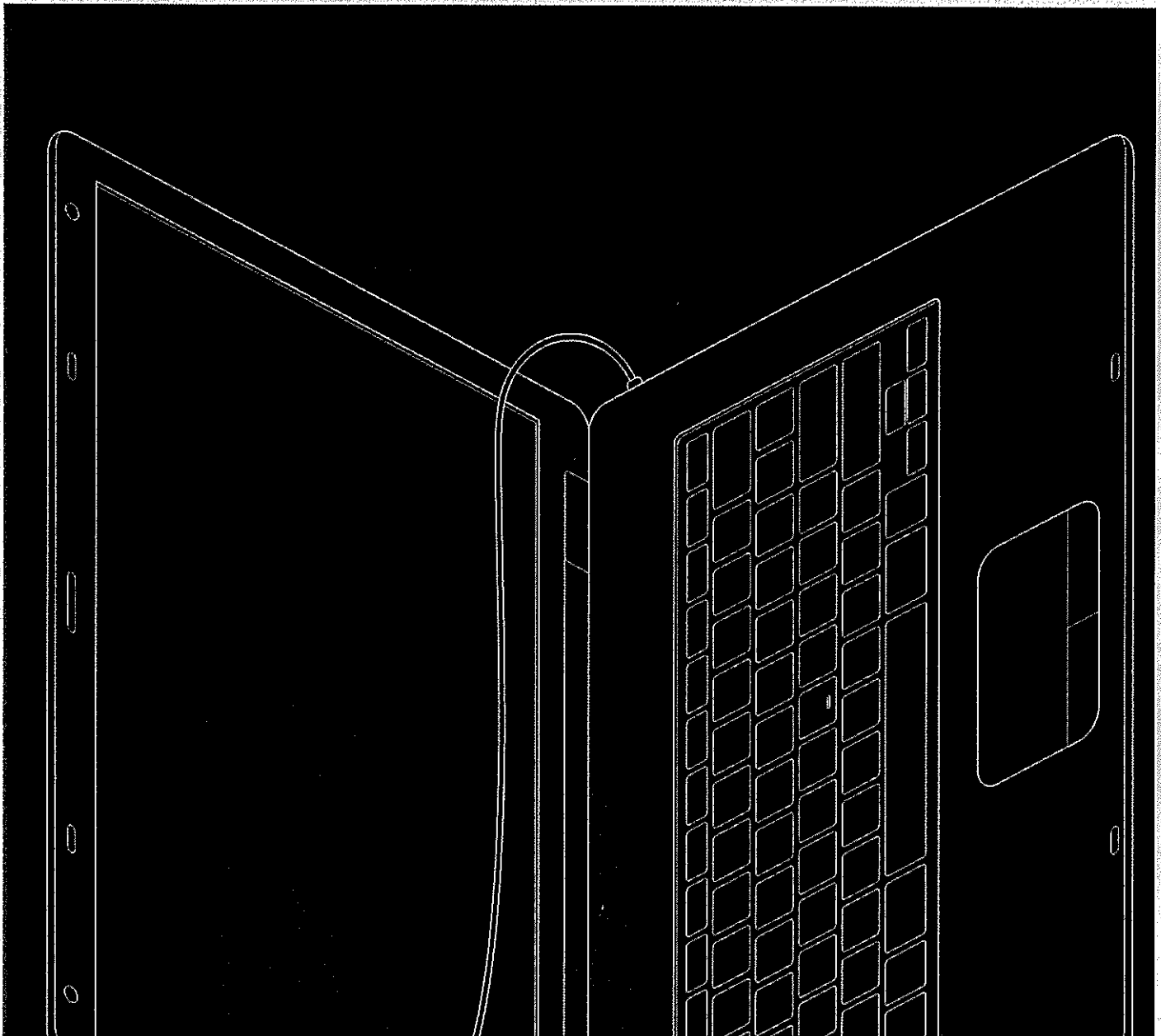


# JUNOS SERVICE PROVIDER SWITCHING

Student Guide

Revision 10.



# **Junos Service Provider Switching**

---

10.a

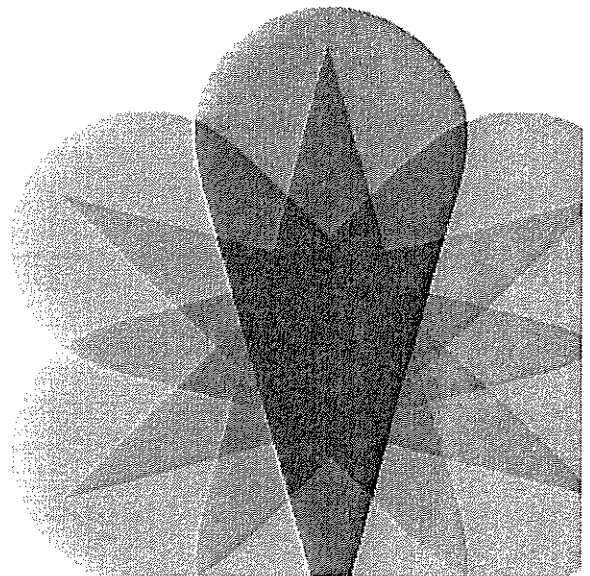
## ***Student Guide***

**JUNIPER**  
NETWORKS

Worldwide Education Services

1194 North Mathilda Avenue  
Sunnyvale, CA 94089  
USA  
408-745-2000  
[www.juniper.net](http://www.juniper.net)

Course Number: EDU-JUN-JSPX



# Contents

---

<b>Chapter 1:</b>	<b>Course Introduction</b> .....	<b>1-1</b>
<b>Chapter 2:</b>	<b>Carrier Ethernet</b> .....	<b>2-1</b>
	Ethernet in the WAN .....	2-3
	Ethernet Standards Organizations .....	2-10
	MX Series Layer 2 Features .....	2-29
<b>Chapter 3:</b>	<b>Ethernet Switching and Virtual LANs</b> .....	<b>3-1</b>
	Ethernet LANs .....	3-3
	Bridging .....	3-7
	Configuring and Monitoring VLANs .....	3-18
	Configuring and Monitoring IRB .....	3-33
	Layer 2 Address Learning and Forwarding .....	3-38
	Layer 2 Firewall Filtering .....	3-44
	Lab 1: Ethernet Switching and VLANs .....	3-57
<b>Chapter 4:</b>	<b>Virtual Switches</b> .....	<b>4-1</b>
	Routing Instances Overview .....	4-3
	Configuring and Monitoring Virtual Switches .....	4-8
	Interconnecting Routing Instances .....	4-16
	Lab 2: Virtual Switches .....	4-24
<b>Chapter 5:</b>	<b>Provider Bridging</b> .....	<b>5-1</b>
	Expanding the Bridged Network .....	5-3
	Provider Bridging .....	5-9
	Configuring and Monitoring Provider Bridging .....	5-19
	Provider Backbone Bridging .....	5-33
	Configuring and Monitoring Provider Backbone Bridging .....	5-43
	Lab 3: Provider Bridging .....	5-50
<b>Chapter 6:</b>	<b>Spanning-Tree Protocols</b> .....	<b>6-1</b>
	Overview of STP .....	6-3
	Overview of RSTP .....	6-18
	Overview of MSTP .....	6-30
	Overview of VSTP .....	6-38
	Configuring and Monitoring Spanning Tree Protocols .....	6-43
	Understanding BPDU, Loop, and Root Protection .....	6-55
	Lab 4: MSTP .....	6-68

## Course Overview

---

This two-day course is designed to provide students with intermediate switching knowledge and configuration examples. The course includes an overview of switching concepts such as LANs, Layer 2 address learning, bridging, virtual LANs (VLANs), provider bridging, provider backbone bridging, VLAN translation, spanning-tree protocols, and Ethernet Operation, Administration, and Maintenance (OAM). This course also covers Junos operating system-specific implementations of Integrated Routing and Bridging (IRB) interfaces, routing instances, virtual switches, load balancing, and port mirroring. This course is based on the Junos OS Release 10.0R2.10.

Through demonstrations and hands-on labs, students will gain experience in configuring and monitoring the Junos OS and in device operations.

### Objectives

After successfully completing this course, you should be able to:

- Describe carrier Ethernet.
- Describe the different Ethernet standards organizations.
- Describe the Layer 2 services that are available on the MX Series 3D Ethernet Universal Edge Routers.
- Describe the function of an Ethernet LAN.
- Describe learning and forwarding in a bridging environment.
- Describe Ethernet frame filtering.
- Implement VLAN tagging.
- Implement IRB.
- Implement a Layer 2 firewall filter.
- Describe the usage of a routing instance.
- Describe the function of a virtual router.
- Describe the function of a virtual switch.
- Implement a virtual switch.
- Describe interconnecting routing instances.
- Describe the different Institute of Electrical and Electronics Engineers (IEEE) VLAN stacking models.
- Describe the components of provider bridging.
- Configure and monitor provider bridging.
- Describe the components of provider backbone bridging.
- Configure and monitor provider backbone bridging.
- Explain the purpose of the Spanning Tree Protocol (STP).
- Describe the basic operation of the STP, the Rapid Spanning Tree Protocol (RSTP), the Multiple Spanning Tree Protocol (MSTP), and the VLAN Spanning Tree Protocol (VSTP)
- Configure and monitor the STP, the RSTP, the MSTP, and the VSTP.



# Course Agenda

---

## Day 1

- Chapter 1: Course Introduction
- Chapter 2: Carrier Ethernet
- Chapter 3: Ethernet Switching and Virtual LANs
  - Lab 1: Ethernet Switching and VLANs
- Chapter 4: Virtual Switches
  - Lab 2: Virtual Switches

## Day 2

- Chapter 5: Provider Bridging
  - Lab 3: Provider Bridging
- Chapter 6: Spanning-Tree Protocols
  - Lab 4: MSTP
- Chapter 7: Ethernet OAM
  - Lab 5: Ethernet OAM
- Chapter 8: Ethernet Ring Protection
  - Lab 6: Ethernet Ring Protection

## Defined and Undefined Syntax Variables

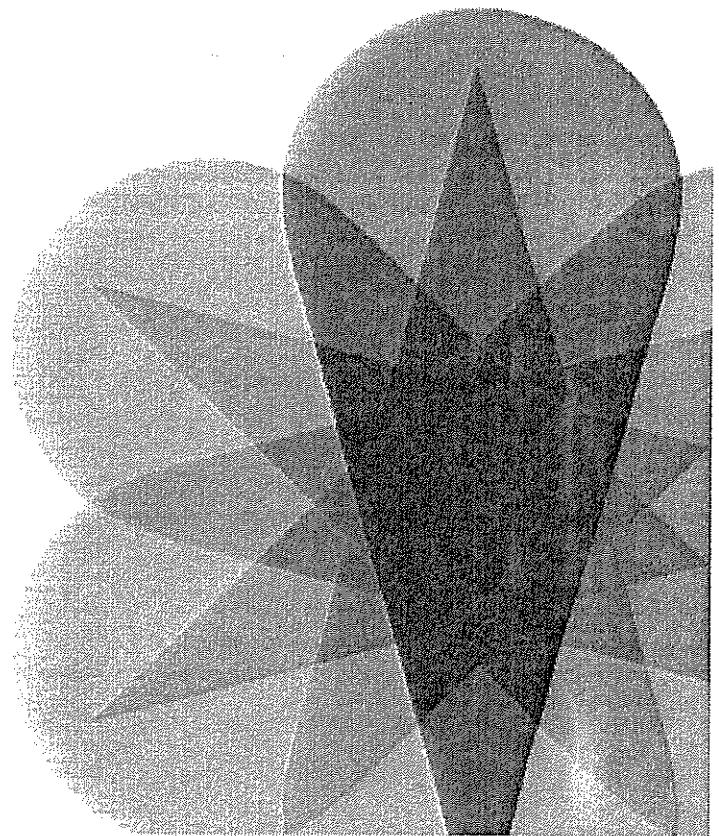
Finally, this course distinguishes between regular text and syntax variables, and it also distinguishes between syntax variables where the value is already assigned (defined variables) and syntax variables where you must assign the value (undefined variables). Note that these styles can be combined with the input style as well.

Style	Description	Usage Example
<i>CLI Variable</i>	Text where variable value is already assigned.	<code>policy my-peers</code>
<i>GUI Variable</i>		Click on <i>my-peers</i> in the dialog.
<u><i>CLI Undefined</i></u>	Text where the variable's value is the user's discretion and text where the variable's value as shown in the lab guide might differ from the value the user must input.	Type <code>set policy <u>policy-name</u></code> .
<u><i>GUI Undefined</i></u>		<code>ping 10.0.<u>x.y</u></code> Select File > Save, and type <u>filename</u> in the Filename field.



# **Junos Service Provider Switching**

## **Chapter 1: Course Introduction**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Get to know one another
  - Identify the objectives, prerequisites, facilities, and materials used during this course
  - Identify additional Juniper Networks courses
  - Describe the Juniper Networks Technical Certification Program

© 2013 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 1-2

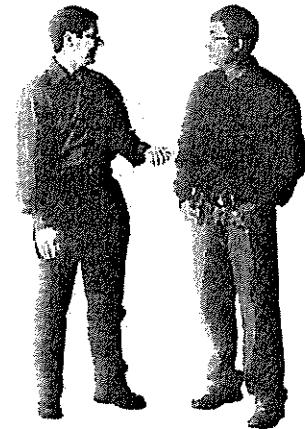
### This Chapter Discusses:

- Objectives and course content information;
- Additional Juniper Networks, Inc. courses; and
- Juniper Networks Technical Certification Program (JNTCP).

## Introductions

### ■ Before we get started...

- What is your name?
- Where do you work?
- What is your primary role in your organization?
- What kind of network experience do you have?
- What is the most important thing for you to learn in this training session?



© 2010 Juniper Networks, Inc. All rights reserved.

Juniper Worldwide Education Services

www.juniper.net | 1-3

## Introductions

The slide asks several questions for you to answer during class introductions.

## Course Contents

### ■ Contents:

- Chapter 1: Course Introduction
- Chapter 2: Carrier Ethernet
- Chapter 3: Ethernet Switching and Virtual LANs
- Chapter 4: Virtual Switches
- Chapter 5: Provider Bridging
- Chapter 6: Spanning-Tree Protocols
- Chapter 7: Ethernet OAM
- Chapter 8: Ethernet Ring Protection

© 2008 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 1-4

### Course Contents

The slide lists the topics we discuss in this course.

## Prerequisites

- The prerequisites for this course are the following:
  - Knowledge of the Open Systems Interconnection model
  - Experience with TCP/IP protocols
  - Junos OS configuration experience—the *Introduction to Junos Software* (IJS) course or equivalent
  - Routing knowledge—the *Junos Routing Essentials* (JRE) course or equivalent

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 1-5

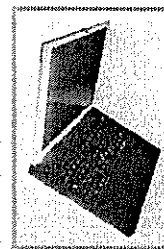
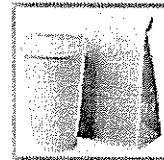
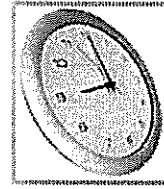
### Prerequisites

The slide lists the prerequisites for this course.

## Course Administration

### ■ The basics:

- Sign-in sheet
- Schedule
  - Class times
  - Breaks
  - Lunch
- Break and restroom facilities
- Fire and safety procedures
- Communications
  - Telephones and wireless devices
  - Internet access



© 2014 Juniper Networks, Inc. All rights reserved.

**JUNIPER**

Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 1-6

### General Course Administration

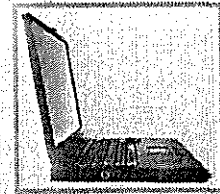
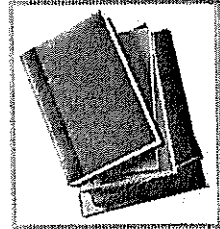
The slide documents general aspects of classroom administration.



## Education Materials

### ■ Available materials:

- In class:
  - Lecture material
  - Lab guide
  - Lab equipment
- Online:
  - eLearning courses



© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-7

### Training and Study Materials

The slide describes Education Services materials that are available for reference both in the classroom and online.

## Additional Resources

### ■ For those who want more:

- Juniper Networks Technical Assistance Center (JTAC)
  - <http://www.juniper.net/support/requesting-support.html>
- Juniper Networks books
  - <http://www.juniper.net/training/jnbooks/>
- Hardware and software technical documentation
  - Online: <http://www.juniper.net/techpubs/>
  - Image files for offline viewing:  
<http://www.juniper.net/techpubs/resources/cdrom.html>
- Certification resources
  - <http://www.juniper.net/training/certification/resources.html>



© 2003 Juniper Networks, Inc. All rights reserved.

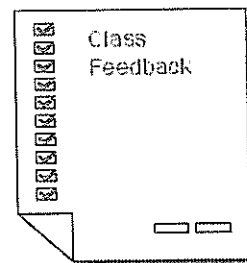
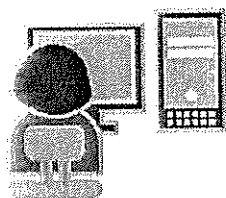
JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 1-8

### Additional Resources

The slide describes additional resources available to assist you in the installation, configuration, and operation of Juniper Networks products.

## Satisfaction Feedback



- To receive your certificate, you must complete the survey
  - Either you will receive a survey to complete at the end of class, or we will e-mail it to you within two weeks
  - Completed surveys help us serve you better!

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 1-9

### Satisfaction Feedback

Juniper Networks uses an electronic survey system to collect and analyze your comments and feedback. Depending on the class you are taking, please complete the survey at the end of the class, or be sure to look for an e-mail about two weeks from class completion that directs you to complete an online survey form. (Ensure that you provide us with your current e-mail address.)

Submitting your feedback entitles you to a certificate of class completion. We thank you in advance for taking the time to help us improve our educational offerings.

## Juniper Networks Education Services Curriculum

### ■ Formats:

- Classroom-based instructor-led technical courses
- Online instructor-led technical courses
- Hardware installation eLearning courses as well as technical eLearning courses

### ■ Complete list of courses:

- [http://www.juniper.net/training/technical\\_education/](http://www.juniper.net/training/technical_education/)

©2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 1-10

## Juniper Networks Education Services Curriculum

Juniper Networks Education Services can help ensure that you have the knowledge and skills to deploy and maintain cost-effective, high-performance networks for both enterprise and service provider environments. We have expert training staff with deep technical and industry knowledge, providing you with instructor-led hands-on courses in the classroom and online, as well as convenient, self-paced eLearning courses.

### Course List

You can access the latest Education Services offerings covering a wide range of platforms at [http://www.juniper.net/training/technical\\_education/](http://www.juniper.net/training/technical_education/).

## Technical Certification Programs

- Demonstrate competence with Juniper Networks technology
  - Multiple tracks
  - Multiple certification levels
  - Written proficiency exams
  - Hands-on configuration and troubleshooting exams
  - For more information and details on how to prepare for the exams
    - <http://www.juniper.net/training/certification/>

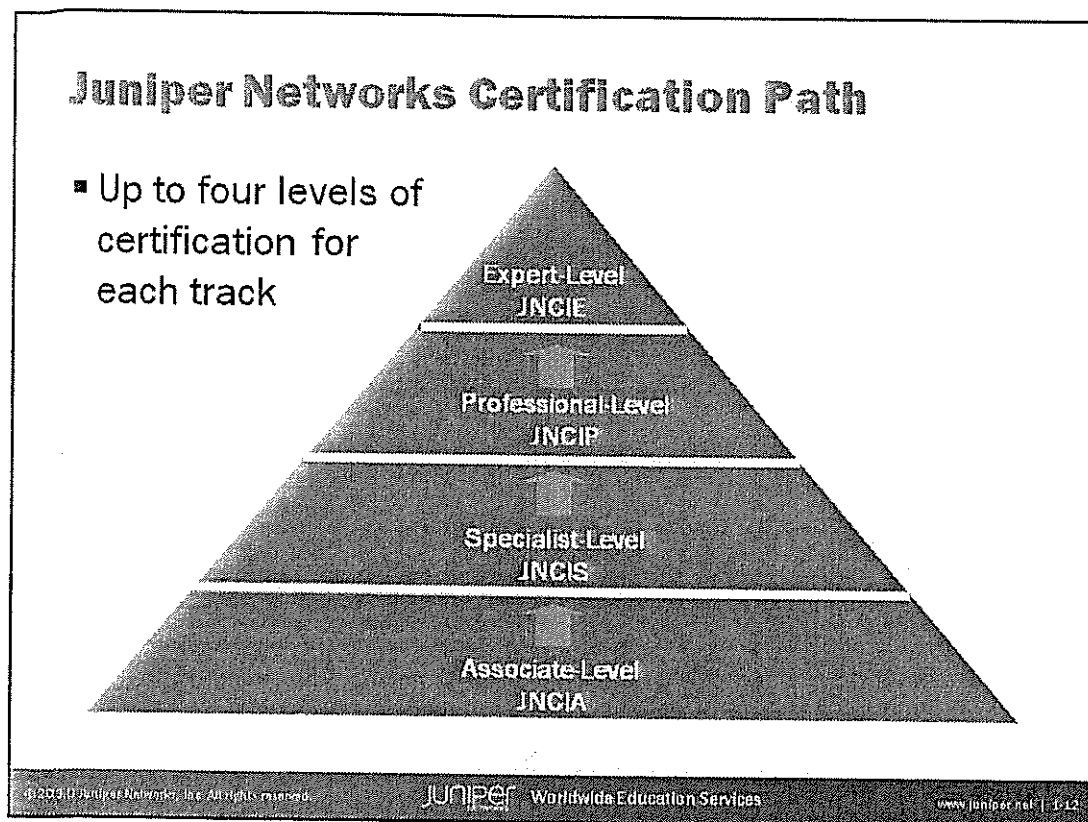
© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) 1-1-11

### JNTCP

The Juniper Networks Technical Certification Program (JNTCP) consists of platform-specific, multitiered tracks that enable participants to demonstrate, through a combination of written proficiency exams and hands-on configuration and troubleshooting exams, competence with Juniper Networks technology. Successful candidates demonstrate thorough understanding of Internet and security technologies and Juniper Networks platform configuration and troubleshooting skills. You can learn more information about the JNTCP at <http://www.juniper.net/training/certification/>.



### Certification Levels

Each JNTCP track has one to four certification levels. Juniper Networks Certified Internet Associate (JNCIA) and Juniper Networks Certified Internet Specialist (JNCIS) exams are computer-based exams composed of multiple choice questions. These computer-based exams are administered at Prometric testing centers worldwide and have no prerequisite certification requirements.

Juniper Networks Certified Internet Professional (JNCIP) and Juniper Networks Certified Internet Expert (JNCIE) exams are composed of hands-on lab exercises that are administered at select Juniper Networks testing centers. Professional-level and Expert-level exams require that you first obtain the next lower certification in the track. Please visit the JNTCP Web site at <http://www.juniper.net/training/certification/> for detailed exam information, exam pricing, and exam registration.

## Certification Preparation

### ■ How to prepare:

- Training and study resources
  - JNTCP Web site  
<http://www.juniper.net/training/certification/>
  - Education Services training classes  
[http://www.juniper.net/training/technical\\_education/](http://www.juniper.net/training/technical_education/)
  - Juniper Networks documentation and white papers  
<http://www.juniper.net/techpubs/>
- Practical exams: lots of hands-on practice
  - On-the-job experience
  - Education Services training classes
  - Equipment access

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 1-13

### Prepping and Studying

The slide lists some options for those interested in prepping for Juniper Networks certification.

## Questions



©2010 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

www.juniper.net | 1-14

### Any Questions?

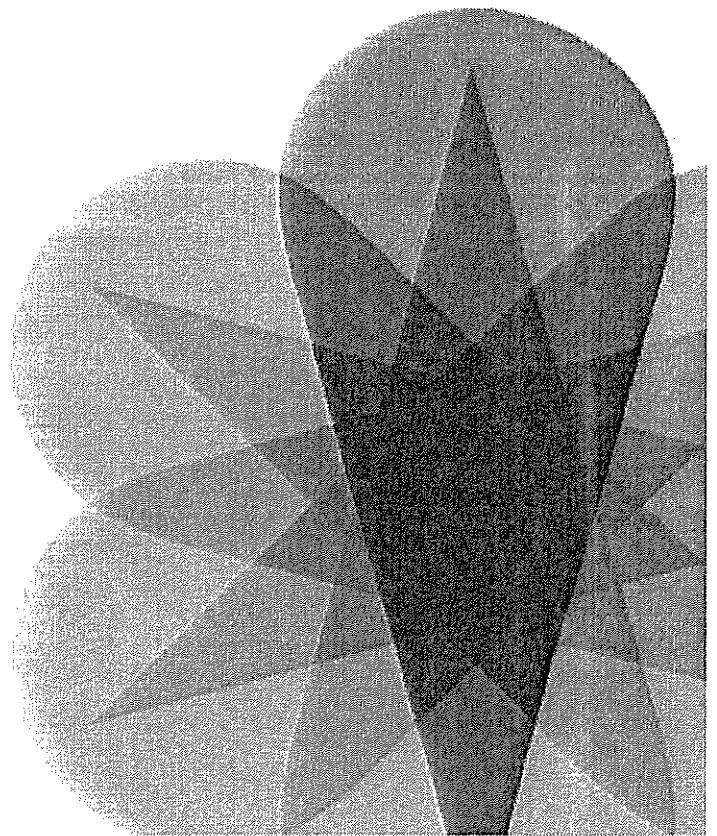
If you have any questions or concerns about the class you are attending, we suggest that you voice them now so that your instructor can best address your needs during class.





# **Junos Service Provider Switching**

## **Chapter 2: Carrier Ethernet**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Describe carrier Ethernet
  - Describe the different Ethernet standards organizations
  - Describe the Layer 2 services that are available on the Juniper Networks MX Series 3D Universal Edge Routers

©2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

[www.junipernet.com](http://www.junipernet.com) | 2-2

### This Chapter Discusses:

- Carrier Ethernet;
- Different Ethernet standards organizations; and
- Layer 2 services that are available on the Juniper Networks MX Series 3D Universal Edge Routers.

## Agenda: Carrier Ethernet

→ Ethernet in the WAN

- Ethernet Standards Organizations
- MX Series Layer 2 Features

©2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-3

### Ethernet in the WAN

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## Networking Terms

### ■ LAN:

- A network that is usually located at a single customer site
- Connects devices that are very close to one another

### ■ MAN:

- A network that is located within a metropolitan area (city or town)
- Connects devices that are within a few miles of each other

### ■ WAN:

- A network that can extend far beyond the MAN
- Connects devices that can be hundreds and thousands of miles away from each other

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-4

### Local Area Network

A LAN is usually a network of Ethernet switches and bridges that provides connectivity between end stations that in general are very close together. In most cases, the end stations and switches are within the same building.

### Metropolitan Area Network

A Metropolitan Area Network (MAN) is located within the confines of a city or town where a service provider might have a fiber infrastructure or a cable infrastructure. A MAN provides the ability to connect customer sites that are located near each other.

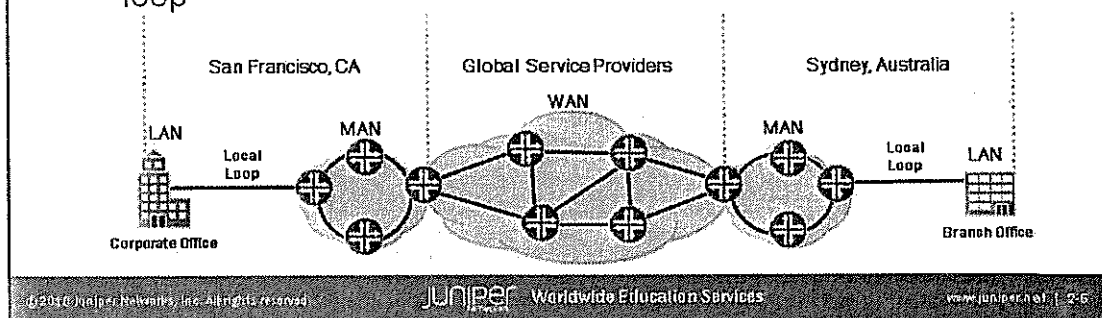
### Wide Area Network

A WAN allows for connectivity that extends far beyond the MAN. A WAN typically connects devices that are hundreds and thousands of miles away from each other.

## Service Providers

- Service providers (in most cases a phone or cable company) typically own and operate MANs and WANs:

- Service providers usually sell the MAN and WAN service to their customers for a monthly fee
- In general, MAN and WAN networks are made up of vast underground, underwater, and overhead networks of fiber
- A customer gains access to the MAN or WAN through a local loop

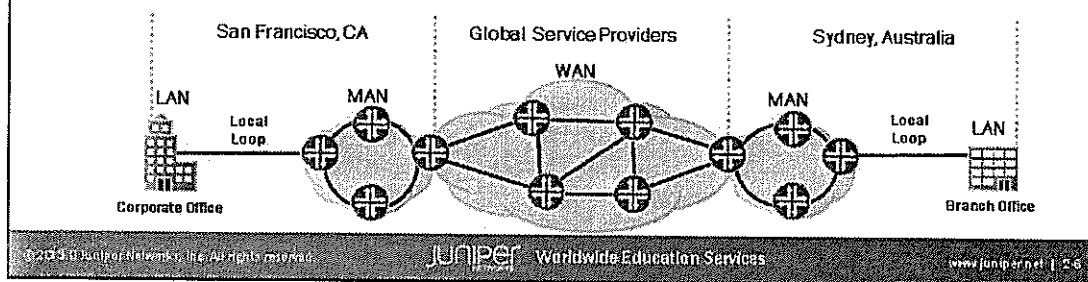


## Service Providers

When a business decides to interconnect two or more sites that are not physically near each other, a service provider usually provides MAN or WAN connectivity between those sites at a price. A service provider (cable company or telco) has the facilities—such as the miles and miles of fiber—that are necessary to transfer data around the world. A customer of the service provider gains access to the MAN or WAN through a local loop or access circuit that the service provider delivers to each site.

