:12
```

EMISTP Configuration Example

Figure 13-7 is an example of EMISTP.



EW_083

Figure 13-7: EMISTP configuration example

The following commands configure the switch located between S1 and S2:

```

create vlan red
config red tag 100
config red add ports 1-4 tagged

create vlan green
config green tag 200
config green add ports 1-2 tagged

create vlan yellow
config yellow tag 300
config yellow add ports 3-4 tagged

create stpd s1
config stpd s1 add green
config stpd s1 tag 200
config stpd s1 add red ports 1-2 emistp

create stpd s2
config stpd s2 add yellow
config stpd s2 tag 300
config stpd s2 add red ports 3-4 emistp

```

Displaying STP Settings

To display STP settings, use the following command:

```
show stpd {detail}
```

This command displays the following information:

- STPD name
- Tag
- Flags
- Bridge ID
- STPD configuration information

To display the STP state of a port, use the following command:

```
show stpd <stpd_name> ports {detail}
```

This command displays the following information:

- STPD port configuration
- STPD state (root bridge, and so on)
- STPD port state (forwarding, blocking, and so on)

14

Extreme Standby Router Protocol

This chapter covers the following topics:

- Overview on page 14-1
- ESRP Basics on page 14-2
- Determining the ESRP Master on page 14-3
- Grouping Blocks of 10/100 Ports on page 14-10
- ESRP Options on page 14-13
- ESRP and VLAN Aggregation on page 14-18
- ESRP Examples on page 14-19
- Displaying ESRP Information on page 14-23

Overview

ESRP is a feature of ExtremeWare that allows multiple switches to provide redundant routing services to users. From the workstation's perspective, there is only one default router (that has one IP address and one MAC address), so ARP cache entries in client workstations do not need to be refreshed or aged-out.

In addition to providing layer 3 routing redundancy for IP and IPX, ESRP also provides for layer 2 redundancy. These "layered" redundancy features can be used in combination or independently. You do not have to configure the switch for routing to make valuable use of ESRP. The layer 2 redundancy features of ESRP offer fast failure recovery and provide for dual-homed system design. In some instances, depending on

network system design, ESRP can provide better resiliency than using the Spanning Tree Protocol (STP).

It is highly recommended all switches participating in ESRP run the same version of ExtremeWare. Not all ESRP features are available in all ExtremeWare software releases.

ESRP-Aware Switches

Extreme switches that are not running ESRP, but are connected on a network that has other Extreme switches running ESRP are *ESRP-aware*. When ESRP-aware switches are attached to ESRP-enabled switches, the ESRP-aware switches reliably perform fail-over and fail-back scenarios in the prescribed recovery times. No configuration of this feature is necessary.

If Extreme switches running ESRP are connected to layer 2 switches that are not manufactured by Extreme Networks (or Extreme switches that are not running ExtremeWare 4.0 or above), the fail-over times seen for traffic local to the segment may appear longer, depending on the application involved and the FDB timer used by the other vendor's layer 2 switch. As such, ESRP can be used with layer 2 switches from other vendors, but the recovery times vary.

The VLANs associated with the ports connecting an ESRP-aware switch to an ESRP-enabled switch must be configured using an 802.1Q tag on the connecting port, or, if only a single VLAN is involved, as untagged using the protocol filter *any*. ESRP will not function correctly if the ESRP-aware switch interconnection port is configured for a protocol-sensitive VLAN using untagged traffic.

ESRP Basics

ESRP is configured on a per-VLAN basis on each switch. A maximum of four switches can participate in providing redundant layer 3 or layer 2 services to a single VLAN. The switches exchange keep-alive packets for each VLAN independently. Only one switch can actively provide layer 3 routing and/or layer 2 switching for each VLAN. The switch performing the forwarding for a particular VLAN is considered the "master" for that VLAN. Other participating switches for the VLAN are in standby mode.

For a VLAN with ESRP enabled, each participating switch uses the same MAC address and must be configured with the same IP address or IPX NetID. It is possible for one switch to be master for one or more VLANs while being in standby for others, thus allowing the load to be split across participating switches.



Note: If you configure OSPF and ESRP, you must manually configure an OSPF router identifier (ID). Be sure that you configure a unique OSPF router ID on each switch running ESRP. For more information on configuring OSPF, refer to Chapter 17.

To have two or more switches participate in ESRP, the following must be true:

- For each VLAN to be made redundant, the switches must have the ability to exchange packets on the same layer 2 broadcast domain for that VLAN. Multiple paths of exchange can be used, and typically exist in most network system designs that take advantage of ESRP.
- For a VLAN to be recognized as participating in ESRP, the assigned IP address or the IPX NETid for the separate switches must be *identical*. Other aspects of the VLAN, including its name, are ignored.
- ESRP must be enabled on the desired VLANs for each switch.



Note: ESRP cannot be enabled on the VLAN default.

- Extreme Discovery Protocol (EDP) must be enabled on the ports that are members of the ESRP VLANs (The default setting is enabled.).

To verify EDP status, use the following command:

```
show ports <portlist> info {detail}
```

Determining the ESRP Master

The ESRP master switch (providing layer 3 routing and/or layer 2 switching services for a VLAN) is determined by the following factors:

- **Active ports**—The switch that has the greatest number of active ports takes highest precedence.
- **Tracking information**—Various types of tracking are used to determine if the switch performing the master ESRP function has connectivity to the outside world. ExtremeWare supports the following types of tracking:
 - VLAN - Tracks any active port connectivity to one or more designated VLANs
 - IP route table entry - Tracks specific learned routes from the IP route table
 - Ping - Tracks ICMP ping connectivity to specified devices

If any of the configured tracking mechanisms fail, the master ESRP switch relinquishes status as master, and remains in standby mode for as long as the tracking mechanism continues to fail.

- **ESRP priority**—This is a user-defined field. The range of the priority value is 0 to 254; a higher number has higher priority. The default priority setting is 0. A priority setting of 255 loses the election and remains in standby mode.
- **System MAC address**—The switch with the higher MAC address has priority.

ESRP Tracking

Tracking information is used to track various forms of connectivity from the ESRP switch to the outside world. This section describes the following ESRP tracking options:

- ESRP Environment and Diagnostic Tracking
- ESRP VLAN Tracking
- ESRP Route Table Tracking

ESRP Environment and Diagnostic Tracking

You can configure ESRP to track hardware status. If a power supply or fan fails, if the chassis is overheating, or if the diagnostics fail, the priority for the ESRP VLAN will change to the failover settings.

To configure the failover priority for ESRP VLAN, follow these steps:

- 1 Assign a priority to each ESRP VLAN, using the following command:

```
config vlan <vlan name> esrp priority
```

The range of the priority value is 0 to 254; a higher number has a higher priority. The default priority setting is 0.



Note: If you set the priority to 255, the ESRP VLAN will remain in standby mode even if the master ESRP VLAN fails.

You will typically configure both ESRP VLANs with the same priority.

- 2 Assign the priority flag precedence over the active ports count, using the following command:

```
config vlan <vlan name> esrp esrp-election priority-ports-track-mac
```

Because the priority of both ESRP VLANs are set to the same value, ESRP will use the active ports count to determine the master ESRP VLAN.

3 Set the failover priority, using the following command:

```
config vlan <vlan name> add [track-rip | track-bgp | track-ospf]
failover <priority>
```

Where:

- track-rip tracks for any available RIP route.
- track-bgp tracks for any available BGP route.
- track-ospf tracks for any available OSPF route.

The range of the priority value is 0 to 254; a higher number has a higher priority. The default priority setting is 0.



Note: If you set the priority to 255, the ESRP VLAN experiencing hardware failure will become the standby VLAN and will remain in standby mode even if the master ESRP VLAN fails.

Typically you will set the failover priority lower than the configured priority. Then, if one of the ESRP VLANs experiences a hardware or diagnostics failure, it will become the standby VLAN.

ESRP VLAN Tracking

You can configure ESRP to track port connectivity to a specified VLAN as criteria for failover. If no active ports remain on the specified VLANs, the switch automatically relinquishes master status and remains in standby mode.

To add or delete a tracked VLAN, use the following command:

```
config vlan <name> [add | delete] track-vlan <vlan_tracked>
```

ESRP Route Table Tracking

You can configure ESRP to track specified routes in the route table as criteria for failover. If any of the configured routes are not available within the route table, the switch automatically relinquishes master status and remains in standby mode.

To participate in ESRP route table tracking, all ESRP switches must run ExtremeWare version 6.0 or above.

To add or delete a tracked route, use the following command:

```
config vlan <name> [add | delete] track-route <ipaddress/mask_length>
```

ESRP Ping Tracking

You can configure ESRP to track connectivity using a simple ping to any outside responder. The responder may represent the default route of the switch, or any device meaningful to network connectivity of the master ESRP switch. The switch automatically relinquishes master status and remains in standby mode if a ping keepalive fails three consecutive times.

To participate in ESRP ping tracking, all ESRP switches must run ExtremeWare version 6.0 or above.

To view the status of tracked devices, use the following command:

```
show esrp
```

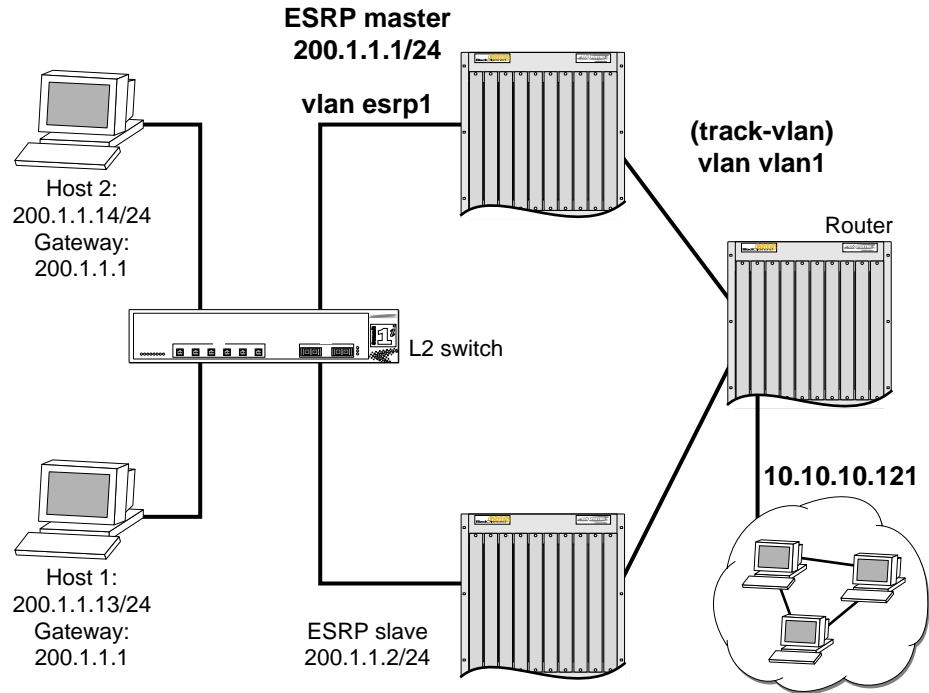
ESRP Multiple Ping Tracking

You can configure ESRP to track connectivity to up to four outside responders using a simple ping. To configure ping tracking, use the following command:

```
config vlan <vlan name> add track-ping <ip address> frequency <seconds>  
miss <number>
```

ESRP Tracking Example

Figure 14-1 is an example of ESRP tracking.



EW_080

Figure 14-1: ESRP tracking

To configure VLAN tracking, use the following command:

```
Configure vlan esrp1 add track-vlan vlan1
```

Using the tracking mechanism, if VLAN1 fails, the ESRP master realizes that there is no path to the upstream router via the Master switch and implements a failover to the slave.

To configure route table tracking, use the following command:

```
Config vlan esrp1 add track-ipvroute 10.10.10.0/24
```

The route specified in this command must exist in the IP routing table. When the route is no longer available, the switch implements a failover to the slave.

To configure ping tracking, use the following command:

```
Config vlan esrp1 add track-ping 10.10.10.121 2 2
```

The specified IP address is tracked. If the fail rate is exceeded the switch implements a failover to the slave.

To configure RIP tracking, use the following command:

```
Config vlan esrp1 add track-rip failover 20
```

The switch tracks RIP routes in its IP routing table. If no RIP routes are available, the switch implements a failover to failover priority 20.

To configure OSP tracking, use the following command:

```
Config vlan esrp1 add track-ospf failover 20
```

The switch tracks OSPF routes in its IP routing table. If no OSPF routes are available, the switch implements a failover to failover priority 20.

To configure BGP tracking, use the following command:

```
Config vlan esrp1 add track-bgp failover 20
```

The switch tracks BGP routes in its IP routing table. If no BGP routes are available, the switch implements a failover to failover priority 20.

ESRP Election Algorithms

You configure the switch to use one of seven different election algorithms to select the ESRP master. Each algorithm considers the election factors in a different order of precedence, as follows:

- `ports-track-priority-mac` — Active ports, tracking information, ESRP priority, MAC address (Default)
- `ports-track-priority` — Active ports, tracking information, ESRP priority
- `track-ports-priority-mac` — Tracking information, active ports, ESRP priority, MAC address
- `track-ports-priority` — Tracking information, active ports, ESRP priority

- `priority-ports-track-mac` — ESRP priority, active ports, tracking information, MAC address
- `priority-track-ports-mac` — ESRP priority, tracking information, active ports, MAC address
- `priority-mac-only` — ESRP priority, MAC address



Caution: All switches in the ESRP network must use the same election algorithm, otherwise loss of connectivity, broadcast storms, or other unpredictable behavior may occur.



Note: Only the `ports-track-priority-mac` election algorithm is compatible with ExtremeWare releases prior to version 6.0.

Master Switch Behavior

If a switch is master, it actively provides layer 3 routing services to other VLANs, and layer 2 switching between all the ports of that VLAN. Additionally, the switch exchanges ESRP packets with other switches that are in standby mode.

Standby Switch Behavior

If a switch is in standby mode, it exchanges ESRP packets with other switches on that same VLAN. When a switch is in standby, it does not perform layer 3 routing or layer 2 switching services for the VLAN. From a layer 3 routing protocol perspective (for example, RIP or OSPF), when in standby for the VLAN, the switch marks the router interface associated with the VLAN as down. From a layer 2 switching perspective, no forwarding occurs between the member ports of the VLAN; this prevents loops and maintains redundancy.

Electing the Master Switch

A new master can be elected in one of the following ways:

- A communicated parameter change
- Loss of communication between master and slave(s)

If a parameter that determines the master changes (for example, link loss or priority change), the election of the new master typically occurs within one timer cycle (2 seconds by default). If a switch in standby mode loses its connection with the master, a

new election (using the same precedence order indicated previously) occurs. The new election typically takes place in three times the defined timer cycle (6 seconds by default).

Failover Time

Failover time is largely determined by the following factors:

- The ESRP timer setting.
- The routing protocol being used for inter-router connectivity if layer 3 redundancy is used. OSPF fail-over time is faster than RIP fail-over time.

The failover time associated with the ESRP protocol is dependent on the timer setting and the nature of the failure. The default timer setting is 2 seconds; the range is 1 to 255 seconds.

If routing is configured, the failover of the particular routing protocol (such as RIP V1, RIP V2, or OSPF) is added to the failover time associated with ESRP.

Grouping Blocks of 10/100 Ports

Restrictions on port groupings apply only to switches that do not use the “i” chipset.

If you enable ESRP on a VLAN that contains 10/100 ports, a specific block of neighboring ports must also be participating in a VLAN running ESRP, or must not be used. The blocks of ports are physically adjacent, regardless of the switch module. For example, the blocks on a BlackDiamond F32T module consist of the following:

- Ports 1-4 and 17-20
- Ports 5-8 and 21-24
- Ports 9-12 and 25-28
- Ports 13-16 and 29-32

Figure 14-2 through Figure 14-6 illustrate the port blocks for each Extreme switch.

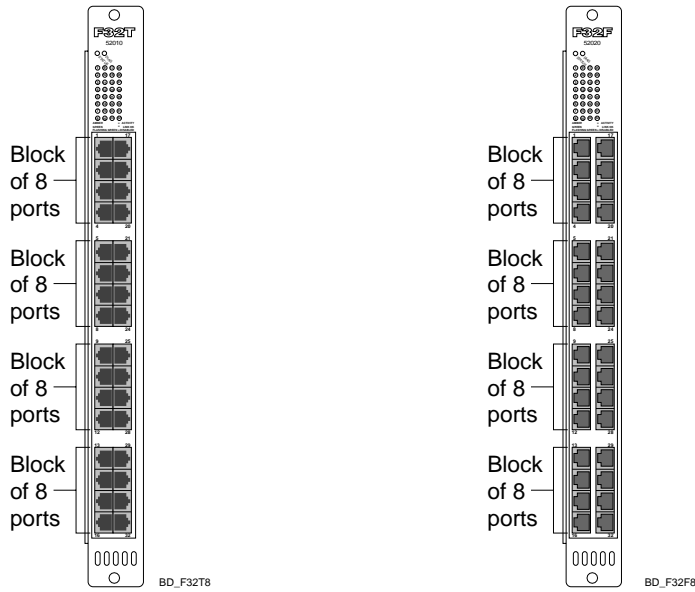
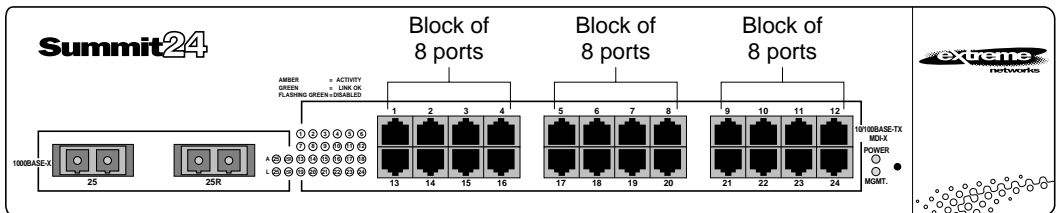


Figure 14-2: F32T and F32F ESRP port blocks



Sum24_8

Figure 14-3: Summit24 switch ESRP port blocks

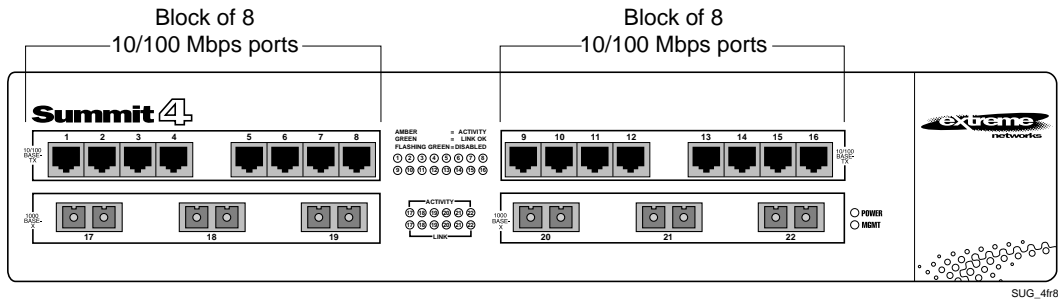


Figure 14-4: Summit4 switch ESRP port blocks

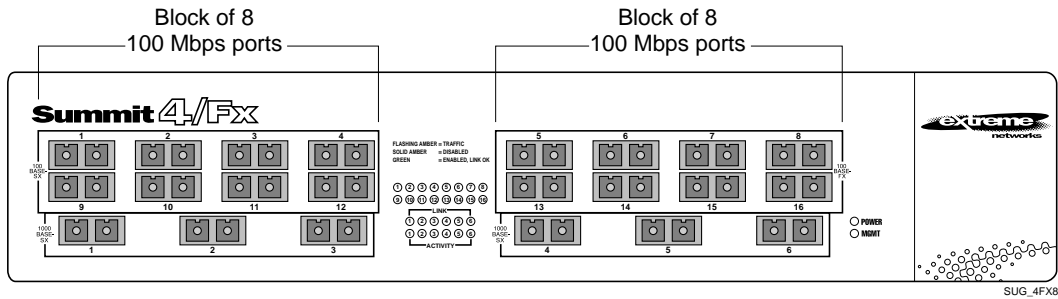


Figure 14-5: Summit4/FX switch ESRP port blocks

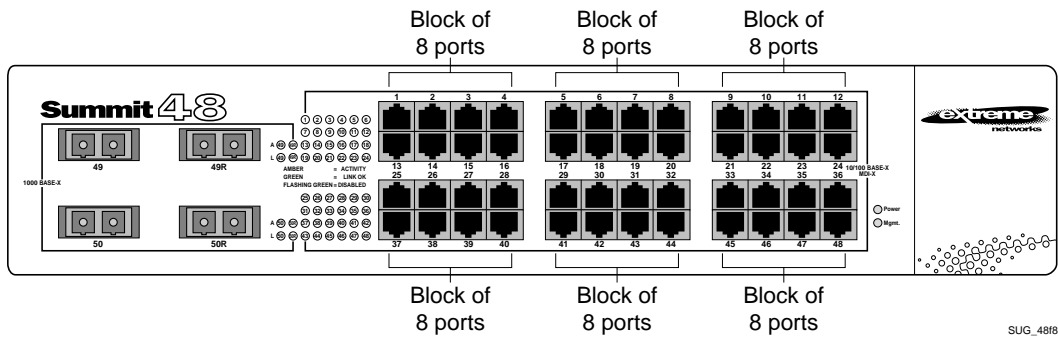


Figure 14-6: Summit48 switch ESRP port blocks



Note: For switches that do not use the “i” chipset, all VLANs using a port or port block must enable ESRP. This requirement does not apply to switches that use the “i” chipset.

ESRP Options

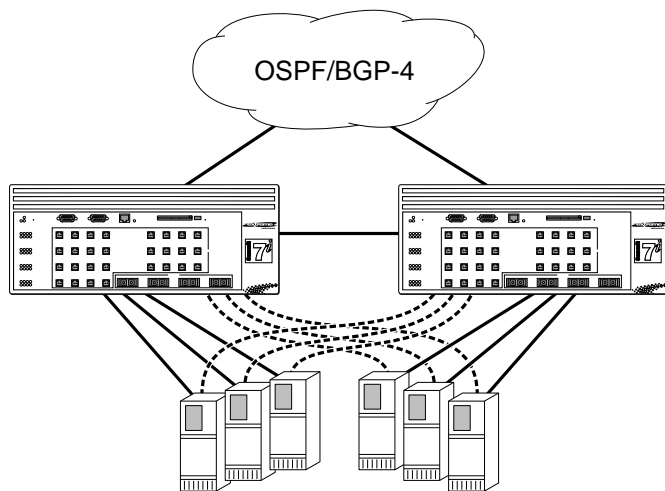
This section discusses the following ESRP options:

- ESRP Host Attach
- ESRP Domains
- ESRP Groups
- Linking ESRP Switches
- Configuring ESRP and Multinetting
- ESRP and Spanning Tree
- ESRP Multiple ping Tracking
- ESRP Port Restart

ESRP Host Attach

ESRP host attach (HA) is an optional ESRP configuration that allows you to connect active hosts directly to an ESRP master or standby switch. Normally, the layer 2 redundancy and loop prevention capabilities of ESRP do not allow packet forwarding from the standby ESRP switch. ESRP HA allows configured ports that do not represent loops to the network to continue layer 2 operation independent of their ESRP status.

The ESRP HA option is useful if you are using dual-homed network interface cards (NICs) for server farms, and in conjunction with high availability server load-balancing (SLB) configurations, as shown in Figure 14-7.



EW_045

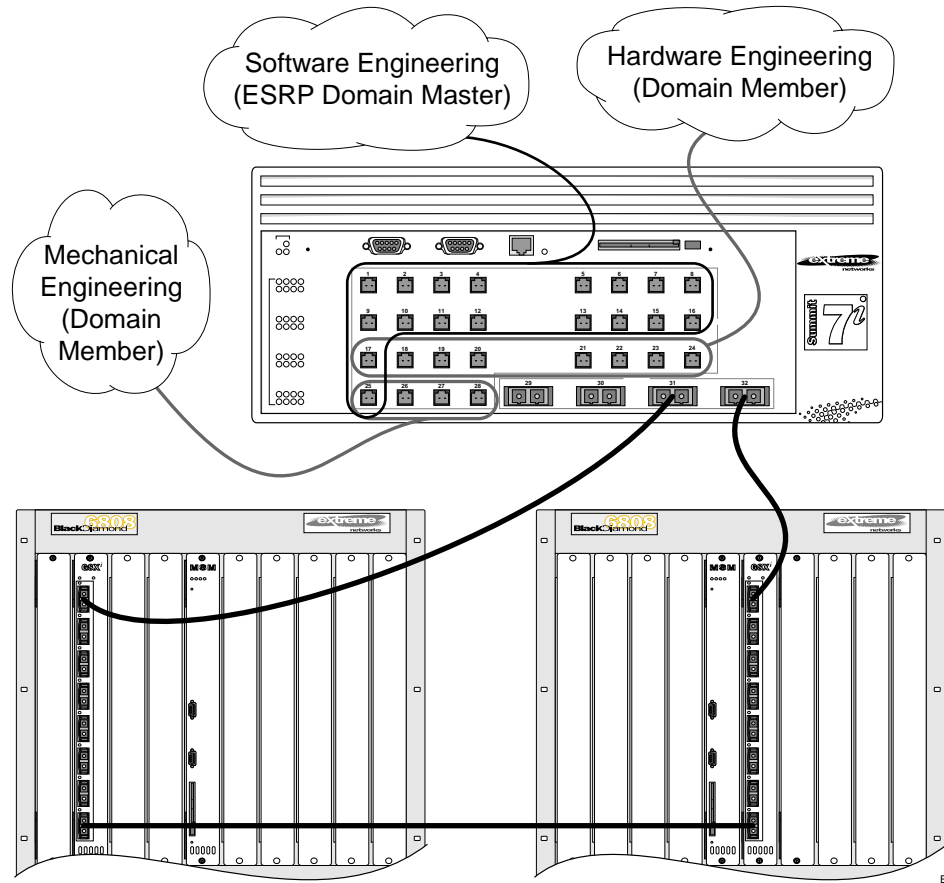
Figure 14-7: ESRP host attach

ESRP VLANs that share ESRP HA ports must be members of different ESRP groups.

Other applications allow lower cost redundant routing configurations, because hosts can be directly attached to the switch involved with ESRP. The ESRP HA feature is used only on switches and I/O modules that have the “i” series chipset. It also requires at least one link between the master and standby ESRP switch for carrying traffic and to exchange ESRP hello packets.

ESRP Domains

ESRP domains is an optional ESRP configuration that allows you to configure multiple VLANs under the control of a single instance of the ESRP protocol. By grouping multiple VLANs under one ESRP domain, the ESRP protocol can scale to provide protection to large numbers of VLANs. All VLANs within an ESRP domain simultaneously share the same active and standby router and failover, providing one port of each member VLAN belongs to the domain master, as shown in Figure 14-8.



EW_065

Figure 14-8: ESRP domains

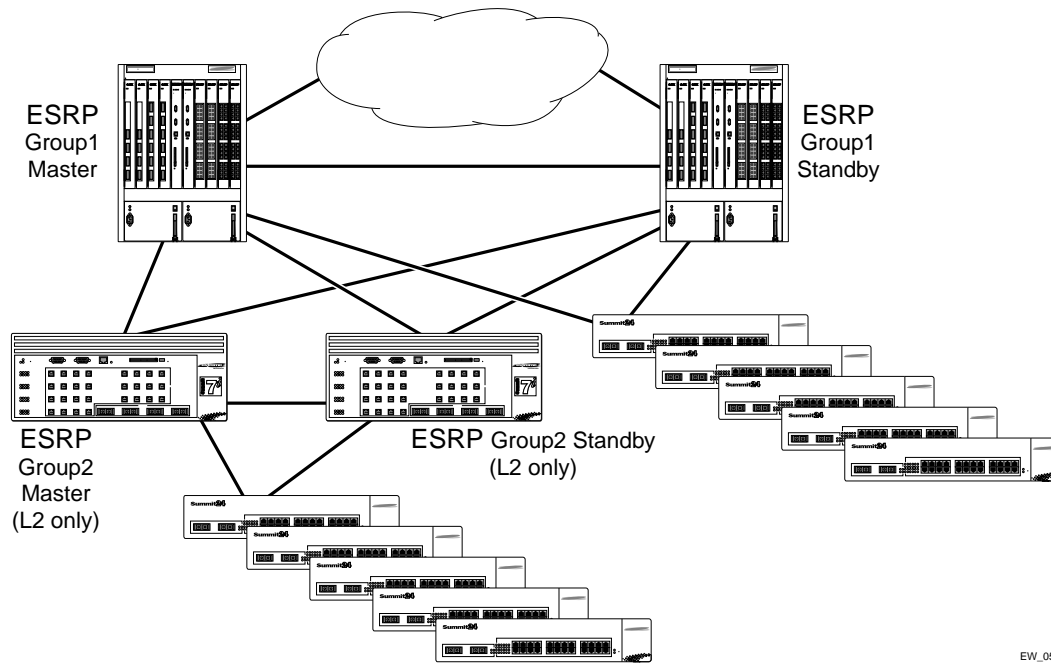
When a port in a member VLAN belongs to the domain master, the member VLAN ports are considered when determining the ESRP master.

The ESRP group feature is used only on switches and I/O modules that have the “i” series chipset.

ESRP Groups

ExtremeWare supports running multiple instances of ESRP within the same VLAN or broadcast domain. This functionality is called an ESRP group. Though other uses exist, the most typical application for multiple ESRP groups is when two or more sets of ESRP switches are providing fast-failover protection within a subnet. A maximum of four distinct ESRP groups can be supported within the same networked broadcast domain.

For example, two ESRP switches provide L2/L3 connectivity and redundancy for the subnet, while another two ESRP switches provide L2 connectivity and redundancy for a portion of the same subnet. Figure 14-9 shows ESRP groups.



EW_056

Figure 14-9: ESRP groups



Note: A switch cannot perform both master and slave functions on the same VLAN for separate instances of ESRP.

An additional user for ESRP groups is ESRP Host Attached, described on page 14-13.

Linking ESRP Switches

When considering system design using ESRP, direct links between ESRP switches are useful under the following conditions:

- A direct link can provide a more direct routed path, if the ESRP switches are routing and supporting multiple VLANs where the master/standby configuration is split such that one switch is master for some VLANs and a second switch is master for other VLANs. The direct link can contain a unique router-to-router VLAN/subnet, so that the most direct routed path between two VLANs with different master switches uses a direct link, instead of forwarding through another set of connected routers.
- A direct link can be used as a highly reliable method to exchange ESRP hellos, so that the possibility of having multiple masters for the same VLAN is lessened, should all downstream layer 2 switches fail.
- A direct link is necessary when the ESRP HA option. The direct link is used to provide layer 2 forwarding services through an ESRP standby switch.

Direct links may contain a router-to-router VLAN, along with VLANs running ESRP. If multiple VLANs are used on the direct links, use 802.1Q tagging. The direct links may be aggregated into a load-shared group, if desired.

Configuring ESRP and Multinetting

When configuring ESRP and IP multinetting on the same switch, the parameters that affect the determination of the ESRP master must be configured identically for all the VLANs involved with IP multinetting. For example, the number of links in your configuration, the priority settings, and timer settings must be identical for all affected VLANs.

ESRP and Spanning Tree

A switch running ESRP should not simultaneously participate in the Spanning Tree Protocol (STP) for the same VLAN(s). Other switches in the VLAN being protected by ESRP may run STP and the switch running ESRP forwards, but does not filter, STP BPDUs. Therefore, you can combine ESRP and STP on a network and a VLAN, but you must do so on separate devices. You should be careful to maintain ESRP connectivity between ESRP master and standby switches when you design a network that uses ESRP and STP.

ESRP Port Restart

You can configure ESRP to restart ports if those ports are members of a VLAN that becomes a slave. To configure port restart, use the following command:

```
config vlan <name> add ports [<portlist> | all] restart
```

To disable port restart, use the following command:

```
config vlan <name> add ports [<portlist> | all] no-restart
```

If a VLAN becomes a slave, ESRP disconnects member ports that have port restart enabled. The disconnection of these ports causes downstream devices to remove the ports from their FDB tables. This feature allows you to use ESRP in networks that include equipment from other vendors. After 3 seconds the ports re-establish connection with the ESRP switch.

To remove a port from the restart configuration, delete the port from the VLAN and re-add it.



Note: The port restart feature is also available for VRRP. For more information on VRRP, see Chapter 15.

ESRP and VLAN Aggregation

ESRP can be used to provide redundant default router protection to VLAN aggregation clients. ESRP is enabled on the super-VLAN *only* (not the sub-VLANs). The procedure is to add ports to the super-VLAN that is shared with the sub VLANs. To do so, the super-VLAN should be configured with an 802.1Q tag, and added as tagged with the sub-VLAN ports to avoid a protocol conflict. Lastly, enable ESRP on the super-VLAN.

The following example combines ESRP and VLAN aggregation for the super-VLAN *vsuper* and two sub-VLANs, *v1sub* and *v2sub*, that have ports 1 and 2 as members, respectively.

1 Create the VLANs and set up the super to sub-VLAN relationship.