## End-to-End Connectivity

- Several options exist to provide connectivity between customer sites:
  - Private line (DS0, T1, E1, or T3)
  - ATM
  - Frame Relay
  - Ethernet
- The customer must purchase and maintain equipment that can support the Layer 1 and Layer 2 interfaces



## Site-to-Site Connectivity Options

Customers have options when it comes to ordering service from the service provider. If the customer sites are relatively close to each other, the customer can purchase a private-line service from the service provider. A private line is a point-to-point circuit that customers can order at varying speeds (DS0, T1, E1, T3, and more). As the distance between sites grows, so does the price for the private-line service. Other options for site-to-site connectivity include Asynchronous Transfer Mode (ATM), Frame Relay, and now Ethernet.

## Equipment and Expertise

To support the site-to-site connectivity, the customer must purchase the correct equipment and have the expertise to be able to support the new circuits. The customer will need Ethernet experts for the LAN, and, in the case of ATM WAN connectivity, they will need ATM experts as well.

## Ethernet as a MAN and WAN Solution

- **Bandwidth needs from the MAN and WAN:**
  - As the bandwidth usage of the customers has increased, ATM and Frame Relay port speeds have not kept up
  - Private-line prices have become far too high
- **Ethernet becomes the MAN or WAN solution:**
  - Ethernet speeds have increased over time
    - 10 Mbps, 100 Mbps, 1 Gbps, and 10 Gbps
  - Ethernet is a technology that the customer understands
    - Customer IT groups are generally Ethernet experts
  - Allows the customer to have a single interface to the service provider
    - Internet access and site-to-site connectivity over a single interface
  - Lower-cost customer premise equipment
    - High-speed ATM equipment—significantly higher cost than Ethernet

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-7

### Bandwidth Becomes a Factor

Over the last 10 years, the need for high-speed access to the Internet as well as for site-to-site connectivity has skyrocketed. With more and more video, voice, and other bandwidth-hogging applications being placed on the network, ATM and Frame Relay networks have not been able to keep up with the demand.

### Ethernet Is the Solution

Ethernet interfaces as fast as 10 Gbps are available. Soon speeds will exceed that limit as well. An Ethernet solution in the WAN benefits both the service provider and the customer in many ways. Using Ethernet as the WAN solution, the customer no longer needs Ethernet and ATM experts to run the network. Service providers can offer multiple services using a single interface to the customer. The list of benefits can go on and on.

## Initial Ethernet Challenges

### ■ Scalability:

- Millions of customer end stations
  - The service provider environment learns a large number of MAC addresses
- Redundant connections between the service provider and the customers for resiliency
  - Spanning tree protocols simply cannot scale

### ■ SLAs:

- Native Ethernet frames do not provide quality of service
  - Best effort had been generally accepted in LAN network

### ■ OAM:

- Fault management and performance monitoring
- Monitoring and troubleshooting Ethernet access links
- Circuit protection

© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-8

## Scalability

Allowing an Ethernet WAN to scale has always posed a challenge to the service provider. For instance, for an Ethernet switch to forward Ethernet frames it must learn the MAC address of each of the end stations on the customer network. For a service provider serving thousands of customers, this need might mean that the service provider-owned switches must potentially learn millions of MAC addresses. Also, when redundant links exist between the service provider and its customers for resiliency purposes, the question arises, "How can you prevent a loop?" The spanning tree protocols of today simply cannot scale to prevent the loops of thousands of customer sites.

## Service-Level Agreements

Usually when a customer purchases WAN service, service-level agreements (SLAs) are in place to ensure that the service provider provides a good service to the customer. Common SLAs would cover frame delay and frame loss.

*Continued on next page.*



## Operation, Administration, and Maintenance

The ability for a service provider to provide and prove the same level of service with Ethernet that a customer could get from ATM, Frame Relay, and private-line service needed to be developed. Ethernet was also lacking Operation, Administration, and Maintenance (OAM) features. For example, in the case of ATM, OAM features would allow administrators to verify the status of ATM permanent virtual circuits (PVCs). This same capability was necessary for Ethernet virtual connections (EVCs).

## **Agenda: Carrier Ethernet**

- Ethernet in the WAN
- Ethernet Standards Organizations
- MX Series Layer 2 Features

© 2013 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 2-10

### **Ethernet Standards Organizations**

The slide highlights the topic we discuss next.

## Internet Standards Organizations

- Ethernet standards organizations are developing standards for Ethernet services, architecture, OAM, and interfaces:
  - Metro Ethernet Forum
  - Institute of Electrical and Electronics Engineers
  - International Telecommunication Union

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER  
NETWORKS

Worldwide Education Services

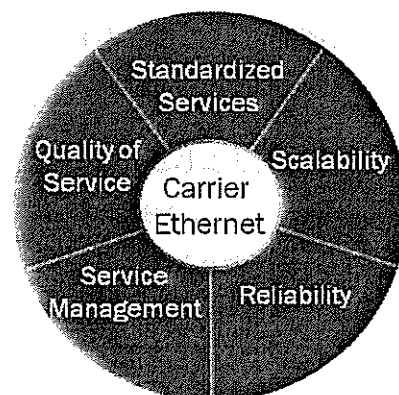
[www.juniper.net](http://www.juniper.net) | 2-11

## Ethernet Standards Organizations

Several organizations have been working to solve the problems that Ethernet poses in the WAN. The three primary organizations that are helping to enable Ethernet WAN services are the Metro Ethernet Forum (MEF), the Institute of Electrical and Electronics Engineers (IEEE), and the International Telecommunication Union (ITU).

## Metro Ethernet Forum

- Nonprofit international industry consortium dedicated to accelerating worldwide adoption of carrier Ethernet networks and services



- Defines carrier Ethernet (also referred to as Metro Ethernet) as a ubiquitous, standardized, carrier-class service defined by five attributes that distinguish carrier Ethernet from LAN-based Ethernet

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-12

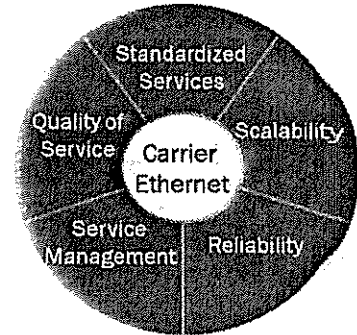
### Metro Ethernet Forum

The MEF, as the defining body of carrier Ethernet, is a global alliance of over 150 organizations including service providers, cable multiple service operators (MSOs), network equipment manufacturers, software manufacturers, semiconductor vendors, and testing organizations. The goal of the MEF is to accelerate the worldwide adoption of carrier Ethernet networks and services. The MEF defines carrier Ethernet as a ubiquitous, standardized, carrier-class service defined by five attributes (illustrated on the slide) that distinguish carrier Ethernet from LAN-based Ethernet. An objective of the MEF is to build a consensus and unite service providers, equipment vendors, and customers on Ethernet service definitions, technical specifications, and interoperability.

## Metro Ethernet Forum Attributes (1 of 2)

### ▪ MEF attributes:

- Standardized services:
  - E-Line, E-LAN, and E-Tree
  - Require no change to customer LAN equipment
  - Suited for converged voice, video, and data networking
  - Wide choice of granularity of bandwidth and quality of service options
- Scalability:
  - The ability for millions of customers to use the service
  - Spans access, metropolitan, national, and global networks with a wide variety of physical infrastructures and service providers



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-13

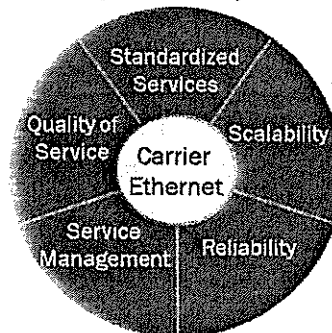
### MEF Attributes: Part 1

The slide discusses the definitions of the Standardized Services and Scalability attributes.

## Metro Ethernet Forum Attributes (2 of 2)

### ■ MEF attributes (contd.):

- Reliability
  - Rapid recovery time
  - The network should be able to detect and recover from outages without impacting users
- Quality of Service
  - Many bandwidth and quality of service options
  - SLAs for end-to-end performance based on committed information rate, frame loss, delay, and delay variation
- Service Management
  - Ability to monitor, diagnose, and centrally manage the network using standards-based implementations
  - Carrier class OAM



© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-14

### MEF Attributes: Part 2

The slide discusses the definitions of the Reliability, Quality of Service, and Service Management attributes.

## MEF Standards (1 of 2)

### ■ Carrier Ethernet technical specifications:

MEF 2	Requirements and Framework for Ethernet Service Protection
MEF 3	Circuit Emulation Service Definitions, Framework and Requirements in Metro Ethernet Networks
MEF 4	Metro Ethernet Network Architecture Framework Part 1: Generic Framework
MEF 6.1	Metro Ethernet Services Definitions Phase 2 (per 604)
MEF 7	EMS-NMS Information Model
MEF 8	Implementation Agreement for the Emulation of PDH Circuits over Metro Ethernet Networks
MEF 9	Abstract Test Suite for Ethernet Services at the UNI
MEF 10.2	MEF 10.2 Ethernet Services Attributes Phase 2 (per 604)
MEF 11	User Network Interface (UNI) Requirements and Framework
MEF 12	Metro Ethernet Network Architecture Framework Part 2: Ethernet Services Layer
MEF 13	User Network Interface (UNI) Type 1 Implementation Agreement
MEF 14	Abstract Test Suite for Traffic Management Phase 1
MEF 15	Requirements for Management of Metro Ethernet Phase 1 Network Elements
MEF 16	Ethernet Local Management Interface
MEF 17	Service OAM Framework and Requirements

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-15

### Technical Specifications: Part 1

The slide shows the MEF-developed carrier Ethernet technical specifications.

## MEF Standards (2 of 2)

### ■ Carrier Ethernet technical specifications (contd.):

- All specifications are available for download at <http://metroethernetforum.org/>

MEF 18	Abstract Test Suite for Circuit Emulation Services
MEF 19	Abstract Test Suite for UNI Type 1
MEF 20	UNI Type 2 Implementation Agreement (pdf 6/08)
MEF 21	Abstract Test Suite for UNI Type 2 Part 1 Link OAM
MEF 22	Mobile Backhaul Implementation Agreement (2/09)
MEF 23	Class of Service Phase 1 Implementation Agreement (supersedes any file posted here before November 3, 2009)
MEF 24	Abstract Test Suite for UNI Type 2 Part 2 E-LMI
MEF 25	Abstract Test Suite for UNI Type 2 Part 3 Service OAM
MEF 26	External Network Network Interface (ENNI)-Phase 1

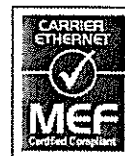
### Technical Specifications: Part 2

The slide shows the continuation of the MEF-developed carrier Ethernet technical specifications table.



## MEF Equipment and Service Provider Certification

- MEF launched a certification program in 2005 to verify compliance of vendor equipment and service-provider services to MEF technical specifications



- Eliminates the need for expensive and complex testing between equipment vendors
- Establishes a solid foundation for carrier Ethernet interoperability
- Provides for a single, universally recognized test and certification process
- Accelerates carrier Ethernet deployment at reduced costs
- Eases making informed decisions about equipment vendors

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-17

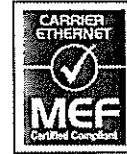
### MEF Certification Program

To help in its objective to promote interoperability between service providers and equipment vendors, the MEF introduced a new certification program in 2005. The certification applies to both service providers and equipment vendors. Having a standardized certification all but eliminates the need for expensive and complex interoperability tests.

## Four MEF Certifications

### ■ MEF 9 (User-to-network interface capabilities):

- Juniper Networks-certified equipment
  - MX Series devices



### ■ MEF 14 (Frame delay, loss, and jitter):

- Juniper Networks-certified equipment
  - MX Series devices

### ■ MEF 18 (CESoETH):

- No Juniper Networks-certified equipment

### ■ MEF 21 (Link OAM):

- Juniper Networks-certified equipment
  - MX Series devices
  - M Series Multiservice Edge Routers
  - T Series Core Routers

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-18

## MEF 9 Certification

The MEF 9 certification tests for compliance with MEF 6.1, 10, and 11. This test ensures the meeting of all requirements at the user-to-network interface (UNI). Some of the tests include:

- Non-looping frame delivery;
- Single copy broadcast and multicast delivery; and
- Customer VLAN (C-VLAN) ID preservation.

## MEF 14 Certification

The MEF 14 certification tests for compliance with MEF 9 and 10. This test ensures the meeting of all requirements for traffic management. Some of the tests include:

- Frame delay service performance;
- Frame delay variation service performance; and
- Frame loss ration service performance.

*Continued on next page.*

## MEF 18 Certification

The MEF 18 certification tests for compliance with MEF 8. This certification ensures the meeting of all requirements for reliable transport of time-division multiplexing (TDM) circuits. This certification includes some of the following tests:

- Encapsulation layers;
- Payload format; and
- Defects.

## MEF 21 Certification

The MEF 21 certification tests for compliance with MEF 20. This certification ensures the meeting of all requirements for UNI Type 2 and link OAM features.

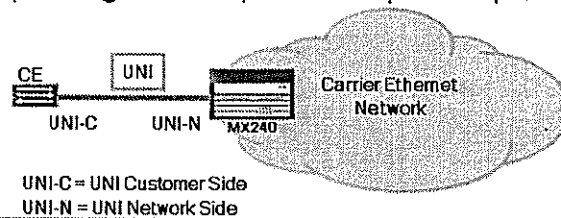
## Carrier Ethernet Terminology (1 of 2)

### ■ Carrier Ethernet terms:

#### • UNI:

- A physical interface or port that is the demarcation between the customer and the service provider
- UNI Type 1: Compliant with MEF 13 and manually configurable
- UNI Type 2: Automatic service discovery through Ethernet-Local Management Interface: supports OAM
- UNI Type 3: Provides for dynamic EVC setup
- An Ethernet interface operating at 10 Mbps, 100 Mbps, 1 Gbps, or 10 Gbps

#### • Customer equipment



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-20

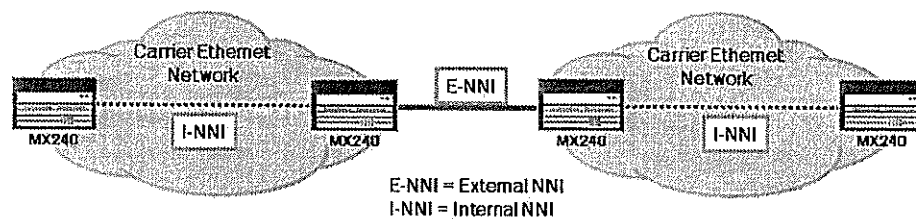
### Carrier Ethernet Terms: Part 1

The slide lists some of the common terms found in a carrier Ethernet network.

## Carrier Ethernet Terminology (2 of 2)

### ▪ Carrier Ethernet terms (contd.):

- Network-to-network interface
  - A physical interface or port that is the demarcation between distinct carrier Ethernet networks, operated by one or more service providers
- Carrier Ethernet network
  - An access, metropolitan, national, or global Ethernet transport network connecting user endpoints



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-21

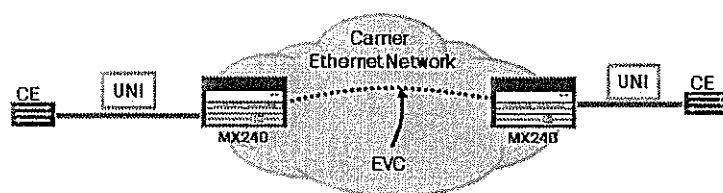
### Carrier Ethernet Terms: Part 2

The slide discusses some of the common terms found in a carrier Ethernet network.

## Ethernet Virtual Connection

### ▪ EVC:

- Connects two or more customer sites or UNIs
- Prevents data transfer between sites that are not part of the same EVC
- Defined in MEF 6.1 and 10.2
  - Point-to-point
  - Multipoint-to-multipoint
  - Rooted multipoint



© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-22

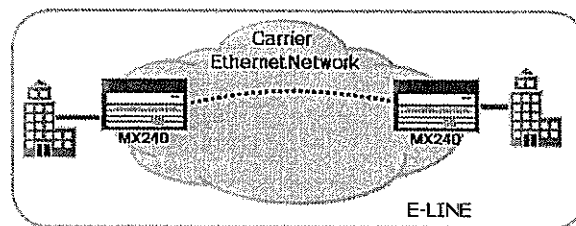
## Ethernet Virtual Connection

An (EVC) is a carrier Ethernet service offered by a service provider. It connects two or more sites. A requirement of an EVC is to prevent data transfer between UNIs that are not part of the same EVC. Three types of EVCs exist: point-to-point, multipoint-to-multipoint, and rooted multipoint.

## Point-to-Point EVC

### ▪ E-Line service EVCs

- Two types:
  - Ethernet Private Line (port-based)
  - Virtual Private Line (VLAN-based)
- Allow for communication between only two UNIs



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-23

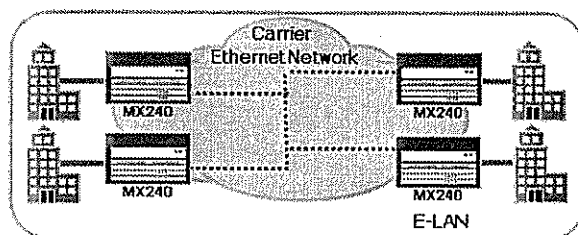
### E-Line Service EVC

A point-to-point EVC is referred to as an Ethernet Line (E-Line) EVC. It provides connectivity between only two UNIs. Two types of E-Line EVCs exist. An Ethernet Private Line EVC is port-based, where each of the UNIs is a dedicated port to a customer. All virtual LANs (VLANs) for the UNI can traverse the EVC. A Virtual Private Line EVC is VLAN-based, such that it allows for the mapping of individual VLANs to the EVC. This mapping allows the service provider to multiplex multiple customers using a single access port.

## Multipoint-to-Multipoint EVC

### ■ E-LAN service EVCs

- Two types:
  - Ethernet Private LAN (port-based)
  - Virtual Private LAN (VLAN-based)
- Allows for communication between two or more UNIs
  - Ingress broadcast or multicast frames at one UNI are forwarded to all other UNIs



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-24

### E-LAN Service EVCs

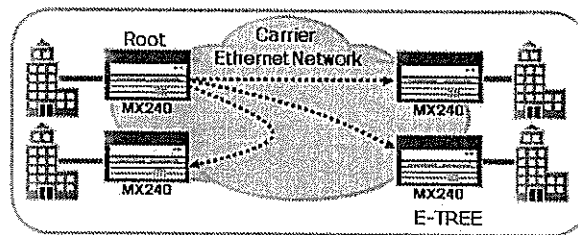
Multipoint-to-multipoint EVCs are referred to as Ethernet LAN (E-LAN) EVCs. Essentially, an E-LAN EVC makes the service provider network appear to be a single broadcast domain to the customer. E-LAN EVCs come in the form of either an Ethernet Private LAN or Virtual Private LAN, similar to the E-Line EVC.



## Rooted Multipoint EVC

### ▪ E-Tree service EVCs

- Two types:
  - Ethernet Private Tree (Port-based)
  - Virtual Private Tree (VLAN-based)
- A root UNI can send ingress frames to one or all leaf UNIs
- A leaf UNI can exchange data only with the root UNI
- Useful for multicast video applications



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-25

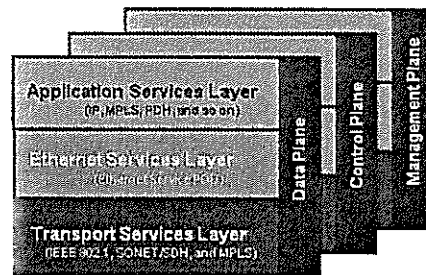
### E-Tree Service EVCs

Rooted multipoint EVCs are referred to as E-Tree EVCs. The slide describes the forwarding properties of an E-Tree EVC. E-Tree EVCs come in the form of either an Ethernet Private Tree or Virtual Private Tree, similar to the E-Line EVC.

## MEF Layered Approach

### ■ Three-layer model:

- Application Layer
  - End-user applications carried by the Ethernet Services Layer
- Ethernet Services Layer
  - EVCs
- Transport Services Layer
  - Various networking and media types that deliver the Ethernet services



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-26

### MEF's Three-Layer Model

The MEF has defined a three-layer model for carrier Ethernet networks. The Application Services Layer supports end-user applications. The Ethernet Services Layer carries the applications. This layer is the main focus of the MEF. Carrier Ethernet resides on the Ethernet Services Layer. To deliver the Ethernet services, the Transport Services Layer uses various networking and media types. This layer includes technologies like provider backbone bridging, virtual private LAN service (VPLS), and SONET. As shown on the slide, each layer of the MEF model has its own data, control, and management planes.

## IEEE Ethernet Standards

- The IEEE Ethernet standards fall into the 802 category:
  - IEEE 802.3—Physical Layer and Data Link MAC sublayer for wired Ethernet
  - IEEE 802.1—Bridging and management
    - 802.1D/802.1Q: Bridges and VLAN
    - 802.1ad: Provider bridging
    - 802.1ah: Provider backbone bridging
    - 802.1ag: Connectivity fault management
    - Many more

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-27

### IEEE Standards

The slide lists some of the important IEEE Ethernet standards.

## ITU-T Recommendations

- ITU-T is the ITU's telecommunication standardization sector:
  - G series—Transmission systems and media, digital systems, and networks
    - G.8010: Architecture of Ethernet Layer networks
    - G.8011.1: Ethernet Private Line Service
    - G.8011.2: Ethernet Virtual Private Line Service
    - G.8032: Ethernet Ring Protection
  - Y series—Global information infrastructure, IP aspects, and next-generation networks
    - Y.1730: Ethernet OAM requirements
    - Y.1731: OAM mechanisms

Note: Information is current as of Feb 11, 2010

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 2-28

### ITU-T Recommendations

The slide shows some of the International Telecommunication Union Telecommunication Standardization (ITU-T) Ethernet recommendations.

## **Agenda: Carrier Ethernet**

- Ethernet in the WAN
- Ethernet Standards Organizations
- MX Series Layer 2 Features

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-29

### **MX Series Layer 2 Features**

The slide highlights the topic we discuss next.

## MX Series Highlights

### ■ MX Series highlights:

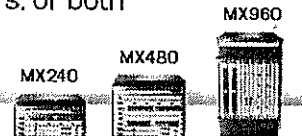
- Designed for next-generation services at the Ethernet edge
  - Function as Layer 2 switches, Layer 3 routers, or both
  - Provider edge for Layer 3 VPN, Layer 2 VPN, VPLS, or the Internet

### • Scalable and reliable

- Fully redundant design
- Distributed packet forwarding
- Full set of Junos OS routing capabilities

### • MX240 and MX480

- Mid-range platforms
- Optimized for sites with space and power restrictions



	MX240	MX480	MX960
Units per 7' rack	9	6	3
Gbps Capacity	240 G	480 G	960 G
10 GE and GE ports	12 / 120	24 / 240	48 / 480
MAC Addresses	1 million		

## MX Series Highlights

The slide shows some of the highlights of the MX Series devices.

## MX Series Layer 2 Features

### ▪ MX Series devices support:

#### • IEEE

- 802.1D: Bridging
- 802.1Q: VLAN tagging
- 802.1ad: Provider bridging
- 802.1ah: Provider backbone bridging
- 802.1ag: CFM
- 802.3, clause 57: LFM

#### • ITU-T

- Y.1731: CFM and Frame Delay Measurement
  - G.8032: Ethernet Ring Protection
- #### • Internet Engineering Task Force
- RFC 4761: VPLS using BGP
  - RFC 4762: VPLS using LDP

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-31

### Layer 2 Features

The slide shows some of the Layer 2 features supported on MX Series devices.

## Summary

■ In this chapter, we:

- Described carrier Ethernet
- Described the different Ethernet standards organizations
- Described the Layer 2 services that are available on the MX Series devices

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 2-32

### This Chapter Discussed:

- Carrier Ethernet;
- Different Ethernet standards organizations; and
- Layer 2 services that are available on MX Series devices.



## Review Questions

1. List two properties that make carrier Ethernet more desirable than older WAN methods like Frame Relay and ATM.
2. List the three prominent Ethernet standards organizations.
3. List three Layer 2 services that an MX Series device can provide.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 2-33

## Review Questions

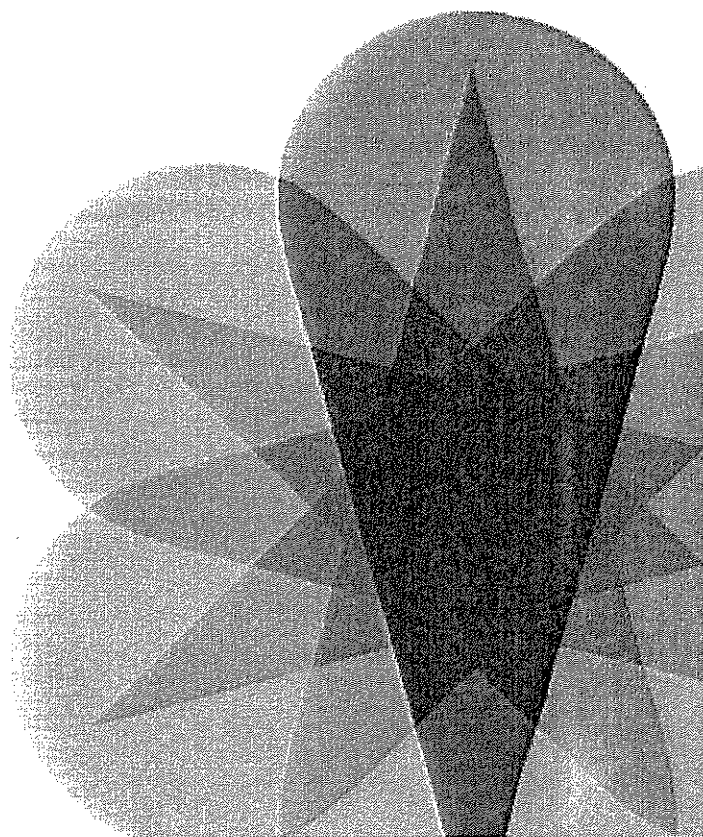
- 1.
- 2.
- 3.





# **Junos Service Provider Switching**

## **Chapter 3: Ethernet Switching and Virtual LANs**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Describe the functions of an Ethernet LAN
  - Describe learning and forwarding in a bridging environment
  - Implement VLAN tagging
  - Implement IRB
  - Implement Layer 2 address learning and forwarding
  - Implement Layer 2 firewall filters

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-2

### This Chapter Discusses:

- The functions of an Ethernet LAN;
- Learning and forwarding in a bridging environment;
- Implementation of virtual LAN (VLAN) tagging;
- Implementation of integrated routing and bridging (IRB);
- Implementation of Layer 2 address learning and forwarding; and
- Implementation of Layer 2 firewall filters.

## **Agenda: Ethernet Switching and Virtual LANs**

### → Ethernet LANs

- Bridging
- Configuring and Monitoring VLANs
- Configuring and Monitoring IRB
- Layer 2 Address Learning and Forwarding
- Layer 2 Firewall Filtering

© 2010 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

www.juniper.net | 3-3

### **Ethernet LANs**

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## Overview of Ethernet

### ■ Ethernet defined:

- Family of LAN specifications, standardized in IEEE 802.3
  - 10Base-T (802.3i)—10 Mbps
  - 100Base-TX (802.3u)—100 Mbps
  - 1000Base-T (802.3ab)—1000 Mbps
- Uses Data Link Layer technology to create LANs
  - Shared medium—a single broadcast and collision domain
  - Uniquely identifies all nodes on the LAN with a 48-bit MAC address
- Uses CSMA/CD to avoid and manage frame collisions

© 2003 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-4

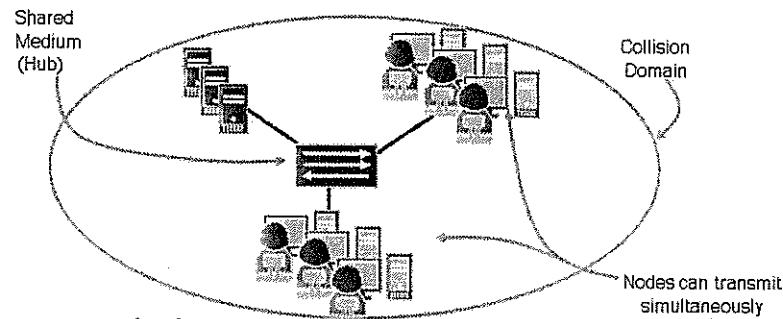
### Ethernet Defined

Ethernet is a family of LAN specifications defined in the Institute of Electrical and Electronics Engineers (IEEE) 802.3 standard. The slide lists some common examples, including the 803.3i, 802.3u, and 802.3ab specifications. Each Ethernet implementation uses a unique wiring and signaling standard—typically a copper-based medium or fiber optics—for the Physical Layer. Although the various implementations of Ethernet can use various wiring and signaling standards, they all use a common addressing format.

Ethernet is a Data Link Layer technology, as defined by Layer 2 of the Open Systems Interconnection (OSI) model of communications. An Ethernet LAN consists of a shared medium that encompasses a single broadcast and collision domain. Network devices, referred to as nodes, on the Ethernet LAN transmit data in bundles that are generally referred to as frames. Each node on a LAN has a unique identifier so that it can be unambiguously located on the network. Ethernet uses the Layer 2 media access control (MAC) address for this purpose. MAC addresses are 48-bit hardware addresses programmed into the Ethernet processor of each node.

Ethernet uses the carrier-sense multiple access with collision detection (CSMA/CD) protocol to avoid and manage frame collisions.

## Ethernet LANs (1 of 2)



### ■ Characteristics:

- Shared medium
- Single collision domain
- Nodes can transmit simultaneously

### Problems:

1. All devices see traffic
2. Collisions can occur
3. Unwanted resource consumption

©2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

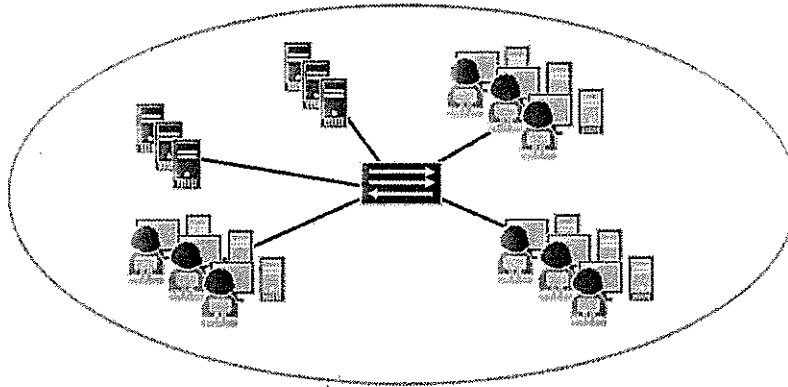
www.juniper.net | 3-5

## Ethernet LANs: Part 1

Ethernet LANs consist of a shared medium that defines a single collision domain. As previously mentioned, Ethernet uses the CSMA/CD protocol to help avoid and manage frame collisions. The sample topology on the slide shows a series of nodes connected through a hub using a copper-based physical medium. This type of implementation allows only a single stream of data at a time. All nodes participating in this shared Ethernet LAN listen to verify that the line is idle before transmitting. If the line is idle, the nodes begin transmitting data frames. If multiple nodes listen and detect that the line is idle and then begin transmitting data frames simultaneously, a collision occurs. When collisions occur, an error is generated and travels back to the transmitting devices. When a node receives a collision error message, it stops transmitting immediately and waits for a period of time before trying to send the frame again. If the node continues to detect collisions, it progressively increases the time between retransmissions in an attempt to find a time when no other data is being transmitted on the LAN. The node uses a backoff algorithm to calculate the increasing retransmission time intervals. When a node does successfully transmit traffic, that traffic replicates out all ports on the hub and all other nodes on the shared Ethernet segment see it. This traffic-flooding approach, coupled with collisions, consumes network resources.

## Ethernet LANs (2 of 2)

- As the network grows, the likelihood of collisions increases:
  - As collisions increase, overall LAN efficiency decreases



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-8

### Ethernet LANs: Part 2

Ethernet LANs were originally implemented for small, simple networks. Over time, LANs have become larger and more complex. As an Ethernet LAN grows, the likelihood of collisions on that LAN also grows. As more users join a shared Ethernet segment, each participating node receives an increase of traffic from all other participating nodes for which it is not the actual destination. This unwanted consumption of network resources, along with an increase of collisions, inevitably decreases the overall efficiency on the LAN.



## **Agenda:**

### **Ethernet Switching and Virtual LANs**

- Ethernet LANs

#### **→ Bridging**

- Configuring and Monitoring VLANs
- Configuring and Monitoring IRB
- Layer 2 Address Learning and Forwarding
- Layer 2 Firewall Filtering

© 2010 Juniper Networks, Inc. All rights reserved.

**JUNIPER**  
NETWORKS

Worldwide Education Services

www.junipernet | 3-7

## **Bridging**

The slide highlights the topic we discuss next.

## Overview of Bridging

### ■ Bridging:

- Defined in the IEEE 802.1D-2004 standard
- Segments a single collision domain
- Isolates the Physical Layer
- Learns and maintains a forwarding table (bridge table)
- Performs intelligent forwarding decisions based on the bridge table

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.junipernet.com

### Bridging Defined

Defined in the IEEE 802.1D-2004 standard, bridging addresses some of the inherent problems of large, shared Ethernet LANs. Bridging uses microsegmentation to divide a single-collision domain into multiple, smaller, bridged collision domains. Reducing the size of a collision domain effectively reduces the likelihood that collisions might occur. This approach also enhances performance by allowing multiple streams of data to flow through the switch within a common LAN or broadcast domain.

Bridging allows a mixed collection of interface types and speeds to be logically grouped within the same bridged LAN. The ability to logically group dissimilar interfaces in a bridged LAN environment provides design flexibility not found in a shared Ethernet LAN environment.

Bridging builds and maintains a forwarding table, known as a bridge table, for all destinations within the bridged LAN. The bridge table is based on the source MAC addresses for all devices participating in the bridged LAN. The bridge table can aid in intelligent forwarding decisions. This approach reduces unnecessary traffic on the LAN.

## Bridging: How Does it Work?

- Transparent bridging builds and maintains bridge tables using the following mechanisms:
  - Learning:
    - Learns MAC addresses and associated ports
  - Forwarding:
    - Forwards packets out the proper egress interface toward the destination
  - Flooding:
    - Replicates packets out *other* ports for unknown destination MAC addresses; also used when passing multicast and broadcast traffic
  - Filtering:
    - Limits traffic to its associated network segment
  - Aging:
    - Ensures bridge table entries are current

© 2014 Juniper Networks, Inc. All rights reserved.

Juniper Worldwide Education Services

www.juniper.net/1-3-9

### Bridging Mechanics

The transparent bridging protocol allows a switch to learn information about all nodes on the LAN. The switch uses this information to create the address-lookup tables, referred to as bridge tables, that it consults when forwarding traffic to (or toward) a destination on the LAN.

When a switch first connects to an Ethernet LAN or VLAN, it has no information about other nodes on the network. *Learning* is a process the switch uses to obtain the MAC addresses of all the nodes on the network. It stores these addresses in a bridge table. To learn MAC addresses, the switch reads all frames that it detects on the LAN or on the local VLAN, looking for MAC addresses of sending nodes. It places these addresses into its bridge table, along with two other pieces of information—the interface (or port) on which the traffic was received and the time it learned the address.

The switch uses the *forwarding* mechanism to deliver traffic, passing it from an incoming interface to an outgoing interface that leads to (or toward) the destination. To forward frames, the switch consults the bridge table to determine whether the table contains the MAC address corresponding to the destination of the frames. If the bridge table contains an entry for the desired destination address, the switch sends the traffic out the interface associated with the MAC address. The switch also consults the bridge table in the same way when transmitting frames that originate on devices connected directly to the switch.

*Continued on next page.*

## Bridging Mechanics (contd.)

*Flooding* is a transparent mechanism used to deliver packets to unknown MAC addresses. If the bridging table has no entry for a particular destination MAC address, or if the packet received is a broadcast or multicast packet, the switch floods the traffic out all interfaces except the interface on which it was received. (If traffic originates on the switch, the switch floods that traffic out all interfaces.) When the unknown destination host responds to traffic that has been flooded through a switch, the switch learns the MAC address of that node and updates its bridge table with the source MAC address of the host and ingress port.

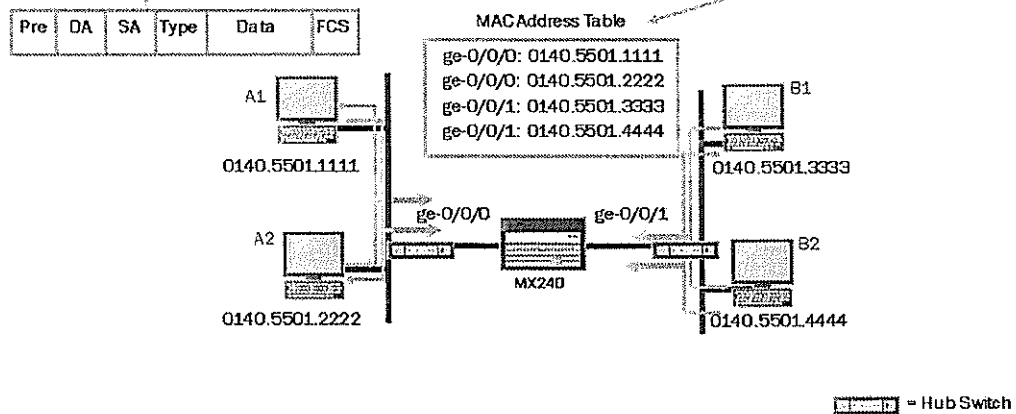
The *filtering* mechanism limits traffic to its associated network segment or VLAN. As the number of entries in the bridge table grows, the switch pieces together an increasingly complete picture of the individual network segments—the picture clarifies which nodes belong to which network. The switch uses this information to filter traffic. Filtering prevents the switch from forwarding traffic from one network segment to another.

Finally, the switch uses *aging* to ensure that only active MAC address entries are in the bridge table. For each MAC address in the bridge table, the switch records a timestamp of when it learned the information about the network node. Each time the switch detects traffic from a MAC address, it updates the timestamp. A timer on the switch periodically checks the timestamp; if the timestamp is older than the `global-mac-table-aging-time` value (discussed later in this chapter), the switch removes the node's MAC address from the bridge table.

## MAC Address Learning

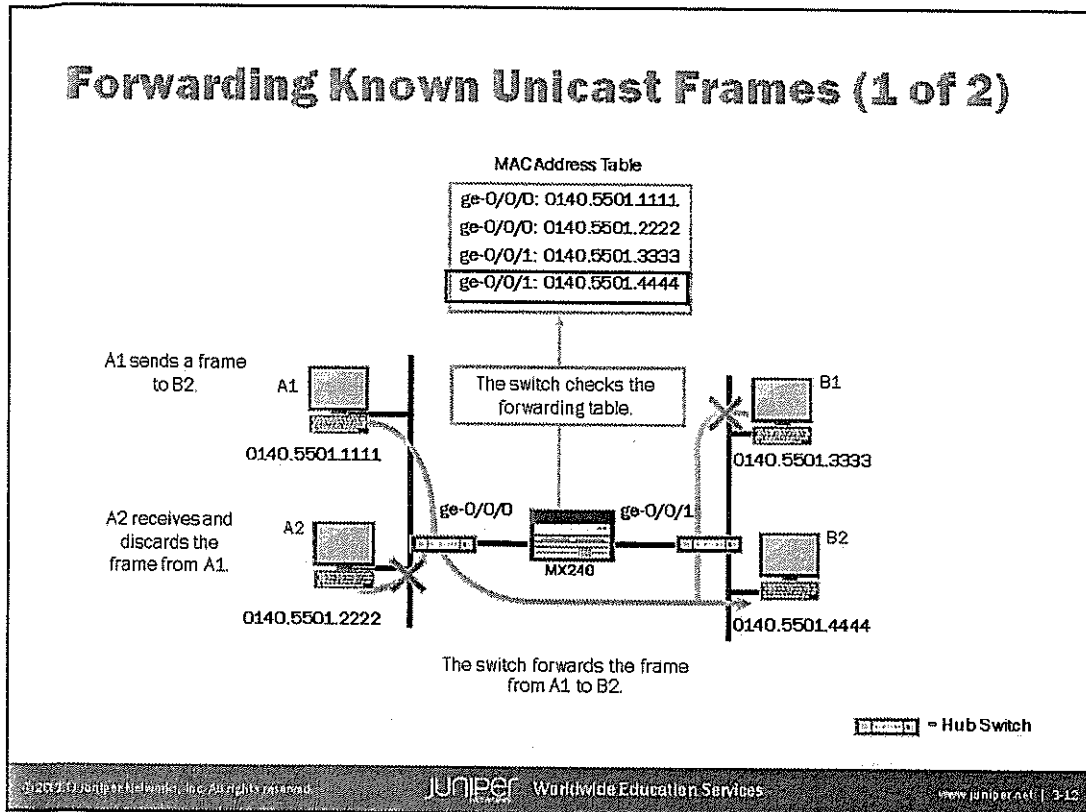
Source MAC addresses are learned for all incoming Ethernet frames

Each MAC address is associated with the frame's incoming interface



### MAC Address Learning

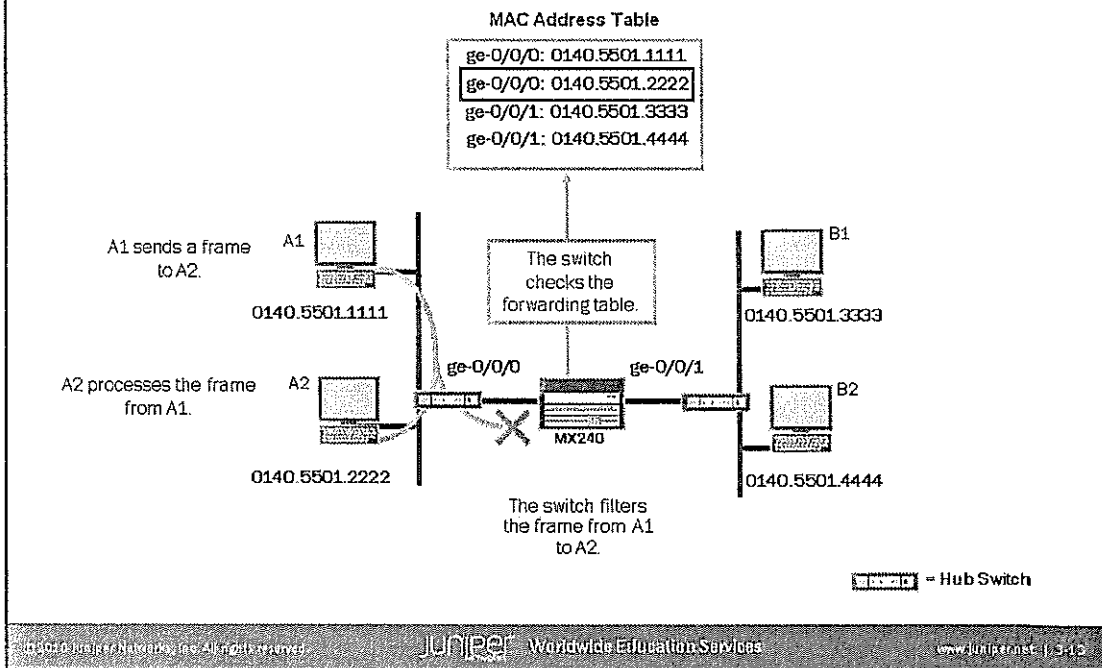
The slide illustrates a basic view of the MAC address learning process. In this example, each switch port connects to a hub and the individual hubs have multiple connected nodes. As each node sends traffic toward the other nodes on the bridged LAN, the switch reviews that traffic and creates a MAC address table (a bridge table) based on the source address of the sender along with the switch port on which it received the traffic. In this example, we see that the MAC addresses for A1 and A2 are associated with port ge-0/0/0, whereas the MAC addresses for B1 and B2 are associated with port ge-0/0/1.



### Forwarding Known Unicast Frames: Part 1

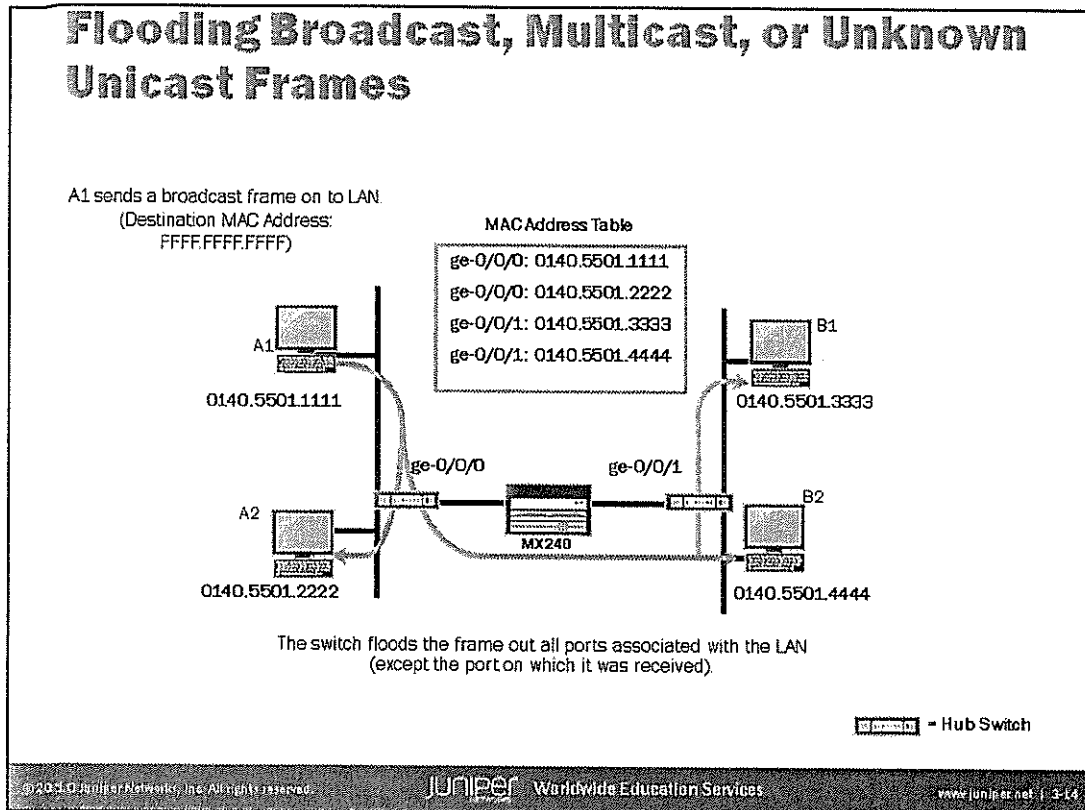
In the example on the slide, A1 sends a frame to B2. The frame is repeated out all ports on the attached hub, which results in frames traveling to both A2 as well as the switch shown in the middle of the illustration. A2 receives the frame and detects that the destination MAC address does not match its own MAC address, at which time A2 discards the frame. The switch receives the frame, checks the MAC address table for a matching entry, and forwards the frame out the associated port based on the lookup results. Ultimately, B2 receives and processes the frame while B1 receives and discards the frame.

## Forwarding Known Unicast Frames (2 of 2)



### Forwarding Known Unicast Frames: Part 2

In this example, A1 sends a frame to A2. The attached hub receives the frame and sends it out all ports, which results in duplicate frames sent to A2 as well as to the switch. A2 receives the frame and detects that the destination MAC address matches its own MAC address, at which time A2 processes the frame. The switch receives the frame and checks the MAC address table for a matching entry. The entry in the MAC address table shows the egress port, which, in this example, is the same port on which the switch received the frame. Because the egress port in the MAC address table is the same port on which the frame was received, the switch filters the frame.



### Flooding Frames

Flooding is used to learn a MAC address not recorded in the bridge table. This mechanism is also used when sending broadcast and, in many cases, multicast frames. The example on the slide shows A1 sending a broadcast frame with a destination MAC address of FFFF.FFFF.FFFF to the LAN. The attached hub sends the frame out all ports. The switch floods the broadcast frame out all ports associated with the LAN, except for the port on which it received the frame. The slide shows that, ultimately, all nodes on the LAN receive the frame.



## Viewing the MAC Address Table

- Use the `show bridge mac-table` command to view MAC address table entries

```
user@switch> show bridge mac-table
```

```
MAC flags (S -static MAC, D -dynamic MAC,  
SE -Statistics enabled, NM -Non configured MAC)
```

```
Routing instance : default-switch
```

```
Bridging domain : vlan_100, VLAN : 100
```

MAC address	MAC flags	Logical interface
00:21:59:ab:8a:95	D	ge-1/0/0.0
00:21:59:ab:8a:96	D	ge-1/0/1.0

Entries are organized based on associated VLAN.

© 2006 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-15

## Viewing the MAC Address Table

Use the `show bridge mac-table` command to view all entries within the MAC address table. This command generates a list of learned MAC addresses along with the corresponding VLANs and interfaces. All entries are organized based on their associated VLANs.

## Clearing the MAC Address Table

- Use the `clear bridge mac-table` command to clear the MAC address table contents

```
user@switch> clear bridge mac-table ?
Possible completions:
<[Enter]>      Execute this command
<address>      MAC address
bridge-domain  Name of bridging domain, or 'all'
instance       Display information for a specified instance
interface      Clear media access control table for specified interface
logical-system Name of logical system, or 'all'
vlan-id        Clear MAC address learned on a specified VLAN (0..4095)
|              Pipe through a command
```

Clear all entries in the table or entries based on a specific property.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-18

## Clearing MAC Address Table Entries

Use the `clear bridge mac-table` command to clear all entries within the MAC address table. Optionally, you can use the `interface` option to clear only those MAC table entries learned through the specified interface. The following example highlights the use of the `interface` option:

```
user@switch> show bridge mac-table
MAC flags (S -static MAC, D -dynamic MAC,
          SE -Statistics enabled, NM -Non configured MAC)
```

```
Routing instance : default-switch
Bridging domain : vlan_100, VLAN : 100
  MAC          MAC          Logical
  address      flags        interface
00:21:59:ab:8a:95 D          ge-1/0/0.0
00:21:59:ab:8a:99 D          ge-1/0/3.0
```

```
...
Routing instance : default-switch
Bridging domain : vlan_200, VLAN : 200
  MAC          MAC          Logical
  address      flags        interface
00:21:59:ab:8a:97 D          ge-1/0/2.0
00:21:59:ab:8a:99 D          ge-1/0/3.0
```

*Continued on next page.*

## Clearing MAC Address Table Entries (contd.)

```
user@switch> clear bridge mac-table interface ge-1/0/3.0
```

```
user@switch> show bridge mac-table
```

```
MAC flags (S -static MAC, D -dynamic MAC,
           SE -Statistics enabled, NM -Non configured MAC)
```

```
Routing instance : default-switch
Bridging domain : vlan_100, VLAN : 100
  MAC          MAC      Logical
  address      flags    interface
  00:21:59:ab:8a:95  D      ge-1/0/0.0
```

```
MAC flags (S -static MAC, D -dynamic MAC,
           SE -Statistics enabled, NM -Non configured MAC)
```

```
Routing instance : default-switch
Bridging domain : vlan_200, VLAN : 200
  MAC          MAC      Logical
  address      flags    interface
  00:21:59:ab:8a:97  D      ge-1/0/2.0
```

## **Agenda: Ethernet Switching and Virtual LANs**

- Ethernet LANs
- Bridging
- Configuring and Monitoring VLANs
- Configuring and Monitoring IRB
- Layer 2 Address Learning and Forwarding
- Layer 2 Firewall Filtering

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-18

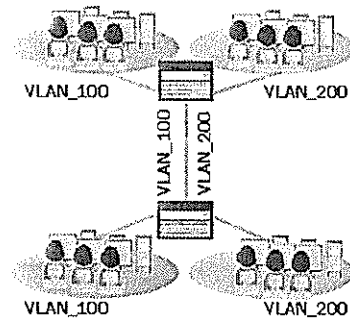
### **Configuring and Monitoring VLANs**

The slide highlights the topic we discuss next.

## Overview of VLANs

### ■ VLANs:

- Segment a single broadcast domain into multiple broadcast domains
- Allow for grouping users based on business needs, regardless of physical location



© 2019 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 3-19

### VLANs Defined

A VLAN is a collection of network nodes that are logically grouped together to form separate broadcast domains. A VLAN has the same general attributes as a physical LAN, but it allows all nodes for a particular VLAN to be grouped together, regardless of physical location. One advantage of using VLANs is design flexibility. VLANs allow grouping of individual users based on business needs. You can establish and maintain connectivity within a VLAN through software configuration, which makes VLANs such a dynamic and flexible option in today's networking environments.

## Switch Port Modes

### ■ Switch ports operate in either access or trunk mode

- Access mode:
  - Connects to network devices (desktops, IP phones, printers, and so forth)
  - Typically transmit untagged Ethernet frames for a single VLAN
- Trunk mode:
  - Connects to other switches or a router
  - Typically transmits tagged Ethernet frames for multiple VLANs—the exceptions are when the native VLAN option is configured or when control traffic is sent

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-20

## Switch Port Modes

Switch ports operate in either access mode or trunk mode.

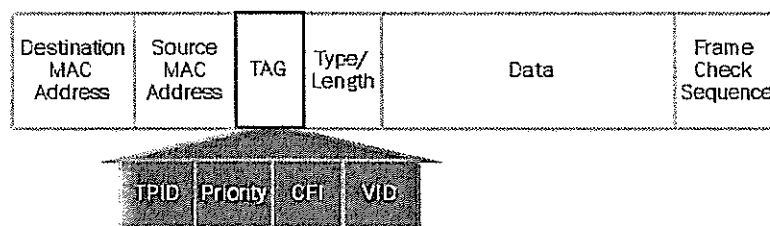
An access port connects to network devices such as desktop computers, IP phones, printers, or file servers. Access ports typically belong to a single VLAN and transmit and receive untagged Ethernet frames.

A trunk port typically connects to another switch or to a customer edge router. Interfaces configured for trunk mode handle traffic for multiple VLANs, multiplexing the traffic for all configured VLANs over the same physical connection, and separating the traffic by tagging it with the appropriate VLAN ID. Trunk ports can also carry untagged traffic when configured with the **native-vlan-id** statement. Furthermore, trunk ports send control traffic untagged.

## 802.1Q—Ethernet Frame

### ▪ 4-byte tag inserted into the Ethernet frame (max 1522 bytes)

- Tag Protocol Identifier: 16 bits, default 0x8100
- Priority: 3 bits, 802.1p
- Canonical Format Indicator: 1 bit, default 0
- Unique VLAN identifier: 12 bits



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-21

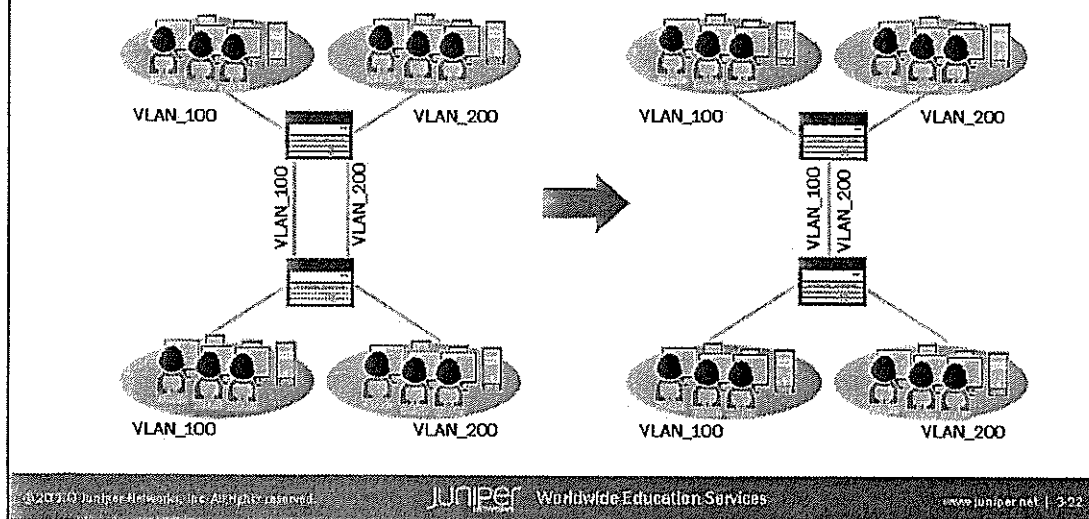
## 802.1Q—Ethernet Frame

To consistently associate traffic with a particular VLAN, the individual frames must be tagged as they pass throughout a network. The slide illustrates an 802.1Q-tagged Ethernet frame along with the key components of the tag:

- Tag Protocol Identifier (TPID);
- Priority;
- Canonical Format Indicator (CFI); and
- Unique VLAN identifier (VID).

## 802.1Q—Trunk Links

- A trunk is a single Ethernet link that can carry traffic for multiple VLANs



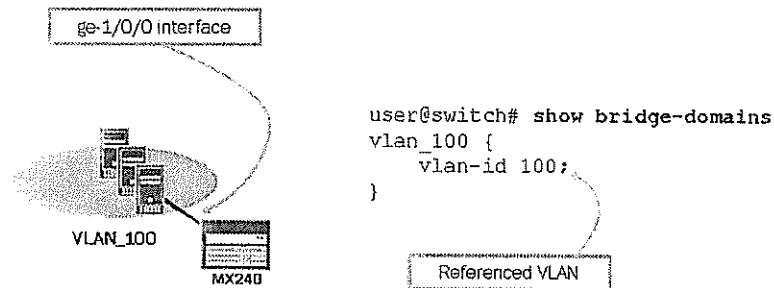
### 802.1Q Trunk Links

A trunk is a single Ethernet link used to carry traffic for multiple VLANs. A trunk link typically interconnects multiple switches or a switch with a customer edge router. As shown on the slide, interfaces configured as trunk ports handle traffic for multiple VLANs, multiplexing the traffic for all configured VLANs over a single physical connection rather than using separate physical links for each configured VLAN.