```
create vlan v1sub
create vlan v2sub
create vlan vsuper
config vsuper ipaddress 10.1.2.3/24
enable ipforwarding
enable ospf
```



```

config ospf add vsuper
config vlsub add port 1
config v2sub add port 2
config vsuper add subvlan vlsub
config vsuper add subvlan v2sub

```

2 Turn on ESRP for the VLAN *vsuper*.

```

config vsuper tag 1234
config vsuper add port 1,2 tagged
enable esrp vlan vsuper

```

Use the following commands to verify the configuration:

- `show vlan {detail}`—Displays super- and sub-VLAN relationships, IP addresses, and port membership.
- `show esrp {detail}`—Verifies ESRP is enabled and operational.

ESRP Examples

This section provides examples of ESRP configurations.

Single VLAN Using Layer 2 and Layer 3 Redundancy

This example, shown in Figure 14-10, uses a number of Summit switches that perform layer 2 switching for VLAN *Sales*. The Summit switches are dual-homed to the BlackDiamond switches. The BlackDiamond switches perform layer 2 switching between the Summit switches and layer 3 routing to the outside world. Each Summit switch is dual-homed using active ports to two BlackDiamond switches (as many as four could be used). ESRP is enabled on each BlackDiamond switch only for the VLAN that interconnects to the Summit switches. Each BlackDiamond switch has the VLAN *Sales* configured using the identical IP address. The BlackDiamond switches then connect to the routed enterprise normally, using the desired routing protocol (for example RIP or OSPF).

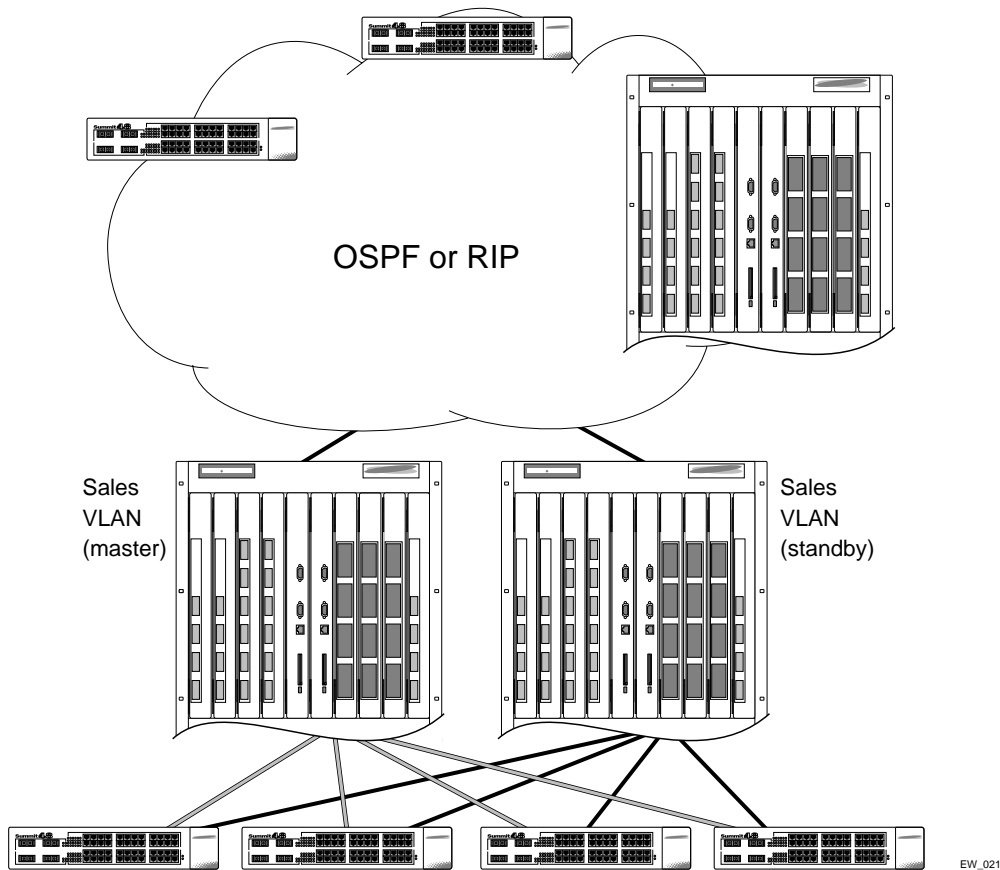


Figure 14-10: ESRP example using layer 2 and layer 3 redundancy

The BlackDiamond switch, acting as master for VLAN *Sales*, performs both layer 2 switching and layer 3 routing services for VLAN *Sales*. The BlackDiamond switch in standby mode for VLAN *Sales* performs neither, thus preventing bridging loops in the VLAN. The BlackDiamond switch in standby mode does, however, exchange ESRP packets with the master BlackDiamond switch.

There are four paths between the BlackDiamond switches on VLAN *Sales*. All the paths are used to send ESRP packets, allowing for four redundant paths for ESRP communication. The Summit switches, being ESRP-aware, allow traffic within the VLAN to fail-over quickly, as they will sense when a master/slave transition occurs and

flush FDB entries associated with the uplinks to the ESRP-enabled BlackDiamond switches.

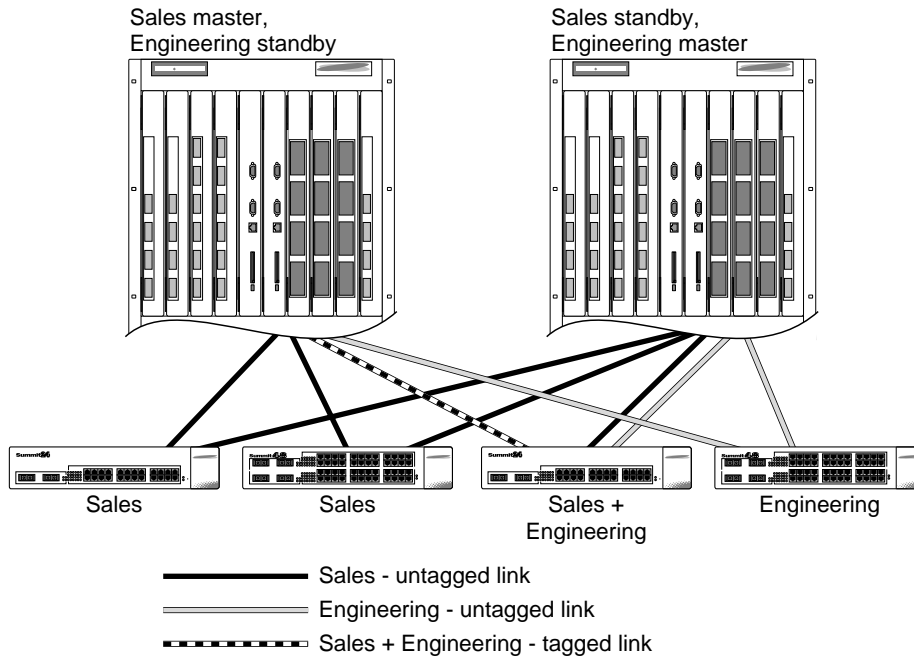
The following commands are used to configure both BlackDiamond switches. The assumption is that the inter-router backbone is running OSPF, with other routed VLANs already properly configured. Similar commands would be used to configure a switch on a network running RIP. The primary requirement is that the IP address for the VLAN(s) running ESRP must be identical. In this scenario, the master is determined by the programmed MAC address of the switch, because the number of active links for the VLAN and the priority are identical to both switches.

The commands used to configure the BlackDiamond switches are as follows:

```
create vlan sales
config sales add port 1:1-1:4
config sales ipaddr 10.1.2.3/24
enable ipforwarding
enable esrp sales
enable edp ports all
config ospf add vlan sales
enable ospf
```

Multiple VLANs Using Layer 2 Redundancy

The example shown in Figure 14-11 illustrates an ESRP configuration that has multiple VLANs using layer 2 redundancy.



EW_022

Figure 14-11: ESRP example using layer 2 redundancy

This example builds on the previous example, but eliminates the requirement of layer 3 redundancy. It has the following features:

- An additional VLAN, *Engineering*, is added that uses layer 2 redundancy.
- The VLAN *Sales* uses three active links to each BlackDiamond switch.
- The VLAN *Engineering* has two active links to each BlackDiamond switch.
- The third Summit switch carries traffic for both VLANs.
- The link between the third Summit switch and the first BlackDiamond switch uses 802.1Q tagging to carry traffic from both VLANs traffic on one link. The BlackDiamond switch counts the link active for each VLAN.
- The second BlackDiamond switch has a separate physical port for each VLAN connected to the third Summit switch.

In this example, the BlackDiamond switches are configured for ESRP such that the VLAN *Sales* normally uses the first BlackDiamond switch and the VLAN *Engineering* normally uses the second BlackDiamond switch. This is accomplished by manipulating the ESRP priority setting for each VLAN for the particular BlackDiamond switch.

Configuration commands for the first BlackDiamond switch are as follows:

```
create vlan sales
config sales tag 10
config sales add port 1:1-1:2
config sales add port 1:3 tagged
config sales ipaddr 10.1.2.3/24
create vlan eng
config eng tag 20
config eng add port 1:4
config eng add port 1:3 tagged
config eng ipaddr 10.4.5.6/24
enable esrp sales
enable esrp eng
enable edp ports all
config sales esrp priority 5
```

Configuration commands for the second BlackDiamond switch are as follows:

```
create vlan sales
config sales add port 1:1-1:3
config sales ipaddr 10.1.2.3/24
create vlan eng
config eng add port 1:4, 2:1
config eng ipaddr 10.4.5.6/24
enable esrp sales
enable esrp eng
config eng esrp priority 5
```

Displaying ESRP Information

To verify the operational state of an ESRP VLAN and the state of its neighbor, use the following command:

```
show esrp
```

To view tracking information about a particular VLAN, including the VLANs tracked by it and a list of the VLANs tracking it, use the `show vlan` command.

15

Virtual Router Redundancy Protocol

This chapter covers the following topics:

- Overview on page 15-1
- Determining the VRRP Master on page 15-3
- Additional VRRP Highlights on page 15-7
- VRRP Operation on page 15-7
- VRRP Configuration Parameters on page 15-11
- VRRP Examples on page 15-12

This chapter assumes that you are already familiar with the Virtual Router Redundancy Protocol (VRRP). If not, refer to the following publications for additional information:

- RFC 2328 — *Virtual Router Redundancy Protocol (VRRP)*
- RFC 2787 — *Definitions of Managed Objects for the Virtual Router Redundancy Protocol*

Overview

Like ESRP, VRRP is a protocol that allows multiple switches to provide redundant routing services to users. VRRP is used to eliminate the single point of failure associated with manually configuring a default gateway address on each host in a network. Without using VRRP, if the configured default gateway fails, you must reconfigure each

host on the network to use a different router as the default gateway. VRRP provides a redundant path for the hosts. If the default gateway fails, the backup router assumes forwarding responsibilities.

VRRP Terms

Table 15-1 describes terms associated with VRRP.

Table 15-1: VRRP Terms

Term	Description
virtual router	A VRRP router is a group of one or more physical devices that acts as the default gateway for hosts on the network. The virtual router is identified by a virtual router identifier (VRID) and an IP address.
VRRP router	Any router that is running VRRP. A VRRP router can participate in one or more virtual routers. A VRRP router can be a backup router for one more master routers.
IP address owner	A single VRRP router that has the IP address of the virtual router configured as its real interface address. The IP address owner responds to TCP/IP packets addressed to the virtual router IP address. The IP address owner is optional in a VRRP configuration.
master router	The physical device (router) in the virtual router that is responsible for forwarding packets sent to the virtual router, and responding to ARP requests. The master router sends out periodic advertisements that let backup routers on the network know that it is alive. If the IP address owner is identified, it always becomes the master.
backup router	Any VRRP router in the virtual router that is not elected as the master. The backup router is available to assume forwarding responsibility if the master becomes unavailable.
VRID	Virtual router identifier. Each virtual router is given a unique VRID. All of the VRRP routers that participate in the virtual router are assigned the same VRID.
virtual router MAC address	RFC 2338 assigns a static MAC address for the first 5 octets of the virtual router. These octets are set to 00-00-5E-00-01. When you configure the VRID, the last octet of the MAC address is dynamically assigned the VRID number.

Determining the VRRP Master

The VRRP master is determined by the following factors:

- **IP address**—If a router is configured with the IP address of the virtual IP address, it becomes the master.
- **VRRP priority**—This is a user-defined field. The range of the priority value is 1 to 254; a higher number has higher priority. The value of 255 is reserved for a router that is configured with the virtual router IP address. A value of 0 is reserved for the master router, to indicate it is releasing responsibility for the virtual router. The default value is 100.
- **Higher IP address**—If the routers have the same configured priority, the router with the higher IP address becomes the master.

VRRP Tracking

Tracking information is used to track various forms of connectivity from the VRRP router to the outside world. This section describes the following VRRP tracking options:

- VRRP VLAN tracking
- VRRP route table tracking
- VRRP ping tracking

VRRP VLAN Tracking

You can configure VRRP to track connectivity to one or more specified VLANs as criteria for failover. If no active ports remain on the specified VLANs, the router automatically relinquishes master status and remains in backup mode.

To add or delete a tracked VLAN, use the following command:

```
config vlan <name> [add | delete] track-vlan <name>
```

VRRP Route Table Tracking

You can configure VRRP to track specified routes in the route table as criteria for failover. If any of the configured routes are not available within the route table, the router automatically relinquishes master status and remains in backup mode.

To add or delete a tracked route, use the following command:

```
config vlan <name> [add | delete] track-route <ipaddress/mask_length>
```

VRRP Ping Tracking

You can configure VRRP to track connectivity using a simple ping to any outside responder. The responder may represent the default route of the router, or any device meaningful to network connectivity of the master VRRP router. The router automatically relinquishes master status and remains in backup mode if a ping keepalive fails three consecutive times.

To add or delete a tracked route, use the following command:

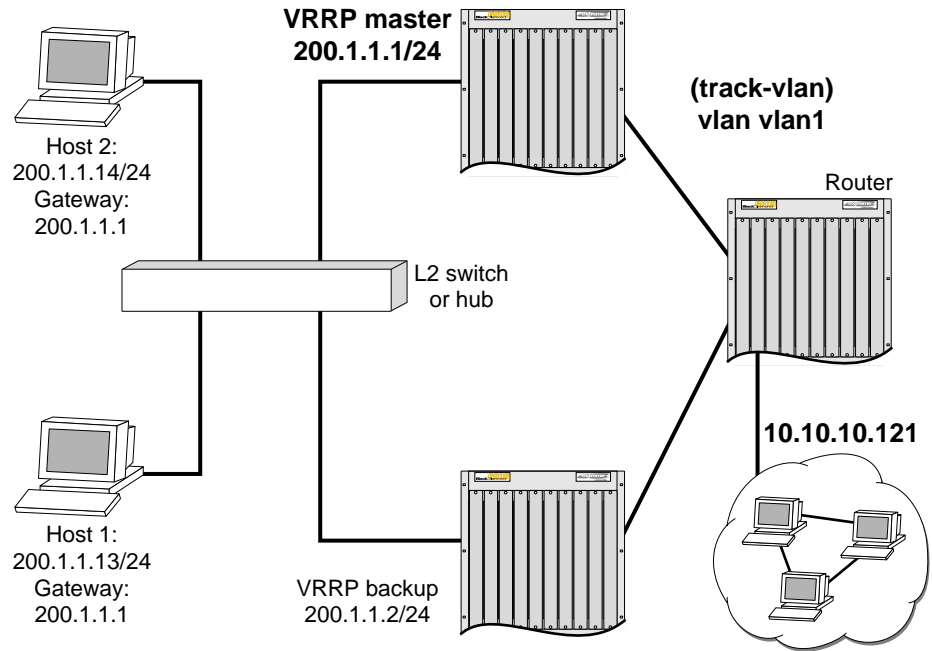
```
config vlan <name> [add | delete] track-ping <ipaddress> frequency  
<number> miss <number>
```

To view the status of tracked devices, use the following command:

```
show vrrp {vlan <name>} {detail}
```

VRRP Tracking Example

Figure 15-1 is an example of VRRP tracking.



EW_079

Figure 15-1: VRRP tracking

To configure VLAN tracking, as shown in Figure 15-1, use the following command:

```
Configure vlan vrrp1 add track-vlan vlan1
```

Using the tracking mechanism, if VLAN1 fails, the VRRP master realizes that there is no path to upstream router via the Master switch and implements a failover to the backup.

To configure route table tracking, as shown in Figure 15-1, use the following command:

```
Config vlan vrrp1 add track-iproute 10.10.10.0/24
```

The route specified in this command must exist in the IP routing table. When the route is no longer available, the switch implements a failover to the backup.

To configure ping tracking, as shown in Figure 15-1, use the following command:

```
Config vlan vrrp1 add track-ping 10.10.10.121 2 2
```

The specified IP address is tracked. If the fail rate is exceeded the switch implements a failover to the backup.

Electing the Master Router

VRRP uses an election algorithm to dynamically assign responsibility for the master router to one of the VRRP routers on the network. A VRRP router is elected master if one of the following is true:

- The router is the IP address owner.
- The router is configured with the highest priority (the range is 3 - 255).

If the master router becomes unavailable, the election process provides dynamic failover and the backup router that has the highest priority assumes the role of master.

A new master is elected when one of the following things happen:

- VRRP is disabled on the master router.
- Loss of communication between master and backup router(s).

When VRRP is disabled on the master interface, the master router sends an advertisement with the priority set to 0 to all backup routers. This signals the backup routers that they do not need to wait for the master down interval to expire, and the master election process for a new master can begin immediately.

The master down interval is set as follows:

$3 * \text{advertisement interval} + \text{skew time}$

Where:

- The advertisement interval is a user-configurable option.
- The skew time is $(256 - \text{priority}) / 256$.

Additional VRRP Highlights

The following additional points pertain to VRRP:

- VRRP packets are encapsulated IP packets.
- The VRRP multicast address is 224.0.0.18.
- The virtual router MAC address is 00 00 5E 00 01 <vrid>
- Duplicate virtual router IDs are allowed on the router, but not on the same interface.
- The maximum number of supported VRIDs per interface is 4.
- An interconnect link between VRRP routers should not be used, except when VRRP routers have hosts directly attached.
- A maximum of 64 VRID instances are supported on the router.
- Up to 4 unique VRIDs can be configured on the router. VRIDs can be re-used, but not on the same interface.
- VRRP and Spanning Tree can be simultaneously enabled on the same switch.
- VRRP and ESRP cannot be simultaneously enabled on the same switch.

VRRP Port Restart

You can configure VRRP to restart ports if those ports are members of a VLAN that becomes a backup. To configure port restart, use the following command:

```
config vlan <name> add ports [<portlist> | all] restart
```

To disable port restart, use the following command:

```
config vlan <name> add ports [<portlist> | all] no-restart
```

If a VLAN becomes a backup, VRRP disconnects member ports that have port restart enabled. The disconnection of these ports causes downstream devices to remove the ports from their FDB tables. This feature allows you to use VRRP in networks that include equipment from other vendors. After 3 seconds the ports re-establish connection with the VRRP switch.

To remove a port from the restart configuration, delete the port from the VLAN and re-add it.



Note: The port restart feature is also available for ESRP. For more information on ESRP, see Chapter 14.

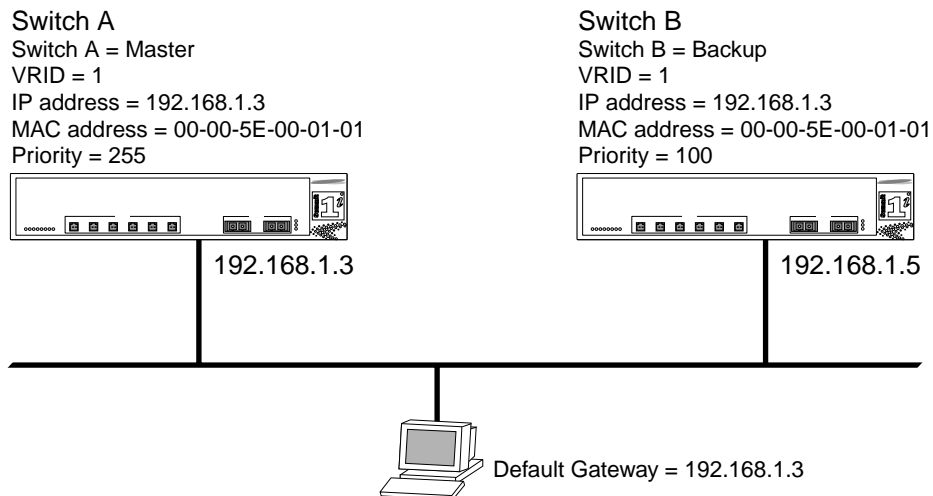
VRRP Operation

This section describes two VRRP network configuration:

- A simple VRRP network
- A fully-redundant VRRP network

Simple VRRP Network Configuration

Figure 15-2 shows a simple VRRP network.



EW_067

Figure 15-2: Simple VRRP network

In Figure 15-2, a virtual router is configured on Switch A and Switch B using these parameters:

- VRID is 1.
- MAC address is 00-00-5E-00-01-01.
- IP address is 192.168.1.3.

Switch A is configured with a priority of 255. This priority indicates that it is the master router. Switch B is configured with a priority of 100. This indicates that it is a backup router.

The master router is responsible for forwarding packets sent to the virtual router. When the VRRP network becomes active, the master router broadcasts an ARP request that contains the virtual router MAC address (in this case, 00-00-5E-00-01-01) for each IP address associated with the virtual router. Hosts on the network use the virtual router MAC address when they send traffic to the default gateway.

The virtual router IP address is configured to be the real interface address of the IP address owner. The IP address owner is usually the master router. The virtual router IP address is also configured on each backup router. However, in the case of the backup router, this IP address is not associated with a physical interface. Each physical interface on each backup router must have a unique IP address. The virtual router IP address is also used as the default gateway address for each host on the network.

If the master router fails, the backup router assumes forwarding responsibility for traffic addressed to the virtual router MAC address. However, because the IP address associated with the master router is not physically located on the backup router, the backup router cannot reply to TCP/IP messages (such as pings) sent to the virtual router.

Fully-Redundant VRRP Network

You can use two or more VRRP-enabled switches to provide a fully-redundant VRRP configuration on your network. Figure 15-3 shows a fully-redundant VRRP configuration.

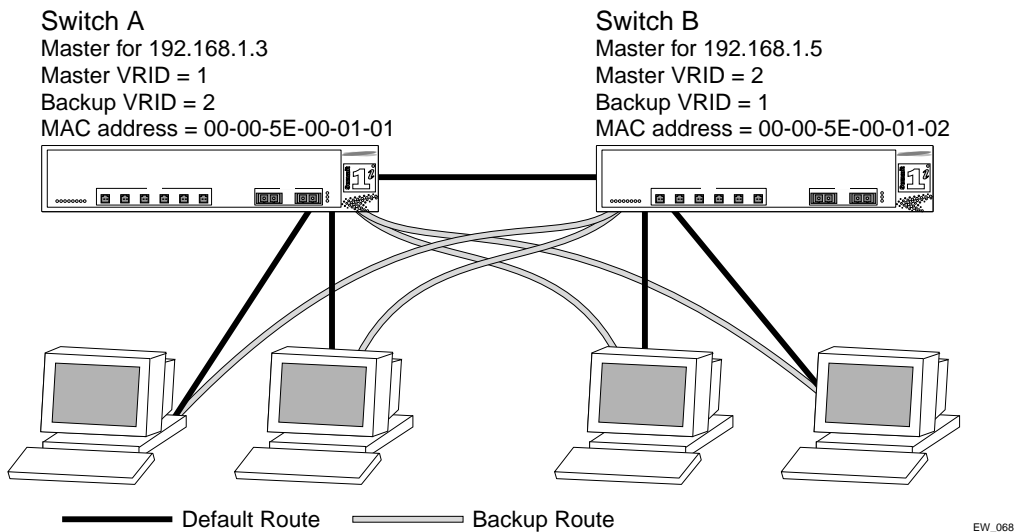


Figure 15-3: Fully-redundant VRRP configuration

In Figure 15-3, switch A is configured as follows:

- IP address 192.168.1.3
- Master router for VRID 1
- Backup router for VRID 2
- MAC address 00-00-5E-00-01-01

Switch B is configured as follows:

- IP address 192.168.1.5
- Master router for VRID 2
- Backup router for VRID 1
- MAC address 00-00-5E-00-01-02

Both virtual routers are simultaneously operational. The traffic load from the four hosts is split between them. Host 1 and host 2 are configured to use VRID 1 on switch A as their default gateway. Host 3 and host 4 are configured to use VRID 2 on switch B as their default gateway. In the event that either switch fails, the backup router configured is standing by to resume normal operation.

VRRP Configuration Parameters

Table 15-2 lists the parameters that are configured on a VRRP router.

Table 15-2: VRRP Configuration Parameters

Parameter	Description
vrid	Virtual router identifier. Configured item in the range of 1- 255. This parameter has no default value.
priority	Priority value to be used by this VRRP router in the master election process. A value of 255 is reserved for a router that is configured with the virtual router IP address. A value of 0 is reserved for the master router to indicate it is releasing responsibility for the virtual router. The range is 1 - 254. The default value is 100.
ip_address	One or more IP addresses associated with this virtual router. This parameter has no default value.
advertisement_interval	Time interval between advertisements, in seconds. The range is 1 - 255. The default value is 1 second.
skew_time	Time to skew master_down_interval, in seconds. This value is calculated as $((256 - \text{priority}) / 256)$.
master_down_interval	Time interval for backup router to declare master down, in seconds. This value is calculated as $((3 * \text{advertisement_interval}) + \text{skew_time})$.
preempt_mode	Controls whether a higher priority backup router preempts a lower priority master. A value of true allows preemption. A value of false prohibits preemption. The default setting is true.



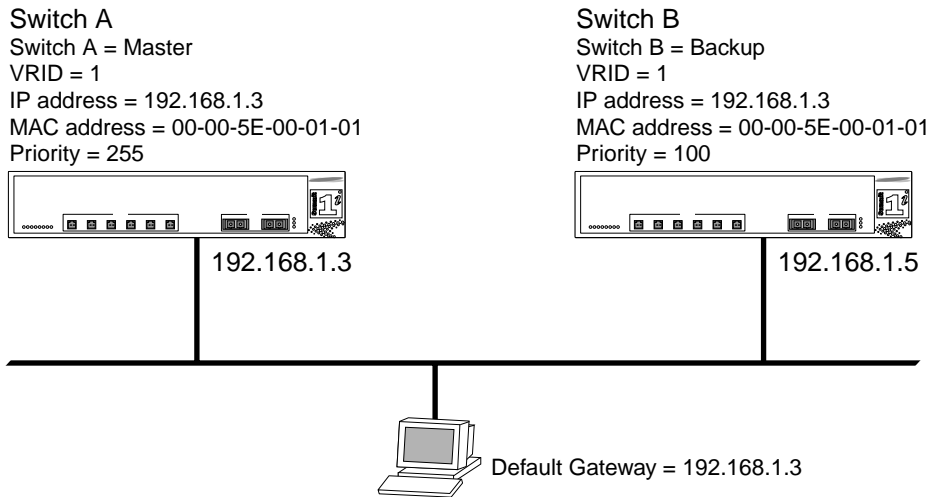
Note: The router that owns the virtual router IP address always preempts, independent of the setting of this parameter.

VRRP Examples

This section provides the configuration syntax for the two VRRP networks discussed in this chapter.

Configuring the Simple VRRP Network

The following illustration shows the simple VRRP network described in Figure 15-2.



The configuration commands for switch A are as follows:

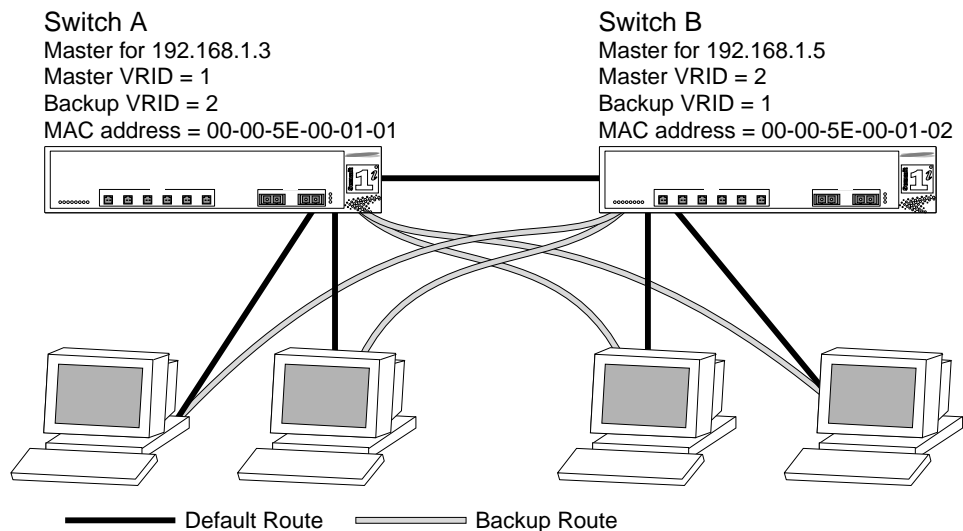
```
config vlan vlan1 ipaddress 192.168.1.3
config vrrp add vlan vlan2
config vrrp vlan vlan1 add master vrid 1 192.168.1.3
enable vrrp
```

The configuration commands for switch B are as follows:

```
config vlan vlan1 ipaddress 192.168.1.5
config vrrp add vlan vlan1
config vrrp vlan vlan1 add backup vrid 1 192.168.1.3
enable vrrp
```

Configuring the Fully-Redundant VRRP Network

The following illustration shows the fully-redundant VRRP network configuration described in Figure 15-3.



EW_068

The configuration commands for switch A are as follows:

```
config vlan vlan1 ipaddress 192.168.1.3
config vrrp vlan vlan1 add master vrid 1 192.168.1.3
config vrrp vlan vlan1 add backup vrid 2 192.168.1.5
config vrrp add vlan vlan1
enable vrrp
```

The configuration commands for switch B are as follows:

```
config vlan vlan1 ipaddress 192.168.1.5
config vrrp vlan vlan1 add master vrid 2 192.168.1.5
config vrrp vlan vlan1 add backup vrid 1 192.168.1.3
config vrrp add vlan vlan1
enable vrrp
```


16

IP Unicast Routing

This chapter describes the following topics:

- Overview of IP Unicast Routing on page 16-2
- Proxy ARP on page 16-6
- Relative Route Priorities on page 16-8
- Configuring IP Unicast Routing on page 16-8
- Routing Configuration Example on page 16-9
- IP Multinetting on page 16-11
- Configuring DHCP/BOOTP Relay on page 16-14
- UDP-Forwarding on page 16-15
- VLAN Aggregation on page 16-17

This chapter assumes that you are already familiar with IP unicast routing. If not, refer to the following publications for additional information:

- RFC 1256 — *ICMP Router Discovery Messages*
- RFC 1812 — *Requirements for IP Version 4 Routers*



Note: For more information on interior gateway protocols, refer to Chapter 17. For information on exterior gateway protocols, refer to Chapter 18.

Overview of IP Unicast Routing

The switch provides full layer 3, IP unicast routing. It exchanges routing information with other routers on the network using either the Routing Information Protocol (RIP) or the Open Shortest Path First (OSPF) protocol. The switch dynamically builds and maintains a routing table, and determines the best path for each of its routes.

Each host using the IP unicast routing functionality of the switch must have a unique IP address assigned. In addition, the default gateway assigned to the host must be the IP address of the router interface.

Router Interfaces

The routing software and hardware routes IP traffic between router interfaces. A router interface is simply a VLAN that has an IP address assigned to it.

As you create VLANs with IP addresses belonging to different IP subnets, you can also choose to route between the VLANs. Both the VLAN switching and IP routing function occur within the switch.



Note: Each IP address and mask assigned to a VLAN must represent a unique IP subnet. You cannot configure the same IP address and subnet on different VLANs.

In Figure 16-1, a BlackDiamond switch is depicted with two VLANs defined; *Finance* and *Personnel*. All ports on slots 1 and 3 are assigned to *Finance*; all ports on slots 2 and 4 are assigned to *Personnel*. *Finance* belongs to the IP network 192.207.35.0; the router interface for *Finance* is assigned the IP address 192.206.35.1. *Personnel* belongs to the IP network 192.207.36.0; its router interface is assigned IP address 192.207.36.1. Traffic within each VLAN is switched using the Ethernet MAC addresses. Traffic between the two VLANs is routed using the IP addresses.

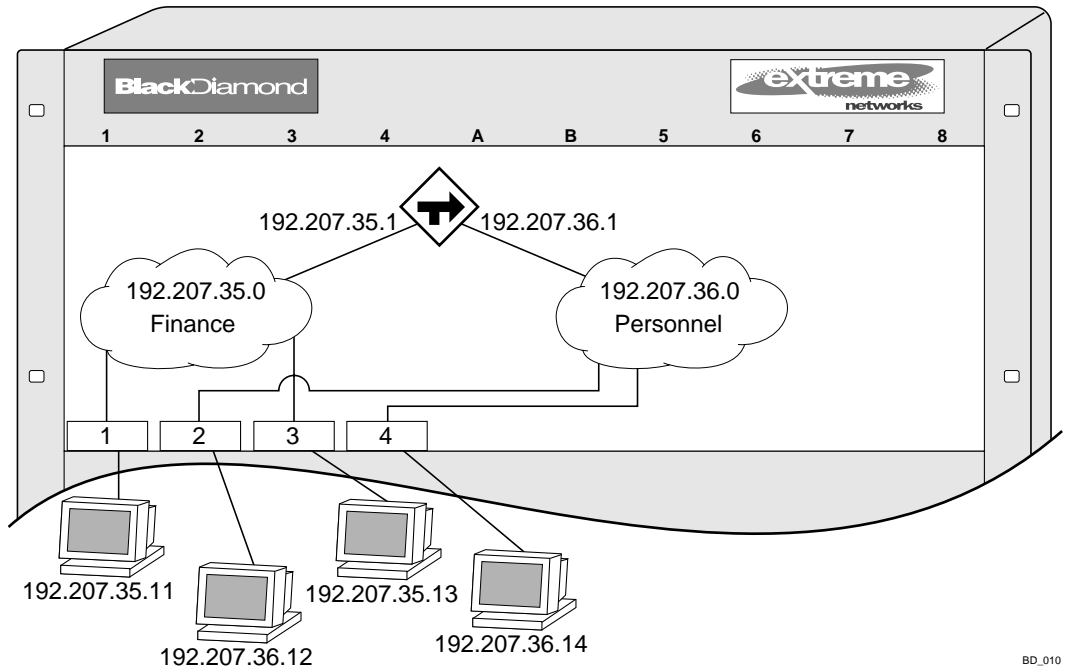


Figure 16-1: Routing between VLANs

Populating the Routing Table

The switch maintains an IP routing table for both network routes and host routes. The table is populated from the following sources:

- Dynamically, by way of routing protocol packets or by ICMP redirects exchanged with other routers
- Statically, by way of routes entered by the administrator
 - Default routes, configured by the administrator
 - Locally, by way of interface addresses assigned to the system
 - By other static routes, as configured by the administrator



Note: If you define a default route, and subsequently delete the VLAN on the subnet associated with the default route, the invalid default route entry remains. You must manually delete the configured default route.

Dynamic Routes

Dynamic routes are typically learned by way of RIP or OSPF. Routers that use RIP or OSPF exchange information in their routing tables in the form of advertisements. Using dynamic routes, the routing table contains only networks that are reachable.

Dynamic routes are aged out of the table when an update for the network is not received for a period of time, as determined by the routing protocol.

Static Routes

Static routes are manually entered into the routing table. Static routes are used to reach networks not advertised by routers.

Static routes can also be used for security reasons, to control which routes you want advertised by the router. You can decide if you want all static routes to be advertised, using one of the following commands:

```
[enable | disable] rip export static
[enable | disable] ospf export static
```

The default setting is disabled. Static routes are never aged out of the routing table.

A static route must be associated with a valid IP subnet. An IP subnet is associated with a single VLAN by its IP address and subnet mask. If the VLAN is subsequently deleted, the static route entries using that subnet must be deleted manually.

Multiple Routes

When there are multiple, conflicting choices of a route to a particular destination, the router picks the route with the longest matching network mask. If these are still equal, the router picks the route using the following criteria (in the order specified):

- Directly attached network interfaces
- ICMP redirects
- Static routes
- Directly attached network interfaces that are not active.



Note: If you define multiple default routes, the route that has the lowest metric is used. If multiple default routes have the same lowest metric, the system picks one of the routes.

You can also configure *blackhole* routes — traffic to these destinations is silently dropped.

IP Route Sharing

IP route sharing allows multiple equal-cost routes to be used concurrently. IP route sharing can be used with static routes or with OSPF routes. In OSPF, this capability is referred to as *equal cost multipath* (ECMP) routing. To use IP route sharing, use the following command:

```
enable iproute sharing
```

Next, configure static routes and/or OSPF as you would normally. ExtremeWare supports unlimited route sharing across static routes and up to eight ECMP routes for OSPF.

Route sharing is useful only in instances where you are constrained for bandwidth. This is typically not the case using Extreme switches. Using route sharing makes router troubleshooting more difficult because of the complexity in predicting the path over which the traffic will travel.

Route Maps

Route maps for IP routing can be configured based on the route origin. When routes are added to the IP routing table from various source, the route map configured for the origin of the route is applied to the route. After matching on specified characteristics, the characteristics for the route can be modified using the route maps. The characteristics that can be matched and modified are dependent on the origin of the route. Route maps for IP routing can be dynamically changed. In the case of direct and static route origins, the changes are reflected immediately. In the case of routes that are sourced from other origin, the changes are reflected within 30 seconds.

To configure route maps for IP routing, use the following command:

```
config iproute route-map [bgp | direct | e-bgp | i-bgp | ospf |
ospf-extern1 | ospf-extern2 | ospf-inter | ospf-intra | rip | static]
[<route map> | none]
```

To view the route maps for IP routing, use the following command:

```
show iproute route-map
```

Subnet-Directed Broadcast Forwarding

You can enable or disable the hardware forwarding of subnet-directed broadcast IP packets. This allows the switch to forward subnet-directed broadcast packets at wire-speed.

To enable or disable hardware forwarding, use the following command:

```
[enable | disable] ipforwarding fast-direct-broadcast [vlan <vlan_name>]
```

The entries are added to the IP forwarding table as standard entries and you can view them using the `show ipfdb` command.

You can also configure the VLAN router interface to either forward and process all subnet-directed broadcast packets, or to simply forward these packets after they have been added to the IP forwarding database. The latter option allows you to improve CPU forwarding performance by having upper layers, such as UDP and TCP, ignore broadcast packet processing (for example, if the packets have IP-options configured).

To enable or disable broadcast packet processing, use the following command:

```
[enable | disable] ipforwarding ignore-broadcast vlan <vlan_name>
```

Using these commands together, you can achieve a 30-50% reduction in system processing cycles in forwarding subnet-directed broadcast traffic on a BlackDiamond switch, and a 100% reduction on the Alpine and Summit switches.



Note: Although forwarding performance is improved in the BlackDiamond switch, the CPU continues to observe the subnet-directed broadcast packets and does not ignore such packets when traversing modules in a BlackDiamond. Only “i” series modules support this command on the BlackDiamond switch.

Proxy ARP

Proxy Address Resolution Protocol (ARP) was first invented so that ARP-capable devices could respond to ARP Request packets on behalf of ARP-incapable devices. Proxy ARP can also be used to achieve router redundancy and simplify IP client configuration. The switch supports proxy ARP for this type of network configuration. The section describes some example of how to use proxy ARP with the switch.

ARP-Incapable Devices

To configure the switch to respond to ARP Requests on behalf of devices that are incapable of doing so, you must configure the IP address and MAC address of the ARP-incapable device using the use the following command:

```
config iparp add proxy <ipaddress> {<mask>} <mac_address> {always}
```

Once configured, the system responds to ARP Requests on behalf of the device as long as the following conditions are satisfied:

- The valid IP ARP Request is received on a router interface.
- The target IP address matches the IP address configured in the proxy ARP table.
- The proxy ARP table entry indicates that the system should always answer this ARP Request, regardless of the ingress VLAN (the `always` parameter must be applied).

Once all the proxy ARP conditions are met, the switch formulates an ARP Response using the configured MAC address in the packet.

Proxy ARP Between Subnets

In some networks, it is desirable to configure the IP host with a wider subnet than the actual subnet mask of the segment. Proxy ARP can be used so that the router answers ARP Requests for devices outside of the subnet. As a result, the host communicates as if all devices are local. In reality, communication with devices outside of the subnet are proxied by the router.

For example, an IP host is configured with a class B address of 100.101.102.103 and a mask of 255.255.0.0. The switch is configured with the IP address 100.101.102.1 and a mask of 255.255.255.0. The switch is also configured with a proxy ARP entry of IP address 100.101.0.0 and mask 255.255.0.0, *without* the `always` parameter.

When the IP host tries to communicate with the host at address 100.101.45.67, the IP hosts communicates as if the two hosts are on the same subnet, and sends out an IP ARP Request. The switch answers on behalf of the device at address 100.101.45.67, using its own MAC address. All subsequent data packets from 100.101.102.103 are sent to the switch, and the switch routes the packets to 100.101.45.67.

Relative Route Priorities

Table 16-1 lists the relative priorities assigned to routes depending upon the learned source of the route.



Note: Although these priorities can be changed, do not attempt any manipulation unless you are expertly familiar with the possible consequences.

Table 16-1: Relative Route Priorities

Route Origin	Priority
Direct	10
BlackHole	50
Static	1100
ICMP	1200
OSPFIntra	2200
OSPFInter	2300
RIP	2400
OSPFEextern1	3200
OSPFEextern2	3300
BOOTP	5000

To change the relative route priority, use the following command:

```
config iproute priority [rip | bootp | icmp | static | ospf-intra |
ospf-inter | ospf-as-external | ospf-extern1 | ospf-extern2] <priority>
```

Configuring IP Unicast Routing

This section describes the commands associated with configuring IP unicast routing on the switch. To configure routing, follow these steps:

- 1 Create and configure two or more VLANs.
- 2 Assign each VLAN that will be using routing an IP address using the following command:

```
config vlan <name> ipaddress <ipaddress> {<mask>}
```

Ensure that each VLAN has a unique IP address.

3 Configure a default route using the following command:

```
config iproute add default <gateway> {<metric>}
```

Default routes are used when the router has no other dynamic or static route to the requested destination.

4 Turn on IP routing for one or all VLANs using the following command:

```
enable ipforwarding {vlan <name>}
```

5 Turn on RIP or OSPF using one of the following commands:

```
enable rip
enable ospf
```

Verifying the IP Unicast Routing Configuration

Use the `show iproute` command to display the current configuration of IP unicast routing for the switch, and for each VLAN. The `show iproute` command displays the currently configured routes, and includes how each route was learned.

Additional verification commands include:

- `show iparp` — Displays the IP ARP table of the system.
- `show ipfdb` — Displays the hosts that have been transmitting or receiving packets, and the port and VLAN for each host.
- `show ipconfig` — Displays configuration information for one or more VLANs.

Routing Configuration Example

Figure 16-2 illustrates a BlackDiamond switch that has three VLANs defined as follows:

- *Finance*
 - Protocol-sensitive VLAN using the IP protocol.
 - All ports on slots 1 and 3 have been assigned.
 - IP address 192.207.35.1.
- *Personnel*
 - Protocol-sensitive VLAN using the IP protocol.

- All ports on slots 2 and 4 have been assigned.
- IP address 192.207.36.1.
- *MyCompany*
 - Port-based VLAN.
 - All ports on slots 1 through 4 have been assigned.

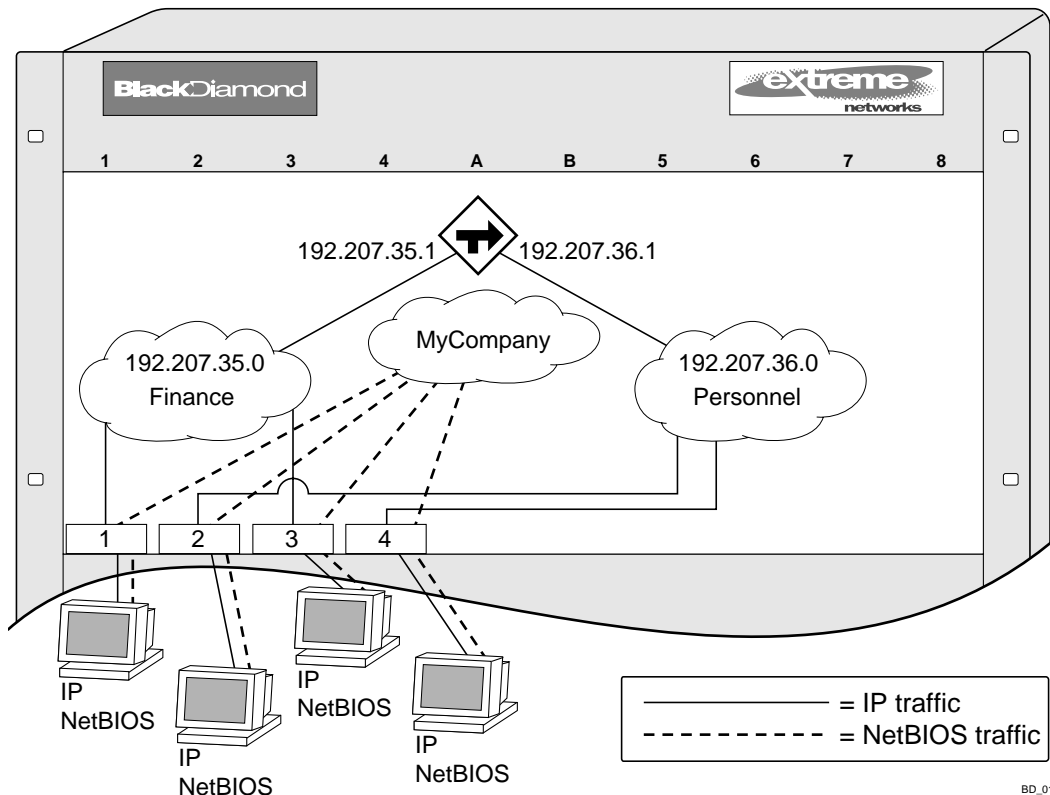


Figure 16-2: Unicast routing configuration example

The stations connected to the system generate a combination of IP traffic and NetBIOS traffic. The IP traffic is filtered by the protocol-sensitive VLANs. All other traffic is directed to the VLAN *MyCompany*.

In this configuration, all IP traffic from stations connected to slots 1 and 3 have access to the router by way of the VLAN *Finance*. Ports on slots 2 and 4 reach the router by way of the VLAN *Personnel*. All other traffic (NetBIOS) is part of the VLAN *MyCompany*.

The example in Figure 16-2 is configured as follows:

```
create vlan Finance
create vlan Personnel
create vlan MyCompany

config Finance protocol ip
config Personnel protocol ip

config Finance add port 1:*,3:*
config Personnel add port 2:*,4:*
config MyCompany add port all

config Finance ipaddress 192.207.35.1
config Personnel ipaddress 192.207.36.1

config rip add vlan Finance
config rip add vlan Personnel

enable ipforwarding
enable rip
```

IP Multinetting

IP multinetting is used in many legacy IP networks when there is need to overlap multiple subnets onto the same physical segment. Though it can be a critical element in a transition strategy, due to the additional constraints introduced in troubleshooting and bandwidth, it is recommended that multinetting be used as a transitional tactic, and not as a long-term network design strategy.

On the switch, each subnet is represented by a different VLAN, and each of those VLANs has its own IP address. All of the VLANs share the same physical port(s). The switch routes IP traffic from one subnet to another, all within the same physical port(s).

The following rules and comments apply when you are configuring IP multinetting:

- Multiple VLANs share the same physical ports; each of the VLANs is configured with an IP address.
- A maximum of four subnets (or VLANs) on multinetted ports is recommended.
- All VLANs used in the multinetting application must share the same port assignment.

- One VLAN is configured to use an IP protocol filter. This is considered the "primary" VLAN interface for the multinetted group.
- The "secondary" multinetted VLANs can be exported using the `export direct` command.
- The FDB aging timer is automatically set to 3,000 seconds (50 minutes).
- If you are using a UDP or DHCP relay function, only the "primary" VLAN that is configured with the IP protocol filter is capable of servicing these requests.
- The VLAN *default* should not be used for multinetting.

IP Multinetting Operation

To use IP multinetting, follow these steps:

- 1 Select a slot (modular switches only) and port on which IP multinetting is to run.
For example, slot 1, port 2 on a modular switch, or port 2 on a stand-alone switch.

- 2 Remove the port from the default VLAN using the following command:

```
config default delete port 1:2 (modular switch)
```

or

```
config default delete port 2 (stand-alone switch)
```

- 3 Create a dummy protocol by using the following command:

```
create protocol mnet
```

- 4 Create the multinetted subnets using the following commands:

```
create vlan net21
```

```
create vlan net22
```

- 5 Assign IP addresses to the net VLANs using the following commands:

```
config net21 ipaddress 123.45.21.1 255.255.255.0
```

```
config net22 ipaddress 192.24.22.1 255.255.255.0
```

- 6 Assign one of the subnets to the IP protocol using the following command:

```
config net21 protocol ip
```

- 7 Assign the other subnets to the dummy protocol using the following command:

```
config net22 protocol mnet
```


- 8** Assign the subnets to a physical port using the following commands:

```
config net21 add port 1:2
config net22 add port 1:2
```

- 9** Enable IP forwarding on the subnets using the following command:

```
enable ipforwarding
```

- 10** Enable IP multinetting using the following command:

```
enable multinetting
```

- 11** If you are using RIP, disable RIP on the dummy VLANs using the following command:

```
config rip delete net22
```



Note: Multineted VLAN groups must contain identical port assignments.

IP Multinetting Examples

The following example configures a modular switch to have one multineted segment (slot 5, port 5) that contains three subnets (192.67.34.0, 192.67.35.0, and 192.67.37.0).

```
config default delete port 5:5
create protocol mnet
create vlan net34
create vlan net35
create vlan net37
config net34 ipaddress 192.67.34.1
config net35 ipaddress 192.67.35.1
config net37 ipaddress 192.67.37.1
config net34 protocol ip
config net35 protocol mnet
config net37 protocol mnet
config net34 add port 5:5
config net35 add port 5:5
config net37 add port 5:5
enable ipforwarding
enable multinetting
```

The following example configures a modular switch to have one multineted segment (slot 5: port 5) that contains three subnets (192.67.34.0, 192.67.35.0, and 192.67.37.0). It also configures a second multineted segment consisting of two subnets (192.67.36.0 and

192.99.45.0). The second multinetted segment spans three ports (slot1:port 8, slot2:port 9, and slot3:port 10). RIP is enabled on both multinetted segments.

```
config default delete port 5:5
create protocol mnet
create vlan net34
create vlan net35
create vlan net37
config net34 ipaddress 192.67.34.1
config net35 ipaddress 192.67.35.1
config net37 ipaddress 192.67.37.1
config net34 protocol ip
config net35 protocol mnet
config net37 protocol mnet
config net34 add port 5:5
config net35 add port 5:5
config net37 add port 5:5
config default delete port 1:8, 2:9, 3:10
create vlan net36
create vlan net45
config net36 ipaddress 192.67.36.1
config net45 ipaddress 192.99.45.1
config net36 protocol ip
config net45 protocol mnet
config net36 add port 1:8, 2:9, 3:10
config net45 add port 1:8, 2:9, 3:10
config rip add vlan net34
config rip add vlan net36
enable rip
enable ipforwarding
enable multinetting
```

Configuring DHCP/BOOTP Relay

Once IP unicast routing is configured, you can configure the switch to forward Dynamic Host Configuration Protocol (DHCP) or BOOTP requests coming from clients on subnets being serviced by the switch and going to hosts on different subnets. This feature can be used in various applications, including DHCP services between Windows NT servers and clients running Windows 95. To configure the relay function, follow these steps:

- 1 Configure VLANs and IP unicast routing.
- 2 Enable the DHCP or BOOTP relay function, using the following command:


```
enable bootprelay
```
- 3 Configure the addresses to which DHCP or BOOTP requests should be directed, using the following command:


```
config bootprelay add <ipaddress>
```

To delete an entry, use the following command:

```
config bootprelay delete {<ipaddress> | all}
```

Verifying the DHCP/BOOTP Relay Configuration

To verify the DHCP/BOOTP relay configuration, use the following command:

```
show ipconfig
```

This command displays the configuration of the BOOTP relay service, and the addresses that are currently configured.

UDP-Forwarding

UDP-forwarding is a flexible and generalized routing utility for handling the directed forwarding of broadcast UDP packets. UDP-forwarding allows applications, such as multiple DHCP relay services from differing sets of VLANs, to be directed to different DHCP servers. The following rules apply to UDP broadcast packets handled by this feature:

- If the UDP profile includes BOOTP or DHCP, it is handled according to guidelines in RFC 1542.
- If the UDP profile includes other types of traffic, these packets have the IP destination address modified as configured, and changes are made to the IP and UDP checksums and decrements to the TTL field, as appropriate.

If the UDP-forwarding is used for BOOTP or DHCP forwarding purposes, do not configure or use the existing `bootprelay` function. However, if the previous `bootprelay` functions are adequate, you may continue to use them.



Note: UDP-forwarding only works across a layer 3 boundary.

Configuring UDP-Forwarding

To configure UDP-forwarding, the first thing you must do is create a UDP-forward destination profile. The profile describes the types of UDP packets (by port number) that are used, and where they are to be forwarded. You must give the profile a unique name, in the same manner as a VLAN, protocol filter, or Spanning Tree Domain.

Next, configure a VLAN to make use of the UDP-forwarding profile. As a result, all incoming traffic from the VLAN that matches the UDP profile is handled as specified in the UDP-forwarding profile.

A maximum of ten UDP-forwarding profiles can be defined. Each named profile may contain a maximum of eight “rules” defining the UDP port, and destination IP address or VLAN. A VLAN can make use of a single UDP-forwarding profile. UDP packets directed toward a VLAN use an all-ones broadcast on that VLAN.

UDP-Forwarding Example

In this example, the VLAN *Marketing* and the VLAN *Operations* are pointed toward a specific backbone DHCP server (with IP address 10.1.1.1) and a backup server (with IP address 10.1.1.2). Additionally, the VLAN *LabUser* is configured to use any responding DHCP server on a separate VLAN called *LabSvrs*.

The commands for this configuration are as follows:

```
create udp-profile backbonedhcp
create udp-profile labdhcp
config backbonedhcp add 67 ipaddress 10.1.1.1
config backbonedhcp add 67 ipaddress 10.1.1.2
config labdhcp add 67 vlan labsvrs
config marketing udp-profile backbonedhcp
config operations udp-profile backbonedhcp
config labuser udp-profile labdhcp
```

ICMP Packet Processing

As ICMP packets are routed or generated, you can take various actions to control distribution. For ICMP packets typically generated or observed as part of the routing function, you can assert control on a per-type, per-VLAN basis. You would alter the default settings for security reasons: to restrict the success of tools that can be used to find an important application, host, or topology information. The controls include the disabling of transmitting ICMP messages associated with unreachable,

port-unreachables, time-exceeded, parameter-problems, redirects, time-stamp, and address-mask requests.

For ICMP packets that are typically routed, you can apply access lists to restrict forwarding behavior. Access lists are described in Chapter 8.

VLAN Aggregation

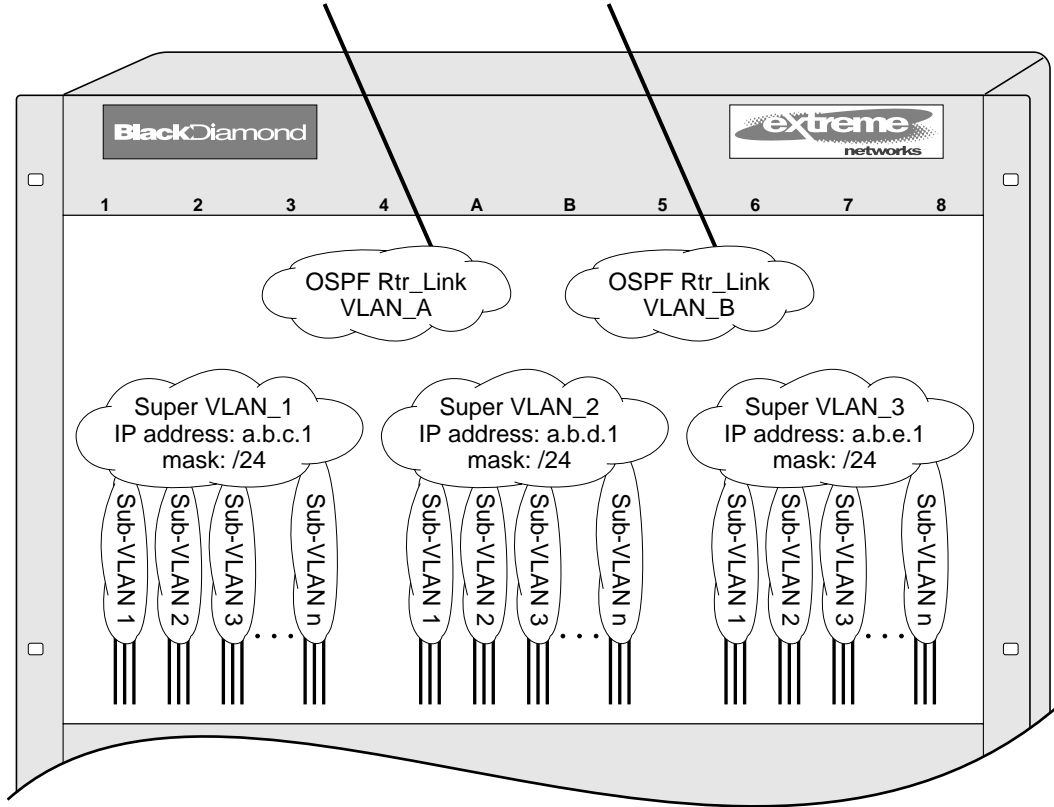
VLAN aggregation is an ExtremeWare feature aimed primarily at service providers. The purpose of VLAN aggregation is to increase the efficiency of IP address space usage. It does this by allowing clients within the same IP subnet to use different broadcast domains while still using the same default router.

Using VLAN aggregation, a *super-VLAN* is defined with the desired IP address, but without any member ports (unless it is running ESRP). The sub-VLANs use the IP address of the super-VLAN as the default router address. Groups of clients are then assigned to sub-VLANs that have no IP address, but are members of the super-VLAN. In addition, clients can be informally allocated any valid IP addresses within the subnet. Optionally, you can prevent communication between sub-VLANs for isolation purposes. As a result, sub-VLANs can be quite small, but allow for growth without re-defining subnet boundaries.

Without using VLAN aggregation, each VLAN has a default router address, and you need to use large subnet masks. The result of this is more unused IP address space.

Multiple secondary IP addresses can be assigned to the super-VLAN. These IP addresses are *only* used to respond to ICMP ping packets to verify connectivity.

Figure 16-3 illustrates VLAN aggregation.



EW_026

Figure 16-3: VLAN aggregation

In Figure 16-3, all stations are configured to use the address 10.3.2.1 for the default router.

VLAN Aggregation Properties

VLAN aggregation is a very specific application, and the following properties apply to its operation:

- All broadcast and unknown traffic remains local to the sub-VLAN and does not cross the sub-VLAN boundary. All traffic within the sub-VLAN is switched by the

sub-VLAN, allowing traffic separation between sub-VLANs (while using the same default router address among the sub-VLANs).

- Hosts are located on the sub-VLAN. Each host can assume any IP address within the address range of the super-VLAN router interface. Hosts on the sub-VLAN are expected to have the same network mask as the super-VLAN and have their default router set to the IP address of the super-VLAN.
- All traffic (IP unicast and IP multicast) between sub-VLANs is routed through the super-VLAN. For example, no ICMP redirects are generated for traffic between sub-VLANs, because the super-VLAN is responsible for sub-VLAN routing. Unicast IP traffic across the sub-VLANs is facilitated by the automatic addition of an ARP entry (similar to a proxy ARP entry) when a sub-VLAN is added to a super-VLAN. This feature can be disabled for security purposes.
- IP multicast traffic between sub-VLANs is routed when an IP multicast routing protocol is enabled on the super-VLAN.

VLAN Aggregation Limitations

The following limitations apply to VLAN aggregation:

- No additional routers may be located in a sub-VLAN. This feature is only applicable for “leaves” of a network.
- A sub-VLAN cannot be a super-VLAN, and vice-versa.
- Sub-VLANs are not assigned an IP address.
- Typically, a super-VLAN has no ports associated with it, except in the case of running ESRP.
- If a client is moved from one sub-VLAN to another, you must clear the IP ARP cache at the client and the switch to resume communication.

VLAN Aggregation SubVLAN Address Range Checking

You can configure subVLAN address ranges on each subVLAN to prohibit the entry of IP addresses from hosts outside of the configured range.

To configure a subVLAN range, use the following command:

```
config vlan <name> subvlan-address-range <ip_address> - <ip_address>
```

To remove a subVLAN address range, use the following command:

```
config vlan <name> subvlan-address-range 0.0.0.0 - 0.0.0.0
```

To view the subVLAN address range, use the following command:

```
show vlan <vlan_name>
```

There is no error checking to prevent the configuration of overlapping subVLAN address ranges between multiple subVLANs. Doing so can result in unexpected behavior of ARP within the superVLAN and associated subVLANs.

Isolation Option for Communication Between Sub-VLANs

To facilitate communication between sub-VLANs, by default, an entry is made in the IP ARP table of the super-VLAN that performs a proxy ARP function. This allows clients on one sub-VLAN to communicate with clients on another sub-VLAN. In certain circumstances, intra-sub-VLAN communication may not be desired for isolation reasons.

To prevent normal communication between sub-VLANs, disable the automatic addition of the IP ARP entries on the super-VLAN using the following command:

```
disable subvlan-proxy-arp vlan <super-vlan name>
```



Note: The isolation option works for normal, dynamic, ARP-based client communication.

VLAN Aggregation Example

The follow example illustrates how to configure VLAN aggregation. The VLAN *vsuper* is created as a super-VLAN, and sub-VLANs, *vsub1*, *vsub2*, and *vsub3* are added to it.

- 1 Create and assign an IP address to a VLAN designated as the super-VLAN. This VLAN should have no member ports. Be sure to enable IP forwarding, and any desired routing protocol, on the switch.