## Access Port Configuration (1 of 3)

- Define the bridge domains (broadcast domains) and VLANs to be used for switching



© 2009 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

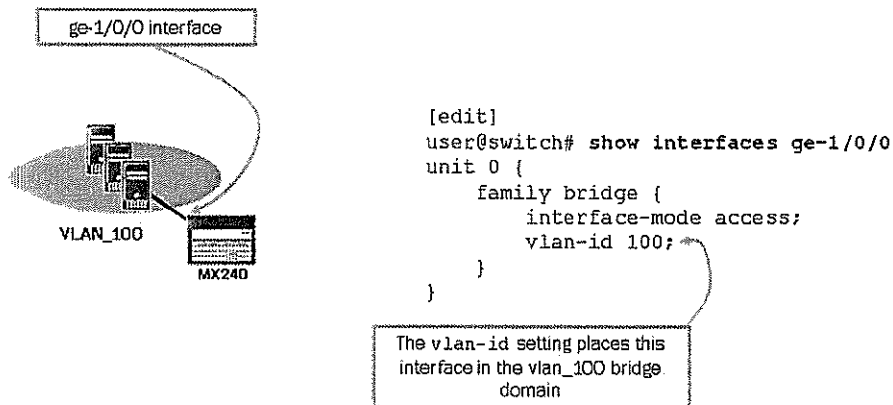
www.juniper.net | 3-23

### Define a Bridge Domain

To allow an MX Series 3D Ethernet Universal Edge Router to act as a switch and build a MAC address table, you must first specify the particular VLAN IDs that it will use for the purpose of switching. To do so, specify the appropriate VLAN ID as part of a named bridge domain. This method requires that you configure each VLAN as part of a single bridge domain. On a following slide, we cover how we can specify several VLANs within a single bridge domain using the `vlan-id-list` statement.

## Access Port Configuration (2 of 3)

- Assign the interface to the bridge domain and set the interface mode to **access**

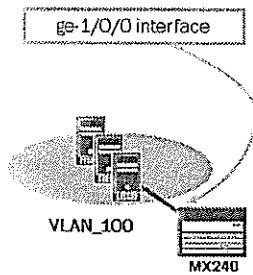


### Assign an Interface to a Bridge Domain

To allow an interface to act as an access port for a particular VLAN, you must specify its interface mode as **access** and you must specify the VLAN to which it belongs. For access ports, you must use **0** as the unit number.

## Access Port Configuration (3 of 3)

- The following method is another way to add an access port to a bridge domain:
  - Accomplishes the same objective as the process on the previous two slides



```
[edit]
user@switch# show

...
interfaces {
    ge-1/0/0 {
        encapsulation ethernet-bridge;
        unit 0;
    }
}
...
bridge-domains {
    vlan_100 {
        vlan-id 100;
        interface ge-1/0/0.0;
    }
}
...
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-25

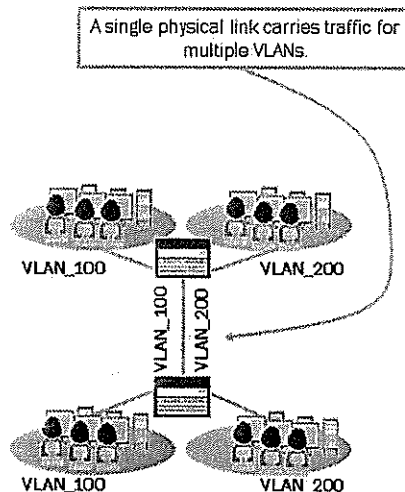
### Another Method

The slide shows a secondary (original style) method to accomplish adding the ge-1/0/0 interface as an access port to VLAN 100. Either method is acceptable, but we recommend that you choose one method to prevent confusion. However, you can use both methods simultaneously.

## 802.1Q Trunk Example (1 of 2)

```
[edit]
user@switch# show interfaces
...
ge-1/0/3 {
  native-vlan-id 100;
  vlan-tagging;
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list [ 100 200 ];
    }
  }
}

[edit]
user@switch# show bridge-domains
vlan_100 {
  vlan-id 100;
}
vlan_200 {
  vlan-id 200;
}
```



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-26

### 802.1Q Trunk Configuration Example: Part 1

The slide illustrates an 802.1Q trunk configuration example. In this case, the interface is configured as a trunk port and is associated with the `vlan_100` and `vlan_200` bridge domains. The partnering switch would have a similar configuration for the interface functioning as a trunk.

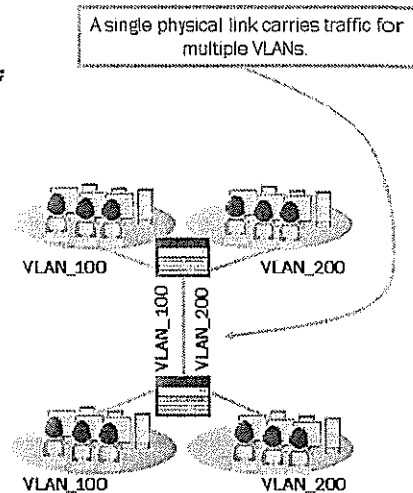
The slide also illustrates the usage of the `native-vlan-id` statement. This configuration statement does two things. First, if interface `ge-1/0/3` receives any untagged frames, it associates those frames to VLAN 100. Second, if interface `ge-1/0/3` transmits any outgoing frames that associate with VLAN 100, they transmit as untagged frames.

Notice the `vlan-id-list` statement. It specifies the VLANs to which the interface will be a member. The following statements are examples of how you can use the `vlan-id-list` statement:

- `vlan-id-list [100]`: VLAN 100 only;
- `vlan-id-list [100-200]`: All VLANs between 100 and 200, inclusive;
- `vlan-id-list [100-109 111-200]`: All VLANs between 100 and 200, except VLAN 110; or
- `vlan-id-list [100-109 111 113-200]`: All VLANs between 100 and 200, except VLAN 110 and 112.

## 802.1Q Trunk Example (2 of 2)

```
[edit]
user@switch# show interfaces
...
ge-1/0/3 {
  vlan-tagging;
  encapsulation flexible-ethernet-services;
  unit 100 {
    encapsulation vlan-bridge;
    vlan-id 100;
  }
  unit 200 {
    encapsulation vlan-bridge;
    vlan-id 200;
  }
}
[edit]
user@switch# show bridge-domains
vlan_100 {
  vlan-id 100;
  interface ge-1/0/3.100;
}
vlan_200 {
  vlan-id 200;
  interface ge-1/0/3.200;
}
```



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-27

### 802.1Q Trunk Configuration Example: Part 2

The slide shows a secondary (original style) method to accomplish adding the ge-1/0/3 interface as a trunk port using VLAN ID 100 and VLAN ID 200. Either method is acceptable, but we recommend that you choose one method to prevent confusion. However, you can use both methods simultaneously.

## Dealing with Large Numbers of VLANs

- Service providers typically deal with thousands of bridge domains and VLANs for each switch:
  - Use a single statement to create multiple bridge domains
    - Bridge domain names take the form prefix-vlan-number

```
[edit]
user@switch# show bridge-domains
customer {
    vlan-id-list [ 2-400 405 409-650 ];
}
user@switch> show bridge domain
```

Routing instance	Bridge domain	VLAN ID	Interfaces
...			
default-switch	customer-vlan-0099	99	
default-switch	customer-vlan-0100	100	ge-1/0/0.0 ge-1/0/1.0 ge-1/0/3.0
default-switch	customer-vlan-0101	101	
...			

© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-28

## Dealing with Many VLANs

As opposed to configuring individual bridge domains for each VLAN used for switching, the Junos operating system allows for the configuration of many VLANs within a single bridge domain. The slide shows that instead of using the `vlan-id` statement, you would use the `vlan-id-list` statement. The usage of this statement is similar to the usage described on the previous page. When using the `vlan-id-list` statement, the switch automatically configures the appropriate bridge domains, which have names that take the form prefix-vlan-number, where the prefix is the configured bridge domain name.

## Monitoring VLAN Assignments

```
user@switch> show bridge domain
```

Routing instance	Bridge domain	VLAN ID	Interfaces
default-switch	vlan_100	100	ge-1/0/0.0 ge-1/0/1.0 ge-1/0/3.0
default-switch	vlan_200	200	ge-1/0/2.0 ge-1/0/3.0 ge-1/1/4.0

```
user@switch> show bridge domain vlan_100 detail
```

```
Routing instance: default-switch
Bridge domain: vlan_100
Bridge VLAN ID: 100
Interfaces:
  ge-1/0/0.0
  ge-1/0/1.0
  ge-1/0/3.0
Total mac count: 2
State: Active
```

© 2006 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 329

## Monitoring VLAN Assignments

The slide shows some key commands used to monitor VLAN assignments. In this example, the ge-1/0/3 interface belongs to the bridge domain named `vlan_100`, which has an 802.1Q tag of 100. Because this interface is configured as an access port, it receives and transmits only untagged frames. If a trunk port were also configured to pass traffic for the `vlan_100` bridge domain, it would add and remove the 802.1Q tag value of 100 for all traffic for the `vlan_100` bridge domain. We look at a trunk port configuration and monitoring example next.

## Monitoring 802.1Q Trunks (1 of 2)

- Use the `show interfaces` command to determine the interface mode

```

user@switch> show interfaces ge-1/0/3
Physical interface: ge-1/0/3, Enabled, Physical link is Up
  Interface index: 135, SNMP ifIndex: 175
  Link-level type: Ethernet, MTU: 1514, Speed: 1000mbps, BPDU Error: None,
  MAC-REWRITE Error: None, Loopback: Disabled,
  Source filtering: Disabled, Flow control: Enabled, Auto-negotiation:
  Enabled, Remote fault: Online,
  Speed-negotiation: Disabled, Auto-MDIX: Enabled
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000
  ...

Logical interface ge-1/0/3.0 (Index 75) (SNMP ifIndex 226)
  Flags: SNMP-Traps 0x24020000 Encapsulation: Ethernet-Bridge
  Input packets : 335
  Output packets: 345
  Protocol bridge, MTU: 1514
  Flags: Trunk-Mode

```

The interface is in trunk mode.

### Monitoring 802.1Q Trunks: Part 1

The `show interfaces` command shows that the `ge-1/0/3` interface is configured for trunk mode, meaning it will transmit VLAN tags.



## Monitoring 802.1Q Trunks (2 of 2)

- Use the `show bridge domain` command to determine the interface VLAN assignments

```
user@switch> show bridge domain
```

Routing instance	Bridge domain	VLAN ID	Interfaces
default-switch	vlan_100	100	ge-1/0/0.0 ge-1/0/1.0 ge-1/0/3.0
default-switch	vlan_200	200	ge-1/0/2.0 ge-1/0/3.0 ge-1/1/4.0

The interface is in trunk mode using vlan-id 100 and 200.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-31

## Monitoring 802.1Q Trunks: Part 2

The `show bridge domain` command shows the interfaces and their VLAN assignments.

## Monitoring Bridge Statistics

```

user@switch> show bridge statistics
  Local interface: ge-1/0/0.111, Index: 81
    Broadcast packets:          31
    Broadcast bytes   :          1984
    Multicast packets:           0
    Multicast bytes   :           0
    Flooded packets   :          31
    Flooded bytes     :          2362
    Unicast packets   :         24093
    Unicast bytes     :        2553438
    Current MAC count:           1 (Limit 1024)
  Local interface: ge-1/0/0.112, Index: 80
    Broadcast packets:           0
    Broadcast bytes   :           0
    Multicast packets:           0
    Multicast bytes   :           0
    Flooded packets   :           0
    Flooded bytes     :           0
    Unicast packets   :           0
    Unicast bytes     :           0
    Current MAC count:           0 (Limit 1024)
  ...

```

Current number of MACs  
learned by the interface

### Monitor Bridge Statistics

The **show bridge statistics** command displays traffic statistics and MAC count information related to each logical interface of the switch.

## **Agenda:**

### **Ethernet Switching and Virtual LANs**

- Ethernet LANs
- Bridging
- Configuring and Monitoring VLANs
- Configuring and Monitoring IRB
- Layer 2 Address Learning and Forwarding
- Layer 2 Firewall Filtering

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

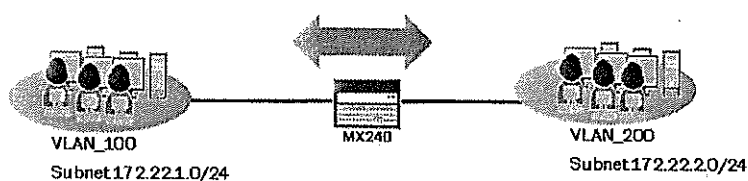
www.juniper.net | 3-33

### **Configuring and Monitoring IRB**

The slide highlights the topic we discuss next.

## IRB Interface

- IRB allows for both Layer 2 bridging and Layer 3 routing in a bridge domain
  - An IRB interface is an IP gateway for the hosts of a bridge domain



© 2011 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-34

### IRB Allows for Bridging and Routing

If you use a Layer 2-only Ethernet switch (no Layer 3 functionality), then you must add a separate router to your environment to provide routing between the subnets shown on the slide. However, an MX Series router can act as both a Layer 2 Ethernet switch and a router at the same time. An IRB interface is a logical Layer 3 interface used as an IP gateway for a VLAN. The following slides provide configuration and monitoring examples for an IRB interface.

## IRB Configuration Example

```
(edit)
user@switch# show interfaces
ge-1/0/0 {
    unit 0 {
        family bridge {
            interface-mode access;
            vlan-id 100;
        }
    }
}
...
irb {
    unit 0 {
        description "GW for VLAN_100";
        family inet {
            address 172.22.1.254/24;
        }
    }
    unit 1 {
        description "GW for VLAN_200";
        family inet {
            address 172.22.2.254/24;
        }
    }
}
...
}
```

```
(edit)
user@switch# show bridge-domains
vlan_100 {
    vlan-id 100;
    routing-interface irb.0;
}
vlan_200 {
    vlan-id 200;
    routing-interface irb.1;
}
```

This example facilitates routing between the vlan\_100 and vlan\_200 bridge domains.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-35

### IRB Configuration Example

The slide provides a configuration example for an IRB interface. In this example, the switch performs a Layer 3 lookup when it receives traffic with a destination MAC address that matches its own MAC address. For the switch to perform this routing operation, the attached devices must have configured gateway addresses that match the IP address associated with the corresponding IRB interface.

## Monitoring an IRB Interface

- Use the `show interfaces` command to verify the status of an IRB interface

```

user@switch> show interfaces terse irb*
Interface      Admin Link Proto Local      Remote
irb            up    up
irb.0          up    up    inet  172.22.1.254/24
irb.1          up    up    inet  172.22.2.254/24
  
```

At least one port must be active for IRB state to be up.

IRB state and IP address details

### Monitoring IRB

The slide lists a key command used to monitor an IRB interface, and shows the output from the `show interfaces terse` command. This command shows the state and IP address information for an IRB interface. As indicated on the slide, at least one active port must associate with the bridge domain for the IRB interface to be administratively up.

## Verifying Routing

- Use the `show route` command to verify the router's ability to route between the appropriate subnets

```
user@switch> show route
```

```
inet.0: 6 destinations, 6 routes (6 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both
```

```
10.210.12.0/27      *{Direct/0} 15:08:57
> via fxp0.0
10.210.12.1/32      *{Local/0} 15:08:57
Local via fxp0.0
172.22.1.0/24       *{Direct/0} 00:00:01
> via irb.0
172.22.1.254/32     *{Local/0} 00:21:31
Local via irb.0
172.22.2.0/24       *{Direct/0} 00:21:31
> via irb.1
172.22.2.254/32     *{Local/0} 00:21:31
Local via irb.1
```

Routes were added to the routing table as a result of configuring the IRB interfaces.

© 2010 Juniper Networks, Inc. All rights reserved.

Juniper Worldwide Education Services

www.juniper.net | 3-37

## Verifying Routing

As with any router, when you configure an IP address for an interface on that router, routes are automatically added to the routing table. In the Junos OS, for each configured IP interface, two routes are added to the routing table. One route is a host route (32-bit mask) that is used to forward traffic to the Routing Engine (RE) when locally destined packets arrive. The other route is a route to the network subnet to which that interface belongs. This route allows the router to route packets to other hosts on that same subnet. The slide shows that four routes were added to the inet.0 table as a result of configuring two IRB interfaces.

## **Agenda: Ethernet Switching and Virtual LANs**

- Ethernet LANs
- Bridging
- Configuring and Monitoring VLANs
- Configuring and Monitoring IRB
- Layer 2 Address Learning and Forwarding
- Layer 2 Firewall Filtering

©2010 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 3-38

### **Layer 2 Address Learning and Forwarding**

The slide highlights the topic we discuss next.



## MAC Address Learning and Forwarding

- A switch learns the source MAC addresses from incoming frames and learns destination MAC addresses as a result of the flooding processes:
  - By default, the Junos OS performs MAC learning, but it allows for you to change the default Layer 2 learning properties globally, per virtual switch, per bridge domain, and per interface level
    - Timeout interval for MAC entries (default is 300s)
    - MAC statistics (default is disabled)
    - Maximum number of MAC address learned (default is 393.215)
    - Turn off MAC learning

© 2020 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-35

## MAC Address Learning and Forwarding

As we discussed previously in this chapter, a switch learns MAC addresses from incoming frames as well as from the flooding process. The Junos OS allows you to override the default MAC learning behavior. The slide lists the things that you can change as well as where the changes can apply to a switch. The following list provides configurable values for each of the MAC learning properties:

- MAC timeout interval: 10s–1000000s (300s is the default);
- MAC statistics: Can be enabled (disabled by default);
- Global MAC limit: 20–1048575 (393215 is the default);
- Switch MAC limit: 16–1048575 (5120 is the default);
- Bridge domain MAC limit: 16–1048575 (5120 is the default); and
- Interface MAC limit: 1–131071 (1024 is the default).

To view MAC statistics once you enable the feature, issue the **show bridge mac-table extensive** command.

## Global and Switch Address Settings

### ■ Global level (all virtual switches and bridge domains)

```
[edit protocols l2-learning]
user@switch# set ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except  Don't inherit configuration data from these groups
> global-mac-limit      System level MAC limit options
  global-mac-statistics Enable MAC address statistics at system level
  global-mac-table-aging-time System level MAC table aging time (10..10000000
seconds)
  global-no-mac-learning Disable dynamic MAC address learning at system
level
```

### ■ Switch level (all bridge domains for the virtual switch)

```
[edit switch-options]
user@switch# set ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except  Don't inherit configuration data from these groups
> interface            Interface for configuring bridge-options
> interface-mac-limit  Maximum MAC address learned per interface
  mac-statistics        Enable MAC address statistics
> mac-table-size       Size of MAC address forwarding table
  no-mac-learning       Disable dynamic MAC address learning
```

### Global Level Settings

Global level settings apply to all virtual switches (discussed in a later chapter) and all bridge domains.

### Switch Level Settings

Switch level settings apply to all bridge domains associated with a virtual switch.

## Bridge Domain and Interface Address Settings

### ■ Bridge domain (all associated logical interfaces)

```
[edit bridge-domains bd]
user@switch# set bridge-options ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except   Don't inherit configuration data from these groups
> interface             Interface that connect this site to the VPN
> interface-mac-limit   Maximum MAC address learned per interface
  mac-statistics        Enable MAC address statistics
> mac-table-size        Size of MAC address forwarding table
  no-mac-learning        Disable dynamic MAC address learning
```

### ■ Interface level (single interface)

```
[edit bridge-domains bd]
user@switch# set bridge-options interface ge-1/0/0.0 ?
Possible completions:
<[Enter]>              Execute this command
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except   Don't inherit configuration data from these groups
> interface-mac-limit   Maximum number of MAC addresses learned on the
  interface
  no-mac-learning        Disable dynamic MAC address learning
> static-mac            Static MAC addresses assigned to this interface
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-41

## Bridge Domain Level Settings

Settings at this level affect all interfaces associated with the bridge domain.

## Interface Level

Settings at this level affect only the interface specified in the configuration.

## MAC Table Size Example

- Specify the `mac-table-size number` option to limit the number of learned MAC addresses for the bridge domain:
  - The default action when the limit is reached is to flood any frames with unknown MAC addresses
  - Optionally, you can specify `packet-action drop` to discard frames with unknown MAC addresses when the MAC table is full

```
[edit bridge-domains bd]
user@switch# show
vlan-id 100;
bridge-options {
  mac-table-size {
    4000;
    packet-action drop;
  }
}
```

## Layer 2 Learning Example

The example on the slide shows that the MAC table size limit for the bridge domain changed from the default of 5120 to 4000. By default, when the bridge domain MAC learning limit is reached, the device does not learn any more MAC addresses but still forwards or floods traffic in the case of unknown destinations. The slide shows that this default behavior was overridden so that Ethernet frames with unknown destinations will drop when the configured limit is reached.

## Monitor Layer 2 Information

- Use `show l2-learning` commands to view Layer 2 information

```
lab@switch> show l2-learning global-mac-count
47 dynamic and static MAC addresses learned globally
```

```
user@switch> show l2-learning global-information
Global Configuration:
```

```
MAC aging interval      : 300
MAC learning            : Enabled
MAC statistics          : Enabled
MAC limit Count         : 393215
MAC limit hit           : Disabled
MAC packet action drop: Disabled
```

```
user@switch> show l2-learning interface
Routing Instance Name : Default
Logical Interface flags (DL -disable learning, AD -packet action drop,
                        LH - MAC limit hit, DN - Interface Down )
Logical      BD      MAC      STP      Logical
Interface    Name    Limit   State   Interface flags
ge-1/0/4.0   bd      0      Forwarding
                        bd      1024
Routing Instance Name : Default
Logical Interface flags (DL -disable learning, AD -packet action drop,
                        LH - MAC limit hit, DN - Interface Down )
Logical      BD      MAC      STP      Logical
Interface    Name    Limit   State   Interface flags
ge-1/0/0.114 bd      0      Forwarding
                        bd      1024
```

© 2016 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-43

### Layer 2 Learning and Forwarding Status

The slide shows some of the commands that you can use to view the Layer 2 learning and forwarding status for the switch.

## **Agenda: Ethernet Switching and Virtual LANs**

- Ethernet LANs
- Bridging
- Configuring and Monitoring VLANs
- Configuring and Monitoring IRB
- Layer 2 Address Learning and Forwarding
- Layer 2 Firewall Filtering

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 3-44

### **Layer 2 Firewall Filtering**

The slide highlights the topic we discuss next.

## Layer 2 Firewall Filters

- Filter frames based on their contents and perform an action on frames that match the filter
- Filters can accept or discard packets based on:
  - Address fields
  - Protocol type
  - VLAN ID
  - 802.1p bits
  - IP address of the packet carried within an Ethernet frame
  - Many more factors

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 3-45

### Actions

You can use filters to control the frames destined to the RE as well as control frames passing through the router.

### Accept or Discard

You can define input filters that affect only inbound traffic and output filters that affect only outbound traffic. Filters can accept or discard frames based on the contents of the frame's address fields, protocol type, VLAN ID, or even the 802.1p bit field in the frame header.

## Layer 2 Firewall Filter Syntax

- Syntax is similar to that for policy statements

- Defined under the [edit firewall family bridge] hierarchy level

- Named filters, one or more terms

- Terms process sequentially
- All frames match a term when a *from* condition is not specified
- An implicit *discard all* exists for frames that do not match any term

- Actions: Accept or discard

- Modifiers: count, police, and so on

```
[edit firewall family bridge]
filter filter-name {
  term term-name {
    from {
      match-conditions;
    }
    then {
      action;
      action-modifiers;
    }
    term implicit-rule {
      then discard;
    }
  }
}
```

© 2003 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-16

### Syntax

A Junos OS filter consists of one or more named terms, similar to a policy statement. Each term has a set of match conditions preceded by the keyword *from*, and a set of actions or action modifiers preceded by the keyword *then*.

### Hierarchy Level

Layer 2 firewall filters are defined under the [edit firewall family bridge] section of the configuration hierarchy.

### One or More Terms

Firewall filter terms (at least one term is necessary) are processed sequentially. If no *from* condition is present, then all frames match. If no frames match any term, the default action is to discard the frame silently! Take care to ensure that wanted frames are not discarded. Use the command-line interface (CLI) **insert**, **copy**, and **rename** functions to assist in the management of your multiterm firewall filters.

*Continued on next page.*



## Actions and Modifiers

A filter can accept a frame for normal forwarding or discard a frame silently. You can modify these actions by applying a modifier. For example, you can apply the **count** modifier to increment a counter. We discuss other modifiers on the following slides.

## Applying Layer 2 Firewall Filters

- You can apply Layer 2 firewall filters to either an individual interface, a bridge domain, or to both
  - Interface level
    - You can apply a single filter for each interface (input, output, or both)
    - Apply a chain of filters using the **input-list** or **output-list** statements
  - Bridge domain level
    - You can apply a single filter for each bridge domain (input only)
  - Interface and bridge domain at the same time (input only)
    - The interface filter is processed first, followed by the bridge domain filter

```
[edit]
user@switch# show interfaces ge-1/0/0
unit 0 {
  family bridge {
    filter {
      input example;
    }
  }
  interface-mode access;
}
```

```
[edit]
user@switch# show bridge-domains
vlan_100
vlan-id 100;
routing-interface irb.0;
forwarding-options {
  filter {
    input example;
  }
}
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-48

## Applying Layer 2 Firewall Filters

Once you configure a firewall filter, you must apply it to one or more interfaces. You can accomplish this task in several different ways. The easiest way to apply a firewall filter to an individual Layer 2 interface is to specify the filter as an input or output filter at the [edit interface interface interface-name unit number family bridge filter] level of the configuration hierarchy. To apply a filter to all interfaces that belong to a particular bridge domain, you can apply a firewall filter at the [edit bridge-domain name forwarding-options filter] level of the configuration hierarchy. If firewall filters are applied as input filters to both the interface and bridge-domain levels, the Junos OS logically concatenates the bridge-domain-level filter to the end of the interface-level filter.

Note that you cannot use bridge-domain-level filters when the **vlan-id-list** statement was used to create the bridge domain.

## How Filters Are Evaluated

- **Single-term filters:**
  - If the frame matches all the conditions, the device takes the action in the `then` statement
  - If the frame does not match all the conditions, the device discards it
- **Multiple-term filters (terms evaluated sequentially):**
  - The Junos OS sequentially evaluates the frame against the conditions in each term's `from` statement, beginning with the first term
    - If the frame matches, the device takes the action in the `then` statement
    - If a frame passes through all the terms in the filter without matching any of them, the device discards it
- **Filter lists (chains) are concatenated logically into a single filter**

© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 3-19

### Single Terms

When a firewall filter consists of a single term, the filter is evaluated as follows: if the frame matches all the conditions, the device takes the action in the `then` statement; if the frame does not match all the conditions, the device discards it.

### Multiple Terms

When a firewall filter consists of more than one term, the filter is evaluated sequentially. First, the frame is evaluated against the conditions in the `from` statement in the first term. If the frame matches, the device takes the action in the `then` statement. If it does not match, it is evaluated against the conditions in the `from` statement in the second term. This process continues until either the frame matches the `from` condition in one of the subsequent terms or until no more terms remain.

If a frame passes through all the terms in the filter without matching any of them, the device discards it.

If a term does not contain a `from` statement, the frame is considered to match, and the device takes the action in the term's `then` statement.

If a term does not contain a `then` statement, or if you do not configure an action in the `then` statement (that is, the frame is just counted), and if the frame matches the conditions in the term's `from` statement, the device accepts the frame.

*Continued on next page.*

## Filter Lists

Instead of applying a single filter to an interface using filter input or filter output, you can apply a list of up to 16 filters. You perform this action with the **input-list** and **output-list** keywords.

## Firewall Filter Match Conditions (1 of 2)

```
[edit]
user@switch# set firewall family bridge filter example term 10 from ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except  Don't inherit configuration data from these groups
> destination-mac-address  Destination MAC address
+ destination-port      Match TCP/UDP destination port
+ destination-port-except Do not match TCP/UDP destination port
+ dscp                  Match Differentiated Services (DiffServ) code point
+ dscp-except           Do not match Differentiated Services (DiffServ) code point
+ ether-type            Match Ethernet type
+ ether-type-except     Do not match Ethernet type
+ forwarding-class      Match forwarding class
+ forwarding-class-except Do not match forwarding class
+ icmp-code             Match ICMP message code
+ icmp-code-except     Do not match ICMP message code
+ icmp-type             Match ICMP message type
+ icmp-type-except     Do not match ICMP message type
+ interface-group       Match interface group
+ interface-group-except Do not match interface group
> ip-address            Match IP source or destination address
> ip-destination-address Match IP destination address
+ ip-precedence          Match IP precedence value
+ ip-precedence-except  Do not match IP precedence value
+ ip-protocol            Match IP protocol type
+ ip-protocol-except    Do not match IP protocol type
> ip-source-address     Match IP source address
```

www.junipernet | 3-51

### Match Conditions: Part 1

The slide shows some of the many match conditions that you can use in a Layer 2 firewall filter.

## Firewall Filter Match Conditions (2 of 2)

+ isid	Match Internet Service ID
+ isid-dei	Match Internet Service ID DEI bit
+ isid-dei-except	Do not match Internet Service ID DEI bit
+ isid-except	Do not match Internet Service ID
+ isid-priority-code-point	Match Internet Service ID Priority Code Point
+ isid-priority-code-point-except	Do not match Internet Service ID Priority Code Point
+ learn-vlan-lp-priority	Match Learned 802.1p VLAN Priority
+ learn-vlan-lp-priority-except	Do not match Learned 802.1p VLAN Priority
+ learn-vlan-dei	Match User VLAN ID DEI bit
+ learn-vlan-dei-except	Do not match User VLAN ID DEI bit
+ learn-vlan-id	Match Learnt VLAN ID
+ learn-vlan-id-except	Do not match Learnt VLAN ID
+ loss-priority	Match Loss Priority
+ loss-priority-except	Do not match Loss Priority
+ port	Match TCP/UDP source or destination port
+ port-except	Do not match TCP/UDP source or destination port
+ source-mac-address	Source MAC address
+ source-port	Match TCP/UDP source port
+ source-port-except	Do not match TCP/UDP source port
+ tcp-flags	Match TCP flags
+ traffic-type	Match Match traffic type
+ traffic-type-except	Do not match Match traffic type
+ user-vlan-lp-priority	Match User 802.1p VLAN Priority
+ user-vlan-lp-priority-except	Do not match User 802.1p VLAN Priority
+ user-vlan-id	Match User VLAN ID
+ user-vlan-id-except	Do not match User VLAN ID
+ vlan-ether-type	Match VLAN Ethernet type
+ vlan-ether-type-except	Do not match VLAN Ethernet type

www.juniper.net | 3-52

### Match Conditions: Part 2

The slide shows some of the many match conditions that you can use in a Layer 2 firewall filter.

## Firewall Filter Match Actions

- Match actions determine what happens to frames once a match occurs:
  - Besides accept and discard for actions, you can apply modifiers to frames like setting the forwarding class, and so on

```

user@switch# set firewall family bridge filter example term 10 then ?
Possible completions:
  accept                Accept the packet
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except  Don't inherit configuration data from these groups
  count                Count the packet in the named counter
  discard              Discard the packet
  forwarding-class      Classify packet to forwarding class
  loss-priority         Packet's loss priority
  next                 Continue to next term in a filter
  next-hop-group        Use specified next-hop group
  policer              Name of policer to use to rate-limit traffic
  port-mirror           Port-mirror the packet
  port-mirror-instance  Port-mirror the packet to specified instance
> three-color-policer  Police the packet using a three-color-policer

```

### Match Actions

accept and discard are the actions that you can apply to a frame. However, you can apply modifiers to the frames as well:

- **count:** This modifier counts the number of matches that occur to a named counter. See the current totals by issuing the **show firewall** command.
- **forwarding-class:** This modifier is used for multifold classification for class of service (CoS). Essentially, this setting specifies the queue in which this frame should be placed.
- **loss-priority:** This modifier allows you to change the packet loss priority bit of the IP packet in the payload of the Ethernet frame.
- **next:** This modifier allows the frame to be evaluated by the next term in the filter.
- **next-hop-group:** This modifier specifies which next-hop group will be applied.
- **policer:** This modifier applies a rate-limiting policer to the matching frames.
- **port-mirror:** This modifier allows copies of the frame to be sent to an outbound interface for analysis. The original frame forwards as normal.

## Firewall—Family Bridge Example

- Configure, apply, and view the effects of a firewall filter

```
[edit]
user@switch# show firewall
family bridge {
  filter example {
    term 10 {
      from {
        learn-vlan-id 100;
      }
      then {
        count learn-vlan-100;
        accept;
      }
    }
    term 1000 {
      then {
        count all-others;
        accept;
      }
    }
  }
}

[edit]
user@switch# show bridge-domains
vlan_100 {
  vlan-id 100;
  routing-interface irb.0;
  forwarding-options {
    filter {
      input example;
    }
  }
}
```

```
user@switch> show firewall
```

```
Filter: __default_bpdu_filter__
```

```
Filter: example
```

```
Counters:
```

Name	Bytes	Packets
all-others	0	0
learn-vlan-100	56300	533

### Example Filter

The slide shows an example of configuring, applying, and viewing the effects of a firewall filter. To clear the counters, use the **clear firewall** command.



## Summary

### ■ In this chapter, we:

- Described the functions of Ethernet LANs
- Described learning and forwarding in a bridging environment
- Discussed implementing VLAN tagging
- Discussed implementing IRB
- Discussed implementing Layer 2 address learning and forwarding
- Discussed implementing Layer 2 firewall filters

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-55

### This Chapter Discussed:

- The functions of an Ethernet LAN;
- Learning and forwarding in a bridging environment;
- Implementation of VLAN tagging;
- Implementation of IRB;
- Implementation of Layer 2 address learning and forwarding; and
- Implementation of Layer 2 firewall filters.

## Review Questions

1. What is the purpose of a bridge domain on an MX Series router?
2. How does a bridge handle multicast Ethernet frames?
3. What is the purpose of an IRB interface?
4. Which match condition is used in a Layer 2 firewall filter to match on 802.1p priority bits?

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-66

## Review Questions

- 1.
- 2.
- 3.
- 4.

## Lab 1: Ethernet Switching and VLANs

- Configure a bridge domain.
- Configure a Layer 2 interface.
- Use operational mode commands to verify the status of a Layer 2 interface.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 3-67

### Lab 1: Ethernet Switching and VLANs

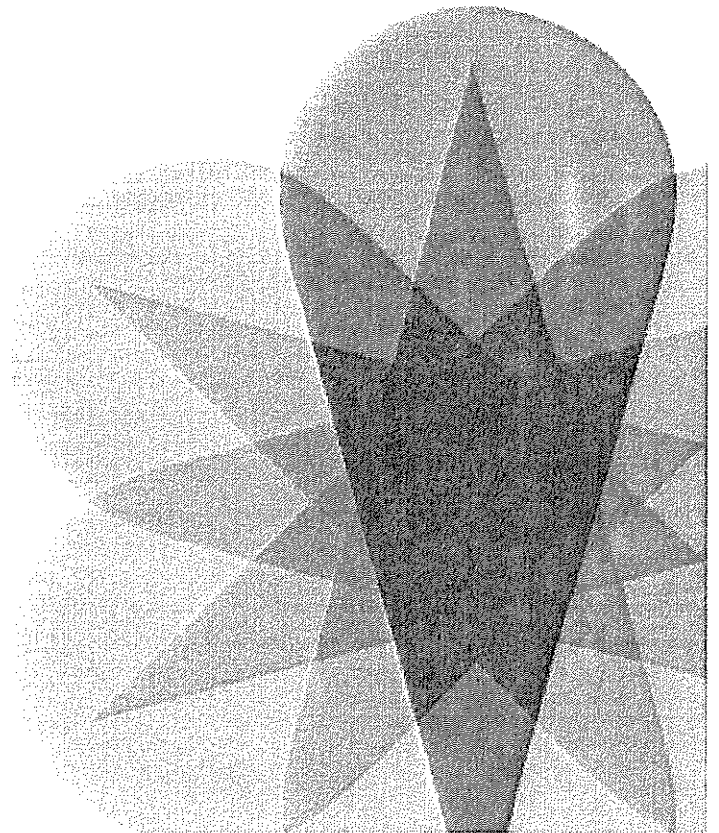
The slide provides the objectives for this lab.





# **Junos Service Provider Switching**

## **Chapter 4: Virtual Switches**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Describe the use of a routing instance
  - Describe the function of a virtual router
  - Describe the function of a virtual switch
  - Implement a virtual switch
  - Interconnect local routing instances

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-2

### This Chapter Discusses:

- The use of a routing instance;
- The function of a virtual router;
- The function of a virtual switch;
- Implementation of a virtual switch, and
- Interconnection of local routing instances.

## Agenda: Virtual Switches

- Routing Instances Overview
- Configuring and Monitoring Virtual Switches
- Interconnecting Routing Instances

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

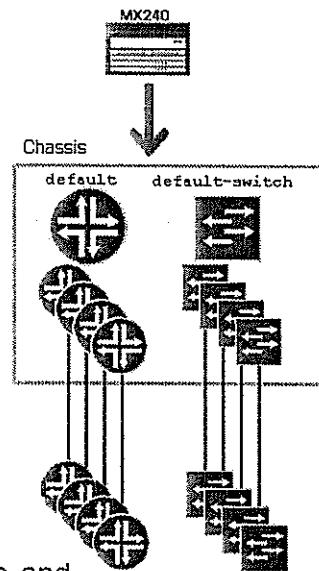
www.juniper.net/j-3

### Routing Instances Overview

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## Routing Instances

- Several different types of routing instances exist:
  - Virtual-router routing instances allow for your single chassis to appear as multiple routers to the outside world
    - Each with their own separate routing tables, protocols, link-state databases, and so on
    - Default instance is named `default`
  - Virtual-switch routing instances allow for your single chassis to appear as multiple switches to the outside world
    - Each has its own MAC tables, VLAN ID space, and spanning-tree domains
    - The default instance is named `default-switch`



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-4

### Routing Instance Types

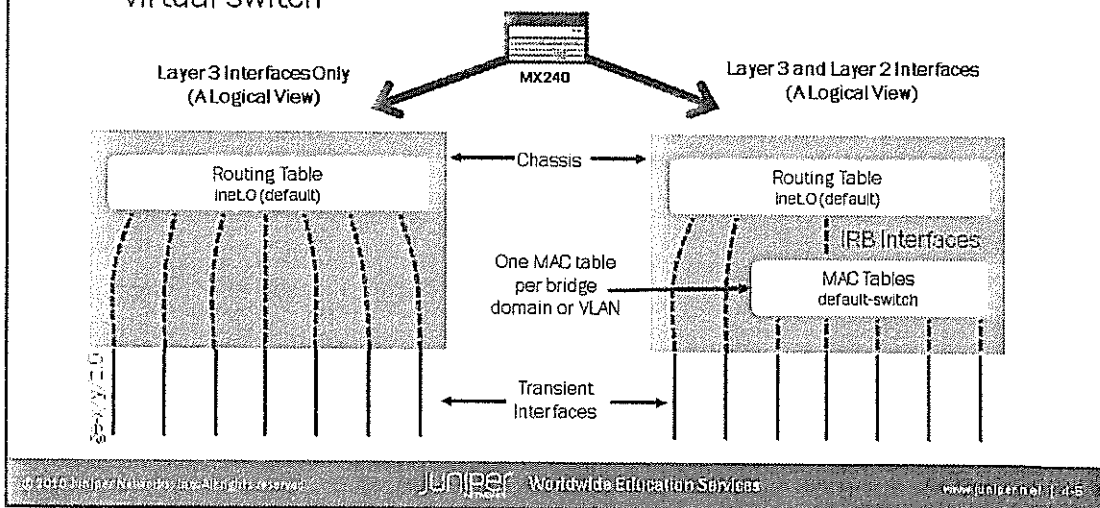
The Junos operating system provides several different routing-instance types with which to work. In this course we work with two types of routing-instances: `virtual-router` and `virtual-switch`. Essentially, these two routing-instance types allow your single chassis to appear as either more than one router or more than one switch, respectively. Each virtual router acts as a standalone router. For example, each virtual router has its own routing table, routing protocols, interfaces, and just about everything that encompasses the typical things that comprise a router. Similarly, each configured virtual switch has its own MAC tables, virtual LAN (VLAN) ID space, bridge domains, spanning-tree domains, and so forth. A Juniper Networks MX Series 3D Universal Edge Router uses two default routing instances. For routing, it uses the `default` virtual router (`inet.0` is its routing table). For switching, it uses the `default-switch` virtual switch.



## A Simplistic View of Routing Instances

### Interface to table default assignment:

- Layer 3 interfaces associate with `inet.0`
- Layer 2 interfaces associate with the default-switch virtual switch



### Routing Instance and Interface Default Relationship

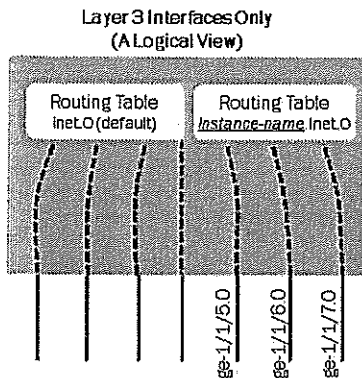
The slide shows a very simplistic view of the default relationship of interfaces to the routing and MAC tables of an MX Series router. Keep in mind that we have left out discussion of the Packet Forwarding Engine (PFE) and the associated forwarding tables. When troubleshooting virtual routers and switches, you generally can spend your time focused on the Routing Engine's (RE's) copy of the routing and MAC tables, while trusting that equivalent copies appear as forwarding tables in the PFEs of your switch. To view the PFE forwarding tables, both for routing and switching, use the **show route forwarding-table** command.

In a routing-only environment, configured interfaces and their associated local and direct routes appear in the default virtual router's routing table, `inet.0`. In a mixed Layer 2 and Layer 3 environment, Layer 3 interfaces continue to work as described, whereas Layer 2 interfaces, having been associated with a bridge domain at the `[edit bridge-domains]` hierarchy, associate with the default virtual switch's MAC tables. Because IRB interfaces are Layer 3 interfaces, their associated local and direct routes appear in `inet.0` as well.

## Routing Instances—Virtual Router

### ■ You must assign interfaces to a virtual router

- Place routes associated with those interfaces only in that virtual router's routing table



```
[edit]
user@switch# show routing-instances
vr1 (
  instance-type virtual-router;
  interface ge-1/1/5.0;
  interface ge-1/1/6.0;
  interface ge-1/1/7.0;
  protocols {
    ospf {
      area 0.0.0.0 {
        interface ge-1/1/5.0;
        interface ge-1/1/6.0;
        interface ge-1/1/7.0;
      }
    }
  }
}
```

```
user@switch> show route table vr1.inet.0
```

```
vr1.inet.0: 2 destinations, 2 routes (2 active, ...
+ = Active Route, - = Last Active, * = Both
```

```
172.22.2.0/24      *[Direct/0] 20:59:24
                   > via ge-1/1/5.0
172.22.2.3/32     *[Local/0] 20:59:24
                   Local via ge-1/1/5.0
-
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

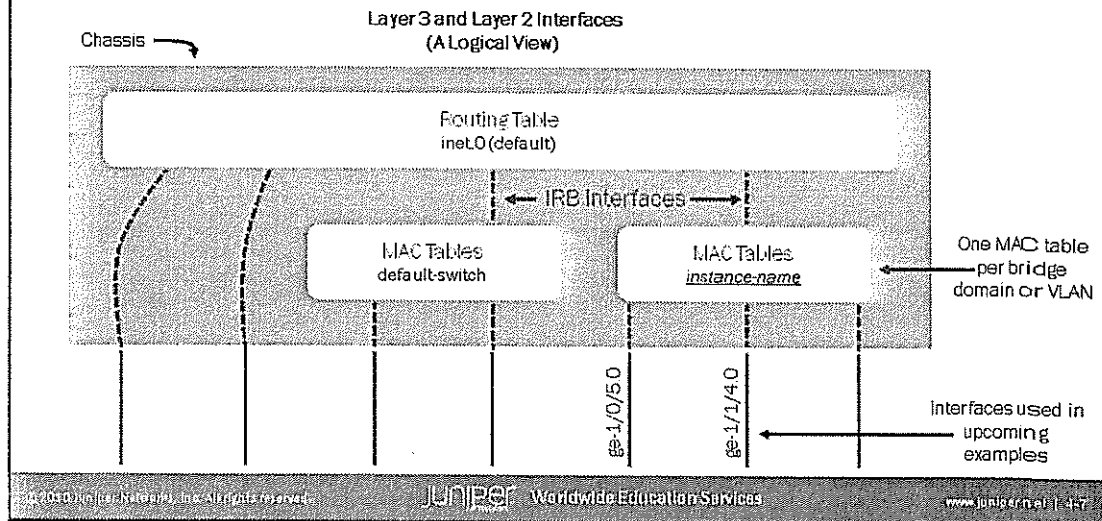
www.junipernet | 4-6

### Assign Interfaces to a Virtual Router

By default, once you configure an interface with properties at the [edit interfaces interface-name unit number family inet] level of the hierarchy, that interface's local and direct routes are placed in the inet.0 routing table. To override that behavior, you simply list the interface at the [edit routing-instances instance-name] level of the hierarchy. The local and direct routes now appear in the instance-name.inet.0 routing table (the virtual router's routing table.)

## Routing Instances—Virtual Switch

- Each virtual-switch routing instance operates independently of the other virtual switches:
  - Routes associated with IRB interfaces are placed in inet.0 regardless of the virtual switch to which they belong



### Virtual Switch

The slide shows the routing and MAC table relationships when using virtual switches. Each virtual switch, including the default switch, has interfaces assigned for bridging. Also, you can configure integrated routing and bridging (IRB) interfaces for each virtual switch. The local and direct routes for all IRB interfaces in all virtual switches are placed in inet.0, by default. However, you can also place them in a virtual router's routing table by listing the IRB interfaces at the [edit routing-instances *instance-name*] level of the hierarchy. The following slides cover the process of configuring a virtual switch.

## Agenda: Virtual Switches

- Routing Instances Overview
- Configuring and Monitoring Virtual Switches
- Interconnecting Routing Instances

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-8

### Configuring and Monitoring Virtual Switches

The slide highlights the topic we discuss next.

## Configure a Routing Instance

- Configure a base virtual-switch routing instance:
  - Define the bridge domains and VLAN IDs that the switch will use

```
[edit]
user@switch# show routing-instances
virtual-sw-1 {
    instance-type virtual-switch;
    bridge-domains {
        vlan_100 {
            vlan-id 100;
        }
        vlan_200 {
            vlan-id 200;
        }
    }
}
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-9

### Virtual-Switch Routing Instance

The configuration on the slide creates a `virtual-sw-1` routing instance and allows for VLAN IDs 100 and 200 to be used for the purpose of Layer 2 switching. MAC tables for these new bridge domains will not be used for learning and forwarding until you assign at least one interface to the virtual switch.

## Configuring an Access Port

- Configure an interface that acts as an access port for the virtual switch:
  - Specify the correct `vlan-id` so that this interface is associated with the correct bridge domain

```
[edit]
user@switch# show interfaces ge-1/0/5
unit 0 {
    family bridge {
        interface-mode access;
        vlan-id 200;
    }
}
```

### WARNING!!!

Committing at this point can cause this interface to be added to the default-switch routing instance, which could introduce a loop into your topology.

## Virtual-Switch Access Port

You configure the interface properties for an access port using the exact same process as when defining it for the default switch. In fact, if you were to commit the configuration, the `ge-1/0/5.0` interface would be placed in the default switch. Be careful not to commit the configuration as it stands, because you might introduce a loop into your switched network. One of the following slides shows how to place the interface in the virtual switch. We highly recommend that you perform that step before committing the configuration.

## Configuring a Trunk Port

- Configure an interface that acts as a trunk port for the virtual switch:

- Specify the correct `vlan-id-list` so that this interface is associated with the correct bridge domains

```
[edit]
user@switch# show interfaces ge-1/1/4
unit 0 {
    family bridge {
        interface-mode trunk;
        vlan-id-list [ 100 200 ];
    }
}
```

### WARNING!!!

Remember to bind the trunk port to the virtual switch to prevent the introduction of a loop into your topology.

## Configure a Trunk Port

You configure the interface properties for a trunk port using the exact same process as when defining it for the default switch. In fact, if you were to commit the configuration on the slide, the `ge-1/1/4.0` interface would be placed in the default switch. Be careful not to commit the configuration as it stands, because you might introduce a loop into your switched network. One of the following slides shows how to place the interface in the virtual switch. We highly recommended that you perform that step before committing the configuration.

## Configuring a Virtual-Switch IRB Interface

- Configure an IRB interface that acts as the IP gateway for a bridge domain within the virtual switch

```
[edit]
user@switch# show interfaces
-
irb {
    unit 1 {
        description "GW for VLAN_200";
        family inet {
            address 172.22.2.254/24;
        }
    }
}
-
```

### Configure an IRB Interface

You configure the interface properties for an IRB interface using the exact same process as when defining it for the default switch. In fact, if you were to commit the configuration on the slide, the irb.1 interface would be placed in the default switch. Be careful not to commit the configuration as it stands, because you might introduce a loop into your switched network. The following slide shows how to place the interface in the virtual switch. We recommended that you perform that step before committing the configuration.



## Bind the Interfaces to the Virtual Switch

### ■ Specify the interfaces that belong to the virtual switch:

- List the trunk and access ports as member interfaces of the virtual switch
- List the IRB as the routing interface for the appropriate bridge domain within the virtual switch

```
[edit]
user@switch# show routing-instances
virtual-sw-1 {
    instance-type virtual-switch;
    interface ge-1/0/5.0;
    interface ge-1/1/4.0;
    bridge-domains {
        vlan_100 {
            vlan-id 100;
        }
        vlan_200 {
            vlan-id 200;
            routing-interface irb.1;
        }
    }
}
```

### Add the Interfaces to the Virtual Switch

After configuring the access and trunk ports as shown on the previous slides, you simply need to list the interface at the [edit routing-instances instance-name] level of the hierarchy. The irb.1 interface should be listed as the routing-interface for the appropriate bridge domain.

## Verify That Ports Belong to the Virtual Switch

- Use the `show bridge domain` command to ensure that the configuration setting accomplished your goal

```
user@switch> show bridge domain
```

Routing instance	Bridge domain	VLAN ID	Interfaces
default-switch	vlan_100	100	ge-1/0/0.0 ge-1/0/4.0
default-switch	vlan_200	200	ge-1/0/2.0 ge-1/0/4.0
virtual-sw-1	vlan_100	100	ge-1/1/4.0
virtual-sw-1	vlan_200	200	ge-1/0/5.0 ge-1/1/4.0

ge-1/1/4.0 and ge-1/0/5.0 have been added to the correct routing instance, bridge domains, and VLANs.

© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-14

### Verify Settings

Looking at the output on the slide, you can see that the `ge-1/1/4.0` interface is now bound to the `virtual-sw-1` routing instance and the bridge domains `vlan_100` and `vlan_200`. Also, `ge-1/0/5.0` is bound to the appropriate routing instance and bridge domain.

## Verify Routing

- Ensure that the appropriate routes appear in the inet.0 routing table

```

user@switch> show route

inet.0: 6 destinations, 6 routes (6 active, 0 holddown, 0
hidden)
+ = Active Route, - = Last Active, * = Both

10.210.12.0/27      *{Direct/0} 1d 22:44:16
                   > via fxp0.0
10.210.12.1/32     *{Local/0} 1d 22:44:16
                   Local via fxp0.0
172.22.2.0/24      *{Direct/0} 00:05:54
                   > via irb.1
172.22.2.254/32    *{Local/0} 00:13:39
                   Local via irb.1

```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-15

### IRB Routes

The local and direct routes that associate with the IRB interface should be in the inet.0 table. Use the **show route** command to verify that the routes were added properly.

## Agenda: Virtual Switches

- Routing Instances Overview
- Configuring and Monitoring Virtual Switches
- Interconnecting Routing Instances

© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER  
NETWORKS

Worldwide Education Services

www.juniper.net | 4-16

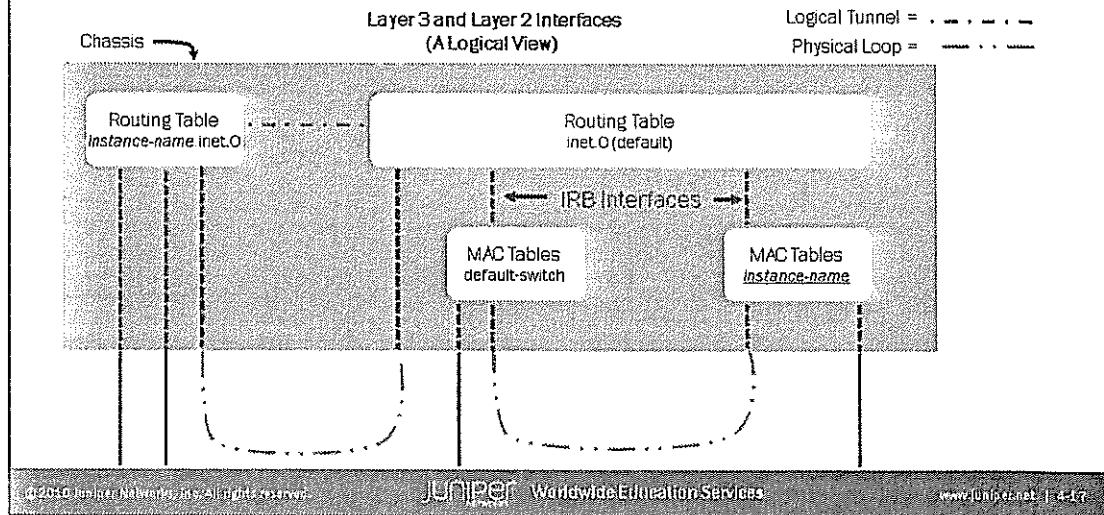
### Interconnecting Routing Instances

The slide highlights the topic we discuss next.

## Interconnection Methods

### Supported methods:

- Internal, logical tunnel interface (Layer 3 only)
- Two external, physically looped interfaces (1 cable)



### Supported Methods of Interconnecting Routing Instances

As mentioned previously, to the outside world virtual routers and virtual switches appear as individual routers and switches. At some point you might want to interconnect the virtual routers and virtual switches that are local to a single chassis. For virtual routers, you can accomplish this task using either a logical tunnel interface or by looping two interfaces together with a single cable. For virtual switches, this process works only using the external cable method. The reason why spanning tree protocols do not function properly between virtual switches is because all virtual switches use the same MAC address as part of their bridge ID in the bridge protocol data units (BPDUs). Unfortunately, you cannot change a virtual switch's MAC address.

## Enabling Tunnel Support

- You must enable tunneling on the PFE of MX Series routers

```

user@switch> show interfaces terse
Interface           Admin Link Proto Local Remote
...
ge-1/0/8             up    down
ge-1/0/9             up    down
ge-1/1/0             up    down

user@switch> show interfaces terse
Interface           Admin Link Proto Local Remote
...
ge-1/0/8             up    down
ge-1/0/9             up    down
gr-1/0/10            up    up
ip-1/0/10            up    up
lt-1/0/10            up    up
mt-1/0/10            up    up
pd-1/0/10            up    up
pe-1/0/10            up    up
vt-1/0/10            up    up
ge-1/1/0             up    down

```

(edit)

```

user@switch# show chassis
fpc 1 {
  pic 0 {
    tunnel-services {
      bandwidth lg;
    }
  }
}

```

© 2010 Juniper Networks, Inc. All rights reserved.

Juniper Worldwide Education Services

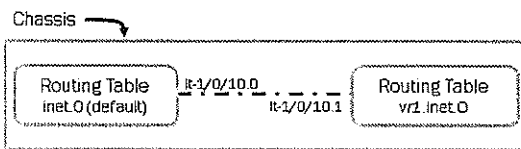
www.juniper.net | 4-18

### Tunnel Services

Anytime you need to use layer tunneling, you must enable tunnel services on the MX Series router. For example, you must enable tunnel services for a generic routing encapsulation (GRE) tunnel, an IP over IP (IP-IP) tunnel, Physical Interface Module (PIM) encapsulation or decapsulation of register messages, and for our case, using logical tunnel interfaces. Each Dense Port Concentrator (DPC) on a switch has either 40 Gigabit Ethernet ports (10 ports per PFE) or 4 10-Gigabit Ethernet ports (1 port per PFE.) Each PFE on an MX Series DPC can provide tunneling services but you must enable it. The slide shows how to enable tunnel services on the first PFE (serving ge-1/0/0 through ge-1/0/9) on the 40 1-Gigabit Ethernet DPC in slot number 1. Once you enable this feature, you will notice that you have several tunnel type interfaces that become available for your use. Notice that the tunnel interfaces use the logical PIC port number of 10 (normally PIC port numbers stop at 9.) When enabling tunnel services on a PFE of a 4-port 10 Gigabit Ethernet DPC, the Ethernet interface for that PFE is removed from service and is no longer visible in the command-line interface (CLI).

## Layer 3 Logical Tunnel Interface Configuration

- Configure and assign the logical tunnel interfaces to the appropriate virtual routers



```
[edit]
user@switch# show routing-instances vr1
instance-type virtual-router;
interface lt-1/0/10.1;
interface ge-1/1/5.0;
```

```
[edit]
user@switch# show interfaces lt-1/0/10
unit 0 {
    encapsulation vlan;
    vlan-id 100;
    peer-unit 1;
    family inet {
        address 172.22.10.1/30;
    }
}
unit 1 {
    encapsulation vlan;
    vlan-id 100;
    peer-unit 0;
    family inet {
        address 172.22.10.2/30;
    }
}
```

```
user@switch> ping routing-instance vr1 172.22.10.1
PING 172.22.10.1 (172.22.10.1): 56 data bytes
64 bytes from 172.22.10.1: icmp_seq=0 ttl=64 time=0.838 ms
64 bytes from 172.22.10.1: icmp_seq=1 ttl=64 time=4.913 ms
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

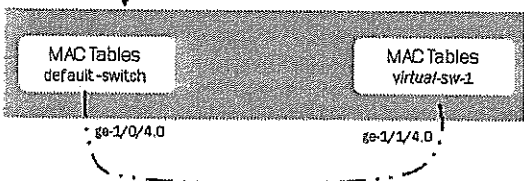
www.juniper.net | 4-19

### Configure and Assign Logical Tunnel Interfaces

You configure the logical tunnel interfaces similar to how you would for any other Layer 3 interface. You configure each logical tunnel Layer 3 interface as a logical unit. To map one logical unit to another, use the **peer-unit** statement. By default, logical tunnel interfaces are placed in the default virtual router. To place a logical tunnel interface in a virtual router, specify the logical tunnel interface at the [edit routing instance instance-name] level of the hierarchy.