```
create vlan vsuper
config vsuper ipaddress 192.201.3.1/24
enable ipforwarding
enable ospf
config ospf add vsuper
```


2 Create and add ports to the sub-VLANs.

```
create vlan vsub1
con vsub1 add port 10-12
create vlan vsub2
config vsub2 add po 13-15
create vlan vsub3
config vsub3 add po 16-18
```

3 Configure the super-VLAN by adding the sub-VLANs.

```
config vsuper add subvlan vsub1
config vsuper add subvlan vsub2
config vsuper add subvlan vsub3
```

4 Optionally, disable communication among sub-VLANs.

```
disable subvlan-proxy-arp <super-VLAN name>
```

Verifying the VLAN Aggregation Configuration

The following commands can be used to verify proper VLAN aggregation configuration.

- `show vlan` — Indicates the membership of a sub-VLANs in a super-VLAN.
- `show iparp` — Indicates an ARP entry that contains sub-VLAN information. Communication with a client on a sub-VLAN must have occurred in order for an entry to be made in the ARP table.



Interior Gateway Routing Protocols

This chapter describes the following topics:

- Overview on page 17-2
- Overview of RIP on page 17-3
- Overview of OSPF on page 17-5
- Route Re-Distribution on page 17-12
- RIP Configuration Example on page 17-15
- Configuring OSPF on page 17-17
- OSPF Configuration Example on page 17-18
- Displaying OSPF Settings on page 17-21

This chapter assumes that you are already familiar with IP unicast routing. If not, refer to the following publications for additional information:

- RFC 1058 — *Routing Information Protocol (RIP)*
- RFC 1723 — *RIP Version 2*
- RFC 2178 — *OSPF Version 2*
- *Interconnections: Bridges and Routers*
by Radia Perlman
ISBN 0-201-56332-0
Published by Addison-Wesley Publishing Company

Overview

The switch supports the use of two interior gateway protocols (IGPs); the Routing Information Protocol (RIP) and the Open Shortest Path First (OSPF) protocol for IP unicast routing.

RIP is a distance-vector protocol, based on the Bellman-Ford (or distance-vector) algorithm. The distance-vector algorithm has been in use for many years, and is widely deployed and understood.

OSPF is a link-state protocol, based on the Dijkstra link-state algorithm. OSPF is a newer Interior Gateway Protocol (IGP), and solves a number of problems associated with using RIP on today's complex networks.



Note: Both RIP and OSPF can be enabled on a single VLAN.

RIP Versus OSPF

The distinction between RIP and OSPF lies in the fundamental differences between distance-vector protocols and link-state protocols. Using a distance-vector protocol, each router creates a unique routing table from summarized information obtained from neighboring routers. Using a link-state protocol, every router maintains an identical routing table created from information obtained from all routers in the autonomous system. Each router builds a shortest path tree, using itself as the root. The link-state protocol ensures that updates sent to neighboring routers are acknowledged by the neighbors, verifying that all routers have a consistent network map.

The biggest advantage of using RIP is that it is relatively simple to understand and implement, and it has been the *de facto* routing standard for many years.

RIP has a number of limitations that can cause problems in large networks, including:

- A limit of 15 hops between the source and destination networks.
- A large amount of bandwidth taken up by periodic broadcasts of the entire routing table.
- Slow convergence.
- Routing decisions based on hop count; no concept of link costs or delay.
- Flat networks; no concept of areas or boundaries.

OSPF offers many advantages over RIP, including:

- No limitation on hop count.
- Route updates multicast only when changes occur.
- Faster convergence.
- Support for load balancing to multiple routers based on the actual cost of the link.
- Support for hierarchical topologies where the network is divided into areas.

The details of RIP and OSPF are explained later in this chapter.

Overview of RIP

RIP is an Interior Gateway Protocol (IGP) first used in computer routing in the Advanced Research Projects Agency Network (ARPANet) as early as 1969. It is primarily intended for use in homogeneous networks of moderate size.

To determine the best path to a distant network, a router using RIP always selects the path that has the least number of hops. Each router that data must traverse is considered to be one hop.

Routing Table

The routing table in a router using RIP contains an entry for every known destination network. Each routing table entry contains the following information:

- IP address of the destination network
- Metric (hop count) to the destination network
- IP address of the next router
- Timer that tracks the amount of time since the entry was last updated

The router exchanges an update message with each neighbor every 30 seconds (default value), or if there is a change to the overall routed topology (also called *triggered updates*). If a router does not receive an update message from its neighbor within the route timeout period (180 seconds by default), the router assumes the connection between it and its neighbor is no longer available.

Split Horizon

Split horizon is a scheme for avoiding problems caused by including routes in updates sent to the router from which the route was learned. Split horizon omits routes learned from a neighbor in updates sent to that neighbor.

Poison Reverse

Like split horizon, poison reverse is a scheme for eliminating the possibility of loops in the routed topology. In this case, a router advertises a route over the same interface that supplied the route, but the route uses a hop count of 16, defining it as unreachable.

Triggered Updates

Triggered updates occur whenever a router changes the metric for a route, and it is required to send an update message immediately, even if it is not yet time for a regular update message to be sent. This will generally result in faster convergence, but may also result in more RIP-related traffic.

Route Advertisement of VLANs

VLANs that are configured with an IP address, but are configured to not route IP or are not configured to run RIP, do not have their subnets advertised by RIP. Only those VLANs that are configured with an IP address and are configured to route IP and run RIP have their subnets advertised.

RIP Version 1 Versus RIP Version 2

A new version of RIP, called RIP version 2, expands the functionality of RIP version 1 to include:

- Variable-Length Subnet Masks (VLSMs).
- Support for next-hop addresses, which allows for optimization of routes in certain environments.
- Multicasting.

RIP version 2 packets can be multicast instead of being broadcast, reducing the load on hosts that do not support routing protocols.



Note: If you are using RIP with supernetting/Classless Inter-Domain Routing (CIDR), you must use RIPv2 only. In addition, RIP route aggregation must be turned off.

Overview of OSPF

OSPF is a link-state protocol that distributes routing information between routers belonging to a single IP domain, also known as an *autonomous system* (AS). In a link-state routing protocol, each router maintains a database describing the topology of the autonomous system. Each participating router has an identical database maintained from the perspective of that router.

From the link-state database (LSDB), each router constructs a tree of shortest paths, using itself as the root. The shortest path tree provides the route to each destination in the autonomous system. When several equal-cost routes to a destination exist, traffic can be distributed among them. The cost of a route is described by a single metric.

Link-State Database

Upon initialization, each router transmits a link-state advertisement (LSA) on each of its interfaces. LSAs are collected by each router and entered into the LSDB of each router. Once all LSAs are received, the router uses the LSDB to calculate the best routes for use in the IP routing table. OSPF uses flooding to distribute LSAs between routers. Any change in routing information is sent to all of the routers in the network. All routers within an area have the exact same LSDB. Table 17-1 describes LSA type numbers.

Table 17-1: LSA Type Numbers

Type Number	Description
1	Router LSA
2	Network LSA
3	Summary LSA
4	AS summary LSA
5	AS external LSA
7	NSSA external LSA
9	Link local
10	Area scoping

Table 17-1: LSA Type Numbers

Type Number	Description
11	AS scoping

Database Overflow

The OSPF database overflow feature allows you to limit the size of the LSDB and to maintain a consistent LSDB across all the routers in the domain, which ensures that all routers have a consistent view of the network.

Consistency is achieved by:

- Limiting the number of external LSAs in the database of each router.
- Ensuring that all routers have identical LSAs.

To configure OSPF database overflow, use the following command:

```
config ospf ase-limit <number> {timeout <seconds>}
```

where:

- `<number>` – Specifies the number of external LSAs (excluding the default LSAs) that the system supports before it goes into overflow state. A limit value of zero disables the functionality.

When the LSDB size limit is reached, OSPF database overflow flushes LSAs from the LSDB. OSPF database overflow flushes the same LSAs from all the routers, which maintains consistency.

- `timeout` – Specifies the timeout, in seconds, after which the system ceases to be in overflow state. A timeout value of zero leaves the system in overflow state until OSPF is disabled and re-enabled.

Opaque LSAs

Opaque LSAs are a generic OSPF mechanism used to carry auxiliary information in the OSPF database. Opaque LSAs are most commonly used to support OSPF traffic engineering.

Normally, support for opaque LSAs is auto-negotiated between OSPF neighbors. In the event that you experience interoperability problems, you can disable opaque LSAs across the entire system using the following command:


```
disable ospf capability opaque-lsa
```

To re-enable opaque LSAs across the entire system, use the following command:

```
enable ospf capability opaque-lsa
```

If your network uses opaque LSAs, we recommend that all routers on your OSPF network support opaque LSAs. Routers that do not support opaque LSAs do not store or flood them. At minimum a well-interconnected subsection of your OSPF network needs to support opaque LSAs to maintain reliability of their transmission.

On an OSPF broadcast network, the designated router (DR) must support opaque LSAs or none of the other routers on that broadcast network will reliably receive them. You can use the OSPF priority feature to give preference to an opaque-capable router, so that it becomes the elected DR.

For transmission to continue reliably across the network, the backup designated router (BDR) must also support opaque LSAs.



Note: Opaque LSAs are supported in ExtremeWare version 6.2 and above.

Areas

OSPF allows parts of a network to be grouped together into *areas*. The topology within an area is hidden from the rest of the autonomous system. Hiding this information enables a significant reduction in LSA traffic, and reduces the computations needed to maintain the LSDB. Routing within the area is determined only by the topology of the area.

The three types of routers defined by OSPF are as follows:

- **Internal Router (IR)**
An internal router has all of its interfaces within the same area.
- **Area Border Router (ABR)**
An ABR has interfaces in multiple areas. It is responsible for exchanging summary advertisements with other ABRs. You can create a maximum of 7 non-zero areas.
- **Autonomous System Border Router (ASBR)**
An ASBR acts as a gateway between OSPF and other routing protocols, or other autonomous systems.

Backbone Area (Area 0.0.0.0)

Any OSPF network that contains more than one area is required to have an area configured as area 0.0.0.0, also called the *backbone*. All areas in an autonomous system must be connected to the backbone. When designing networks, you should start with area 0.0.0.0, and then expand into other areas.

The backbone allows summary information to be exchanged between ABRs. Every ABR hears the area summaries from all other ABRs. The ABR then forms a picture of the distance to all networks outside of its area by examining the collected advertisements, and adding in the backbone distance to each advertising router.

When a VLAN is configured to run OSPF, you must configure the area for the VLAN. If you want to configure the VLAN to be part of a different OSPF area, use the following command:

```
config ospf vlan <name> area <areaid>
```

If this is the first instance of the OSPF area being used, you must create the area first using the following command:

```
create ospf area <areaid>
```

Stub Areas

OSPF allows certain areas to be configured as *stub areas*. A stub area is connected to only one other area. The area that connects to a stub area can be the backbone area. External route information is not distributed into stub areas. Stub areas are used to reduce memory consumption and computation requirements on OSPF routers.

Not-So-Stubby-Areas (NSSA)

NSSAs are similar to the existing OSPF stub area configuration option, but have the following two additional capabilities:

- External routes originating from an ASBR connected to the NSSA can be advertised within the NSSA.
- External routes originating from the NSSA can be propagated to other areas, including the backbone area.

The CLI command to control the NSSA function is similar to the command used for configuring a stub area, as follows:

```
config ospf area <area_id> nssa {summary | nosummary} stub-default-cost
<cost> {translate}
```

The `translate` option determines whether type 7 LSAs are translated into type 5 LSAs. When configuring an OSPF area as an NSSA, the `translate` should only be used on NSSA border routers, where translation is to be enforced. If `translate` is not used on any NSSA border router in a NSSA, one of the ABRs for that NSSA is elected to perform translation (as indicated in the NSSA specification). The option should not be used on NSSA internal routers. Doing so inhibits correct operation of the election algorithm.

Normal Area

A normal area is an area that is not:

- Area 0.
- Stub area.
- NSSA.

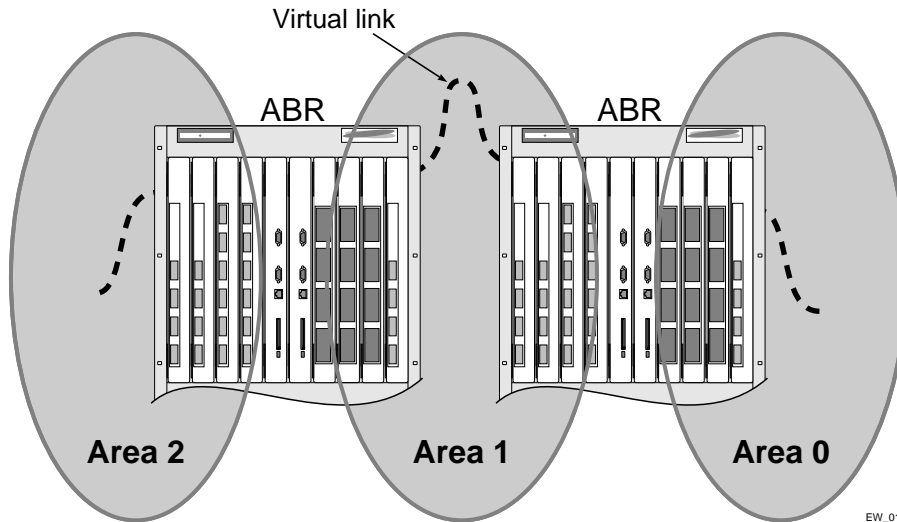
Virtual links can be configured through normal areas. External routes can be distributed into normal areas.

Virtual Links

In the situation when a new area is introduced that does not have a direct physical attachment to the backbone, a *virtual link* is used. A virtual link provides a logical path between the ABR of the disconnected area and the ABR of the normal area that connects to the backbone. A virtual link must be established between two ABRs that have a common area, with one ABR connected to the backbone. Figure 17-1 illustrates a virtual link.



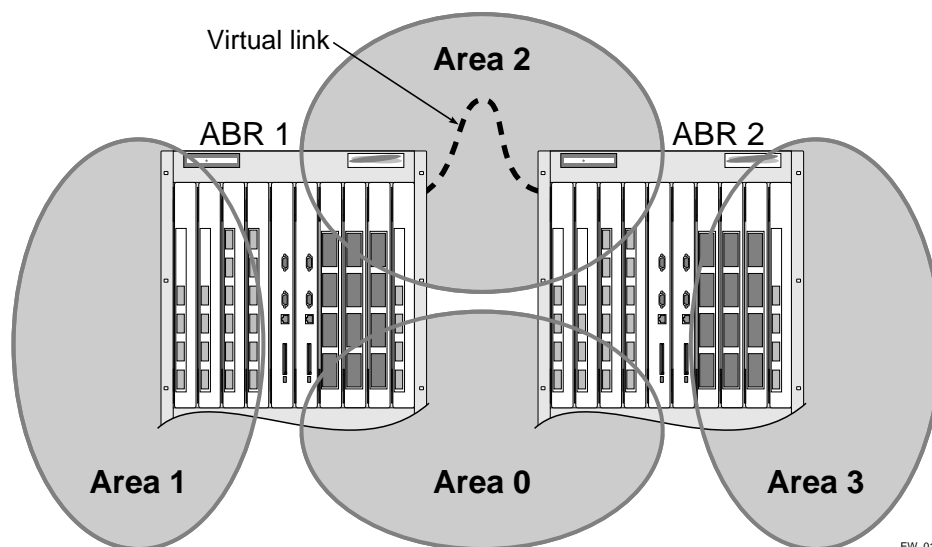
Note: Virtual links can not be configured through a stub or NSSA area.



EW_016

Figure 17-1: Virtual link using Area 1 as a transit area

Virtual links are also used to repair a discontinuous backbone area. For example, in Figure 17-2, if the connection between ABR1 and the backbone fails, the connection using ABR2 provides redundancy so that the discontinuous area can continue to communicate with the backbone using the virtual link.



EW_017

Figure 17-2: Virtual link providing redundancy

Point-to-Point Support

You can manually configure the OSPF link type for a VLAN. Table 17-2 describes the link types.

Table 17-2: OSPF Link Types

Link Type	Number of Routers	Description
Auto	Varies	ExtremeWare automatically determines the OSPF link type based on the interface type. This is the default setting.
Broadcast	Any	Routers must elect a designated router (DR) and a backup designated router (BDR) during synchronization. Ethernet is an example of a broadcast link.

Table 17-2: OSPF Link Types (continued)

Link Type	Number of Routers	Description
Point-to-point	Up to 2	Synchronizes faster than a broadcast link because routers do not elect a DR or BDR. Does not operate with more than two routers on the same VLAN. PPP is an example of a point-to-point link. An OSPF point-to-point link supports only zero to two OSPF routers and does not elect a DR or BDR. If you have three or more routers on the VLAN, OSPF will fail to synchronize if the neighbor is not configured.



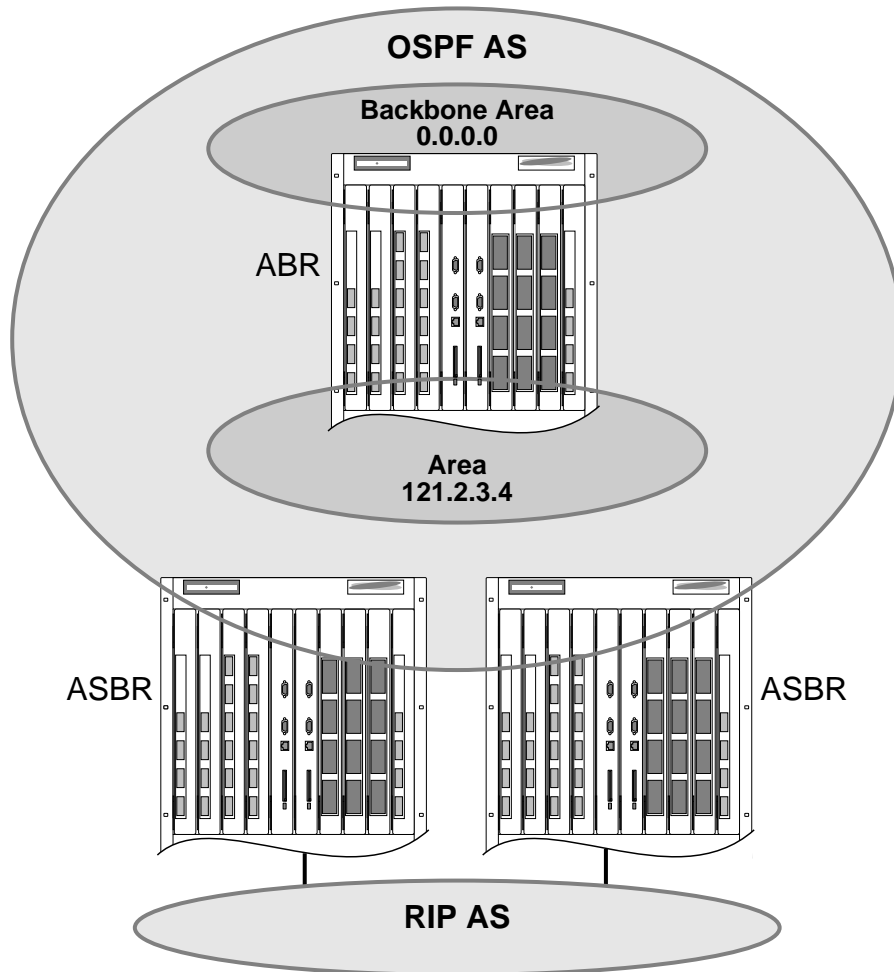
Note: The number of routers in an OSPF point-to-point link is determined per-VLAN, not per-link.



Note: All routers in the VLAN must have the same OSPF link type. If there is a mismatch, OSPF attempts to operate, but may not be reliable.

Route Re-Distribution

Both RIP and OSPF can be enabled simultaneously on the switch. Route re-distribution allows the switch to exchange routes, including static routes, between the two routing protocols. Figure 17-3 is an example of route re-distribution between an OSPF autonomous system and a RIP autonomous system.



EW_019

Figure 17-3: Route re-distribution

Configuring Route Re-Distribution

Exporting routes from OSPF to RIP, and from RIP to OSPF, are discreet configuration functions. To run OSPF and RIP simultaneously, you must first configure both protocols and then verify the independent operation of each. Then you can configure the routes to export from OSPF to RIP and the routes to export from RIP to OSPF.

Re-Distributing Routes into OSPF

Enable or disable the exporting of RIP, static, and direct (interface) routes to OSPF using the following commands:

```
enable ospf export [static | rip | direct] [cost <metric> [ase-type-1 |  
ase-type-2] {tag <number>} | <route map>]
```

```
disable ospf export [static | rip | direct]
```

These commands enable or disable the exporting of RIP, static, and direct routes by way of LSA to other OSPF routers as AS-external type 1 or type 2 routes. The default setting is disabled.

The cost metric is inserted for all RIP-learned, static, and direct routes injected into OSPF. If the cost metric is set to 0, the cost is inserted from the route. The tag value is used only by special routing applications. Use 0 if you do not have specific requirements for using a tag. The tag value in this instance has no relationship with 802.1Q VLAN tagging.

The same cost, cost-type, and tag values can be inserted for all the export routes, or route maps can be used for selective insertion. When a route map is associated with the export command, the route map is applied on every exported route. The exported routes can also be filtered using route maps.

Enable or disable the export of virtual IP addresses to other OSPF routers using the following commands:

```
enable ospf export vip [cost <metric> [ase-type-1 | ase-type-2] {tag  
<number>} | <route map>]
```

```
disable ospf export vip
```

Verify the configuration using the command:

```
show ospf
```

Previous Release Issues with OSPF Re-Distribution

In versions of ExtremeWare prior to release 6.0, direct routes corresponding to the interfaces on which RIP was enabled were exported into OSPF as part of RIP routes, using the command `enable ospf export rip`. Using ExtremeWare 6.0 and above, you

must configure ExtremeWare to export these direct routes to OSPF. You can use an access profile to filter unnecessary direct routes, using the command `config ospf direct-filter [<access-profile> | none]`.

Re-Distributing Routes into RIP

Enable or disable the exporting of static, direct, and OSPF-learned routes into the RIP domain using the following commands:

```
enable rip export [static | direct | ospf | ospf-intra | ospf-inter |
ospf-extern1 | ospf-extern2 | vip] cost <metric> tag <number>
```

```
disable rip export [static | direct | ospf | ospf-intra | ospf-inter |
ospf-extern1 | ospf-extern2 | vip]
```

These commands enable or disable the exporting of static, direct, and OSPF-learned routes into the RIP domain. You can choose which types of OSPF routes are injected, or you can simply choose `ospf`, which will inject all learned OSPF routes regardless of type. The default setting is disabled.

OSPF Timers and Authentication

Configuring OSPF timers and authentication on a per-area basis is a shorthand for applying the timers and authentication to each VLAN in the area at the time of configuration. If you add more VLANs to the area, you must configure the timers and authentication for the new VLANs explicitly.

RIP Configuration Example

Figure 17-4 illustrates a BlackDiamond switch that has three VLANs defined as follows:

- *Finance*
 - Protocol-sensitive VLAN using the IP protocol.
 - All ports on slots 1 and 3 have been assigned.
 - IP address 192.207.35.1.
- *Personnel*
 - Protocol-sensitive VLAN using the IP protocol.
 - All ports on slots 2 and 4 have been assigned.

- IP address 192.207.36.1.
- *MyCompany*
 - Port-based VLAN.
 - All ports on slots 1 through 4 have been assigned.

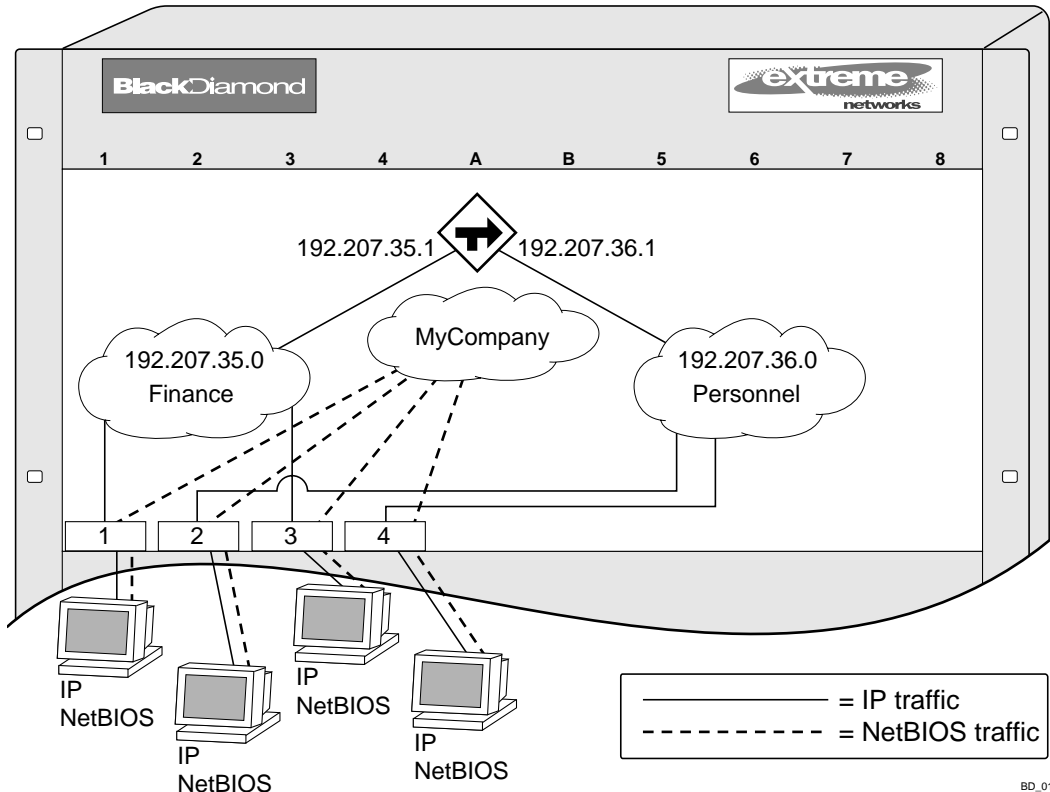


Figure 17-4: RIP configuration example

The stations connected to the system generate a combination of IP traffic and NetBIOS traffic. The IP traffic is filtered by the protocol-sensitive VLANs. All other traffic is directed to the VLAN *MyCompany*.

In this configuration, all IP traffic from stations connected to slots 1 and 3 have access to the router by way of the VLAN *Finance*. Ports on slots 2 and 4 reach the router by way of the VLAN *Personnel*. All other traffic (NetBIOS) is part of the VLAN *MyCompany*.

The example in Figure 17-4 is configured as follows:

```

create vlan Finance
create vlan Personnel
create vlan MyCompany

config Finance protocol ip
config Personnel protocol ip

config Finance add port 1:*,3:*
config Personnel add port 2:*,4:*
config MyCompany add port all

config Finance ipaddress 192.207.35.1
config Personnel ipaddress 192.207.36.1

enable ipforwarding
config rip add vlan all
enable rip

```

Configuring OSPF

Each switch that is configured to run OSPF must have a unique router ID. It is recommended that you manually set the router ID of the switches participating in OSPF, instead of having the switch automatically choose its router ID based on the highest interface IP address. Not performing this configuration in larger, dynamic environments could result in an older link state database remaining in use.

Configuring OSPF Wait Interval

ExtremeWare allows you to configure the OSPF wait interval, rather than using the router dead interval.



Caution: Do not configure OSPF timers unless you are comfortable exceeding OSPF specifications. Non-standard settings might not be reliable under all circumstances.

To specify the timer intervals, use the following command:

```

config ospf vlan <vlan> timer <rxmtinterval> <transitdelay>
<hellointerval> <routerdeadinterval> [<waitinterval>]

```

You can configure the following parameters:

- Retransmit interval (RxmtInterval) — The length of time that the router waits before retransmitting an LSA that is not acknowledged. If you set an interval that is too short, unnecessary retransmissions will result. The default value is 5 seconds.
- Transit delay (TransitDelay) — The length of time it takes to transmit an LSA packet over the interface. The transit delay must be greater than 0.
- Hello interval (HelloInterval) — The interval at which routers send hello packets. Smaller times allow routers to discover each other more quickly, but also increase network traffic. The default value is 10 seconds.
- Dead router wait interval (RouterDeadInterval) — The interval after which a neighboring router is declared down due to the fact that hello packets are no longer received from the neighbor. This interval should be a multiple of the hello interval. The default value is 40 seconds.
- Router wait interval (WaitInterval) — The interval between the interface coming up and the election of the DR and BDR. This interval should be greater than the hello interval. If it is close to the hello interval, the network synchronizes very quickly, but might not elect the correct DR or BDR. The default value is equal to the dead router wait interval.



Note: The OSPF standard specifies that wait times are equal to the dead router wait interval.

OSPF Configuration Example

Figure 17-5 is an example of an autonomous system using OSPF routers. The details of this network follow.

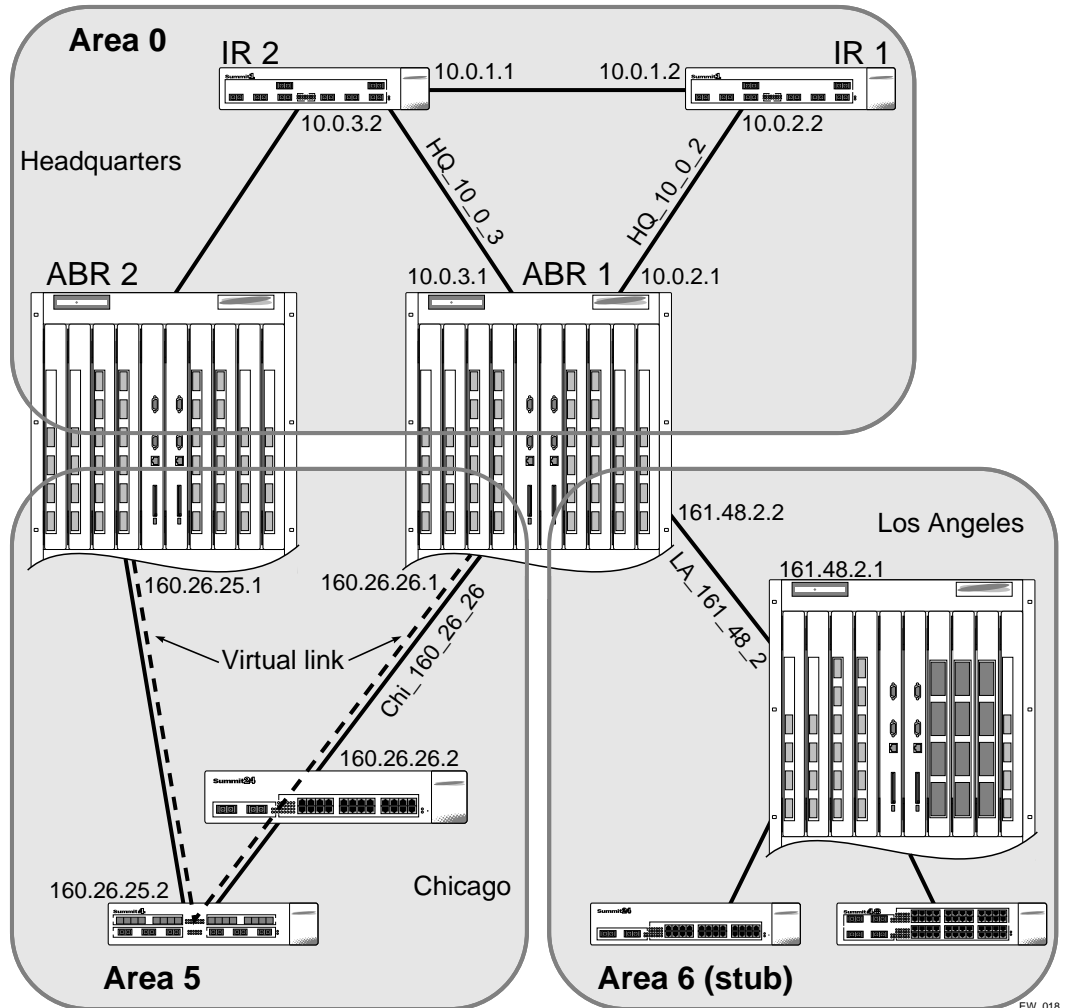


Figure 17-5: OSPF configuration example

Area 0 is the backbone area. It is located at the headquarters and has the following characteristics:

- Two internal routers (IR1 and IR2)
- Two area border routers (ABR1 and ABR2)

- Network number 10.0.x.x
- Two identified VLANs (HQ_10_0_2 and HQ_10_0_3)

Area 5 is connected to the backbone area by way of ABR1 and ABR2. It is located in Chicago and has the following characteristics:

- Network number 160.26.x.x
- One identified VLAN (Chi_160_26_26)
- Two internal routers

Area 6 is a stub area connected to the backbone by way of ABR1. It is located in Los Angeles and has the following characteristics:

- Network number 161.48.x.x
- One identified VLAN (LA_161_48_2)
- Three internal routers
- Uses default routes for inter-area routing

Two router configurations for the example in Figure 17-5 are provided in the following section.

Configuration for ABR1

The router labeled ABR1 has the following configuration:

```
create vlan HQ_10_0_2
create vlan HQ_10_0_3
create vlan LA_161_48_2
create vlan Chi_160_26_26

config vlan HQ_10_0_2 ipaddress 10.0.2.1 255.255.255.0
config vlan HQ_10_0_3 ipaddress 10.0.3.1 255.255.255.0
config vlan LA_161_48_26 ipaddress 161.48.2.26 255.255.255.0
config vlan Chi_160_26_26 ipaddress 160.26.2.1 255.255.255.0

create ospf area 0.0.0.5
create ospf area 0.0.0.6

enable ipforwarding

config ospf area 0.0.0.6 stub nosummary stub-default-cost 10
```

```

config ospf vlan LA_161_48_2 area 0.0.0.6
config ospf vlan Chi_160_26_26 area 0.0.0.5
config ospf add vlan all

enable ospf

```

Configuration for IR1

The router labeled IR1 has the following configuration:

```

config vlan HQ_10_0_1 ipaddress 10.0.1.2 255.255.255.0
config vlan HQ_10_0_2 ipaddress 10.0.2.2 255.255.255.0
config ospf add vlan all
enable ipforwarding
enable ospf

```

Displaying OSPF Settings

There are a number of commands you can use to display settings for OSPF. To show global OSPF information, use the `show ospfshow ospf` command with no options.

To display information about one or all OSPF areas, use the following command:

```
show ospf area {<area-id> | detail}
```

The `detail` option displays information about all OSPF areas in a detail format.

To display information about OSPF interfaces for an area, a VLAN, or for all interfaces, use the following command:

```
show ospf interfaces {vlan <name> | area <areaid> | detail}
```

The `detail` option displays information about all OSPF interfaces in a detail format.

OSPF LSD Display

ExtremeWare provides several filtering criteria for the `show ospf lsdb` command. You can specify multiple search criteria and only results matching all of the criteria are displayed. This allows you to control the displayed entries in large routing tables.

To display the current link-state database, use the following command:

```
show ospf lsdb [detail | summary | stats] area [all | <areaid>[</len>]]  
lstype [all | as-external | external-type7 | network | router |  
summary-asb | summary-net] [lsid <id>[</len>]] [routerid <id>[</len>]]
```

The `detail` option displays all fields of matching LSAs in a multi-line format. The `summary` option displays several important fields of matching LSAs, one line per LSA. The `stats` option displays the number of matching LSAs, but not any of their contents. If not specified, the default is to display in the summary format.

A common use of this command is to omit all optional parameters, resulting in the following shortened form:

```
show ospf lsdb
```

The shortened form displays all areas and all types in a summary format.

18

Exterior Gateway Routing Protocols

This chapter covers the following topics:

- Overview on page 18-2
- BGP Attributes on page 18-2
- BGP Communities on page 18-3
- BGP Features on page 18-3

This chapter describes how to configure the Border Gateway Protocol (BGP), an exterior routing protocol available on the switch.

For more information on BGP, refer to the following documents:

- RFC 1771 – *Border Gateway Protocol version 4 (BGP-4)*
- RFC 1965 – *Autonomous System Confederations for BGP*
- RFC 1966 – *BGP Route Reflection*
- RFC 1997 – *BGP Communities Attribute*
- RFC 1745 – *BGP/OSPF Interaction*



Note: ExtremeWare supports BGP version 4 only.

Overview

BGP is an exterior routing protocol that was developed for use in TCP/IP networks. The primary function of BGP is to allow different autonomous systems (ASs) to exchange network reachability information.

An autonomous system is a set of routers that are under a single technical administration. This set of routers uses a different routing protocol (such as OSPF) for intra-AS routing. One or more routers in the AS are configured to be border routers, exchanging information with other border routers (in different autonomous systems) on behalf of all of the intra-AS routers.

BGP can be used as an exterior gateway protocol (E-BGP), or it can be used within an AS as an interior gateway protocol (I-BGP).

BGP Attributes

The following well-known BGP attributes are supported by the switch:

- **Origin** – Defines the origin of the route. Possible values are IGP, EGP, and incomplete.
- **AS_Path** – The list of ASs that are traversed for this route.
- **Next_hop** – The IP address of the next hop BGP router to reach the destination listed in the NLRI field.
- **Multi_Exist_Discriminator** – Used to select a particular border router in another AS when multiple border routers exist.
- **Local_Preference** – Used to advertise this router's degree of preference to other routers within the AS.
- **Atomic_aggregate** – Indicates that the sending border router is used a route aggregate prefix in the route update.
- **Aggregator** – Identifies the BGP router AS number and IP address that performed route aggregation.
- **Community** – Identifies a group of destinations that share one or more common attributes.
- **Cluster_ID** – Specifies a 4 byte field used by a route reflector to recognize updates from other route reflectors in the same cluster.

BGP Communities

A BGP community is a group of BGP destinations that require common handling. ExtremeWare supports the following well-known BGP community attributes:

- no-export
- no-advertise
- internet

BGP Features

This section describes the following BGP features supported by ExtremeWare:

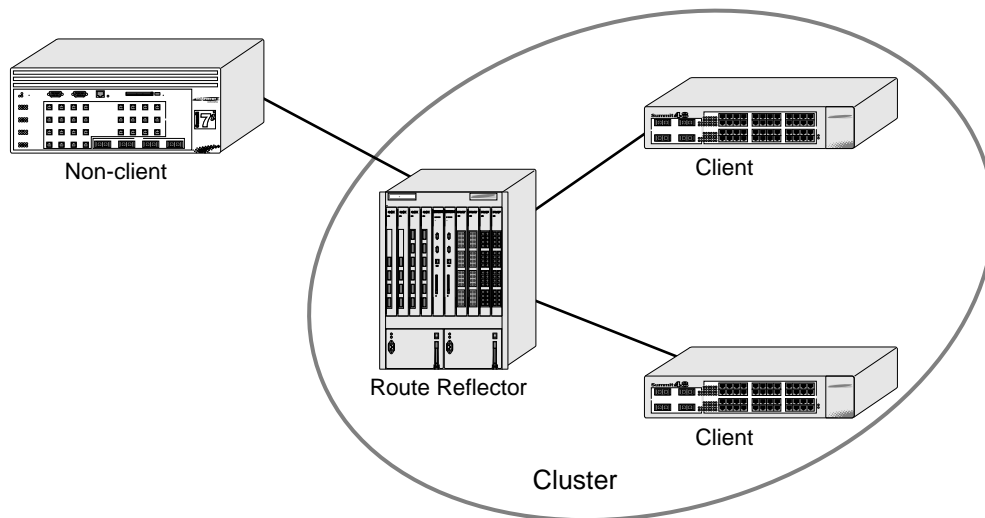
- Route Reflectors
- Route Confederations
- Route Aggregation
- IGP Synchronization
- Using the Loopback Interface
- BGP Peer Groups

Route Reflectors

Another way to overcome the difficulties of creating a fully-meshed AS is to use *route reflectors*. Route reflectors allow a single router to serve as a central routing point for the AS or sub-AS.

A *cluster* is formed by the route reflector and its client routers. Peer routers that are not part of the cluster must be fully meshed according to the rules of BGP.

A BGP cluster, including the route reflector and its clients, is shown in Figure 18-1.



EW_042

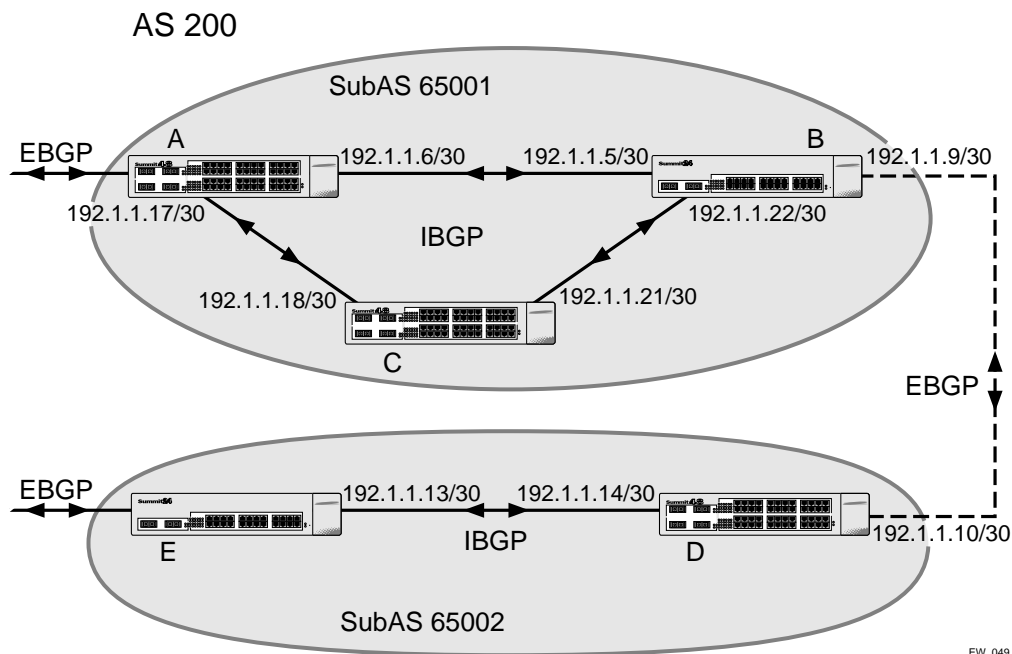
Figure 18-1: Route reflectors

Route Confederations

BGP requires networks to use a fully-meshed router configuration. This requirement does not scale well, especially when BGP is used as an interior gateway protocol. One way to reduce the size of a fully-meshed AS is to divide the AS into multiple sub-autonomous systems and group them into a *routing confederation*. Within the confederation, each sub-AS must be fully-meshed. The confederation is advertised to other networks as a single AS.

Route Confederation Example

Figure 18-2 shows an example of a confederation.



EW_049

Figure 18-2: Routing confederation

In this example, AS 200 has five BGP speakers. Without a confederation, BGP would require that the routes in AS 200 be fully meshed. Using the confederation, AS 200 is split into two sub-ASs: AS65001 and AS65002. Each sub-AS is fully meshed, and IBGP is running among its members. EBGP is used between sub-AS 65001 and sub-AS 65002. Router B and router D are EBGP peers. EBGP is also used between the confederation and outside ASs.

To configure router A, use the following commands:

```
create vlan ab
config vlan ab add port 1
config vlan ab ipaddress 192.1.1.6/30
enable ipforwarding vlan ab
config ospf add vlan ab area 0.0.0.0
```

```
create vlan ac
config vlan ac add port 2
config vlan ac ipaddress 192.1.1.17/30
enable ipforwarding vlan ac
config ospf add vlan ac area 0.0.0.0
```

```
disable bgp
config bgp as-number 65001
config bgp routerid 192.1.1.17
config bgp confederation-id 200
enable bgp
```

```
create bgp neighbor 192.1.1.5 as-number remote-AS-number 65001
create bgp neighbor 192.1.1.18 as-number remote-AS-number 65001
enable bgp neighbor all
```

To configure router B, use the following commands:

```
create vlan ba
config vlan ba add port 1
config vlan ba ipaddress 192.1.1.5/30
enable ipforwarding vlan ba
config ospf add vlan ba area 0.0.0.0
```

```
create vlan bc
config vlan bc add port 2
config vlan bc ipaddress 192.1.1.22/30
enable ipforwarding vlan bc
config ospf add vlan bc area 0.0.0.0
```

```
create vlan bd
config vlan bd add port 3
config vlan bd ipaddress 192.1.1.9/30
enable ipforwarding vlan bd
config ospf add vlan bd area 0.0.0.0
```

```
disable bgp
config bgp as-number 65001
config bgp routerid 192.1.1.22
config bgp confederation-id 200
enable bgp
```

```
create bgp neighbor 192.1.1.6 as-number remote-AS-number 65001
```

```
create bgp neighbor 192.1.1.21 as-number remote-AS-number 65001
create bgp neighbor 192.1.1.10 as-number remote-AS-number 65002
enable bgp neighbor all
config bgp add confederation-peer sub-AS-number 65002
```

To configure router C, use the following commands:

```
create vlan ca
config vlan ca add port 1
config vlan ca ipaddress 192.1.1.18/30
enable ipforwarding vlan ca
config ospf add vlan ca area 0.0.0.0
```

```
create vlan cb
config vlan cb add port 2
config vlan cb ipaddress 192.1.1.21/30
enable ipforwarding vlan cb
config ospf add vlan cb area 0.0.0.0
```

```
disable bgp
config bgp as-number 65001
config bgp routerid 192.1.1.21
config bgp confederation-id 200
enable bgp
```

```
create bgp neighbor 192.1.1.22 as-number remote-AS-number 65001
create bgp neighbor 192.1.1.17 as-number remote-AS-number 65001
enable bgp neighbor all
```

To configure router D, use the following commands:

```
create vlan db
config vlan db add port 1
config vlan db ipaddress 192.1.1.10/30
enable ipforwarding vlan db
config ospf add vlan db area 0.0.0.0
```

```
create vlan de
config vlan de add port 2
config vlan de ipaddress 192.1.1.14/30
enable ipforwarding vlan de
config ospf add vlan de area 0.0.0.0
```

```
disable bgp
```

```
config bgp as-number 65002
config bgp routerid 192.1.1.14
config bgp confederation-id 200
enable bgp

create bgp neighbor 192.1.1.9 as-number remote-AS-number 65001
create bgp neighbor 192.1.1.13 as-number remote-AS-number 65002
enable bgp neighbor all
config bgp add confederation-peer sub-AS-number 65001
```

To configure router E, use the following commands:

```
create vlan ed
config vlan ed add port 1
config vlan ed ipaddress 192.1.1.13/30
enable ipforwarding vlan ed
config ospf add vlan ed area 0.0.0.0

disable bgp
config bgp as-number 65002
config bgp routerid 192.1.1.13
config bgp confederation-id 200
enable bgp

create bgp neighbor 192.1.1.14 as-number remote-AS-number 65002
enable bgp neighbor 192.1.1.14
```

Route Aggregation

Route aggregation is the process of combining the characteristics of several routes so that they are advertised as a single route. Aggregation reduces the amount of information that a BGP speaker must store and exchange with other BGP speakers. Reducing the information that is stored and exchanged also reduces the size of the routing table.

Using Route Aggregation

To use BGP route aggregation, follow these steps:

- 1 Enable aggregation using the following command:

```
enable bgp aggregation
```


2 Create an aggregate route using the following commands:

```
config bgp add aggregate-address <ipaddress>/<masklength> {as-set}
{summary-only} {advertise-route-map <route-map>} {attribute-route-map
<route-map>}
```

IGP Synchronization

You can configure an AS to be a transit AS, so that it can pass traffic through from one AS to a third AS. When you configure a transit AS, it is important that the routes advertised by BGP are consistent with the routes that are available within the AS using its interior gateway protocol. To ensure consistency, BGP should be synchronized with the IGP used within the AS. This will ensure that the routes advertised by BGP are, in fact, reachable within the AS. IGP synchronization is enabled by default.

Using the Loopback Interface

If you are using BGP as your interior gateway protocol, you may decide to advertise the interface as available, regardless of the status of any particular interface. The loopback interface can also be used for EBGp multihop. Using the loopback interface eliminates multiple, unnecessary route changes.

BGP Peer Groups

You can use BGP peer groups to group together up to 128 BGP neighbors. All neighbors within the peer group inherit the parameters of the BGP peer group. The following mandatory parameters are shared by all neighbors in a peer group:

- remote AS
- source-interface
- out-nlri-filter
- out-asp-path-filter
- out-route-map
- send-community
- next-hop-self

Each BGP peer group is assigned a unique name when it is created. To create or delete peer groups, use the following command:

```
[create | delete] bgp peer-group <peer group>
```

Changes made to the parameters of a peer group are applied to all neighbors in the peer group. Modifying the following parameters will automatically disable and enable the neighbors before changes take effect:

- remote-as
- timer
- source-interface
- soft-in-reset
- password

Adding Neighbors to a BGP Peer Group

To create a new neighbor and add it to a BGP peer group, use the following command:

```
create bgp neighbor <ip address> peer-group <peer group> {multi-hop}
```

The new neighbor is created as part of the peer group and inherits all of the existing parameters of the peer group. The peer group must have remote AS configured.

To add an existing neighbor to a peer group, use the following command:

```
config bgp neighbor [<ip address> | all] peer-group <peer group>
{acquire-all}
```

If you do not specify acquire-all, only the mandatory parameters are inherited from the peer group. If you specify acquire-all, all of the parameters of the peer group are inherited. This command disables the neighbor before adding it to the peer group.

BGP Route Selection

BGP will select routes based on the following precedence (from highest to lowest):

- weight
- local preference
- shortest length (shortest AS path)
- lowest origin code
- lowest MED
- route from external peer

- lowest cost to Next Hop
- lowest routerID

Stripping Out Private AS Numbers from Route Updates

Private AS numbers are AS numbers in the range 64512 through 65534. You can remove private AS numbers from the AS path attribute in updates that are sent to external BGP (EBGP) neighbors. Possible reasons for using private AS numbers include:

- The remote AS does not have officially allocated AS numbers.
- You want to conserve AS numbers if you are multi-homed to the local AS.

Private AS numbers should not be advertised on the Internet. Private AS numbers can only be used locally within an administrative domain. Therefore, when routes are advertised out to the Internet, the routes can be stripped out from the AS Paths of the advertised routes using this feature.

To configure private AS numbers to be removed from updates, use the following command:

```
enable bgp neighbor [<ipaddress> | all] remove-private-as-numbers
```

To disable this feature, use the following command:

```
disable bgp neighbor [<ipaddress> | all] remove-private-as-number
```

Route Re-Distribution

BGP, OSPF, and RIP can be enabled simultaneously on the switch. Route re-distribution allows the switch to exchange routes, including static, direct, and VIP routes, between any two routing protocols.

Exporting routes from OSPF to BGP, and from BGP to OSPF, are discreet configuration functions. To run OSPF and BGP simultaneously, you must first configure both protocols and then verify the independent operation of each. Then you can configure the routes to export from OSPF to BGP and the routes to export from BGP to OSPF.

Configuring Route Re-Distribution

Exporting routes between any two routing protocols are discreet configuration functions. For example, you must configure the switch to export routes from OSPF to BGP and, if desired, you must configure the switch to export routes from BGP to OSPF. You must first configure both protocols and then verify the independent operation of each. Then you can configure the routes to export from OSPF to BGP and the routes to export from BGP to OSPF.

You can use route maps to associate BGP attributes including Community, NextHop, MED, Origin, and Local Preference with the routes. Route maps can also be used to filter out exported routes.

To enable or disable the exporting of OSPF, RIP, static, direct (interface), and VIP routes to BGP, use the following commands:

```
enable bgp export [direct | static | rip | ospf | ospf-intra |  
ospf-inter | ospf-extern1 | ospf-extern2 | vip] {<route map>}
```

```
disable bgp export [direct | static | rip | ospf | ospf-intra |  
ospf-inter | ospf-extern1 | ospf-extern2 | vip]
```

Using the `export` command to redistribute routes complements the redistribution of routes using the `config bgp add network` command. The `config bgp add network` command adds the route to BGP only if the route is present in the routing table. The `enable bgp export` command redistributes an individual route from the routing table to BGP. If you use both commands to redistribute routes, the routes redistributed using the `network` command take precedence over routes redistributed using the `export` command.

19

IP Multicast Routing

This chapter covers the following topics:

- Overview on page 19-2
- Configuring IP Multicasting Routing on page 19-5
- Configuration Examples on page 19-6

For more information on IP multicasting, refer to the following publications:

- RFC 1112 – *Host Extension for IP Multicasting*
- RFC 2236 – *Internet Group Management Protocol, Version 2*
- DVMRP Version 3 – *draft_ietf_dvmrp_v3_07*
- PIM-DM Version 2 – *draft_ietf_pim_v2_dm_03*
- RFC 2362 – *Protocol Independent Multicast-Sparse Mode*

The following URLs point to the Web sites for the IETF Working Groups:

- IETF DVMRP Working Group – <http://www.ietf.org/html.charters/idmr-charter.html>
- IETF PIM Working Group – <http://www.ietf.org/html.charters/pim-charter.html>

Overview

IP multicast routing is a function that allows a single IP host to send a packet to a group of IP hosts. This group of hosts can include devices that reside on the local network, within a private network, or outside of the local network.

IP multicast routing consists of the following functions:

- A router that can forward IP multicast packets.
- A router-to-router multicast routing protocol (for example, Distance Vector Multicast Routing Protocol (DVMRP) or Protocol Independent Multicast (PIM)).
- A method for the IP host to communicate its multicast group membership to a router (for example, Internet Group Management Protocol (IGMP)).



Note: You should configure IP unicast routing before you configure IP multicast routing.

DVMRP Overview

DVMRP is a distance vector protocol that is used to exchange routing and multicast information between routers. Like RIP, DVMRP periodically sends the entire routing table to its neighbors.

DVMRP has a mechanism that allows it to prune and graft multicast trees to reduce the bandwidth consumed by IP multicast traffic.

PIM Overview

The switch supports both dense mode and sparse mode operation. You can configure dense mode or sparse mode on a per-interface basis. Once enabled, some interfaces can run dense mode, while others run sparse mode.

PIM Dense Mode

Protocol Independent Multicast-Dense Mode (PIM-DM) is a multicast routing protocol that is similar to DVMRP. PIM-DM routers perform reverse path multicasting (RPM). However, instead of exchanging its own unicast route tables for the RPM algorithm, PIM-DM uses the existing unicast route table for the reverse path. As a result, PIM-DM requires less system memory.

PIM-DM is a broadcast and prune protocol. Using PIM-DM, multicast routes are pruned and grafted in the same way as DVMRP.

PIM Sparse Mode (PIM-SM)

Unlike PIM-DM, PIM-SM is an explicit join and prune protocol, and it supports shared trees as well as shortest path trees (SPTs). The routers must explicitly join the group(s) in which they are interested in becoming a member, which is beneficial for large networks that have group members who are sparsely distributed.

Using PIM-SM, the router sends a join message to the rendezvous point (RP). The RP is a central multicast router that is responsible for receiving and distributing multicast packets. By default, the RP is selected dynamically. You can also define a static RP in your network, using the following command:

```
config pim crp static <rp_address>
```

If you use a static RP, all switches in your network must be configured with the same RP address.

When a router has a multicast packet to distribute, it encapsulates the packet in a unicast message and sends it to the RP. The RP decapsulates the multicast packet and distributes it among all member routers.

When a router determines that the multicast rate from a particular originating router (not the RP) has exceeded a configured threshold, that router can send an explicit join to the originating router. Once this occurs, the receiving router gets the multicast directly from the sending router, and bypasses the RP.



Note: You can run either PIM-DM or PIM-SM per VLAN.

PIM Mode Translation

An Extreme Networks switch functioning can function as a PIM multicast border router (PMBR). A PMBR integrates PIM-SM and PIM-DM traffic.

When forwarding PIM-DM traffic into a PIM-SM network, the PMBR notifies the RP that the PIM-DM network exists. The PMBR forwards PIM-DM multicast packets to the RP, which, in turn, forwards the packets to those routers that have joined the multicast group.

The PMBR also forwards PIM-SM traffic to a PIM-DM network. The PMBR sends a join message to the RP and the PMBR broadcasts traffic from the RP into the PIM-DM network.

No commands are required to enable PIM mode translation. PIM mode translation is automatically enabled when a dense mode interface and a sparse mode interface are enabled on the same switch.

IGMP Overview

IGMP is a protocol used by an IP host to register its IP multicast group membership with a router. Periodically, the router queries the multicast group to see if the group is still in use. If the group is still active, a single IP host responds to the query, and group registration is maintained.

IGMP is enabled by default on the switch. However, the switch can be configured to disable the generation of periodic IGMP query packets. IGMP query should be enabled when the switch is configured to perform IP unicast or IP multicast routing.

IGMP Snooping

IGMP snooping is a layer 2 function of the switch. It does not require multicast routing to be enabled. The feature reduces the flooding of IP multicast traffic. IGMP snooping optimizes the usage of network bandwidth, and prevents multicast traffic from being flooded to parts of the network that do not need it. The switch does not reduce any IP multicast traffic in the local multicast domain (224.0.0.x).

IGMP snooping is enabled by default on the switch. If you are using multicast routing, IGMP snooping must be enabled. If IGMP snooping is disabled, all IGMP and IP multicast traffic floods within a given VLAN. IGMP snooping expects at least one device in the network to periodically generate IGMP query messages. Without an IGMP querier, the switch stops forwarding IP multicast packets to any port. An optional optimization for IGMP snooping is the strict recognition of multicast routers only if the remote devices have joined the DVMRP (224.0.0.4) or PIM (224.0.0.13) multicast groups.

When a port sends an IGMP leave message, the switch removes the IGMP snooping entry after 10 seconds. The switch sends a query to determine which ports want to remain in the multicast group. If other members of the VLAN want to remain in the multicast group, the router ignores the leave message, but the port that requests removal is removed from the IGMP snooping table.

If the last port within a VLAN sends an IGMP leave message, the router does not receive any responses to the query, and the router immediately removes the VLAN from the multicast group.

Performance Enhancements for the BlackDiamond Switch

The BlackDiamond switch can optimize multicast data forwarding performance for modules that use the “i” series chipset. To increase the performance of multicast applications, you can disable I/O modules in the system that do not use the “i” series chipset.

In addition, you can modify the backplane load-sharing policy for more robust support of multicast streams.



Note: The round-robin algorithm is not supported on non-“i” series I/O modules. The default backplane loadsharing policy is “port-based”.

To configure the switch backplane load-sharing policy, use this command:

```
config backplane-ls-policy [address-based | port-based | round-robin]
```



Note: For more information on load sharing, see Chapter 4.

Configuring IP Multicasting Routing

To configure IP multicast routing, you must do the following:

- 1 Configure the system for IP unicast routing.
- 2 Enable multicast routing on the interface using the following command:
- 3 Enable DVMRP or PIM on all IP multicast routing interfaces using one of the following commands:

```
enable ipmcf forwarding {vlan <name>}
config dvmrp add vlan [<name> | all]
config pim add vlan [<name> | all] {dense | sparse}
```

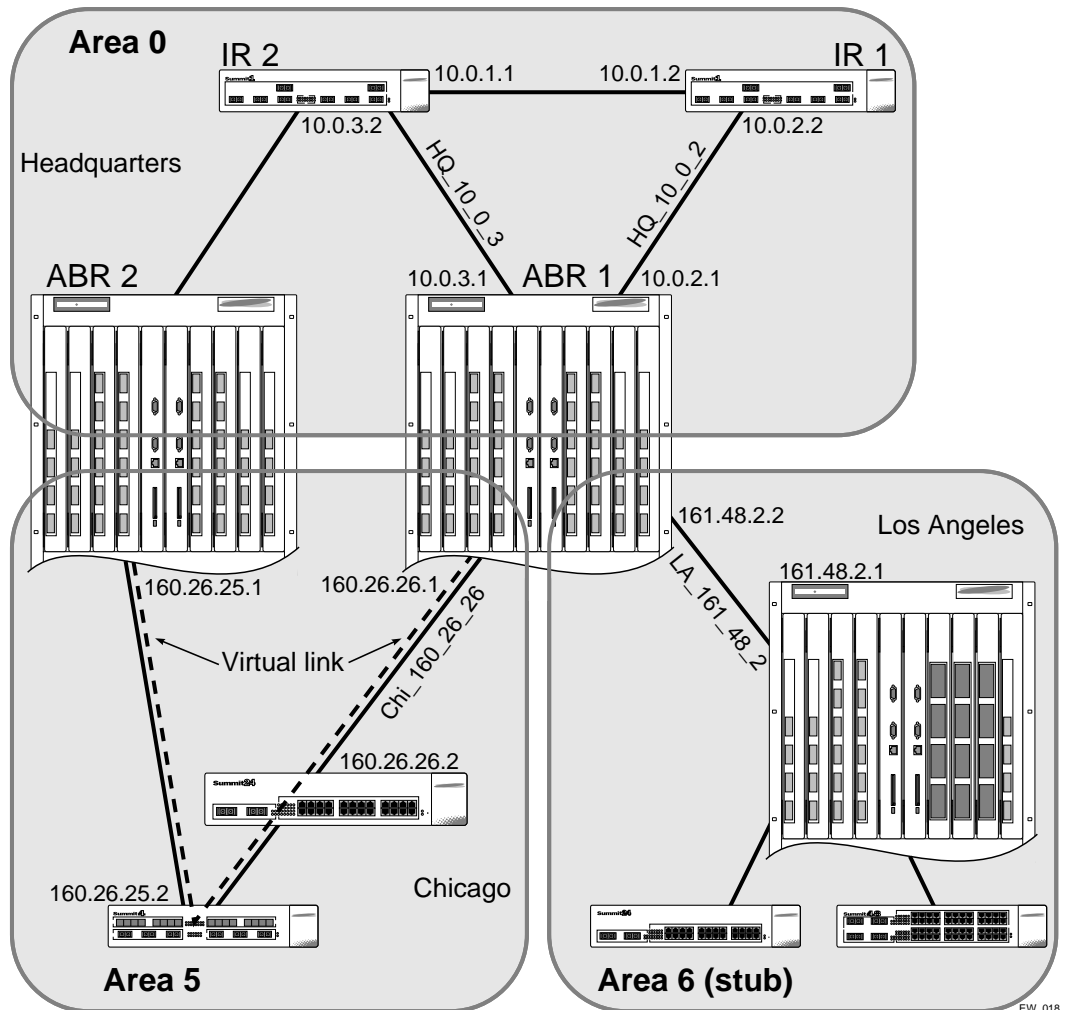
- 4 Enable DVMRP or PIM on the router using one of the following commands:

```
enable dvmrp
enable pim
```

Configuration Examples

Figure 19-1 and Figure 13-2 are used in Chapter 17 to describe the OSPF configuration on a switch. Refer to Chapter 17 for more information about configuring OSPF. In the first example, the system labeled IR1 is configured for IP multicast routing, using PIM-DM. In the second example, the system labeled ABR1 is configured for IP multicast routing using PIM-SM.

PIM-DM Configuration Example



EW_018

Figure 19-1: IP multicast routing using PIM-DM configuration example

Configuration for IR1

The router labeled IR1 has the following configuration:

```
config vlan HQ_10_0_1 ipaddress 10.0.1.2 255.255.255.0
config vlan HQ_10_0_2 ipaddress 10.0.2.2 255.255.255.0
config ospf add vlan all
enable ipforwarding
enable ospf
enable ipmcforwarding
config pim add vlan all dense
config pim spt-threshold 16 8
enable pim
```

The following example configures PIM-SM.

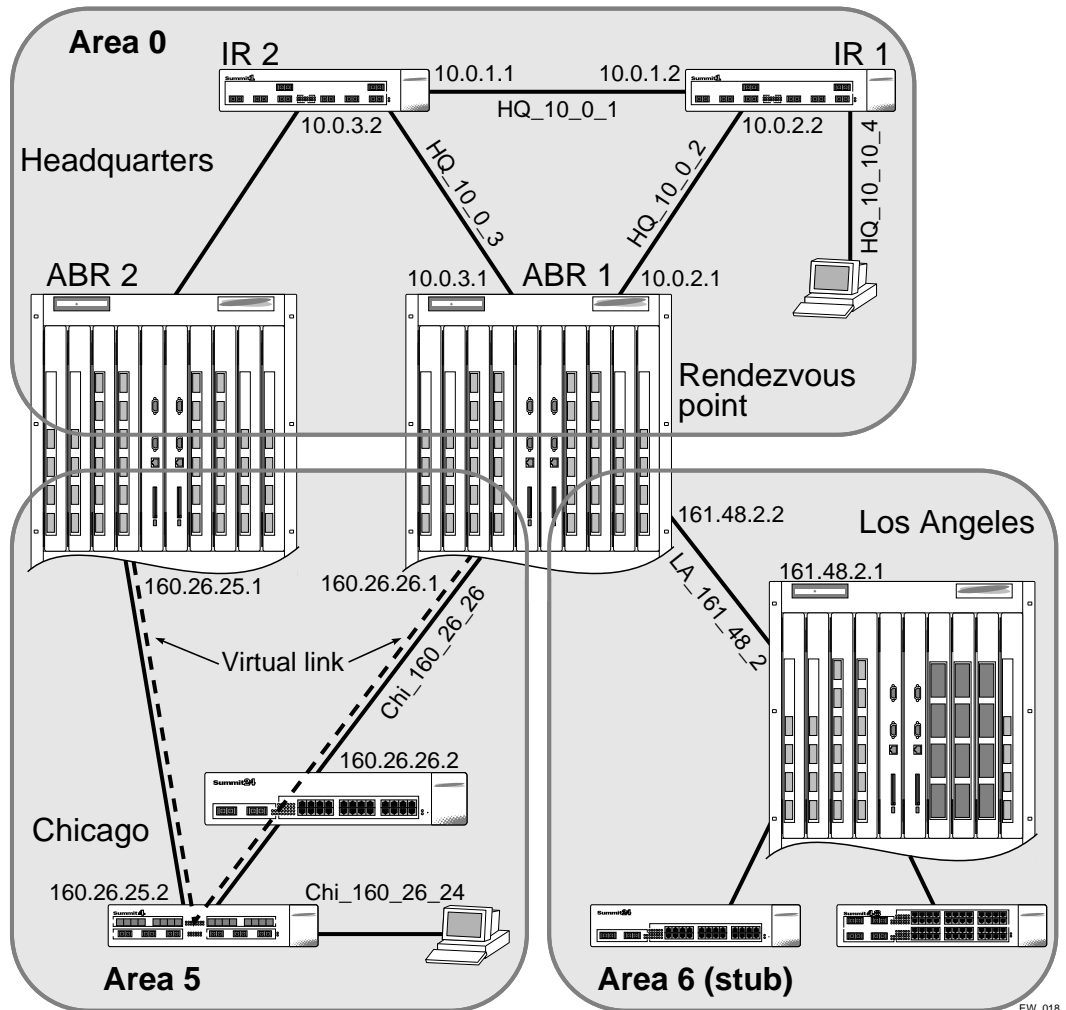


Figure 19-2: IP multicast routing using PIM-SM configuration example

EW_018

Configuration for ABR1

The router labeled ABR1 has the following configuration:

```
config vlan HQ_10_0_2 ipaddress 10.0.2.1 255.255.255.0
config vlan HQ_10_0_3 ipaddress 10.0.3.1 255.255.255.0
config vlan LA_161_48_2 ipaddress 161.48.2.2 255.255.255.0
config vlan CHI_160_26_26 ipaddress 160.26.26.1 255.255.255.0
config ospf add vlan all
enable ipforwarding
enable ipmcforwarding
config pim add vlan all sparse
create access-profile rp-list ipaddress
config rp-list add ipaddress 224.0.0.0 240.0.0.0
enable loopback HQ_10_0_3
config pim crp HQ_10_0_3 rp-list 30
config pim csbr HQ_10_0_3 30

config pim spt-threshold 16 8
```



IPX Routing

This chapter describes the following topics:

- Overview of IPX on page 20-1
- IPX/RIP Routing on page 20-5
- Configuring IPX on page 20-7
- IPX Configuration Example on page 20-8

This chapter assumes that you are already familiar with IPX. If not, refer to your Novell™ documentation.

Overview of IPX

The switch provides support for the IPX, IPX/RIP, and IPX/SAP protocols. The switch dynamically builds and maintains an IPX routing table and an IPX service table.

Router Interfaces

The routing software and hardware routes IPX traffic between IPX router interfaces. A router interface is simply a VLAN that has an IPX network identifier (NetID) and IPX encapsulation type assigned to it.

As you create VLANs with different IPX NetIDs the switch automatically routes between them. Both the VLAN switching and IPX routing function occur within the switch.

Switches that have the “i” chipset support these IPX routing features:

- Separate routing interfaces for IP and IPX traffic on the same VLAN.
- Load sharing of IPX routed traffic.
- 802.1Q tagged packets on a routed IPX VLAN.

Switches that do not have the “i” chipset do not support the features listed above.

Figure 20-1 shows the same BlackDiamond switch discussed in earlier chapters. In Figure 20-1, IPX routing has been added to the BlackDiamond switch, and two additional VLANs have been defined; *Exec* and *Support*. Both VLANs have been configured as protocol-specific VLANs, using IPX.

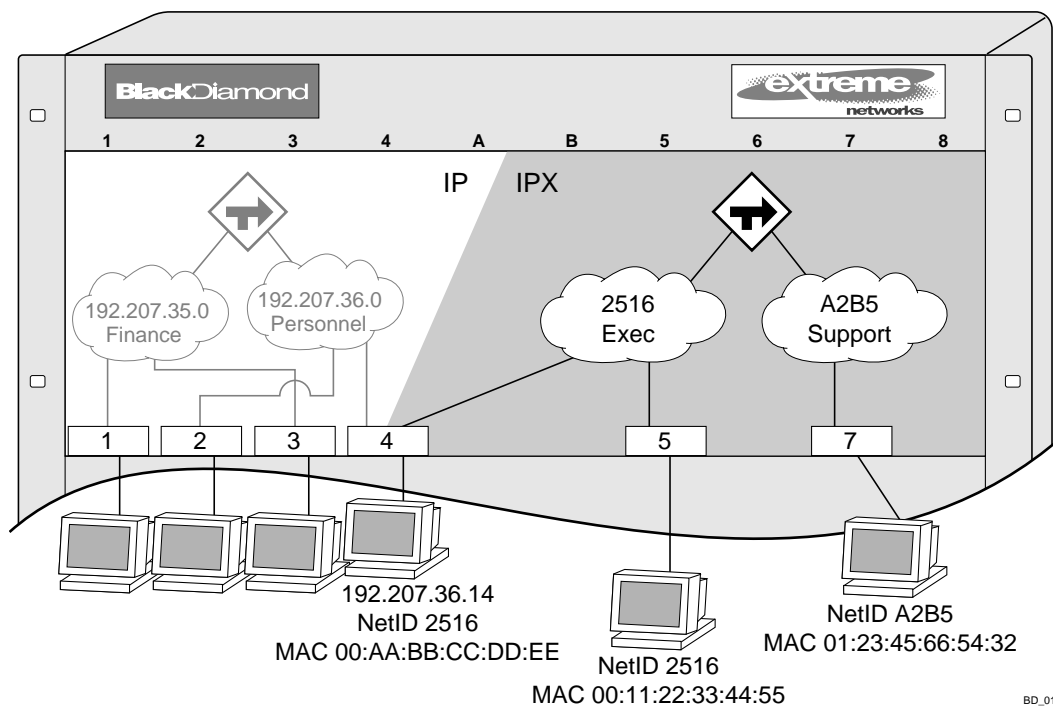


Figure 20-1: IPX VLAN configuration

Exec has been assigned the IPX NetID 2516. *Support* has been assigned the IPX NetID A2B5. All ports on slot 5 are assigned to *Exec*; all ports on slot 7 are assigned to *Support*.

In addition, all ports on slot 4 have been assigned to *Exec*. Thus, the ports on slot 4 belong to both the *Personnel* VLAN (running IP) and the *Exec* VLAN (running IPX).

Traffic within each VLAN is switched using the Ethernet MAC address. Traffic between *Exec* and *Support* is routed using the IPX NetID. Traffic cannot be sent between the IP VLANs (*Finance* and *Personnel*) and the IPX VLANs (*Exec* and *Support*).

IPX Routing Performance

To use IPX routing, you must have a switch that has the “i” chipset. Switches that have the “i” chipset are capable of performing IPX routing at wire-speed.

Switches that do not have the “i” chipset no longer support IPX routing capabilities. Previous versions of ExtremeWare supported CPU-based IPX routing on switches that did not have the “i” chipset. CPU-based IPX routing has been removed on switches that do not use the “i” chipset to support other features in ExtremeWare.

IPX Load Sharing

ExtremeWare supports IPX load sharing on all products that use the “i” chipset. No additional configuration is required to support this function, simply configure load sharing as you would normally.



Note: For more information on load sharing, see Chapter 4.

IPX Encapsulation Types

Novell NetWare™ supports four types of frame encapsulation. The ExtremeWare term for each type is described in Table 20-1.

Table 20-1: IPX Encapsulation Types

Name	Description
ENET_II	The frame uses the standard Ethernet 2 header.
ENET_8023	The frame includes the IEEE 802.3 length field, but does not include the IEEE 802.2 Logical Link Control (LLC) header. This encapsulation is used by NetWare version 2.x and the original 3.x version.

Table 20-1: IPX Encapsulation Types

Name	Description
ENET_8022	The frame uses the standard IEEE format and includes the IEEE 802.2 LLC header. This encapsulation is used by NetWare version 3.12 and 4.x.
ENET_SNAP	The frame adds a Subnetwork Access Protocol (SNAP) header to the IEEE 802.2 LLC header.

To configure a VLAN to use a particular encapsulation type, use the following command:

```
config vlan <name> xnetid <netid> [enet_ii | enet_8023 | enet_8022 |
enet_snap]
```

Tagged IPX VLANs

ExtremeWare supports tagged 802.1Q traffic on an IPX VLAN that is performing routing. Tagging is most commonly used to create VLANs that span multiple switches. Using VLAN tags, multiple VLANs can span multiple switches using one or more trunks. In a port-based VLAN, each VLAN requires its own pair of trunk ports.

Another benefit of tagged VLANs is the ability to have a port be a member of multiple VLANs. This is particularly useful if you have a device (such as a server) that must belong to multiple VLANs. A single port can be a member of only one port-based VLAN. All additional VLAN memberships for that port must be configured with tags.

To configure a tagged IPX VLAN, assign a tag to the VLAN using the following command:

```
config vlan <name> tag <vlanid>
```

The valid range is from 1 to 4095.

To assign tagged ports to the VLAN, use the following command:

```
config vlan <name> add port <portlist> {tagged | untagged}
{nobroadcast}
```

To display your VLAN settings, use the following command:

```
show vlan {<name>} {detail}
```

Populating the Routing Table

The switch builds and maintains an IPX routing table. As in the case of IP, the table is populated using dynamic and static entries.

Dynamic Routes

Dynamic routes are typically learned by way of IPX/RIP. Routers that use IPX/RIP exchange information in their routing tables in the form of advertisements. Using dynamic routes, the routing table contains only networks that are reachable.

Dynamic routes are aged out of the table when an update for the network is not received for a period of time, as determined by the routing protocol.

Static Routes

Static routes are manually entered into the routing table. Static routes are used to reach networks not advertised by routers. You can configure up to 64 static IPX routes on the switch. Static routes are never aged out of the routing table. Static routes are advertised to the network using IPX/RIP.

IPX/RIP Routing

The switch supports the use of IPX/RIP for unicast routing. IPX/RIP is different from IP/RIP. However, many of the concepts are the same. ExtremeWare supports the following IPX/RIP features:

- Split horizon
- Poison reverse
- Triggered Updates

Route information is entered into the IPX route table in one of the following two ways:

- Dynamically, by way of RIP
- Statically, using the command:

```
config ipxroute add [<dest_netid> | default] next_hop_netid
next_hop_node_addr <hops> <ticks>
```

IPX/RIP is automatically enabled when a NetID is assigned to the VLAN. To remove the advertisement of an IPX VLAN, use the command:

```
config ipxrip delete {vlan <name> | all}
```

GNS Support

ExtremeWare supports the Get Nearest Server (GNS) reply function. When a NetID is assigned to the switch, the GNS reply service is automatically enabled. When a station requests a particular service on the network (for example, locating a print server), the station sends a GNS request and the switch responds to the request. If GNS-reply is disabled, the switch drops the request.

To disable GNS-reply, use the following command:

```
disable ipxsap gns-reply {vlan <name>}
```

Routing SAP Advertisements

The switch contains an IPX Service Table, and propagates SAP advertisements to other IPX routers on the network. Each SAP advertisement contains the following:

- Service type
- Server name
- Server NetID
- Server node address

The service information is entered into the IPX Service Table in one of the following two ways:

- Dynamically, by way of SAP
- Statically, using the following command:

```
config ipxservice add <service_type> <service_name> <netid>  
<mac_address> <socket> <hops>
```

Configuring IPX

This section describes the commands associated with configuring IPX, IPX/RIP, and IPX/SAP on the switch. To configure IPX routing, follow these steps:

- 1 Create at least two VLANs.
- 2 If you are combining an IPX VLAN with another VLAN on the same port(s), you must use a protocol filter on one of the VLANs, or use 802.1Q tagging.
- 3 Assign each VLAN a NetID and encapsulation type using the following command:

```
config vlan <name> xnetid <netid> [enet_ii | enet_8023 | enet_8022 |
enet_snap]
```

Ensure that each VLAN has a unique IPX NetID and that the encapsulation type matches the VLAN protocol.

Once you configure the IPX VLAN information, IPX forwarding automatically begins to function. Specifically, configuring the IPX VLAN automatically enables the IPX/RIP, IPX/SAP, and SAP GNS services.

Verifying IPX Router Configuration

You can use the following commands to verify the IPX routing configuration:

- `show vlan` — In addition to other information, this command displays the IPX NetID setting and encapsulation type.
- `show ipxconfig` — This command is analogous to the `show ipconfig` command for the IP protocol. It displays summary global IPX configuration information followed by per-VLAN information. Information includes enable/disable status for IPX/RIP, IPX/SAP, IPX route sharing, IPX service sharing, and so on.
- `show ipxroute` — This command is analogous to the `show iproute` command for the IP protocol. It displays static and learned routes, along with information about the VLAN that uses the route, hop count, age of the route, and so on.
- `show ipxsap` — This command displays the enable status of IPX/SAP for the VLAN, and its operational and administrative status (including the GNS reply service). It also lists any identified IPX/SAP neighbors, SAP packet statistics, and several other timer settings.

- `show ipxrip` — This command displays the enable status of IPX/RIP for the VLAN, including operational and administrative status. It also lists any identified IPX/RIP neighbors, RIP packet statistics, and several other timer settings.
- `show ipxservice` — This command displays the contents of the IPX Service Table.

Protocol-Based VLANs for IPX

When combining IPX VLANs with other VLANs on the same physical port, it may be necessary to assign a protocol filter to the VLAN. This is especially true if it is not possible to use 802.1Q VLAN tagging. For convenience, IPX-specific protocol filters have been defined and named in the default configuration of the switch. Each filter is associated with a protocol encapsulation type. The IPX-specific protocol filters and the associated encapsulation type of each are described in Table 20-2.

Table 20-2: IPX Protocol Filters and Encapsulation Types

Protocol Name	Protocol Filter	Used for Filtering IPX Encapsulation Type
IPX	eypte 0x8137	enet_ii
IPX_8022	llc 0xe0e0	enet_802_2
IPX_snap	SNAP 0x8137	enet_snap

It is not possible to define a protocol-sensitive VLAN for filtering the IPX `enet_8023` encapsulation type. Instead, use a protocol-sensitive filter on the other VLANs that share the same ports, leaving the `enet_8023` encapsulation VLAN configured using the any protocol.

IPX Configuration Example

Figure 20-2 builds on the example showing the IP/RIP configuration that was used in earlier chapters. Now, in addition to having IP VLANs configured, this example illustrates a switch that has the following IPX VLANs defined:

- *Exec*
 - Protocol-sensitive VLAN using the IPX protocol with the filter `IPX_8022`.
 - All ports on slot 4 and slot 5 have been assigned to *Exec*.
 - *Exec* is configured for IPX NetID 2516 and IPX encapsulation type 802.2.

- *Support*
 - All ports on slot 7 have been assigned to *Support*.
 - *Support* is configured for IPX NetID A2B5 and IPX encapsulation type 802.2.

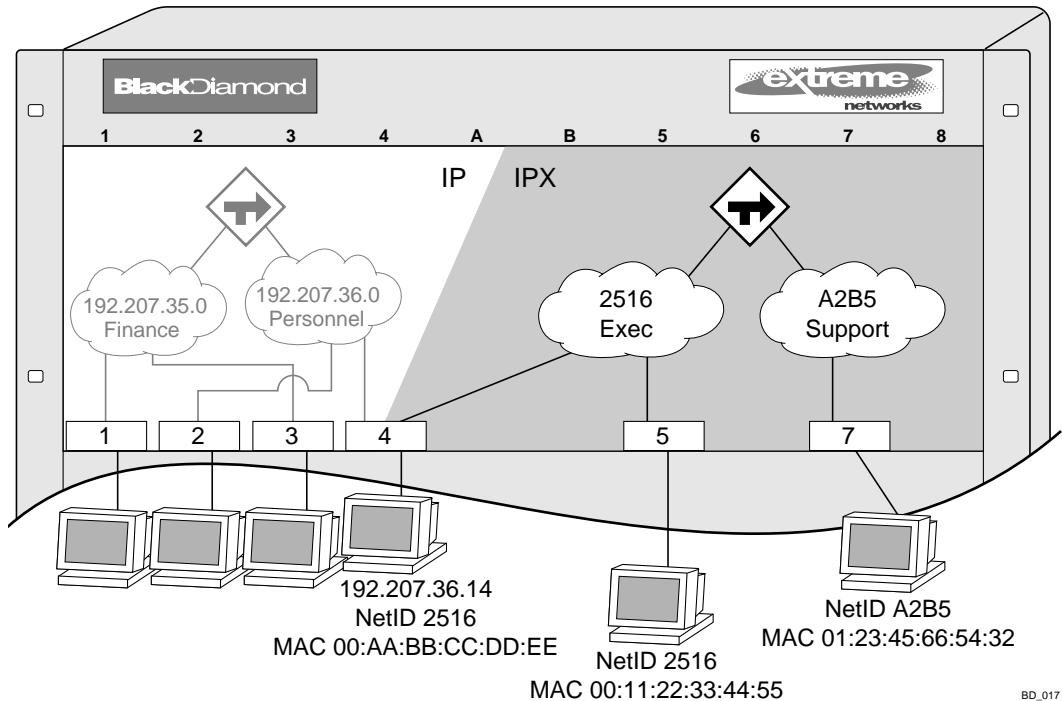


Figure 20-2: IPX routing configuration example

The stations connected to the system generate a combination of IP traffic and IPX traffic. The IP traffic is filtered by the IP VLANs. IPX traffic is filtered by the IPX VLANs.

In this configuration, all IP traffic from stations connected to slots 1 and 3 have access to the IP router by way of the VLAN *Finance*. IP traffic on ports on slots 2 and 4 reach the IP router by way of the VLAN *Personnel*.

Similarly, IPX traffic from stations connected to slots 4 and 5 have access to the IPX router by way of the VLAN *Exec*. IPX traffic on ports on slot 7 reach the IPX router by way of the VLAN *Support*. Both *Exec* and *Support* use enet_8022 as the encapsulation type.

BD_017

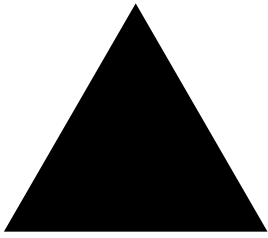
The IPX configuration shown in example in Figure 20-2 is as follows:

```
create vlan Exec
create vlan Support

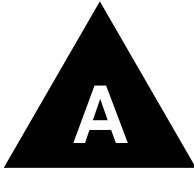
config Exec protocol ipx_8022

config Exec add port 4:*,5:*
config Support add port 7:*

config Exec xnetid 2516 enet_8022
config Support xnetid A2B5 enet_8022
```

Part 3:
Appendixes



Supported Protocols and Standards

The following is a list of software standards and protocols supported by ExtremeWare.

General Routing and Switching

RFC 1812 IPv4 Router Requirements	RFC 793 TCP
RFC 1519 CIDR	RFC 826 ARP
RFC 1256 IPv4 Router Discovery (IRDP)	RFC 2338 VRRP
RFC 783 TFTP	Extreme Standby Router Protocol (ESRP)
RFC 951, 1542 BootP	IPX RIP/SAP Router specification
RFC 2131 BOOTP/DHCP relay agent and DHCP server	IEEE 802.1D-1998 Spanning Tree Protocol
RFC 1591 DNS (client operation)	IEEE 802.1Q - 1998 Virtual Bridged Local Area Networks
RFC 1122 Host Requirements	Ethernet Automatic Protection Switching (EAPS)
RFC 768 UDP	Multiple Instances of Spanning Tree (PVST, 802.1Q interoperable)
RFC 791 IP	Software controlled redundant PHYs
RFC 792 ICMP	

VLANs

IEEE 802.1Q VLAN Tagging	Multiple STP domains per VLAN
IEEE 802.3ad Static ConfigPort-based VLANs	RFC-3069 VLAN Aggregation
MAC-based VLANs	Virtual MANs
Protocol-sensitive VLANs	

Quality of Service

IEEE 802.1D -1998 (802.1p) Packet Priority	RFC 2475 DiffServ Core and Edge Router Functions
RFC 2474 DiffServ Precedence, including 8 queues/port	Layer 1-4, Layer 7 (user name) Policy-Based Mapping
RFC 2598 DiffServ Expedited Forwarding (EF)	Policy-Based Mapping/Overwriting of DiffServ code points, .1p priority
RFC 2597 DiffServ Assured Forwarding (AF)	
Bi-directional Rate Shaping	

RIP

RFC 1058 RIP v1	RFC 2453 RIP v2
-----------------	-----------------

OSPF

RFC 2328 OSPF v2 (including MD5 authentication)	RFC 1765 OSPF Database Overflow
RFC 1587 OSPF NSSA Option	RFC 2370 OSPF Opaque LSA Option

BGP4

RFC 1771 Border Gateway Protocol 4	RFC 1997 BGP Communities Attribute
RFC 1965 Autonomous System Confederations for BGP	RFC 1745 BGP/OSPF Interaction
RFC 1966 BGP Route Reflection	RFC 2385 TCP MD5 Authentication for BGPv4

IP Multicast

RFC 2362 PIM-SM	RFC 2236 IGMP v2
PIM-DM Draft IETF PIM Dense Mode v2-dm-03)DVMRP v3 draft IETF DVMRP v3-07	IGMP Snooping with Configurable Router Registration Forwarding
RFC 1112 IGMP v1	

Management - SNMP & MIBs

RFC 1155 Structure of Mgmt Information (SMIv1)	RFC 1573 Evolution of Interface
RFC 1157 SNMPv1	RFC 1493 Bridge MIB
RFC-1212, RFC-1213, RFC-1215 MIB-II & TRAPs	RFC 1354 IPv4 Forwarding Table MIB
RFC 1901 – 1907 SNMP Version 2c, SMIv2 and Revised MIB-II	RFC 2037 Entity MIB RFC 2233 Interface MIB
RFC 1908 - Coexistence between SNMP Version 1 and Version 2c	RFC 2096 IP Forwarding
RFC 1757 RMON 4 groups: Stats, History, Alarms and Events	RFC 1724 RIPv2 MIB
RFC 2021 RMON2 (probe configuration)	RFC 1850 OSPFv2 MIB
RFC 2613 SMON MIB	RFC 1657 BGPv4 MIB
RFC 2668 802.3 MAU MIB	RFC 2787 VRRP MIB
RFC 1643 Ethernet MIB	RFC 2925 Ping / Traceroute / NSLOOKUP MIB
RFC 1650 Etherlike-MIB	ExtremeWare vendor MIB (includes ACL, MAC FDB, IP FDB, QoS policy and VLAN config)

Management - Other:

RFC 1866 HTML	NetFlow version 1 export
RFC 2068 HTTP	Configuration logging
RFC 854 Telnet	Multiple Images, Multiple Configs
HTML/ HTTP management	BSD System Logging Protocol (SYSLOG), with Multiple Syslog Servers
Secure Shell (SSHv2) and Telnet management, Telnet and SSHv2 clients	999 Local Messages (criticals stored across reboots)
Secure Copy (SCPv2)	RFC 2030 Simple Network Time Protocol v4

Security

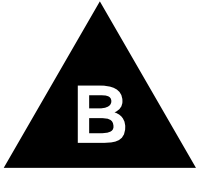
Routing protocol authentication (see above)	Access Profiles on All Routing Protocols
Secure Shell (SSHv2) & Secure Copy (SCPv2) with encryption/authentication	Access Profiles on All Management Methods
RFC 1492 TACACS+	Network Login (including DHCP / RADIUS integration)
RFC 2138 RADIUS Authentication	MAC Address Security / Lockdown
RFC 2139 RADIUS Accounting	Network Address Translation (NAT)
RADIUS Per-command Authentication	Layer 2/3/4/7 Access Control Lists (ACLs)

Denial of Service Protection

RFC 2267 Network Ingress Filtering	ICMP and IP-Option Response Control
RPF (Unicast Reverse Path Forwarding) Control	Server Load Balancing with Layer 3,4 Protection of Servers
Wire-speed ACLs	SYN attack protection
Rate Limiting by ACLs	Uni-directional Session Control
IP Broadcast Forwarding Control	

Robust against common Network Attacks

CERT (http://www.cert.org)	Host Attacks (http://www.rootshell.com)
<ul style="list-style-type: none"> ■ CA--97.28.Teardrop_Land -Teardrop and "LAND" attack ■ IP Options Attack ■ CA--98-13-tcp-denial-of-service ■ CA--98.01.smurf ■ CA--96.26.ping ■ CA--96.21.tcp_syn_flooding ■ CA--96.01.UDP_service_denial ■ CA--95.01.IP_Spoofing_Attacks_and_Hijacked_Terminal_Connections ■ CA-2002-03: SNMP vulnerabilities 	<ul style="list-style-type: none"> ■ Syndrop ■ Nester ■ Latierra ■ Newtear ■ Bonk ■ Winnuke ■ Raped ■ Simping ■ Sping ■ Ascend ■ Stream



Software Upgrade and Boot Options

This appendix describes the following topics:

- Downloading a New Image on page B-1
- Saving Configuration Changes on page B-3
- Using TFTP to Upload the Configuration on page B-4
- Using TFTP to Download the Configuration on page B-5
- Synchronizing MSMs on page B-7
- Upgrading and Accessing BootROM on page B-7

Downloading a New Image

The image file contains the executable code that runs on the switch. It comes preinstalled from the factory. As new versions of the image are released, you should upgrade the software running on your system.

The image is upgraded by using a download procedure from either a Trivial File Transfer Protocol (TFTP) server on the network or from a PC connected to the serial port using the XMODEM protocol. Downloading a new image involves the following steps:

- Load the new image onto a TFTP server on your network (if you will be using TFTP).
- Load the new image onto a PC (if you will be using XMODEM).

- Download the new image to the switch using the following command:

```
download image [<ipaddress> | <hostname>] <filename> {primary | secondary}
```

where the following is true:

`ipaddress` — Is the IP address of the TFTP server.

`hostname` — Is the hostname of the TFTP server. (You must enable DNS to use this option.)

`filename` — Is the filename of the new image.

`primary` — Indicates the primary image.

`secondary` — Indicates the secondary image.

The switch can store up to two images; a primary and a secondary. When you download a new image, you must select into which image space (primary or secondary) the new image should be placed. If not indicated, the primary image space is used.

If two MSMs are installed in the BlackDiamond switch, the downloaded image is saved to the same location on each one.

You can select which image the switch will load on the next reboot by using the following command:

```
use image [primary | secondary]
```

Rebooting the Switch

To reboot the switch, use the following command:

```
reboot { time <date> <time> | cancel}
```

where `date` is the date and `time` is the time (using a 24-hour clock format) when the switch will be rebooted. The values use the following format:

```
mm/dd/yyyy hh:mm:ss
```

If you do not specify a reboot time, the reboot occurs immediately following the command, and any previously schedule reboots are cancelled. To cancel a previously scheduled reboot, use the `cancel` option.

Saving Configuration Changes

The configuration is the customized set of parameters that you have selected to run on the switch. As you make configuration changes, the new settings are stored in run-time memory. Settings that are stored in run-time memory are not retained by the switch when the switch is rebooted. To retain the settings, and have them load when you reboot the switch, you must save the configuration to nonvolatile storage.

The switch can store two different configurations: a primary and a secondary. When you save configuration changes, you can select to which configuration you want the changes saved. If you do not specify, the changes are saved to the configuration area currently in use.

If you have made a mistake, or you must revert to the configuration as it was before you started making changes, you can tell the switch to use the secondary configuration on the next reboot.

To save the configuration, use the following command:

```
save {configuration} {primary | secondary}
```

To use the configuration, use the following command:

```
use configuration [primary | secondary]
```

The configuration takes effect on the next reboot.



Note: If the switch is rebooted while in the middle of a configuration save, the switch boots to factory default settings. The configuration that is not in the process of being saved is unaffected.

Returning to Factory Defaults

To return the switch configuration to factory defaults, use the following command:

```
unconfig switch
```

This command resets the entire configuration, with the exception of user accounts and passwords that have been configured, and the date and time.

To erase the currently selected configuration image and reset all switch parameters, use the following command:

```
unconfig switch all
```

Using TFTP to Upload the Configuration

You can upload the current configuration to a TFTP server on your network. The uploaded ASCII file retains the command-line interface (CLI) format. This allows you to:

- Modify the configuration using a text editor, and later download a copy of the file to the same switch, or to one or more different switches.
- Send a copy of the configuration file to the Extreme Networks Technical Support department for problem-solving purposes.
- Automatically upload the configuration file every day, so that the TFTP server can archive the configuration on a daily basis. Because the filename is not changed, the configured file stored in the TFTP server is overwritten every day.

To upload the configuration, use the following command:

```
upload configuration [<ipaddress> | <hostname>] <filename> {every <time>}
```

where the following is true:

- `ipaddress` — Is the IP address of the TFTP server.
- `hostname` — Is the hostname of the TFTP server. (You must enable DNS to use this option.)
- `filename` — Is the name of the ASCII file. The filename can be up to 255 characters long, and cannot include any spaces, commas, quotation marks, or special characters.
- `every <time>` — Specifies the time of day you want the configuration automatically uploaded on a daily basis. If not specified, the current configuration is immediately uploaded to the TFTP server.

To cancel a previously scheduled configuration upload, use the following command:

```
upload configuration cancel
```

Using TFTP to Download the Configuration

You can download ASCII files that contain CLI commands to the switch to modify the switch configuration. Three types of configuration scenarios that can be downloaded:

- Complete configuration
- Incremental configuration
- Scheduled incremental configuration

Downloading a Complete Configuration

Downloading a complete configuration replicates or restores the entire configuration to the switch. You typically use this type of download in conjunction with the `upload config` command, which generates a complete switch configuration in an ASCII format. As part of the complete configuration download, the switch is automatically rebooted.

To download a complete configuration, use the following command:

```
download configuration [<hostname> | <ip_address>] <filename>
```

After the ASCII configuration is downloaded by way of TFTP, you are prompted to reboot the switch. The downloaded configuration file is stored in current switch memory during the rebooting process, and is not retained if the switch has a power failure.

When the switch completes booting, it treats the downloaded configuration file as a script of CLI commands, and automatically executes the commands. If your CLI connection is through a Telnet connection (and not the console port), your connection is terminated when the switch reboots, but the command executes normally.

Downloading an Incremental Configuration

A partial or incremental change to the switch configuration may be accomplished by downloaded ASCII files that contain CLI commands. These commands are interpreted as a script of CLI commands, and take effect at the time of the download, without requiring a reboot of the switch.

To download an incremental configuration, use the following command:

```
download configuration <hostname | ip_address> <filename> {incremental}
```

Scheduled Incremental Configuration Download

You can schedule the switch to download a partial or incremental configuration on a regular basis. You could use this feature to update the configuration of the switch regularly from a centrally administered TFTP server. As part of the scheduled incremental download, you can optionally configuration a backup TFTP server.

To configure the primary and/or secondary TFTP server and filename, use the following command:

```
config download server [primary | secondary] [<hostname> | <ip_address>]  
<filename>
```

To enable scheduled incremental downloads, use the following command:

```
download configuration every <hour:minute>
```

To display scheduled download information, use the following command:

```
show switch
```

To cancel scheduled incremental downloads, use the following command:

```
download configuration cancel
```

Remember to Save

Regardless of which download option is used, configurations are downloaded into switch runtime memory, only. The configuration is saved only when the `save` command is issued, or if the configuration file, itself, contains the `save` command.

If the configuration currently running in the switch does not match the configuration that the switch used when it originally booted, an asterisk (*) appears before the command line prompt when using the CLI.

Synchronizing MSMs

On the BlackDiamond switch, you can take the master MSM configurations and images and replicate them on the slave MSM using the following command:

```
synchronize
```

In addition to replicating the configuration settings and images, this command also replicates which configuration or image the MSM should use on subsequent reboots. This command does not replicate the run-time configuration. You must use the `save configuration` command to store the run-time configuration first.

Upgrading and Accessing BootROM

The BootROM of the switch initializes certain important switch variables during the boot process. If necessary, BootROM can be upgraded, after the switch has booted, using TFTP. In the event the switch does not boot properly, some boot option functions can be accessed through a special BootROM menu.

Upgrading BootROM

Upgrading BootROM is done using TFTP (from the CLI), after the switch has booted. Upgrade the BootROM only when asked to do so by an Extreme Networks technical representative. To upgrade the BootROM, use the following command:

```
download bootrom [<host_name> | <ip_addr>]
```

Accessing the BootROM menu

Interaction with the BootROM menu is only required under special circumstances, and should be done only under the direction of Extreme Networks Customer Support. The necessity of using these functions implies a non-standard problem which requires the assistance of Extreme Networks Customer Support.

To access the BootROM menu, follow these steps:

- 1 Attach a serial cable to the console port of the switch.

- 2 Attach the other end of the serial cable to a properly configured terminal or terminal emulator, power cycle the switch while depressing the spacebar on the keyboard of the terminal.

As soon as you see the `BOOTROM->` prompt, release the spacebar. You can see a simple help menu by pressing `h`. Options in the menu include

- Selecting the image to boot from
- Booting to factory default configuration
- Performing a serial download of an image

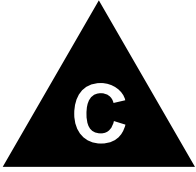
For example, to change the image that the switch boots from in flash memory, press `1` for the image stored in primary or `2` for the image stored in secondary. Then, press the `f` key to boot from newly selected on-board flash memory.

To boot to factory default configuration, press the `d` key for default and the `f` key to boot from the configured on-board flash.

To perform a serial download, you can optionally change the baud rate to 38.4K using the `b` command, and then press the `s` key to prepare the switch for an image to be sent from your terminal using the XMODEM protocol. After this has completed, select the `g` command, to boot the image that is currently in RAM. The switch restores the console port to 9600 bps, and begins the boot process.



Note: Doing a serial download does not store an image into flash, it only allows the switch to boot an operational image so that a normal TFTP upgrade from the CLI can then be performed.



Troubleshooting

If you encounter problems when using the switch, this appendix may be helpful. If you have a problem not listed here or in the release notes, contact your local technical support representative.

LEDs

Power LED does not light:

Check that the power cable is firmly connected to the device and to the supply outlet.

On powering-up, the MGMT LED lights yellow:

The device has failed its Power On Self Test (POST) and you should contact your supplier for advice.

A link is connected, but the Status LED does not light:

Check that:

- All connections are secure.
- Cables are free from damage.
- The devices at both ends of the link are powered-up.

- Both ends of the Gigabit link are set to the same autonegotiation state.

Both sides of the Gigabit link must be enabled or disabled. If the two are different, typically the side with autonegotiation disabled will have the link LED lit, and the side with autonegotiation enabled will not. The default configuration for a Gigabit port is autonegotiation enabled. This can be verified by entering the following command:

```
show port config
```

On power-on, some I/O modules do not boot:

Check if you are using 110V power input. The BlackDiamond switch powers only up to four modules if it is connected to a 110V outlet.

Error LED on the MSM64i turns amber:

Check the syslog message for a "critical" software error.

Status LED on the I/O module turns amber:

Check the syslog message for a related I/O module error. If the error is an inserted I/O module that conflicts with the software configuration, use one of the following commands to reset the slot configuration:

```
clear slot
```

```
config slot <slot> module [f32t | f32f | f48t | g4x | g6x | g8x | g12x]
```

Otherwise, contact Extreme Networks for further assistance.

ENV LED on the MSM64i turns amber:

Check each of the power supplies and all of the fans. Additionally, the status of these should be indicated in the display by entering "show switch" at the CLI. Look for the "Temperature" and "Power Supply" entries in the displayed information.

Switch does not power up:

All products manufactured by Extreme Networks use digital power supplies with surge protection. In the event of a power surge, the protection circuits shut down the power supply. To reset, unplug the switch for 1 minute, plug it back in, and attempt to power up the switch.

If this does not work, try using a different power source (different power strip/outlet) and power cord.

Using the Command-Line Interface

The initial welcome prompt does not display:

Check that your terminal or terminal emulator is correctly configured.

For console port access, you may need to press [Return] several times before the welcome prompt appears.

Check the settings on your terminal or terminal emulator. The settings are 9600 baud, 8 data bits, 1 stop bit, no parity, XON/OFF flow control enabled.

The SNMP Network Manager cannot access the device:

Check that the device IP address, subnet mask, and default router are correctly configured, and that the device has been reset.

Check that the device IP address is correctly recorded by the SNMP Network Manager (refer to the user documentation for the Network Manager).

Check that the community strings configured for the system and Network Manager are the same.

Check that SNMP access was not disabled for the system.

The Telnet workstation cannot access the device:

Check that the device IP address, subnet mask and default router are correctly configured, and that the device has been reset. Ensure that you enter the IP address of the switch correctly when invoking the Telnet facility. Check that Telnet access was not disabled for the switch. If you attempt to log in and the maximum number of Telnet sessions are being used, you should receive an error message indicating so.

Traps are not received by the SNMP Network Manager:

Check that the SNMP Network Manager's IP address and community string are correctly configured, and that the IP address of the Trap Receiver is configured properly on the system.

The SNMP Network Manager or Telnet workstation can no longer access the device:

Check that Telnet access or SNMP access is enabled.

Check that the port through which you are trying to access the device has not been disabled. If it is enabled, check the connections and network cabling at the port.

Check that the port through which you are trying to access the device is in a correctly configured VLAN.

Try accessing the device through a different port. If you can now access the device, a problem with the original port is indicated. Re-examine the connections and cabling.

A network problem may be preventing you accessing the device over the network. Try accessing the device through the console port.

Check that the community strings configured for the device and the Network Manager are the same.

Check that SNMP access was not disabled for the system.

Permanent entries remain in the FDB:

If you have made a permanent entry in the FDB (which requires you to specify the VLAN to which it belongs and then delete the VLAN), the FDB entry will remain. Though causing no harm, you must manually delete the entry from the FDB if you want to remove it.

Default and Static Routes:

If you have defined static or default routes, those routes will remain in the configuration independent of whether the VLAN and VLAN IP address that used them remains. You should manually delete the routes if no VLAN IP address is capable of using them.

You forget your password and cannot log in:

If you are not an administrator, another user having administrator access level can log in, delete your user name, and create a new user name for you, with a new password.

Alternatively, another user having administrator access level can log in and initialize the device. This will return all configuration information (including passwords) to the initial values.

In the case where no one knows a password for an administrator level user, contact your supplier.

Port Configuration

No link light on 10/100 Base port:

If patching from a hub or switch to another hub or switch, ensure that you are using a CAT5 cross-over cable. This is a CAT5 cable that has pins 1&2 on one end connected to pins 3&6 on the other end.

Excessive RX CRC errors:

When a device that has auto-negotiation disabled is connected to a Extreme switch that has auto-negotiation enabled, the Extreme switch links at the correct speed, but in half duplex mode. The Extreme switch 10/100 physical interface uses a method called *parallel detection* to bring up the link. Because the other network device is not participating in auto-negotiation (and does not advertise its capabilities), parallel detection on the Extreme switch is only able to sense 10Mbps versus 100Mbps speed, and not the duplex mode. Therefore, the switch establishes the link in half duplex mode using the correct speed.

The only way to establish a full duplex link is to either force it at both sides, or run auto-negotiation on both sides (using full duplex as an advertised capability, which is the default setting on the Extreme switch).



Note: A mismatch of duplex mode between the Extreme switch and another network device will cause poor network performance. Viewing statistics using the `show port rx` command on the Extreme switch may display a constant increment of CRC errors. This is characteristic of a duplex mismatch between devices. This is NOT a problem with the Extreme switch.

Always verify that the Extreme switch and the network device match in configuration for speed and duplex.

No link light on Gigabit fiber port:

Check to ensure that the transmit fiber goes to the receive fiber side of the other device, and vice-versa. All gigabit fiber cables are of the cross-over type.

The Extreme switch has auto-negotiation set to on by default for gigabit ports. These ports need to be set to auto off (using the command `config port <port #> auto off`) if you are connecting it to devices that do not support auto-negotiation.

Ensure that you are using multi-mode fiber (MMF) when using a 1000BASE-SX GBIC, and single mode fiber (SMF) when using a 1000BASE-LX GBIC. 1000BASE-SX does not work with SMF. 1000BASE-LX works with MMF, but requires the use of a mode conditioning patchcord (MCP).

VLANs

You cannot add a port to a VLAN:

If you attempt to add a port to a VLAN and get an error message similar to

```
localhost:7 # config vlan marketing add port 1:1,1:2
ERROR: Protocol conflict on port 1:5
```

you already have a VLAN using untagged traffic on a port. Only one VLAN using untagged traffic can be configured on a single physical port.

VLAN configuration can be verified by using the following command:

```
show vlan <name>
```

The solution for this error is to remove ports 1 and 2 from the VLAN currently using untagged traffic on those ports. If this were the “default” VLAN, the command would be

```
localhost:23 # config vlan default del port 1:1,1:2
```

which should now allow you to re-enter the previous command without error as follows:

```
localhost:26 # config vlan red add port 1:1,1:2
```

VLAN names:

There are restrictions on VLAN names. They cannot contain whitespaces and cannot start with a numeric value unless you use quotation marks around the name. If a name contains whitespaces, starts with a number, or contains non-alphabetical characters, you must use quotation marks whenever referring to the VLAN name.

802.1Q links do not work correctly:

Remember that VLAN names are only locally significant through the command-line interface. For two switches to communicate across a 802.1Q link, the VLAN ID for the VLAN on one switch should have a corresponding VLAN ID for the VLAN on the other switch.

If you are connecting to a third-party device and have checked that the VLAN IDs are the same, the Ethertype field used to identify packets as 802.1Q packets may differ between the devices. The default value used by the switch is 8100. If the third-party device differs from this and cannot be changed, you may change the 802.1Q Ethertype used by the switch with the following command:

```
config dot1p ethertype <ethertype>
```

Changing this parameter changes how the system recognizes all tagged frames received, as well as the value it inserts in all tagged frames it transmits.

VLANs, IP Addresses and default routes:

The system can have an IP address for each configured VLAN. It is necessary to have an IP address associated with a VLAN if you intend to manage (Telnet, SNMP, ping) through that VLAN or route IP traffic. You can also configure multiple default routes for the system. The system first tries the default route with the lowest cost metric.

STP

You have connected an endstation directly to the switch and the endstation fails to boot correctly:

The switch has STP enabled, and the endstation is booting before the STP initialization process is complete. Specify that STP has been disabled for that VLAN, or turn off STP for the switch ports of the endstation and devices to which it is attempting to connect, and then reboot the endstation.

The switch keeps aging out endstation entries in the switch Forwarding Database (FDB):

Reduce the number of topology changes by disabling STP on those systems that do not use redundant paths.

Specify that the endstation entries are static or permanent.

Debug Tracing

ExtremeWare includes a debug-tracing facility for the switch. The `show debug-tracing` command can be applied to one or all VLANs, as follows:

```
show debug-tracing {vlan <name>}
```

The `debug` commands should only be used under the guidance of Extreme Networks technical personnel.

To reset all debug-tracing to the factory default level, use the following command:

```
clear debug-trace
```

To change the debug tracing facility for a certain module to debug level, use the following command:

```
config debug-trace [access-list | bgp-events | bgp-keepalive | bgp-misc |
| bgp-msgs | bgp-neighbor | bgp-update-in | bgp-update-out |
card-state-change | esrp-system | flow-redirect | health-check |
ipxrip-route | ipxsap-entry | ospf-lsa | ospf-spf | pim-cache |
pim-rp-mgmt | slb-connection | slb-failover | slb-3dns | stp-in-pdu |
stp-out-pdu] <debug level>
```

To change the debug tracing level for a certain module to a particular level for one or more VLANs, use the following command:

```
config debug-trace [bootprelay | dvmrp-cache | dvmrp-hello |
dvmrp-message | dvmrp-neighbor | dvmrp-route | dvmrp-timer |
esrp-message | esrp-state-change | fdb | iparp | ipxgns-messages |
ipxrip-message | ipxsap-message | ospf-hello | ospf-neighbor |
pim-hello | pim-neighbor | pim-message | rip-message | rip-route-change |
| rip-triggered-update | udp-forwarding] <debug level> vlan <name>
```

To display the debug tracing configuration, use the following command:

```
show debug-trace [access-list | bgp-events | bgp-keepalive | bgp-misc |
| bgp-msgs | bgp-neighbor | bgp-update-in | bgp-update-out |
card-state-change | esrp-system | flow-redirect | health-check |
ipxrip-route | ipxsap-entry | ospf-lsa | ospf-spf | pim-cache |
pim-rp-mgmt | slb-connection | slb-failover | slb-3dns | stp-in-pdu |
stp-out-pdu]
```

To display the debug tracing configuration for one or more VLANs, use the following command:

```
show debug-trace [bootprelay | dvmrp-cache | dvmrp-hello |
dvmrp-message | dvmrp-neighbor | dvmrp-route | dvmrp-timer |
esrp-message | esrp-state-change | fdb | iparp | ipxgns-messages |
ipxrip-message | ipxsap-message | ospf-hello | ospf-neighbor |
pim-hello | pim-neighbor | pim-message | rip-message | rip-route-change
| rip-triggered-update | udp-forwarding] vlan <name>
```

TOP Command

The `top` command is a utility that indicates CPU utilization by process.

System Health Check

The system health check tests both the backplane and the CPU by periodically forwarding packets and checking for the validity of these packets. All error messages are logged in the syslog and the diagnostics CLI show output. A “CRIT” message will be posted to the log if any of the packet tests fail. If you observe a failure, please contact Extreme Technical Support.

To enable system health check, use the following command:

```
enable sys-health-check
```

System health check is enabled by default.

To disable the system health checker, use the following command:

```
disable sys-health-check
```

To configure the system health checker, use the following command:

```
configure sys-health-check alarm-level [card-down | default | log |
system-down | traps]
```

This command allows you to configure the switch’s reaction to a failed health check, and provides the following options:

- `log`—posts a CRIT message to the log

- `traps`—posts a CRIT message to the log and sends a trap
- `card-down`—posts a CRIT message to the log, sends a trap, and brings the module down
- `system-down`—posts a CRIT message to the log, sends a trap, and brings the system down

The default option is `log`.

To view the status of the system health checker, use this command:

```
show diag
```



Note: You cannot use both mirroring and the system health checker at the same time. If you configure mirroring with the system health checker enabled, the health checker will indicate that it has been disabled by sending a message to the syslog.

Contacting Extreme Technical Support

If you have a network issue that you are unable to resolve, contact Extreme Networks technical support. Extreme Networks maintains several Technical Assistance Centers (TACs) around the world to answer networking questions and resolve network problems. You can contact technical support by phone at:

- (800) 998-2408
- (408) 579-2826

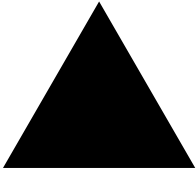
or by email at:

- support@extremenetworks.com

You can also visit the support website at:

- <http://www.extremenetworks.com/extreme/support/techsupport.asp>

to download software updates (requires a service contract) and documentation (including a .pdf version of this manual).



Index

Numerics

1d mode, STP	13-3
3DNS	10-34

A

access levels	2-10
access lists	
BlackDiamond switch maximum entries	8-5
default QoS profile	8-4
default rule	8-3
deleting	8-4
description	8-2
examples	8-6
ICMP filter example	8-10
ICMP traffic	8-5
maximum entries	8-5
permit-established example	8-6
permit-established keyword	8-4
restrictions	8-5
verifying settings	8-6
access policies, description	8-1
access profiles	
ExtremeWare Vista	3-12
reverse mask	8-13
SNMP	3-18
Telnet	3-6
accounts	
creating	2-13
deleting	2-13
viewing	2-13
Address Resolution Protocol. <i>See</i> ARP	
admin account	2-12
aging entries, FDB	6-2

alarm actions	12-26
Alarms, RMON	12-25
area 0, OSPF	17-8
areas, OSPF	17-7
ARP	
and VLAN aggregation	16-19
cache, clearing	16-19
communicating with devices outside subnet	16-7
configuring proxy ARP	16-7
disabling additions on superVLAN	16-20
incapable device	16-7
proxy ARP between subnets	16-7
proxy ARP, description of	16-6
responding to ARP requests	16-7
subVLANs	16-20
superVLANs	16-20
table, displaying	16-9
autonegotiation	
and redundant ports	4-18
description	4-4
autonomous system, description	18-2

B

backbone area, OSPF	17-8
BGP	
attributes	18-2
autonomous system	18-2
autonomous system path	18-2
cluster	18-3
community	18-3
description	18-2
features	18-3
IGP synchronization	18-9

loopback interface	18-9	Command-Line Interface. <i>See</i> CLI	
peer groups		common commands (table)	2-7
creating	18-9	communicating with devices outside subnet	16-7
description	18-9	complete configuration download	B-5
mandatory parameters	18-9	configuration	
neighbors	18-10	downloading	B-5
redistributing to OSPF	18-11	downloading complete	B-5
route aggregation	18-8	downloading incremental	B-5
route maps	8-30	logging	12-12
route reflectors	18-3	primary and secondary	B-3
route selection	18-10	saving changes	B-3
routing access policies	8-23	schedule download	B-6
Bi-directional rate shaping	7-22	uploading to file	B-4
limitations	7-25	console connection	3-2
loopback port	7-23	controlling Telnet access	3-6
maximum bandwidth settings	7-24	conventions	
maximum bandwidth settings (table)	7-24	notice icons, About This Guide	1-xviii
minimum bandwidth settings	7-25	text, About This Guide	1-xix
minimum bandwidth settings (table)	7-25		
BlackDiamond switch			
access list maximum entries	8-5	D	
load sharing group combinations	4-10	database applications, and QoS	7-4
MSMs, synchronizing	B-7	database overflow, OSPF	17-6
port configuration	4-2	default	
blackhole entries, FDB	6-4, 6-7	gateway	15-2
BOOTP		passwords	2-12
and UDP-Forwarding	16-15	settings	1-8
using	3-4	STP domain	13-2
BOOTP relay		users	2-12
configuring	16-14	<i>default</i> VLAN	5-13
BootROM		deleting a session	3-6
menu, accessing	B-7	DHCP and UDP-Forwarding	16-15
prompt	B-8	DHCP relay, configuring	16-14
upgrading	B-7	DHCP server	3-33
Border Gateway Protocol. <i>See</i> BGP		DiffServ, configuring	7-14
BPDU tunneling	13-3	disabling a switch port	4-3
browser		disabling route advertising (RIP)	17-4
controls	3-14	disconnecting a Telnet session	3-6
fonts	3-13	Distance Vector Multicast Routing Protocol. <i>See</i>	
setting up	3-12	DVMRP	
		distance-vector protocol, description	17-2
		DLCS	
		description	7-26
		guidelines	7-26
		limitations	7-27
		DNS	
		description	2-14
		Domain Name Service. <i>See</i> DNS	
		domains, STP	13-2
		downloading incremental configuration	B-5
		DVMRP	
		description	19-2
		routing access policies	8-20
		dynamic entries, FDB	6-2, 6-7
		dynamic routes	16-4, 20-5
C			
CLI			
command history	2-7		
command shortcuts	2-3		
line-editing keys	2-5		
named components	2-4		
numerical ranges, BlackDiamond switch	2-3		
numerical ranges, Summit switch	2-4		
symbols	2-4		
syntax helper	2-2		
using			
command			
history	2-7		
shortcuts	2-3		
syntax, understanding	2-1		

E

EAPS	
domain, creating and deleting	11-7
enabling and disabling a domain	11-11
enabling and disabling on a switch	11-11
polling timers, configuring	11-8
ring port, unconfiguring	11-12
show eaps display fields (table)	11-14
status information, displaying	11-12
switch mode, defining	11-8
ECMP. <i>See</i> IP route sharing	
EDP	
description	4-16
EMISTP	
description	13-3
example	13-7
rules	13-9
enabling a switch port	4-3
Equal Cost Multi-Path (ECMP) routing. <i>See</i> IP route sharing	
errors, port	12-5
ESRP	
and IP multinetting	14-17
and STP	14-17
and VLAN aggregation	14-18
and VRRP	15-7
description	14-1
diagnostic tracking	14-4
direct link	14-17
domains	14-14
environment tracking	14-4
example	14-19
failover time	14-10
groups	14-16
host attach	14-13
linking switches	14-17
master	
behavior	14-9
definition	14-2
determining	14-3
electing	14-9
election algorithms	14-8
ping tracking	14-6
port blocks	14-10
port restart	14-18
restarting ports	14-18
route table tracking	14-5
standby mode	
behavior	14-9
definition	14-2
super-VLAN	14-18
tracking, description	14-4
using 10/100 ports	14-10
VLAN tracking	14-5
establishing a Telnet session	3-3
Events, RMON	12-25

export restrictions	1-8
Extreme Discovery Protocol <i>See</i> EDP	
Extreme Standby Router Protocol. <i>See</i> ESRP	
ExtremeWare	
factory defaults	1-8
features	1-1
ExtremeWare Vista	
accessing	3-13
browser controls	3-14
browser setup	3-12
capturing screen output	3-17
controlling access	3-12
fonts	3-13
home page	3-11, 3-13
navigating	3-14
saving changes	3-15
screen layout	3-14
screen resolution	3-13
status messages	3-15
VLAN configuration	3-11

F

FDB	
adding an entry	6-2
aging entries	6-2
blackhole entries	6-4
contents	6-1
creating a permanent entry example	6-6
displaying	6-10
dynamic entries	6-2
dynamic entries, limiting	6-7
entries	6-1
limiting entries	6-7
non-aging entries	6-3
permanent entries	6-3
prioritizing entries	6-7
QoS profile association	6-5
file server applications, and QoS	7-4
flow control	4-4
flow redirection	10-35
flow statistics	
configuration overview	12-12
configuration parameters, resetting	12-22
enabling and disabling	12-17
flow record filter	
configuring	12-19
enabling and disabling	12-21
flow record timeout, configuring	12-19
ping-check function, configuring	12-21
source IP address, configuring	12-18
status information, displaying	12-22
fonts, browser	3-13
Forwarding Database. <i>See</i> FDB	
forwarding modes, SLB	10-7
full L3 functionality	1-6

<hr/>	
G	
GoGo mode, SLB	10-13
Greenwich Mean Time Offsets (table)	3-37
<hr/>	
H	
history command	2-7
History, RMON	12-24
home page	3-11, 3-13
<hr/>	
I	
ICMP, access lists	8-5
IEEE 802.1Q	5-6
IGMP	
description	19-4
snooping	19-4
image	
downloading	B-1
primary and secondary	B-2
upgrading	B-1
interfaces, router	16-2, 20-1
Internet Group Management Protocol. <i>See</i> IGMP	
IP address, entering	3-4
IP multicast routing	
configuring	19-5
description	1-4, 19-2
DVMRP	
description	19-2
example	19-6
IGMP	
description	19-4
snooping	19-4
PIM mode translation	19-3
PIM multicast border router (PMBR)	19-3
PIM-DM	19-2
PIM-SM	19-3
IP multinetting	
description	16-11
example	16-13
primary VLAN interface	16-12
secondary VLAN interface	16-12
using	16-12
IP route sharing	16-5
IP unicast routing	
BOOTP relay	16-14
configuration examples	16-9
configuring	16-8
default gateway	16-2
description	1-4
DHCP relay	16-14
ECMP	
enabling	16-9
IP route sharing	16-5
multinetting, description	16-11
multinetting, example	16-13
proxy ARP	16-6
route maps	16-5
router interfaces	16-2
routing table	
dynamic routes	16-4
multiple routes	16-4
populating	16-3
static routes	16-4
verifying the configuration	16-9
IPX	
configuration example	20-8
configuring	20-7
load sharing	20-3
protocol filters	20-8
protocol-based VLANs	20-8
router interfaces	20-1
routing access policies	8-18
routing table	
dynamic routes	20-5
populating	20-5
static routes	20-5
tagged VLANs	20-4
verifying router configuration	20-7
IPX/RIP	
configuring	20-7
routing table, populating	20-5
IPX/SAP	
configuring	20-7
<hr/>	
J	
jumbo frames	
description	4-5
enabling	4-5
IP fragmentation	4-6
maximum MTU	4-5
path MTU discovery	4-6
<hr/>	
K	
keys	
line-editing	2-5
port monitoring	12-6
<hr/>	
L	
license keys	1-7
licensing	
basic functionality	1-6
description	1-6
full L3 functionality	1-6
license keys	1-7
ordering	1-7
verifying	1-7
line-editing keys	2-5

link-state database	17-5	port number	4-2
link-state protocol, description	17-2	port-mirroring, virtual port	4-15
load balancing methods, SLB	10-15	slot configuration	4-1
load sharing		verifying load sharing	4-14
algorithms	4-8	monitoring the switch	12-2
and redundant ports	4-19	MSM	3-2
configuring	4-10	multinetting. <i>See</i> IP multinetting	
description	4-8	multiple routes	16-4
group combinations on BlackDiamond switch (table)	4-10		
load-sharing group, description	4-8		
master port	4-10		
port group combinations on Summit switch (table)	4-11		
verifying the configuration	4-14		
local logging	12-10		
log display	12-11		
logging			
and Telnet	12-11		
configuration changes	12-12		
description	12-9		
fault level	12-9		
local	12-10		
message	12-10		
QoS monitor	7-21		
real-time display	12-11		
remote	12-11		
subsystem	12-10		
timestamp	12-9		
logging in	2-12		
loopback port	7-23		

M

MAC-based security	6-7
MAC-based VLANs	
description	5-19
example	5-20
groups	5-19
guidelines	5-19
limitations	5-20
timed configuration download	5-21
maintenance mode, SLB	10-34
management access	2-10
management port	3-2
Management Switch Fabric Module. <i>See</i> MSM	
master port	
load sharing	4-10
maximum MTU	4-5
maximum Telnet session	3-3
<i>mgmt</i> VLAN	3-3
MIBs	3-18
modular switch	
configuring load sharing	4-10
enabling and disabling ports	4-3
jumbo frames	4-5
load sharing example	4-14

N

names, VLANs	5-12
NAT	
creating rules	9-8
rule matching	9-8
native VLAN, PVST+	13-12
Network Address Translation. <i>See</i> NAT	
network login	3-28
campus mode	3-28, 3-29
configuration example	3-30
configuring	3-29
user login	3-30
DHCP server	3-29
disabling	3-34
ISP mode	3-28, 3-32
configuration example	3-32
configuring	3-32
RADIUS server configuration	3-29
settings, displaying	3-33
non-aging entries, FDB	6-3
Not-So-Stubby_Area. <i>See</i> NSSA	
NSSA. <i>See</i> OSPF	

O

opaque LSAs, OSPF	17-6
Open Shortest Path First. <i>See</i> OSPF	
opening a Telnet session	3-3
OSPF	
advantages	17-3
area 0	17-8
areas	17-7
backbone area	17-8
configuration example	17-18
consistency	17-6
database overflow	17-6
description	17-2, 17-5
display filtering	17-21
enabling	16-9
link type	17-11
link-state database	17-5
normal area	17-9
NSSA	17-8
opaque LSAs	17-6
point-to-point links	17-12
redistributing routes	17-13
redistributing to BGP	18-11

router types	17-7
routing access policies	8-18
settings, displaying	17-21
stub area	17-8
virtual link	17-9
wait interval, configuring	17-17

P

passwords	
default	2-12
forgetting	2-13
path MTU discovery	4-6
permanent entries, FDB	6-3
permit-established keyword	8-4
persistence, SLB	10-23
Per-VLAN Spanning Tree. <i>See</i> PVST+	
PIM	
mode translation	19-3
multicast border router (PMBR)	19-3
PIM-DM	
description	19-2
PIM-SM	
description	19-3
rendezvous point	19-3
ping command	2-14
ping-check	10-32
poison reverse	17-4
port	
autonegotiation	4-4
BlackDiamond switch	4-2
configuring on BlackDiamond switch	4-2
enabling and disabling	4-3
errors, viewing	12-5
load-sharing groups	4-12
loopback	7-23
monitoring display keys	12-6
priority, STP	13-14
receive errors	12-5
redundant	
configuring	4-20
description	4-17
operation	4-18
statistics, viewing	12-4
STP state, displaying	13-16
STPD membership	13-2
transmit errors	12-5
port mode	13-3, 13-14
port translation mode, SLB	10-12
port-based VLANs	5-2
port-mirroring	
and protocol analyzers	4-15
description	4-15
modular switch example	4-15
stand-alone switch example	4-16
tagged and untagged frames	4-15
virtual port	4-15

primary image	B-2
<i>private</i> community, SNMP	3-19
profiles, QoS	7-6
protocol analyzers, use with port-mirroring	4-15
protocol filters	5-10
protocol filters, IPX	20-8
Protocol Independent Multicast- Dense Mode. <i>See</i> PIM-DM	
Protocol Independent Multicast- Sparse Mode. <i>See</i> PIM-SM	
protocol-based VLANs	5-9
proxy ARP	
communicating with devices outside subnet	16-7
conditions	16-7
configuring	16-7
MAC address in response	16-7
responding to requests	16-7
subnets	16-7
proxy ARP, description	16-6
<i>public</i> community, SNMP	3-19
PVST+	
description	13-11
native VLAN	13-12
VLAN mapping	13-12
PVST+ mode	13-3

Q

QoS	
802.1p priority	7-12
applications	7-3
blackhole	7-10
buffer	7-6
database applications	7-4
default QoS profiles	7-7
default QoS profiles (table)	7-7
description	1-3, 7-1
DiffServ, configuring	7-14
examples	
MAC address	7-10
source port	7-18
VLAN	7-19
FDB entry association	6-5
file server applications	7-4
maximum bandwidth	7-6
minimum bandwidth	7-6
priority	7-6
profiles	
default	7-7
default (table)	7-7
description	7-5
parameters	7-6
Random Early Detection (RED)	7-2
traffic groupings	7-8
blackhole	7-10
broadcast/unknown rate limiting	7-11

description	7-5	limitations	17-2
explicit packet marking	7-11	poison reverse	17-4
IP address	7-9	redistributing routes	17-13
MAC address	7-9	redistributing to BGP	18-11
source port	7-18	routing access policies	8-16
VLAN	7-19	routing table entries	17-3
traffic groupings (table)	7-9	split horizon	17-4
verifying	7-21	triggered updates	17-4
video applications	7-3	version 2	17-4
voice applications	7-3	RMON	
web browsing applications	7-4	alarm actions	12-26
QoS monitor		Alarms group	12-25
description	7-20	Events group	12-25
logging	7-21	features supported	12-24
real-time display	7-21	History group	12-24
QoS traffic grouping priorities, resetting	7-20	probe	12-24
Quality of Service. <i>See</i> QoS	7-2	Statistics group	12-24
<hr/>			
R		route maps	
RADIUS		BGP	8-30
and TACACS+	3-20, 3-28	changing	8-29
client configuration	3-22	creating	8-24
description	3-20	description	8-2, 8-24
Merit server configuration (example)	3-24	example	8-28
per-command authentication	3-22	goto entries	8-26
per-command configuration (example)	3-25	IP unicast routing	16-5
RFC 2138 attributes	3-22	match entries	8-26
servers	3-20	match operation keywords (table)	8-26
TCP port	3-22	processing	8-28
Random Early Detection (RED)	7-2	set entries	8-26
rapid root failover	13-4	set operation keywords (table)	8-26
Rate shaping, bi-directional. <i>See</i> Bi-directional rate shaping		route sharing. <i>See</i> IP route sharing	
receive errors	12-5	router interfaces	16-2, 20-1
redundant ports		router licensing	
active path	4-19	basic functionality	1-6
and auto-negotiation	4-18	description	1-6
and load sharing	4-19	full L3 functionality	1-6
and smart redundancy	4-19	license keys	1-7
configuring	4-20	ordering	1-7
description	4-17	verifying	1-7
disabling	4-20	router types, OSPF	17-7
operation	4-18	routing access policies	
typical configurations	4-17	access profile	
remote logging	12-11	applying	8-16
Remote Monitoring. <i>See</i> RMON		changing	8-23
renaming a VLAN	5-13	configuring	8-12
reset to factory defaults	B-3	creating	8-12
responding to ARP requests	16-7	types	8-12
reverse mask	8-13	BGP	8-23
RIP		deny	8-12
advantages	17-2	DVMRP	8-20
configuration example	17-15	examples	
description	17-2, 17-3	DVMRP	8-21
disabling route advertising	17-4	OSPF	8-19
enabling	16-9	PIM	8-22
		RIP	8-17
		IPX	8-18
		none	8-12

OSPF	8-18	subnet-route	10-4
permit	8-12	tcp-port-check	10-33
PIM	8-22	traffic type	10-7
removing	8-24	translation mode	10-11
RIP	8-16	transparent mode	10-8
using	8-11	VIPs	10-3
Routing Information Protocol. <i>See</i> RIP		virtual servers	10-3
routing table, populating	16-3	wildcard virtual servers	10-4
routing table, populating IPX	20-5	slot	
routing. <i>See</i> IP unicast routing		automatic configuration	4-1
		clearing	4-2
		manually configuring	4-2
		mismatch	4-2
		smart redundancy, and redundant ports	4-19
		SNAP protocol	5-11
		SNMP	
		community strings	3-18
		configuring	3-18
		controlling access	3-18
		read access	3-18
		read/write access	3-18
		settings, displaying	3-19
		supported MIBs	3-18
		system contact	3-19
		system location	3-19
		system name	3-19
		trap receivers	3-18
		using	3-17
		SNTP	
		configuring	3-34
		Daylight Savings Time	3-34
		description	3-34
		example	3-39
		Greenwich Mean Time offset	3-34
		Greenwich Mean Time Offsets (table)	3-37
		NTP servers	3-34
		software licensing	
		security features	1-8
		SSH2 protocol	1-8
		Spanning Tree Protocol. <i>See</i> STP	
		speed, ports	4-4
		split horizon	17-4
		SSH2 protocol	
		authentication key	3-9
		description	1-8, 3-7
		enabling	3-8
		predefined clients	3-8
		TCP port number	3-8
		stand-alone switch	
		enabling and disabling ports	4-3
		jumbo frames	4-5
		load sharing	4-10
		load sharing example	4-14
		port-mirroring, virtual port	4-15
		verifying load sharing	4-14
		static routes	16-4, 20-5
<hr/>			
S			
saving changes using ExtremeWare Vista	3-15		
saving configuration changes	B-3		
scheduling configuration download	B-6		
screen resolution, ExtremeWare Vista	3-13		
secondary image	B-2		
security licensing			
description	1-7		
obtaining	1-8		
Server Load Balancing <i>See</i> SLB			
service-check	10-33		
sessions, deleting	3-6		
shortcuts, command	2-3		
Simple Network Management Protocol. <i>See</i> SNMP			
SLB			
3DNS support	10-34		
active-active	10-28		
active-active operation	10-28		
client persistence	10-23		
components	10-2		
description	10-1		
failover	10-28		
forwarding mode	10-7		
gateway ping-checking	10-28		
GoGo mode	10-13		
health checking	10-32		
high availability	10-24		
host-route	10-4		
least connections	10-16		
load balancing methods	10-15		
maintenance mode	10-34		
manual fail-back	10-32		
nodes	10-3		
persistence	10-23		
ping-check	10-32		
pools	10-3		
port translation mode	10-12		
priority mode	10-16		
proxy ARP	10-4		
proxy client persistence	10-23		
ratio	10-15		
round-robin	10-15		
service-check	10-33		
standard virtual servers	10-4		
sticky persistence	10-24		

statistics	
port	12-4
VLAN	5-16
VLAN, per port	5-16
Statistics, RMON	12-24
status monitoring	12-2
STP	
1D mode	13-3
advanced example	13-8
and ESRP	14-17
and VLANs	13-2
and VRRP	15-7
basic configuration example	13-4
BPDU tunneling	13-3
bridge priority	13-14
configurable parameters	13-13
configuration examples	13-14
configuring	13-13
default domain	13-2
description	1-3
displaying settings	13-16
domains	13-2
EMISTP	
description	13-3
example	13-7
rules	13-9
forward delay	13-13
hello time	13-13
max age	13-13
overview	13-2
path cost	13-14
port mode	13-14
port modes	13-3
port priority	13-14
port state, displaying	13-16
PVST+	
description	13-11
mode	13-3
rapid root failover	13-4
rules and restrictions	13-12
StpdID	13-3, 13-14
stub area, OSPF	17-8
sub-VLAN	16-17
Summit switch, load sharing group combinations	4-11
super-VLAN	14-18, 16-17
switch	
logging	12-9
monitoring	12-2
RMON features	12-24
synchronizing MSMPs	B-7
syntax, understanding	2-1
syslog host	12-11
system contact, SNMP	3-19
system location, SNMP	3-19
system name, SNMP	3-19

T

TACACS+	
and RADIUS	3-20, 3-28
description	3-27
servers, specifying	3-28
tagging, VLAN	5-6
tcp-port-check	10-33
technical support	C-10
Telnet	
connecting to another host	3-3
controlling access	3-6
disconnecting a session	3-6
logging	12-11
maximum sessions	3-3
opening a session	3-3
using	3-3
Terminal Access Controller Access Control System	
Plus. See TACACS+	
TFTP	
server	B-1
using	B-4
timed configuration download, MAC-based	
VLANs	5-21
traceroute command	2-15
traffic groupings	7-8
translation mode, SLB	10-11
transmit errors	12-5
transparent mode, SLB	10-8
triggered updates	17-4
trunks	5-6
tunneling	5-17

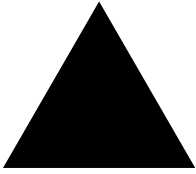
U

UDP-Forwarding	
and BOOTP	16-15
and DHCP	16-15
configuring	16-16
description	16-15
example	16-16
profiles	16-16
VLANs	16-16
upgrading the image	B-1
uploading the configuration	B-4
users	
access levels	2-10
authenticating	3-20
creating	2-13
default	2-12
viewing	2-13

V

video applications, and QoS	7-3
viewing accounts	2-13

VIPs, SLB	10-3	backup router	15-2
Virtual LANs. <i>See</i> VLANs		configuration parameters (table)	15-11
virtual link, OSPF	17-9	default gateway	15-2
virtual router, VRRP	15-2	description	15-1
VLAN aggregation		electing the master	15-6
description	16-17	examples	15-12
limitations	16-19	interfaces	15-2
properties	16-18	IP address	15-2, 15-11
proxy ARP	16-20	IP address owner	15-2
secondary IP address	16-17	MAC address	15-2
sub-VLAN	16-17	master	
super-VLAN	16-17	determining	15-3
VLAN tagging	5-6	master down interval	15-6, 15-11
VLANs		master router	15-2
and ExtremeWare Vista	3-11	multicast address	15-7
and STP	13-2	operation	15-8
assigning a tag	5-6	ping tracking	15-4
benefits	5-2	port restart	15-7
configuration examples	5-14	preempt mode	15-11
configuring	5-13	priority	15-3, 15-6, 15-11
<i>default</i>	5-13	redundancy	15-9
description	1-3	restarting ports	15-7
disabling route advertising	17-4	route table tracking	15-3
displaying settings	5-15	skew time	15-6, 15-11
IP fragmentation	4-7	tracking, description	15-3
MAC-based		virtual router	15-2
description	5-19	virtual router identifier (VRID)	15-2, 15-11
example	5-20	virtual router MAC address	15-2, 15-7, 15-9
groups	5-19	VLAN tracking	15-3
guidelines	5-19	VRRP router	15-2
limitations	5-20		
timed configuration download	5-21		
<i>mgmt</i>	3-3		
mixing port-based and tagged	5-9		
names	5-12		
port-based	5-2		
protocol filters	5-10		
protocol-based	5-9		
protocol-based, IPX	20-8		
renaming	5-13		
routing	16-8, 20-7		
statistics	5-16		
statistics, per port	5-16		
tagged	5-6		
trunks	5-6		
tunneling	5-17		
types	5-2		
UDP-Forwarding	16-16		
vMAN tunneling			
configuring	5-17		
description	5-17		
example	5-18		
voice applications, QoS	7-3		
VRRP			
advertisement interval	15-6, 15-11		
and ESRP	15-7		
and Spanning Tree	15-7		
		W	
		Web access, controlling	3-12
		web browsing applications, and QoS	7-4
		X	
		xmodem	B-1



Index of Commands

C

clear debug-trace	C-8	config dvmrp vlan trusted-gateway	8-21
clear fdb	7-10	config eaps add protect vlan	11-11
clear session	2-7, 3-6	config eaps failtime	11-8
clear slot	4-2	config eaps hellotime	11-8
command	17-21	config eaps mode	11-8
config access-profile add	8-13	config eaps primary port	11-9, 11-10
config access-profile delete	8-16	config eaps secondary port	11-9, 11-10
config access-profile mode	8-13	config fdb agingtime	6-6
config account	2-7	config flowstats export	12-18
config backplane-ls-policy	4-21, 19-5	config flowstats filter ports	12-19
config banner	2-8	config flowstats source	12-19
config bgp add aggregate-address	18-9	config flowstats timeout ports	12-19
config bgp neighbor	18-10	config iparp add proxy	16-7
config bgp neighbor as-path-filter	8-23	config ip-mtu vlan	4-5, 4-7
config bgp neighbor nlri-filter	8-23	config iproute add default	3-6, 16-9
config bootprelay add	16-15	config iproute priority	16-8
config bootprelay delete	16-15	config iproute route-map	16-5
config debug-trace	C-8	config ipxrip delete	20-6
config debug-trace vlan	C-8	config ipxrip vlan export-filter	8-18
config diag	12-2	config ipxrip vlan import-filter	8-18
config diag off	12-2	config ipxroute add	20-5
config diffserv examination code-point	7-16	config ipxsap export-filter	8-18
config diffserv replacement	7-17	config ipxsap import filter	8-18
config dns-client add	2-14	config ipxsap service add	20-6
config dns-client default-domain	2-14	config jumbo-frame size	4-5
config dot1p type	7-13	config log display	12-11
config download server	5-21, B-6	config nat add vlan map source auto-constrain	9-7
config dvmrp add vlan	19-5	config nat add vlan map source destination	9-8
config dvmrp vlan export-filter	8-21	config nat delete vlan map source auto-constrain	9-7
config dvmrp vlan import-filter	8-21	config nat delete vlan map source destination	9-8
		config nat vlan	9-3

config ospf area nssa	17-9	config vlan add ports no-restart	14-18, 15-7
config ospf ase-limit	17-6	config vlan add ports restart	14-18, 15-7
config ospf area external-filter	8-19	config vlan add track-environment failover	14-5
config ospf area interarea-filter	8-19	config vlan add track-ping	14-6, 15-4
config ospf asbr-filter	8-19	config vlan add track-route	14-6, 15-4
config ospf direct-filter	8-19	config vlan add track-vlan	14-5, 15-3
config ospf vlan area	17-8	config vlan delete track-route	14-6, 15-4
config ospf vlan timer	17-17	config vlan delete track-vlan	14-5, 15-3
config pim add vlan	19-5	config vlan esrp esrp-election	14-4
config pim crp static	19-3	config vlan esrp priority	14-4
config pim vlan trusted-gateway	8-22	config vlan ipaddress	2-8, 3-5, 16-8
config ports auto off	2-8, 4-4	config vlan name	5-13
config ports auto on	4-4	config vlan priority	7-13
config ports monitor vlan	5-16	config vlan qosprofile	7-19
config ports qosprofile	7-18	config vlan subvlan-address-range	16-19
config ports redundant	4-20	config vlan tag	20-4
config ports vlan limit-learning	6-7	config vlan xnetid	20-4, 20-7
config ports vlan lock-learning	6-9, 6-10	configure sys-health-check alarm-level	C-9
config ports vlan unlimited-learning	6-8	create access-list icmp	8-10
config protocol add	5-11	create access-profile type	8-12
config qostype priority	7-19	create account	2-8, 2-13
config radius server client-ip	3-20	create bgp neighbor	18-10
config radius shared-secret	3-20, 3-21	create bgp peer-group	18-9
config radius-accounting	3-21	create eaps	11-7
config rip export direct	10-5	create fdbentry	7-9
config rip export vip	10-5	create fdbentry vlan dynamic	6-5
config rip vlan export-filter	8-16	create fdbentry vlan ports	6-5
config rip vlan import-filter	8-16	create ospf area	17-8
config rip vlan trusted-gateway	8-16	create route-map	8-24
config route-map add	8-25	create stpd	13-13
config route-map add goto	8-25	create vlan	2-9
config sharing address-based	4-9		
config slb failover	10-28		
config slot	2-8	D	
config slot module	4-2	delete access-list	8-4
config snmp add community	3-19	delete account	2-9
config snmp add trapreceiver	3-18	delete bgp peer-group	18-9
config snmp community	3-19	delete eaps	11-8
config snmp readonly access-profile	3-18	delete vlan	2-9
config snmp readwrite access-profile	3-18	disable bgp export	18-12
config snmp-client	3-36	disable bgp neighbor remove-private-as-number	18-11
config snmp-client update-interval	3-36	disable bootp	2-9
config ssh2 key	2-8, 3-9	disable cli-config-logging	2-9, 12-12
config ssh2 key pregenerated	3-9	disable clipaging	2-9
config stpd add vlan	13-13	disable eaps	11-11, 11-12
config sys-health-check alarm-level	12-7	disable edp ports	4-16
config sys-health-check auto-recovery	12-8	disable flowstats filter ports	12-21
config syslog	12-11	disable flowstats ping-check	12-22
config sys-recovery-level	2-8, 12-9	disable g1-module support	4-21
config time	2-8	disable idletimeout	2-9
config timezone	2-8, 3-34, 3-35	disable ignore-bpdu	13-4
config vlan add port	20-4		

disable ipforwarding fast-direct-broadcast	16-6	enable ipforwarding	16-9
disable ipforwarding ignore-broadcast	16-6	enable ipforwarding fast-direct-broadcast	16-6
disable ipxsap gns-reply	20-6	enable ipforwarding ignore-broadcast	16-6
disable learning ports	6-4	enable ipmcf forwarding	19-5
disable nat	9-9	enable jumbo-frame ports	4-5
disable ospf capability opaque-lsa	17-7	enable license	2-9
disable ospf export	16-4	enable log display	12-11
disable ospf export rip	17-14	enable nat	9-5
disable ospf export static	17-14	enable ospf	16-9
disable ospf export vip	17-14	enable ospf capability opaque-lsa	17-7
disable ports	2-9, 4-3	enable ospf export	16-4
disable radius	3-21	enable ospf export rip	17-14
disable radius-accounting	3-21	enable ospf export static	17-14
disable rip export	16-4, 17-15	enable ospf export vip	17-14
disable rmon	12-26	enable pim	19-5
disable sharing	4-13	enable ports	4-3
disable ssh2	2-9	enable radius	3-21
disable stpd rapid-root-failover	13-4	enable radius-accounting	3-21
disable sys-health-check	C-9	enable rip	16-9
disable telnet	2-9, 3-7	enable rip export	16-4, 17-15
disable web	2-9, 3-12	enable rmon	12-26
download bootrom	2-14	enable route sharing	16-5
download configuration	2-14, 5-21, B-5	enable sharing grouping	4-13
download configuration cancel	B-6	enable sntp-client	3-36
download configuration every	5-21, B-6	enable ssh2	2-10, 3-8
download configuration incremental	B-6	enable stpd	13-13
download image	2-14, B-2	enable stpd rapid-root-failver	13-4
		enable sys-health-check	C-9
		enable syslog	12-11
		enable telnet	2-10, 3-7
		enable web	2-10, 3-12
<hr/>			
E			
enable bgp aggregation	18-8		
enable bgp export	18-12		
enable bgp neighbor remove-private-as-numbers	18-11		
enable bootp	2-9		
enable bootp vlan	3-4		
enable bootprelay	16-15		
enable cli-config-logging	2-9, 12-12		
enable clipaging	2-9		
enable diffserv examination ports	7-15		
enable diffserv replacement ports	7-17		
enable dot1p replacement ports	7-14, 7-17		
enable dvmrp	19-5		
enable eaps	11-11		
enable edp ports	4-16		
enable flowstats	12-17		
enable flowstats filter ports	12-21		
enable flowstats ping-check	12-21		
enable flowstats ports	12-17		
enable g1-module support	4-20		
enable idletimeout	2-9		
enable ignore-bpdu	13-4		
<hr/>			
		H	
		history	2-7, 2-10
<hr/>			
		L	
		logout	3-6
<hr/>			
		N	
		nslookup	2-14
<hr/>			
		P	
		ping	2-14, 2-15
<hr/>			
		Q	
		quit	3-6

R	
reboot	B-2
run diag	12-2
run diagnostics	12-3
run diagnostics packet-memory	12-3

S	
save	3-6, B-3
scp2	3-11
show access-list	8-6
show access-list-monitor	8-6
show accounts	2-13
show banner	2-10
show debug-trace	C-8
show debug-trace vlan	C-9
show debug-tracing	C-8
show diagnostics	12-4, 12-8
show diagnostics packet-memory slot	12-4
show eaps	11-12
show edp	4-16
show esrp	14-6, 14-19, 14-23
show fdb	6-10
show fdb permanent	7-11, 7-22
show flowstats	12-22
show iparp	16-9, 16-21
show ipconfig	16-9, 16-15
show ipfdb	16-9
show iproute	16-9
show iproute route-map	16-5
show ipxconfig	20-7
show ipxrip	20-8
show ipxroute	20-7
show ipxsap	20-7
show ipxservice	20-8
show log	12-10
show management	3-7, 3-12, 3-19
show nat connections	9-9
show nat rules	9-8
show nat stats	9-9
show ospf	17-14
show ospf area	17-21
show ospf interfaces	17-21
show ospf lsdb	17-22
show ports configuration	4-14
show ports info	7-17, 7-19, 7-22, 14-3
show ports info detail	6-8
show ports qosmonitor	7-21
show ports rxerrors	12-5
show ports stats	12-4
show ports txerrors	12-5

show ports vlan statistics	5-16
show protocol	5-17
show qosprofile	7-11, 7-19, 7-21
show qostype priority	7-20
show session	3-6
show sharing address-based	4-10
show slot	4-2
show snmp client	3-37
show stpd	13-4, 13-16
show stpd port	13-16
show switch	3-37, 5-21, 7-22, B-6
show vlan 5-15, 7-19, 7-22, 14-19, 14-23, 16-20, 16-21, 20-4, 20-7	
show vlan security	6-8
show vlan stats	5-16
show vrrp	15-4
ssh2	3-10
synchronize	B-7

T	
telnet	2-14, 3-3
traceroute	2-14, 2-15

U	
unconfig eaps primary port	11-12
unconfig eaps secondary port	11-12
unconfig flowstats ports	12-22
unconfig ports monitor vlan	5-17
unconfig ports redundant	4-20
unconfig qostype priority	7-20
unconfig switch	2-10, B-3
unconfig switch all	B-3
upload configuration	2-14, B-4
upload configuration cancel	B-4
use configuration	B-3
use image	B-2