## Layer 2 Physical Loop

- Configure and assign the physical interfaces to the appropriate virtual switch



```
[edit]
user@switch# show interfaces
ge-1/0/4 {
    unit 0 {
        family bridge {
            interface-mode trunk;
            vlan-id-list [ 100 200
        ]
    }
}
ge-1/1/4 {
    unit 0 {
        family bridge {
            interface-mode trunk;
            vlan-id-list [ 100 200
        ]
    }
}
...
```

```
[edit]
user@switch# show routing-instances
virtual-sw-1 {
    instance-type virtual-switch;
    ...
    interface ge-1/1/4.0;
    bridge-domains {
        vlan_100 {
            vlan-id 100;
        }
        vlan_200 {
            vlan-id 200;
        }
    }
    ...
}
```

Copyright 2010 Juniper Networks, Inc. All rights reserved. JUNIPER Worldwide Education Services www.juniper.net | 4-20

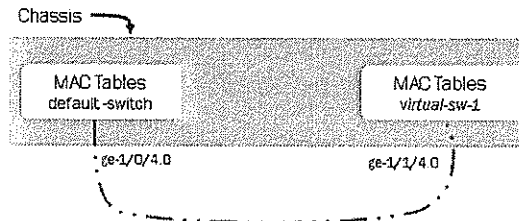
### Configure and Assign Physical Interfaces

The slide shows how to configure and assign Layer 2 interfaces to virtual switches.



## Verify Switch Settings

- Use the `show bridge domain` command to verify settings



```
user@switch> show bridge domain
```

Routing instance	Bridge domain	VLAN ID	Interfaces
default-switch	vlan_100	100	
default-switch	vlan_200	200	ge-1/0/4.0
virtual-sw-1	vlan_100	100	ge-1/0/4.0
virtual-sw-1	vlan_200	200	ge-1/1/4.0

© 2010 Juniper Networks, Inc. All rights reserved.

Juniper Worldwide Education Services

www.juniper.net | 4-21

### Verify Settings

Looking at the output on the slide, you can see that the `ge-1/1/4.0` interface is now bound to the `virtual-sw-1` routing instance and the bridge domains `vlan_100` and `vlan_200`, whereas `ge-1/0/4.0` belongs to the default switch.

## Summary

- In this chapter, we:
  - Described the use of a routing instance
  - Described the function of a virtual router
  - Described the function of a virtual switch
  - Implemented a virtual switch
  - Interconnected local routing instances

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-22

### This Chapter Discussed:

- The use of a routing instance;
- The function of a virtual router;
- The function of a virtual switch;
- Implementation of a virtual switch; and
- Interconnection of local routing instances.

## Review Questions

1. How can you make your MX Series router appear as multiple routers to other devices? How can you make it appear as multiple switches?
2. After configuring an interface under `[edit interfaces]`, which step do you perform next to ensure that the interface appears as part of the `vs1` virtual switch?
3. After configuring an IRB interface as part of the `vs1` virtual switch, in which routing table will you find its associated routes?

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-93

## Review Questions

- 1.
- 2.
- 3.

## Lab 2: Virtual Switches

- Configure virtual-switch routing instances.
- Configure bridge domains for the virtual switches.
- Create a virtual-switched network by configuring virtual-switch routing instances.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 4-24

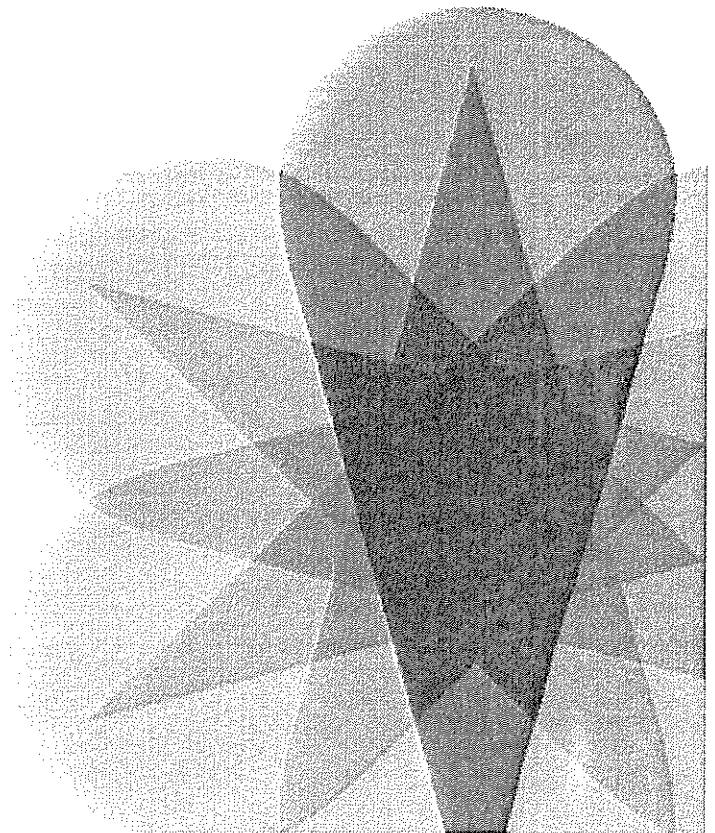
### Lab 2: Virtual Switches

The slide provides the objective for this lab.



# **Junos Service Provider Switching**

## **Chapter 5: Provider Bridging**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Describe the different IEEE VLAN stacking models
  - Describe the components of provider bridging
  - Configure and monitor provider bridging
  - Describe the components of provider backbone bridging
  - Configure and monitor provider backbone bridging

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 5-2

### This Chapter Discusses:

- Institute of Electrical and Electronics Engineers (IEEE) virtual LAN (VLAN) stacking models;
- The components of provider bridging;
- Configuration of provider bridging;
- The components of a provider backbone bridged network (PBBN); and
- Configuration of provider backbone bridging.

## Agenda: Provider Bridging

### → Expanding the Bridged Network

- Provider Bridging
- Configuring and Monitoring Provider Bridging
- Provider Backbone Bridging
- Configuring and Monitoring Provider Backbone Bridging

© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

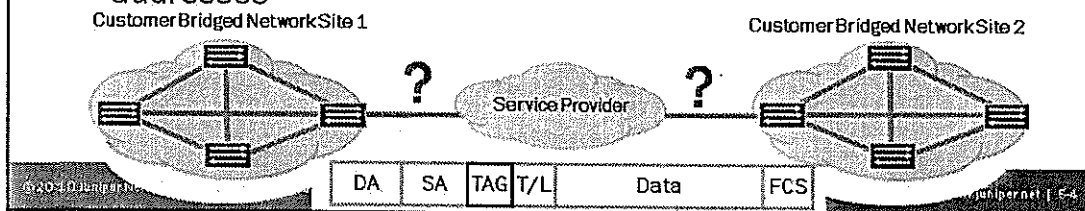
www.juniper.net | 5-3

### Expanding the Bridged Network

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## Customer Bridged Network

- IEEE 802.1Q VLANs allow the customer's local bridged networks to scale:
  - VLAN tags allow for up to 4094 separate broadcast domains
- Service provider scaling issues (for Ethernet virtual connections):
  - Service provider network needs to be aware of customer's bridging (spanning tree) and VLAN administration
  - Problem of overlapping VLAN IDs between service provider customers
  - Service provider bridges learn and store customer MAC addresses



## Scaling Customer Bridged Networks

IEEE 802.1Q VLAN tagging makes it possible for a customer's bridged network to scale. Instead of needing to add more bridging equipment to a growing network, VLAN tagging allows for the logical separation of a bridged network into many broadcast domains (or VLANs). With a 12-bit length VLAN ID, 4094 VLANs are available for use on a single physical Ethernet network.

## Ethernet from Service Providers

Because of its simple nature, service provider customers generally understand Ethernet. For a long time, service providers have searched for ways to deliver Ethernet Virtual Circuits (EVCs) to the customer premises. To a customer, an EVC between two sites should appear as a simple Ethernet link or VLAN through the service provider's network. IEEE 802.1Q VLAN tagging does not provide the scalability (in the service provider network) for a service provider to deliver that type of service.

*Continued on next page.*



## Ethernet from Service Providers (contd.)

From the service provider's point of view, the following is a list of some of the scaling issues that might arise:

- Because only one VLAN tag field exists in an 802.1Q frame, customers and the service provider need to coordinate the use of VLAN ID space. Considering that a service provider might have thousands of customers, this coordination would be an overly extreme effort.
- To pass Ethernet frames between customer sites, the service provider bridges must learn customer MAC addresses.
- To provide redundant links between customers and the service provider, running a form of the Spanning Tree Protocol (STP), which is generally not a viable solution, might be necessary. The STPs of today cannot scale to support all service provider and customer bridges of the world in a single spanning-tree domain.

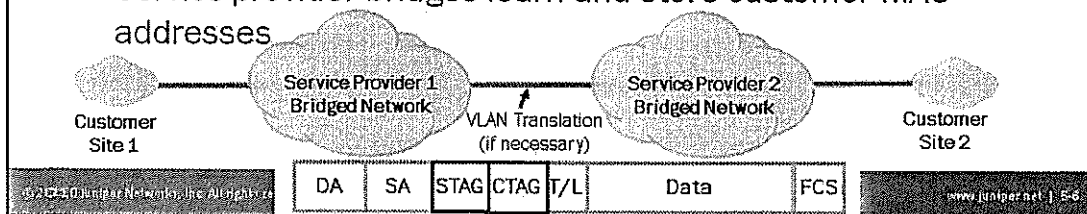
## Provider Bridged Network

### ■ IEEE 802.1ad provides the standard for stacking VLAN tags:

- Allows the service provider to provide LAN service through the service provider network
  - Each outer tag (S-VLAN tag) represents a customer (4094 possible)
  - Inner tag (C-VLAN tag) represents any of a customer's 4094 VLANs
  - The service provider and the customer use unique spanning-tree domains
  - Allows for VLAN translation between service provider bridged networks

### ■ Service provider scaling issues:

- Service provider bridges learn and store customer MAC addresses



## IEEE 802.1ad

IEEE 802.1ad has standardized the methodology of stacking VLAN tags. The slide shows the frame format that the standard introduced. The standard gives a new name to the 802.1Q VLAN tag: the Customer VLAN (C-VLAN) tag (C-TAG). It also introduces a new tag named the Service VLAN (S-VLAN) tag (S-TAG). By adding the S-TAG to the frame, much less coordination is necessary between the customer and the service provider. At the customer site, the customer can continue to use 802.1Q tagging using C-VLAN IDs that are relevant only to their network (not the service provider's network). As 802.1Q-tagged frames arrive at the edge of the service provider's bridged network, the provider edge bridge (PEB) adds an S-TAG to the frame. The S-TAG, using a single S-VLAN ID, can carry any or all of the 4094 C-VLANs that are possibly in use by the customer. In the simplest case, a service provider can allocate a single S-VLAN ID to represent each of its individual customers, which allows the service provider to potentially support up to 4094 customers. IEEE 802.1ad also allows for the translating of S-VLAN IDs at the edge of a service provider's bridged network, which helps in the coordination of VLAN ID usage between service providers.

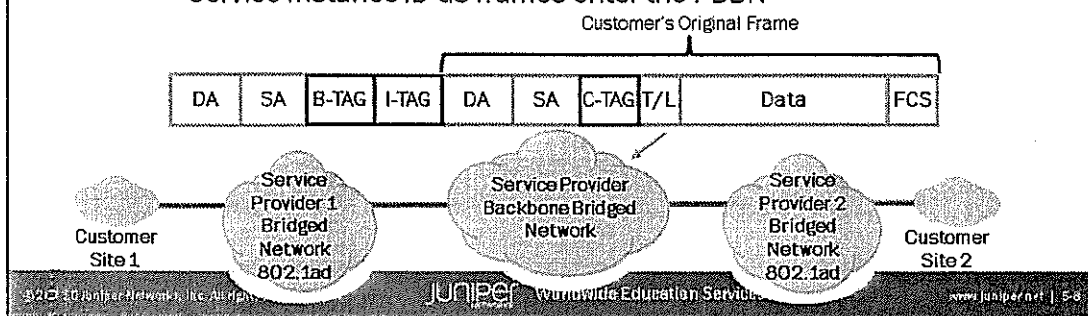
*Continued on next page.*

## Scaling Issues

Although IEEE 802.1ad helps to solve the issue of the limited VLAN ID space that we discussed in relation to IEEE 802.1Q tagging, it does not solve the MAC learning problem. That is, for frames to be forwarded between bridges in the service provider's network, the bridges each must learn and store MAC addresses learned from the customer networks. A service provider can help alleviate this problem by limiting the number of learned MAC addresses or charging the customer more for the EVC service if they exceed the MAC address limit.

## Provider Backbone Bridged Network

- IEEE 802.1ah solves the service provider scaling issues:
  - Specifies a frame format with two new tag types
    - Allows for nearly 16.8 million EVCs (customers)
    - Service provider backbone bridges do not learn and store customer's MAC addresses except at the edge of the PBBN (E-LAN only)
    - Outer MAC addresses exist only within the PBBN
  - IEEE 802.1ad S-TAG VLAN IDs translate into the I-TAG Backbone Service Instance ID as frames enter the PBBN



### IEEE 802.1ah: Provider Backbone Bridging

The IEEE 802.1ah standard attempts to solve all the EVC scaling issues. Customers can continue to use IEEE 802.1Q VLAN tagging at their individual sites using VLAN IDs that are relevant only to their bridged network. The service provider can continue to use IEEE 802.1ad VLAN tagging to provide EVC service to its customers (typically at the edge of the network). IEEE 802.1ah allows the provider to build out a scalable “backbone” network of bridges—the center network on the slide—to provide connectivity between its customers on its individual IEEE 802.1ad bridged networks.

The slide shows the 802.1ah Ethernet frame format. Notice that a new set of source and destination MAC addresses, a B-TAG, and an I-TAG encapsulate the original customer frame. The combination of the B-TAG and the I-TAG allows an individual provider to support up to 16.8 million customers. Also, because forwarding of the frame across the service provider occurs using the new source and destination MAC addresses (the MAC addresses are local to the bridges in the backbone network), the backbone bridges need not learn the customer's MAC addresses at all for Ethernet Line (E-Line) EVCs. For Ethernet LAN (E-LAN) EVCs, only bridges at the edge of the backbone learn customer MAC addresses. We discuss E-Line and E-LAN EVCs in more detail later in this chapter.

## Agenda: Provider Bridging

- Expanding the Bridged Network
- Provider Bridging
- Configuring and Monitoring Provider Bridging
- Provider Backbone Bridging
- Configuring and Monitoring Provider Backbone Bridging

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-9

### Provider Bridging

The slide highlights the topic we discuss next.

## What Is Provider Bridging?

### ▪ Defined by the IEEE 802.1ad standard:

- Allows for service providers to offer the equivalent of separate Ethernet LANs to their customers
  - Easy for the customer to understand (Ethernet)
  - Easy for the service provider to provision (1 VLAN equals 1 customer)
- Requires the use of 2 stacked VLAN tags
  - C-VLAN—typically controlled by the customer
  - S-VLAN—controlled by the service provider

Destination MAC Address	Source MAC Address	S-VLAN TAG	C-VLAN TAG	Type/ Length	Data	Frame Check Sequence
-------------------------------	--------------------------	---------------	---------------	-----------------	------	----------------------------

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 5-10

### Provider Bridging

Provider bridging is defined under IEEE 802.1ad. It was developed to allow a service provider to provide a more scalable EVC service to its customers. A typical provider bridged network (PBN) provides for C-VLAN tagging and forwarding at the edge of the network using the ports that face the customer. For all ports that face the core of the PBN, the provider bridges forward based only on the S-VLAN tag.

## TAG Formats

Destination MAC Address	Source MAC Address	S-VLAN TAG	C-VLAN TAG	Type/ Length	Data	Frame Check Sequence
-------------------------------	--------------------------	---------------	---------------	-----------------	------	----------------------------

### Tag formats:

#### • S-VLAN tag



- Tag Protocol Identifier: 16 bits, default 0x88A8
- Priority: 3 bits, 802.1p
- Drop Eligibility Indicator: 1 bit, default 0
- Unique VLAN identifier: 12 bits

#### • C-VLAN tag



- Tag Protocol Identifier: 16 bits, default 0x8100
- Priority: 3 bits, 802.1p
- Canonical Format Indicator: 1 bit, default 0
- Unique VLAN identifier: 12 bits

© 2008 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

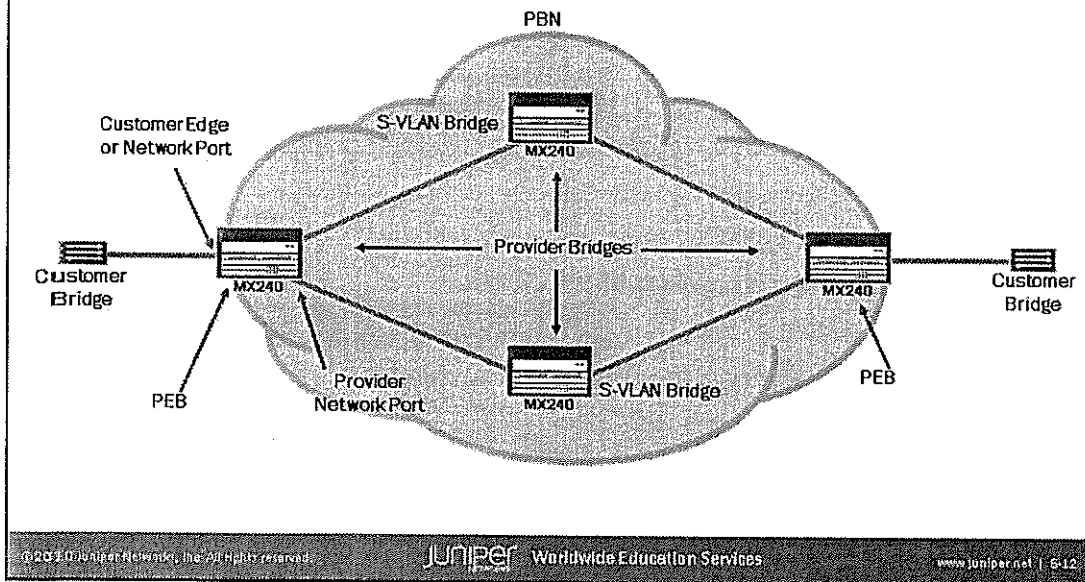
www.juniper.net J-511

## IEEE 802.1ad TAG Formats

The slide shows the S-TAG and C-TAG formats defined under IEEE 802.1ad. Note that the C-TAG remains identical to the IEEE 802.1Q VLAN tag. The S-TAG is similar but a few fields have been redefined. For example, because the canonical format indicator (CFI) field in the C-TAG is rarely used (for use in token ring networks), it has been redefined in the S-TAG to represent a frame's eligibility to be dropped. The Drop Eligibility Indicator (DEI) is used for class of service, which we do not discuss in this course. Also, IEEE 802.1ad has reserved a Tag Protocol Identifier (TPID) of 0x88A8 for the S-TAG, however, the Junos operating system default behavior is to set the TPID equal to 0x8100.

## Provider Bridging Terms

- Each device performs a specific task in a PBN



### PBN Terms

The following terms are used in a PBN network:

- **PBN:** A network of provider bridges that provide for transparent EVC service to the service provider's customers.
- **Provider Bridge:** A bridge in the service provider's network that performs IEEE 802.1ad VLAN tagging and forwarding. These bridges learn and store the MAC addresses of the service provider's customers.
- **Provider Edge Bridge (PEB):** Accepts and forwards IEEE 802.1Q frames to and from customers. PEBs also encapsulate the received customer frames using the IEEE 802.1ad format to forward customer frames across the PBN.
- **S-VLAN Bridge:** A nonedge provider bridge that forwards frames based only on the S-VLAN tag.
- **Provider Network Port:** A port on a provider bridge that forwards frames based on the S-VLAN tag.
- **Customer Edge Port:** A port on a PEB that connects to customer equipment that receives and transmits C-VLAN tagged frames.
- **Customer Network Port:** A port on a PEB that receives and transmits S-VLAN tagged frames.



## VLAN Tag Operations

- Provider bridges make several different types of adjustments to the VLAN stack:
  - These options can be configured explicitly (manually intensive) or using shortcut (implicit) methods that require minimal configuration
    - push: Add an outer tag
    - pop: Remove the outer tag
    - swap: Swap the outer tag with a new one
    - pop-pop: Remove the outer and inner tags
    - push-push: Add two tags
    - swap-swap: Swap the inner and outer tags with new ones
    - pop-swap: Pop the outer tag and swap the inner tag
    - swap-push: Swap the inner tag and add an outer tag
    - rewrite vlan and tag-protocol-id

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

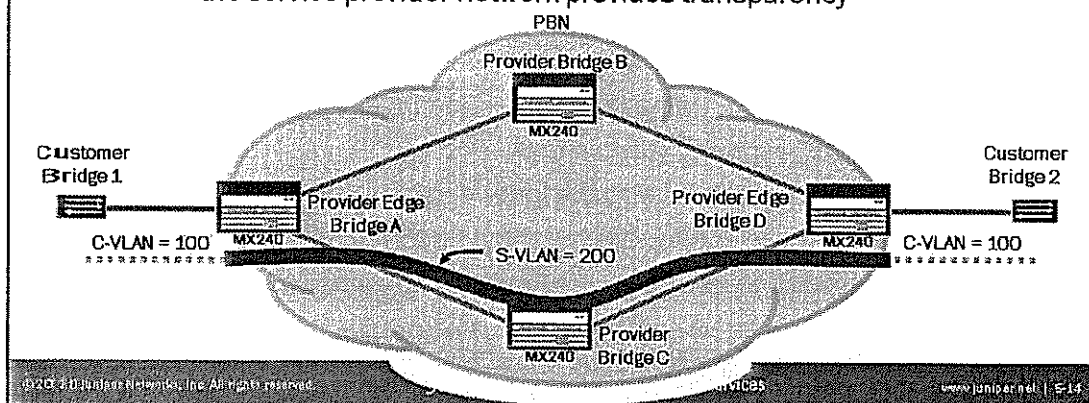
www.juniper.net | 5-13

## VLAN Tag Operations

The slide shows all of the possible operations that a provider bridge can perform on C-tagged frames and S-tagged frames that a port receives and transmits.

## Frame Processing Example (1 of 5)

- Service provider provides an EVC to the customer:
  - Customer uses 802.1Q-tagged frames (C-VLAN 100) to connect to the remote site while the service provider network is transparent
  - S-VLAN tagging of the customer frames during transmission across the service provider network provides transparency

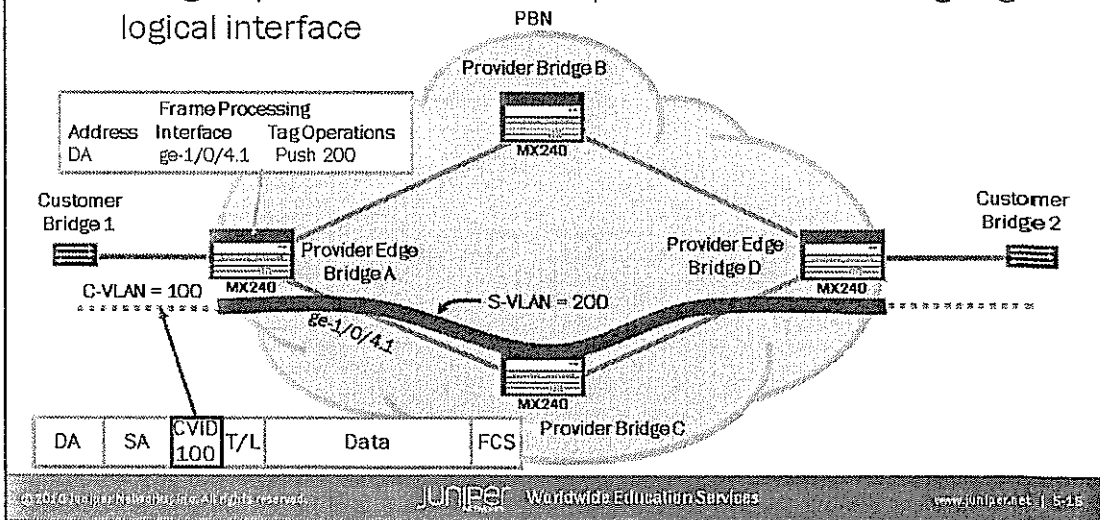


### Service Provider Provides EVC Service to the Customer

In the example, the service provider delivers an Ethernet circuit to each of the customer premises. To provide connectivity between Customer Bridge 1 and Customer Bridge 2, the customer must enable an IEEE 802.1Q VLAN using VLAN ID 100 on the service provider-facing ports. The service provider has allocated an S-VLAN tag of 200 to transparently forward the customer's frames across the PBN. This allocation is performed by configuring a bridge domain on each provider bridge specifically for the customer specifying an S-VLAN ID of 200, and by configuring all possible inbound and outbound interfaces to support the appropriate VLAN tagging for the customer's bridge domain. For example, on Bridge A, the service provider would need to configure a Bridge Domain that accepts C-tagged frames on the customer-facing interface and S-tagged frames (VLAN ID 200) on the core-facing interfaces.

## Frame Processing Example (2 of 5)

- Frames with a single C-VLAN tag with VLAN ID 100 arrive at Bridge A:
  - Bridge A performs a MAC lookup to determine the outgoing logical interface



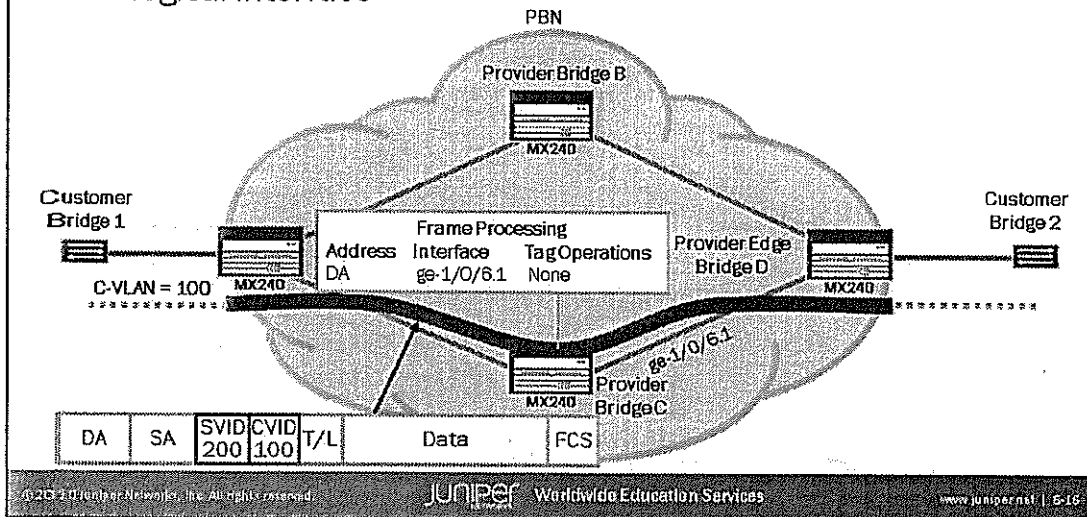
### PEB Processing of Incoming Frames

When C-VLAN-tagged frames arrive at Bridge A (a PEB), Bridge A performs a MAC-table lookup based on the customer's bridge domain. If Bridge A has previously learned the destination MAC address of the frame, it forwards the frame to the appropriate outbound interface (ge-1/0/4.1 in this case) and the interface pushes an S-VLAN of 200 on to the frame before sending the frame to the next bridge.

Note that if Bridge A did not previously learn the destination MAC address of the frames, it floods the frame out of every other interface associated with the customer's bridge domain except for the one that originally received the frame.

## Frame Processing Example (3 of 5)

- Frames with two VLAN tags arrive at Bridge C:
  - Bridge C performs a MAC lookup to determine the outgoing logical interface



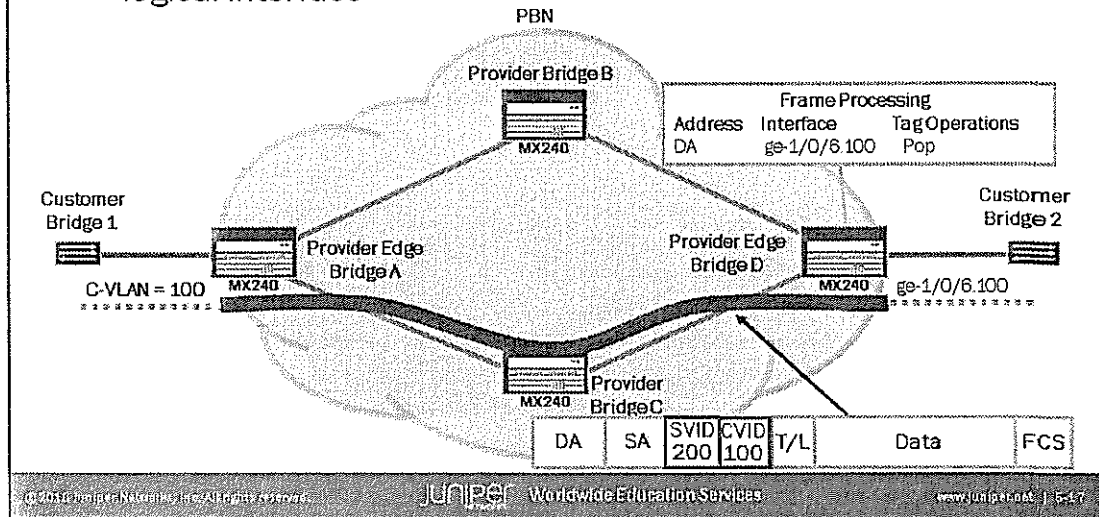
### Bridge C Processes the Frame

When S-VLAN-tagged frames arrive at Bridge C (an S-VLAN bridge), it performs a MAC-table lookup based on the customer's bridge domain. If Bridge C has previously learned the destination MAC address of the frame, it forwards the frame to the appropriate outbound interface (ge-1/0/6.1 in this case) and the interface sends the frame unchanged to the next bridge.

Note that a few ways exist to configure the VLAN operations on an S-VLAN bridge. The inbound interface on Bridge C can possibly also pop the S-VLAN tag on reception and then the outbound interface can push S-VLAN of 200 on transmittal.

## Frame Processing Example (4 of 5)

- Frames with two VLAN tags arrive at Bridge D:
  - Bridge D performs a MAC lookup to determine the outgoing logical interface

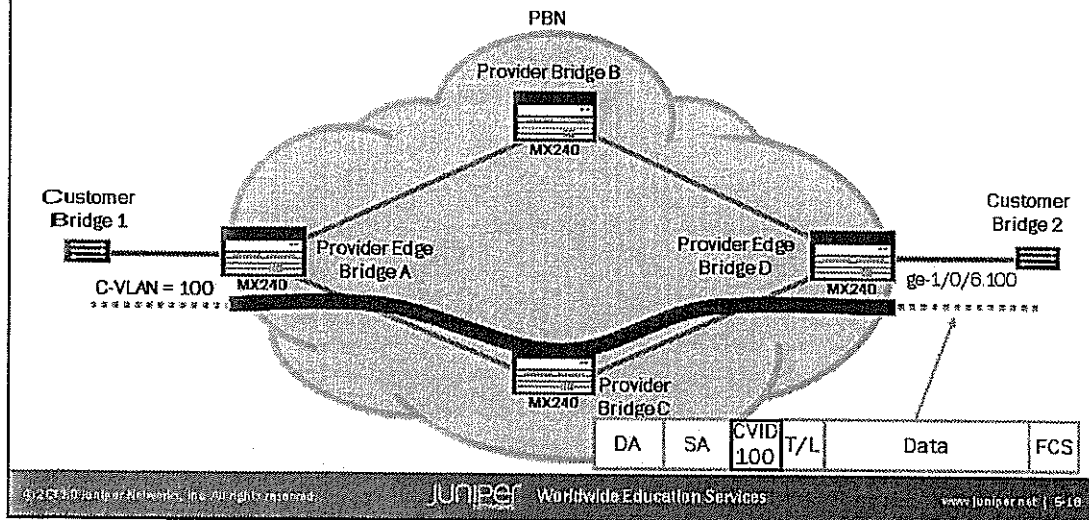


### Bridge D Processes the Frame

When S-VLAN-tagged frames arrive at Bridge D (PEB), the inbound interface pops the S-VLAN tag and Bridge D performs a MAC-table lookup based on the customer's bridge domain. If Bridge D has previously learned the destination MAC address of the frame, it forwards the frame to the appropriate outbound interface (ge-1/0/0.100 in this case) and the interface sends the C-tagged frame to the customer bridge.

## Frame Processing Example (5 of 5)

- Frames with a single VLAN tag arrive at Customer Bridge 2



### Remote Customer Site

The slide shows the frame format of the Ethernet frame as it arrives at Customer Bridge 2. It looks exactly as it did when Customer Bridge 1 transmitted it.

## Agenda: Provider Bridging

- Expanding the Bridged Network
- Provider Bridging
- Configuring and Monitoring Provider Bridging
- Provider Backbone Bridging
- Configuring and Monitoring Provider Backbone Bridging

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

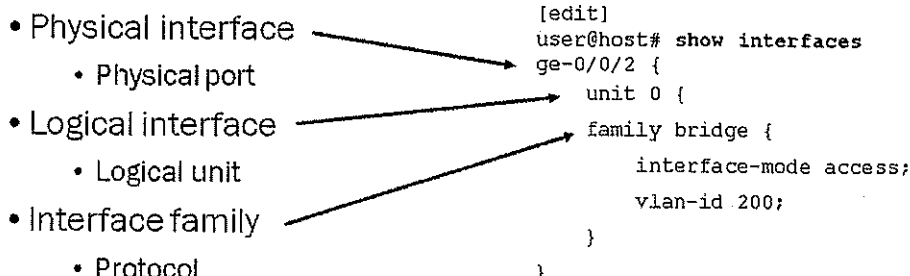
WWW.JUNIPER.NET | 5-19

### Configuring and Monitoring Provider Bridging

The slide highlights the topic we discuss next.

## Terminology Check

- Junos OS interfaces are sometimes referred to in shorthand



- Bridge domains can learn in two modes
  - Independent VLAN learning
    - Learning domain for each VLAN
  - Shared VLAN learning
    - Single learning domain shared by all VLANs in a bridge domain

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-20

## Junos OS Interface Terminology

Shorthand methods of describing the Junos OS interfaces are common. A physical interface refers to a physical port. A logical interface refers to an individual logical unit. An interface family refers to an individual protocol family. Multiple logical interfaces can be configured for each physical interface. Multiple interface families can be configured for each logical interface. In regards to bridging, understanding how a configuration affects the number of logical interfaces on an MX Series 3D Ethernet Universal Edge Router (64 K maximum) is important.

## Bridge Domain Modes

So far, we discussed configuring bridge domains in independent VLAN learning mode (IVL). In this mode, MAC learning occurs on a per VLAN basis. That means, broadcast, unicast with unknown destination, and multicast (BUM) traffic flooded on interfaces is associated with a single VLAN. However, another bridge domain mode exists named shared VLAN learning mode (SVL). This allows for VLANs to share MAC learning. That means, the BUM traffic floods on all interfaces and all VLANs associated with a bridge domain. The following slides show examples of each mode of operation.



## Dual-Stacked VLAN Configuration Examples

### ■ Original style

```
(edit)
user@peb# show interfaces ge-1/0/4
flexible-vlan-tagging;
encapsulation flexible-ethernet-services;
unit 0 {
    encapsulation vlan-bridge;
    vlan-tags outer 200 inner-range 111-114;
}
```

### ■ New style

```
(edit)
user@peb# show interfaces ge-1/0/4
flexible-vlan-tagging;
unit 0 {
    vlan-id 200;
    family bridge {
        interface-mode trunk;
        inner-vlan-id-list 111-114;
    }
}
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-21

### Original Style

The first example on the slide shows one of the original configuration styles used to create dual-stacked VLAN subinterfaces. The `vlan-tags` statement is used to specify the inner and outer VLAN IDs.

### New Style

The second example on the slide shows the “new” method of configuration to create dual-stacked VLAN subinterfaces. To configure the outer VLAN, specify a `vlan-id` at the unit level. To specify one or more inner VLAN IDs use the **`inner-vlan-id-list`** command at the `family bridge` level of the hierarchy.

## Bridge Modes—Original Style Configuration

- VLAN configuration at the [edit bridge-domains name] hierarchy determines the mode
  - If learning domain = 1, SVL is possible

Bridge Domain VLAN	Learning Domains	VLAN Map	Usage
Default (no VLAN specified)	1	Yes	
None (vlan-id none)	1	No	Circuit switching
Single (vlan-id number)	1	No	VLAN switching
Dual (vlan-tags outer number inner number)	1	No	Dual VLAN switching
All (vlan-id all)	4K	No	Tunneling C-VLANs

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | G22

### Bridge Modes

When using the original style of configuring a dual-stacked interface, you still must apply the interface to a bridge domain. The way in which you specify the VLAN IDs for the bridge domain determines the bridge domain's mode of operation. The following list briefly explains the different ways of configuring a bridge domain:

- **Default:** You do not specify a VLAN ID for the bridge domain. The bridge domain is a single learning domain. You configure an input and output VLAN map to explicitly configure push, pop, swap, and other VLAN operations.
- **None:** You specify `vlan-id none` for the bridge domain. The bridge domain is a single learning domain. In this case, all inbound frames have all VLAN IDs popped. All outbound frames take on the VLAN settings of the outbound interfaces.
- **Single:** You specify `vlan-id number` for the bridge domain. The bridge domain is a single learning domain. In this case, all inbound frames have all S-VLAN IDs popped. All inbound C-VLAN IDs are normalized (translated) to the VLAN ID of the bridge domain. All outbound frames take on the VLAN settings of the outbound interface.

*Continued on next page.*

## Bridge Modes (contd.)

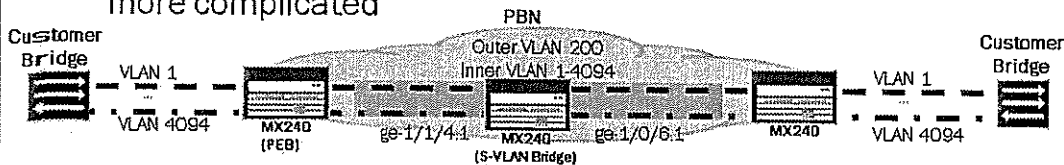
- *Double:* You specify `vlan-tags outer number inner number` for the bridge domain. The bridge domain is a single learning domain. All incoming frames have their VLANs normalized (translated) to the outer and inner VLAN ID that is specified for the bridge domain. All outbound frames take on the VLAN settings of the outbound interface.
- *All:* You specify `vlan-id all` for the bridge domain. The bridge domain has multiple learning domains. One learning domain exists for each C-VLAN configured on interfaces that are associated with the bridge domain. This type of configuration always results in IVL. Inbound frames retain their VLAN tags. All outbound frames take on the VLAN settings of the outbound interface.

Most of these options listed cause a bridge domain to have a single learning domain. If the interfaces assigned to a bridge domain are configured for a unique C-VLAN ID, then the learning mode for the bridge domain will be IVL. If the interfaces assigned to a bridge domain are configured for multiple C-VLANs, then the learning mode for the bridge domain will be SVL.

When using the new style of configuration, IVL is the usual mode of operation. SVL can occur only with the new style of configuration when mixing both old style and new style configurations in a bridge domain.

## S-VLAN Bridge Configuration

- An S-VLAN bridge forwards frames using only the S-TAG:
  - If C-TAG operations are necessary, configuration becomes more complicated



```
(edit)
user@S-VLAN-Bridge# show interfaces
ge-1/0/6 {
    flexible-vlan-tagging;
    encapsulation flexible-ethernet-services;
    unit 1 {
        family bridge {
            interface-mode trunk;
            vlan-id-list 200;
        }
    }
}
ge-1/1/4 {
    flexible-vlan-tagging;
    encapsulation flexible-ethernet-services;
    unit 1 {
        family bridge {
            interface-mode trunk;
            vlan-id-list 200;
        }
    }
}
```

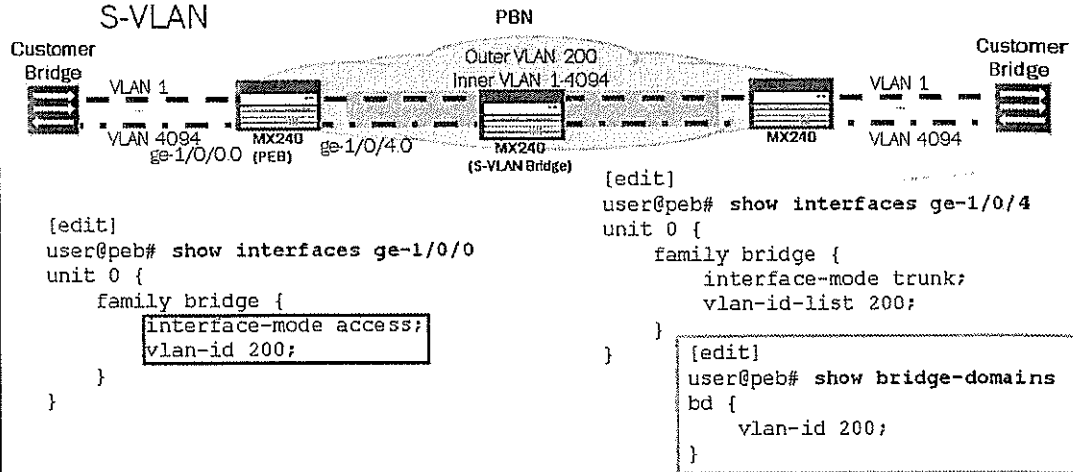
```
(edit)
user@S-VLAN-Bridge# show bridge-domains
customer1 {
    vlan-id 200;
}
```

### S-VLAN Bridge Configuration

The easiest configuration for supporting provider bridging is on an S-VLAN bridge similar to the core (middle) switch on the slide. Because the switch processes only S-VLAN tags, you can configure the bridge domain using the **vlan-id number** statement. We expect only S-tagged frames to arrive on each trunk interface, so you can configure them for a single **vlan-id-list** statement as well. To allow the interfaces to support two VLAN tags, include the **stacked-vlan-tagging** statement or the **flexible-vlan-tagging** statement.

## PEB—Tunnel All C-VLANs (1 of 2)

- The bridge domain references only the outer VLAN ID:
  - Uses one customer-facing logical interface and one bridge domain—uses IVL
  - Adding a second customer is just as easy but uses a different S-VLAN



### Tunnel All C-VLANs: Part 1

The method shown on the slide is the easiest and most elegant method of tunneling all customer C-VLANs across the core of a PBN. The interface and bridge domain configuration require only that you specify the outer S-VLAN ID. To allow single-tagged frames to enter the customer-facing interface, you must specify the `interface-mode access` statement.

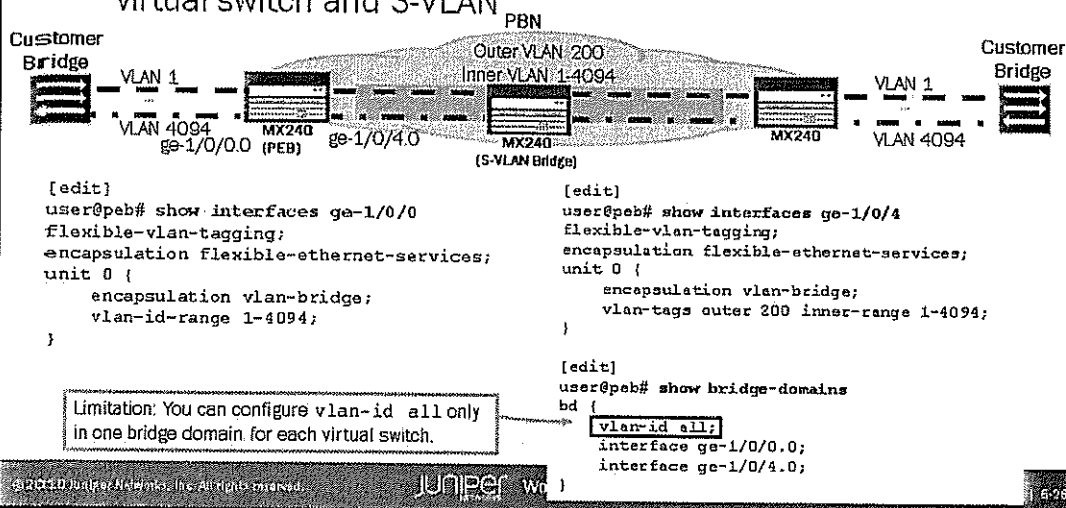
You will see on the next few slides that each configuration method results in some combination of one of the following:

1. A bridge domain mode (IVL or SVL).
2. Customer-facing logical interface usage.
3. Bridge domain usage.
4. Virtual switch usage.

The solution on this slide is so elegant because to support each customer it requires the use of one logical interface, one bridge domain, and also, you can place each customer in the same virtual switch.

## PEB—Tunnel All C-VLANs (2 of 2)

- Configure the bridge domain with `vlan-id all`:
  - Uses one customer-facing logical interface and one bridge domain—uses IVL
  - Adding a second customer requires configuring a second virtual switch and S-VLAN

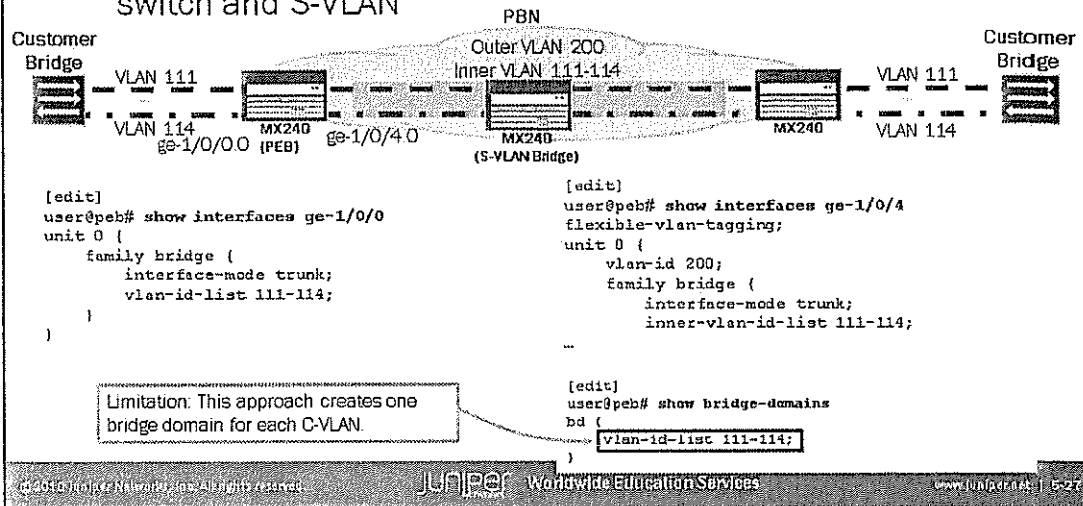


### Tunnel All C-VLANs: Part 2

The slide shows another method of tunneling all customer C-VLANs across the core of a PBN using the original style of configuration. For each customer, this solution requires the use of one logical interface, one bridge domain, and also, you must place each customer in the same virtual switch. The reason that you must place each customer in a different virtual switch is because only one bridge domain in a virtual switch can use the `vlan-id all` statement.

## PEB—Tunnel a Range of C-VLANs (1 of 3)

- Configure the bridge domain with `vlan-id-list`
  - Creates a single logical interface and bridge domain for each C-VLAN—uses IVL
  - Adding a second customer requires configuring a virtual switch and S-VLAN

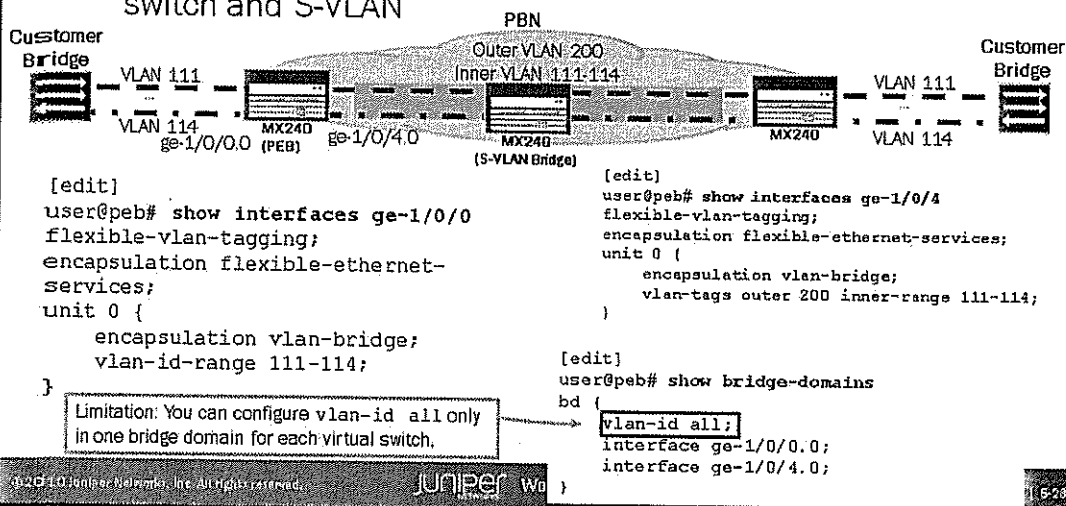


### Range of C-VLANs: Part 1

Allowing only certain C-VLANs to be tunneled across the core might be necessary. Few solutions will allow this tactic. In this solution, the bridge domain references the C-VLAN IDs to be tunneled. Because of this reference, you must add each customer to its own virtual switch (in the case of overlapping C-VLAN space).

## PEB—Tunnel a Range of C-VLANs (2 of 3)

- Configure the bridge domain with `vlan-id all`
  - Creates a single logical interface and one bridge domain—uses IVL
  - Adding a second customer requires configuring a virtual switch and S-VLAN



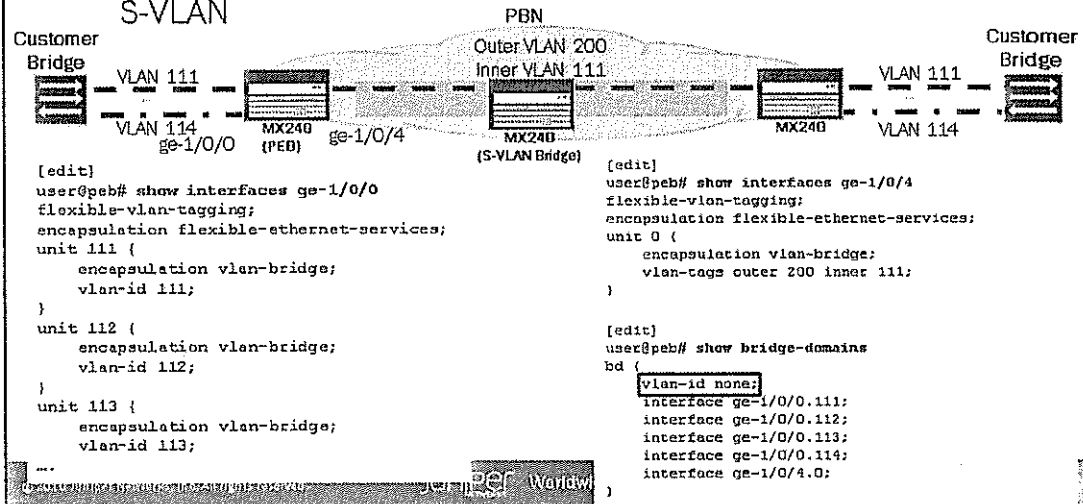
### Range of C-VLANs: Part 2

The slide shows another method of tunneling a range of C-VLANs across the core.



## PEB—Tunnel a Range of C-VLANs (3 of 3)

- Configure the bridge domain with **vlan-id none**
  - Creates multiple logical interfaces and one bridge domain—uses SVL
  - Adding a second customer requires configuring only an S-VLAN



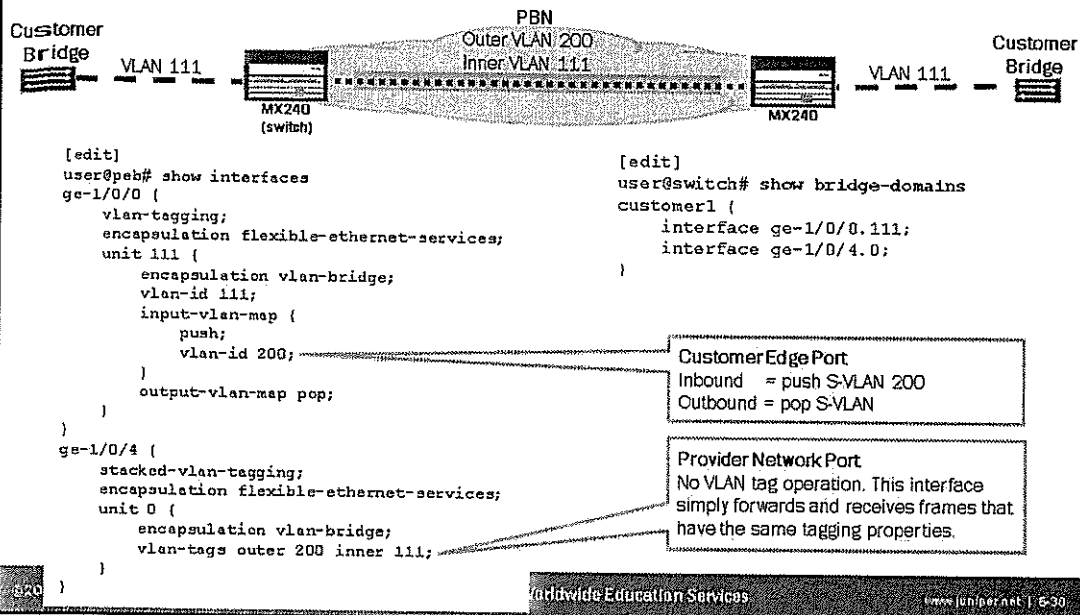
### Range of C-VLANs: Part 3

The slide shows the first example of SVL as well as an example of VLAN normalization (translation). The best way to describe how this solution works is to discuss what happens to a customer frame as it traverses the PBN:

1. A frame with C-VLAN ID 112 arrives on ge-1/0/0.112 destined for a MAC address that exists on the remote side of the network.
2. Because the bridge domain is configured for **vlan-id none**, the C-VLAN tag pops before the MAC-table lookup.
3. If the destination MAC address is unknown, then the frame is flooded out of all interfaces that associate with the bridge domain, including the subinterfaces of ge-1/0/0 (because of SVL). If the destination MAC is known, the frame is forwarded out of the ge-1/0/4.0 interface with a C-VLAN of 111 (normalization) and an S-VLAN of 200.
4. Upon arriving at the remote PEB, assuming the bridge domain is configured for **vlan-id none**, the S-VLAN and the C-VLAN tags are popped before the MAC-table lookup.
5. If the destination MAC address is unknown, then the frame is flooded out of all interfaces that associate with the bridge domain, including the subinterfaces of customer-facing interfaces (because of SVL). If the destination MAC address is known, the frame is forwarded out of the appropriate subinterface using the encapsulation specified on the interface.

## Input and Output VLAN Maps

- Use `input-vlan-map` and `output-vlan-map` to explicitly configure VLAN tag operations

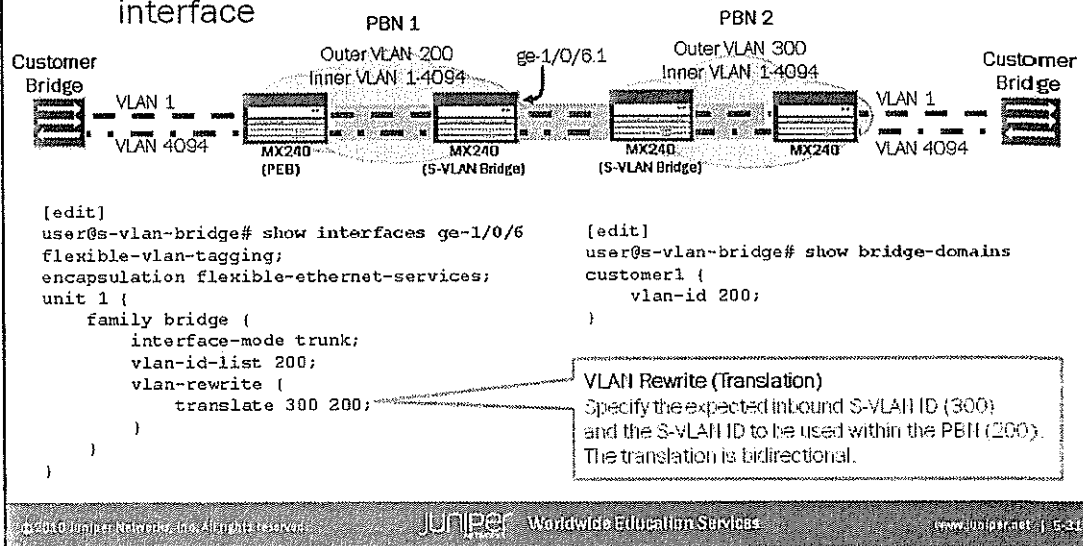


### Explicit Configuration of Tag Operations

The slide shows an example of explicitly configuring the VLAN tag operations to be performed on an interface.

## Typical PBN Network-to-Network Interface Configuration

- Service providers manage their own S-VLAN ID space:
  - A single customer can use multiple service providers
  - S-VLAN translation is used at the network-to-network interface

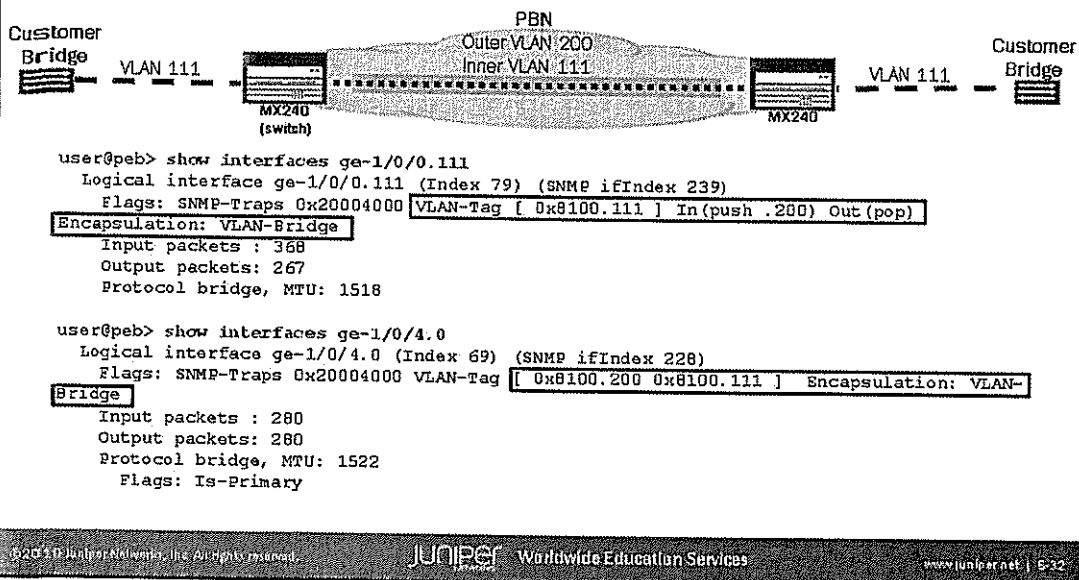


### PBN Network-to-Network Interface

On the slide, two service providers provide an EVC to a single customer. To allow for the interconnection of the two customer sites, the two service providers must exchange S-VLAN-tagged frames between one another. However, the case might be that each service provider is using a different S-TAG to provide the EVC. In the example, PBN 1 uses S-VLAN 200 and PBN 2 uses S-VLAN 300. IEEE 802.1ad provides the ability to perform S-VLAN translation between service providers. The slide shows the configuration necessary for the S-VLAN bridge in PBN 1 to perform VLAN translation.

## Verifying Tag Operation Settings

- Use the `show interface` command to view the tag operations to be performed on interfaces



### View Tag Operation Settings

To view the expected VLAN tag operations that an interface will perform, issue the `show interfaces` command. The `VLAN-tag` field shows the VLAN IDs for which the interface was specifically configured. The `In` and `Out` fields show the VLAN operations that the interface will perform.

## Agenda: Provider Bridging

- Expanding the Bridged Network
- Provider Bridging
- Configuring and Monitoring Provider Bridging
- Provider Backbone Bridging
- Configuring and Monitoring Provider Backbone Bridging

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

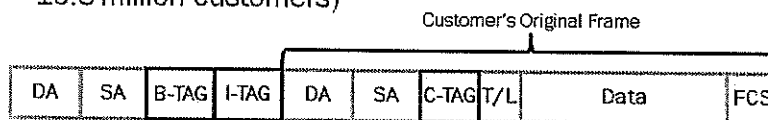
www.juniper.net | 5-33

### Provider Backbone Bridging

The slide highlights the topic we discuss next.

## What Is Provider Backbone Bridging?

- Defined by IEEE 802.1ah standard:
  - Requires the use of two new VLAN tags
    - Backbone Service Instance tag
    - Backbone VLAN tag
  - Allows service providers to offer scalable Ethernet service to their customers
    - E-Line: Point-to-point EVC
    - E-LAN: Multipoint-to-multipoint EVC
    - One B-TAG plus one I-TAG equals one customer (supports almost 16.8 million customers)



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | C-34

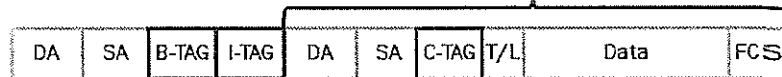
### IEEE 802.1ah

This standard specifies yet another method of stacking VLANs. It allows for the original customer's C-VLAN-tagged frame to be carried across a PBBN while being encapsulated in a new set of MAC addresses and two new tags: the Backbone Service Instance tag (I-TAG) and the Backbone VLAN tag (B-TAG). The MAC addresses are unrelated to the customer's network. Instead, they are the addresses of the service provider's bridges at the edge of the PBBN. These bridges are called backbone edge bridges (BEBs).

Two types of EVCs exist that a PBBN can provide. One type is named an E-Line EVC and the other is an E-LAN EVC. An E-Line EVC emulates a point-to-point Ethernet circuit while the E-LAN EVC emulates a shared Ethernet segment. Because of the addition of the B-TAG and I-TAG, a service provider can support up to 16.8 million EVCs in a PBBN.

## Provider Backbone Bridge Tag Formats

Customer's Original Frame



### Provider backbone bridge tag formats:



#### • I-TAG

- Tag Protocol Identifier: 16 bits, default 0x88E7
- Priority: 3 bits, 802.1p
- Drop Eligibility Indicator: 1 bit, default 0
- Use Customer Address: 1 bit, default 0
- Reserved 1 and Reserved 2: 1 bit and 2 bits, default 0
- Backbone Service Instance ID: 24 bits

#### • B-TAG



- Exactly the same format as the S-VLAN tag

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | E-35

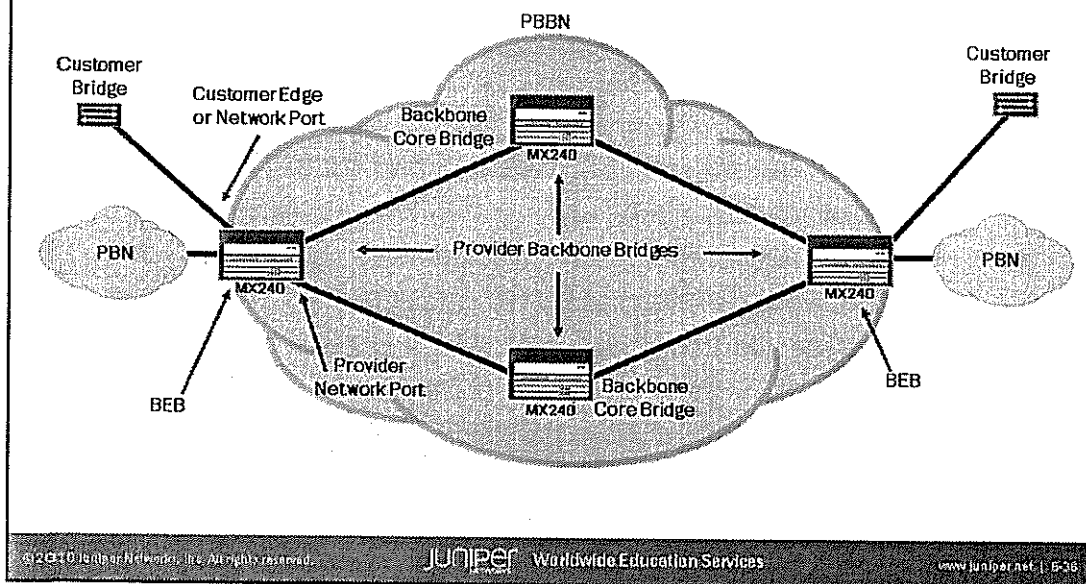
## Provider Backbone Bridge Tag Formats

The I-TAG is 6 octets in length. The C-TAG, S-TAG, and B-TAG are only 4 octets in length. The I-TAG has some similar fields to the other tag-types, such as TPID, Priority, and DEI. These fields serve the same purpose as those described for the other tag types. The User Customer Address (UCA) field signals the Backbone Service Instance—the I-component described later in this chapter—to use the source and destination MAC addresses from the original customer frame. The Backbone Service Instance ID (I-SID) is a 24-bit field that contains an instance ID for the EVC.

The B-TAG not only has the same format as an S-TAG, but it also serves the same purpose as the S-TAG with regard to forwarding frames across the PBBN.

## Provider Backbone Bridging Terms

- Each device performs a specific task in a PBBN



### Provider Backbone Terms

The following are definitions of some of the terms used in a PBBN network:

- **PBBN:** A network of provider backbone bridges that provide for transparent EVC service to the service provider's customers.
- **Provider Backbone Bridges (PBB):** Bridges in the service provider's network that perform IEEE 802.1ah VLAN tagging and forwarding. These bridges learn and store the MAC addresses of the service provider's customers only at the edge of the PBBN.
- **BEB:** Accepts and forwards IEEE 802.1Q or 802.1ad frames to and from its customers. BEBs map S-VLAN-tagged frames to an I-SID and then encapsulate them using the IEEE 802.1ah format to forward frames across the PBBN.
- **Backbone Core Bridge:** A nonedge PBB that forwards frames based only on the Backbone VLAN (B-VLAN) tag.
- **Provider Network Port:** A port on a PBB that forwards frames based on the B-VLAN tag.
- **Customer Edge Port:** A port on a BEB that connects to customer equipment that receives and transmits C-VLAN-tagged frames or S-VLAN-tagged frames.



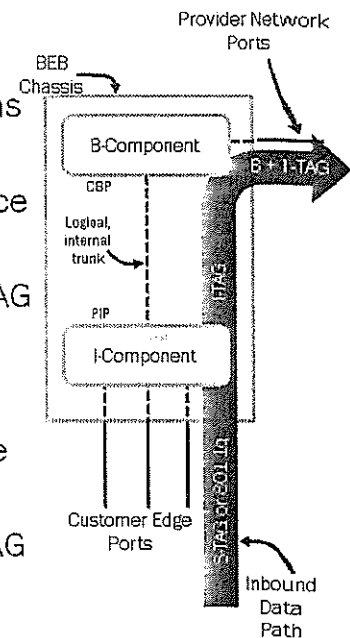
## Backbone Edge Bridge Components

### ■ I-component (a virtual switch):

- Receives frames from customers and learns customer MAC addresses
- Maps inbound frames to a backbone service instance
- Encapsulates customer frames with an I-TAG and a set of backbone MAC addresses

### ■ B-component (a virtual switch):

- Receives I-SID-tagged frames from multiple customers
- Encapsulates customer frames with a B-TAG
- Relays encapsulated frames to other backbone bridges



### I-Component

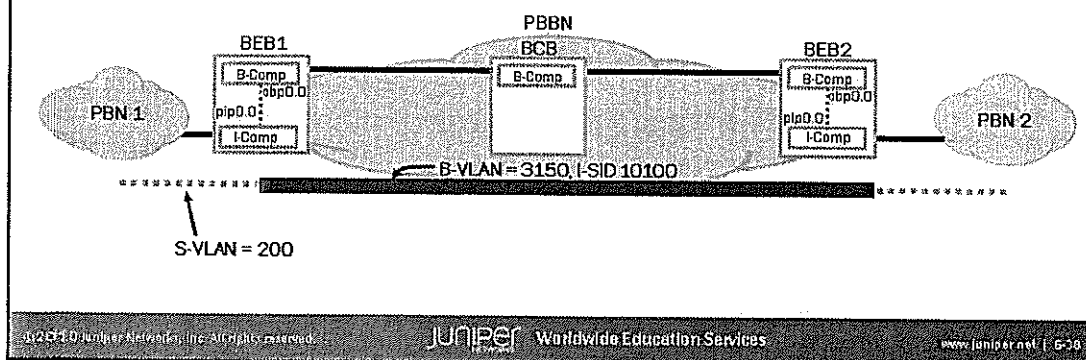
The slide shows the two components of a BEB and the traffic processing that occurs in the inbound direction. A BEB will have at least one I-component virtual switch and one B-component virtual switch. The I-component receives frames from customers and learns the customer's MAC addresses (E-LAN only). For E-Line EVCs, the I-component maps inbound frames to a manually configured I-SID based on the inbound interface. For E-LAN EVCs, the I-component maps inbound frames to a manually configured I-SID based on either the C-TAG (if single-tagged) or S-TAG (if double-tagged). An I-TAG then encapsulates the C-VLAN-tagged frames and a new source MAC address—Provider Instance Port (PIP)—and a new destination address (remote BEB's PIP). The I-SID-tagged frame is then passed internally using a virtual PIP to the B-component virtual switch.

### B-Component

The B-component receives the I-SID-tagged frames from the I-component through an internal virtual interface named the Customer Backbone Port (CBP). The B-component maps the I-SID-tagged frames to a B-VLAN tag. The B-component forwards the IEEE 802.1ah frame using the backbone MAC addresses (outer MAC address) and the B-VLAN ID.

## E-Line Frame Processing Example (1 of 5)

- The service provider (PBBN) provides an E-Line EVC to the customer's PBNs:
  - The customer uses 802.1ad-tagged frames (S-VLAN 200) to connect to the remote network while the service provider network is transparent



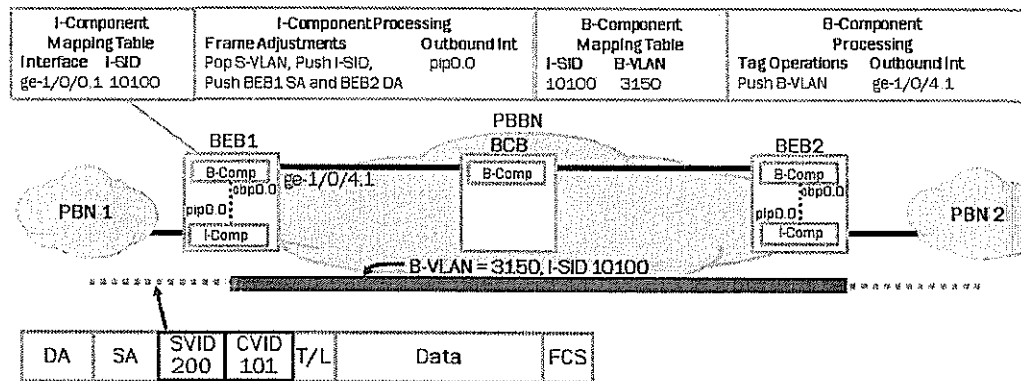
### Service Provider Provides PBBN E-Line EVC Service to the PBN

In the example on the slide, the service provider delivers a physical Ethernet circuit to each of the PBNs. To provide connectivity between PBN 1 and PBN 2, the PBNs use IEEE 802.1ad encapsulation using an S-VLAN ID 200 on the PBBN-facing ports. On the BEBs, the service provider has mapped an I-SID of 10100 to the customer-facing interface. Also, the service provider has mapped the I-SID of 10100 to a B-VLAN ID of 3150. These mappings allow for inbound frames from the PBNs to be forwarded transparently across the PBBN.

## E-Line Frame Processing Example (2 of 5)

### ■ Frames with an S-VLAN tag 200 arrive at BEB1:

- The incoming interface is associated with BEB1's I-component
- The I-component processes the frame and forwards it to the B-component
- The B-component processes the frame and forwards it to the BCB

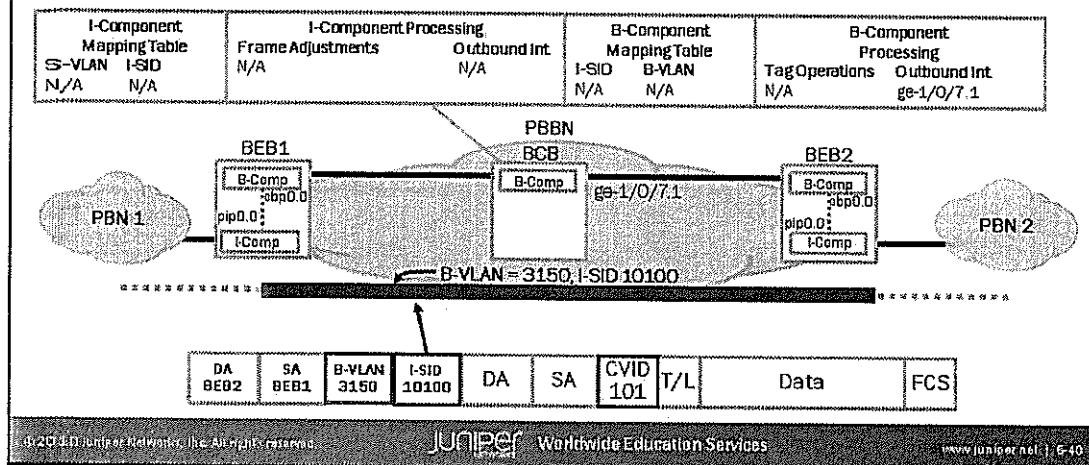


### Inbound S-VLAN-Tagged Frames Arrive

The slide shows all of the processing that occurs on the BEB1 when passing traffic from PBN 1 toward the backbone core bridge (BCB). Because the EVC is an E-Line EVC, the BEBs need not learn the customer's MAC addresses.

## E-Line Frame Processing Example (3 of 5)

- Frames with a B-TAG and an I-TAG arrive at the BCB:
  - The BCB performs a MAC-table lookup to determine the outbound logical interface
  - The BCB learns only the MAC addresses of the PBBN

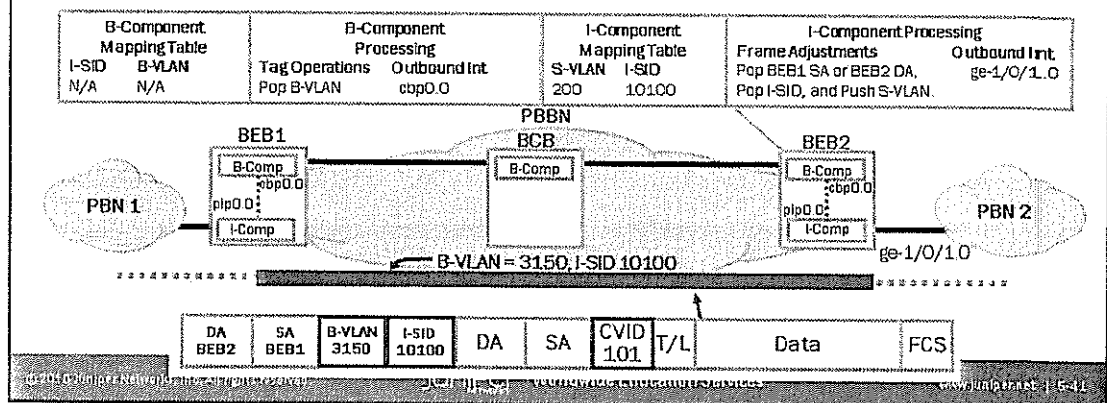


### The BCB Processes IEEE 802.1ah Frames

The slide shows all of the processing that occurs on the BCB when passing traffic from BEB1 toward BEB2.

## E-Line Frame Processing Example (4 of 5)

- Frames with a B-TAG and an I-TAG arrive at BEB2:
  - The incoming interface is associated with the BEB2's B-component
  - The B-component processes the frame and forwards it to the I-component
  - The I-component processes the frame and forwards it to customer

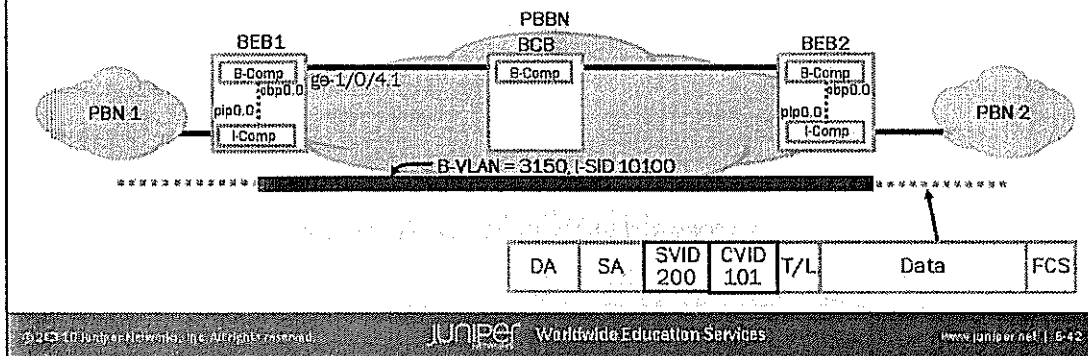


### BEB2 Processes IEEE 802.1ah Frames

The slide shows all of the processing that occurs on the BEB2 when passing traffic from the BCB toward PBN2.

## E-Line Frame Processing Example (5 of 5)

- Frames with an S-VLAN tag 200 arrive at the remote PBN:
- Remember, only the I-component of BEB1 and BEB2 needed to know the MAC addresses of the PBN network!



### Frames Arrive at the Destination

The slide shows the frame format of the Ethernet frame as it arrives at PBN2. It appears exactly as it did when PBN1 transmitted it.

## **Agenda: Provider Bridging**

- Expanding the Bridged Network
- Provider Bridging
- Configuring and Monitoring Provider Bridging
- Provider Backbone Bridging
- Configuring and Monitoring Provider Backbone Bridging

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-43

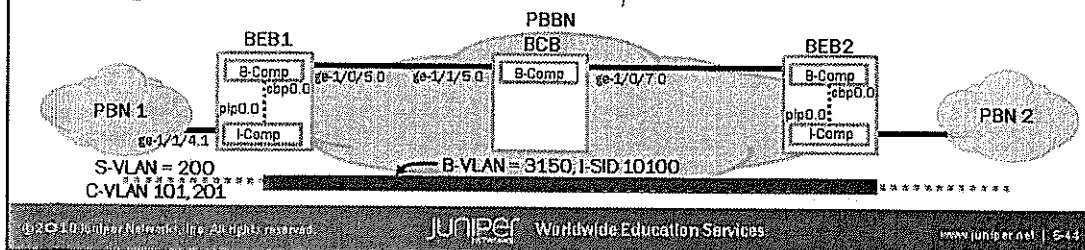
### **Configuring and Monitoring Provider Backbone Bridging**

The slide highlights the topic we discuss next.

## Configuring an E-Line EVC (1 of 3)

### ■ Configure interfaces for BEB1 (similar on BEB2):

```
[edit]
user@beb1# show interfaces
ge-1/0/5 {
    unit 0 {
        family bridge {
            interface-mode trunk;
            vlan-id-list 3150;
        }
    }
}
ge-1/1/4 {
    flexible-vlan-tagging;
    unit 1 {
        family bridge {
            interface-mode trunk;
            vlan-id-list 200;
        }
    }
}
--
cbp0 {
    unit 0 {
        family bridge {
            interface-mode trunk;
            bridge-domain-type bvlan;
            isid-list all;
        }
    }
}
pip0 {
    unit 0 {
        family bridge {
            interface-mode trunk;
            bridge-domain-type svlan;
            isid-list all-service-groups;
        }
    }
}
```



### Configure BEB Interfaces

In this example, BEB1 requires four configured interfaces. Obviously, you must configure the two physical interfaces, but you also must configure the two virtual interfaces (PIP and CBP) that provide the connectivity between the I-component virtual switch and the B-component virtual switch.

You configure all four interfaces as trunk interfaces. For both the PIP and CBP interfaces, you must specify a bridge-domain type. For the PIP interface it is always type **svlan**, and for the CBP interface it is always type **bvlan**. For both the PIP and CBP interfaces, you must specify the list of I-SIDs that the interface supports. For PIP, the only option is **all-service-groups** (as defined in the I-component virtual switch) and for the CBP, the only option is **all**.



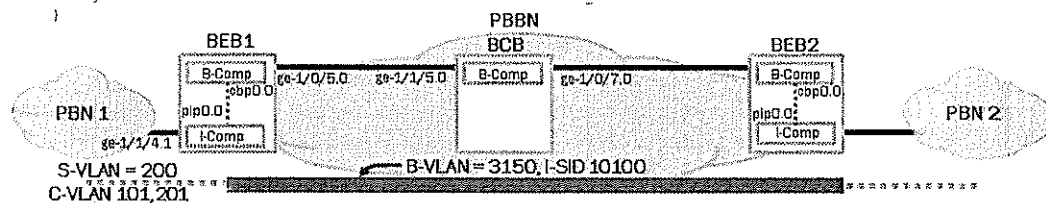
## Configuring an E-Line EVC (2 of 3)

- Configure the I-component and B-component virtual switches for BEB1:

- BEB2 configuration is similar

```
[edit]
user@behl# show routing-instances
b-component {
  instance-type virtual-switch;
  interface ge-1/0/5.0;
  interface cbp0.0;
  bridge-domains {
    eline-bvlan {
      vlan-id 3150;
    }
  }
  pbb-options {
    vlan-id 3150 isid-list 10100;
  }
}
```

```
i-component-for-eline {
  instance-type virtual-switch;
  interface ge-1/1/4.1;
  interface pip0.0;
  bridge-domains {
    eline-svlans {
      vlan-id-list 200;
    }
  }
  pbb-options {
    peer-instance b-component;
  }
  service-groups {
    eline1 {
      service-type eline;
      pbb-service-options {
        isid 10100 interface ge-1/1/4.1;
      }
    }
  }
}
```



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-15

### I-Component and B-Component Configuration

As with any virtual switch, you must specify the interfaces that belong to it. In this case, the two interfaces are pip0.0 and ge-1/1/4.1. You must also configure the bridge domain for the interface that faces PBN1. To allow for the pip0.0 interface to forward traffic internally to the B-component's cbp0.0 interface, you must specify which virtual switch will act as the I-component's peer instance. In this case, the peer instance will be the *b-component* virtual switch. Under the *service-groups* level of the hierarchy, you must also specify the type of EVC that will be used (E-Line in this case) and build the mapping table of the I-SID to the inbound interface. For E-LAN service, your mapping table would map I-SID to an S-VLAN ID.

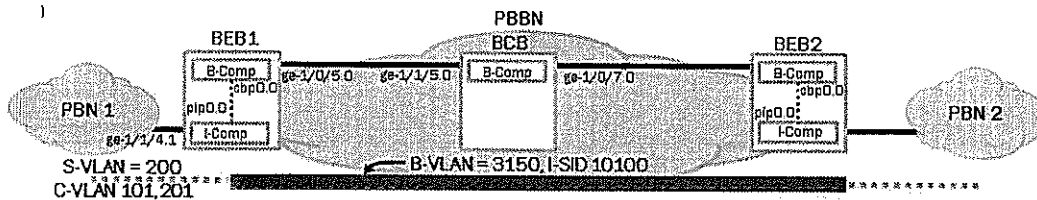
For the B-component, you must again specify the interfaces that belong to the virtual switch. For the bridge domain, you specify that B-VLAN IDs will be used. You must then map the inbound I-SID-tagged frames (from cbp0.0) to the appropriate B-VLAN ID.

## Configuring an E-Line EVC (3 of 3)

- Configure interfaces and a B-component virtual switch for the BCB:

```
[edit]
user@bcb# show interfaces
ge-1/0/7 {
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list 3150;
    }
  }
}
ge-1/1/5 {
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list 3150;
    }
  }
}
```

```
[edit]
user@bcb# show routing-instances
bcb-b-component {
  instance-type virtual-switch;
  interface ge-1/0/7.0;
  interface ge-1/1/5.0;
  bridge-domains {
    eline-bvlan {
      vlan-id 3150;
    }
  }
}
```



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-46

### BCB Configuration

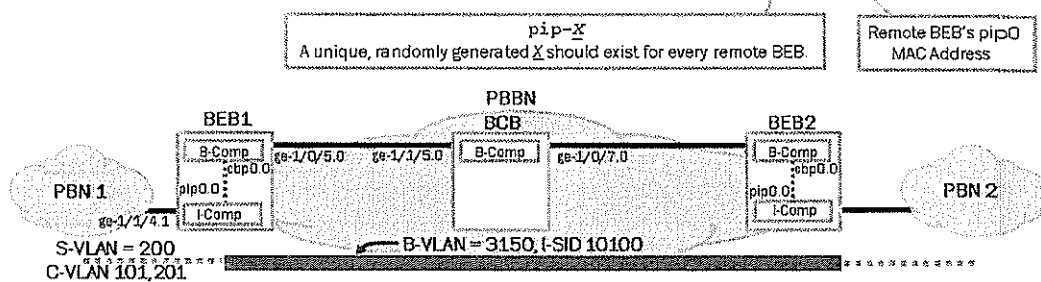
Because the BCB forwards traffic based on the outer MAC addresses and the B-VLAN tag, the configuration for the BCB is very simple. The slide shows the configuration of the BCB.

## Verify Bridge Domain Settings

- Verify that the correct interfaces associate with the B-component and the I-component

```
user@beb1> show bridge domain
```

Routing instance	Bridge domain	VLAN ID	Interfaces
b-component	eline-bvlan	3150	ge-1/0/5.0
i-component-for-eline	eline-svlans-vlan-0200	200	ge-1/1/4.1 pip-2.002159ab8fb0



## Bridge Domain Settings

The slide shows the output of the **show bridge domain** command. For the I-component virtual switch, you should expect to see a logical PIP interface for each remote BEB that the local switch learns.

## Summary

■ In this chapter, we:

- Described the IEEE VLAN stacking models
- Described the components of provider bridging
- Configured and monitored provider bridging
- Described the components of provider backbone bridging
- Configured and monitored provider backbone bridging

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-48

### This Chapter Discussed:

- IEEE VLAN stacking models;
- The components of provider bridging;
- Configuration of provider bridging;
- The components of a PBBN; and
- Configuration of provider backbone bridging.

## Review Questions

1. What are some of the scaling issues that can occur if a service provider were to use IEEE 802.1Q VLANs to provide LAN service to its customers?
2. List three VLAN tag operations that a switch can perform on a frame.
3. Which component of a BEB translates S-VLAN tags to I-SIDs?

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | E-49

## Review Questions

- 1.
- 2.
- 3.

## Lab 3: Provider Bridging

- Configure a PBN.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 5-50

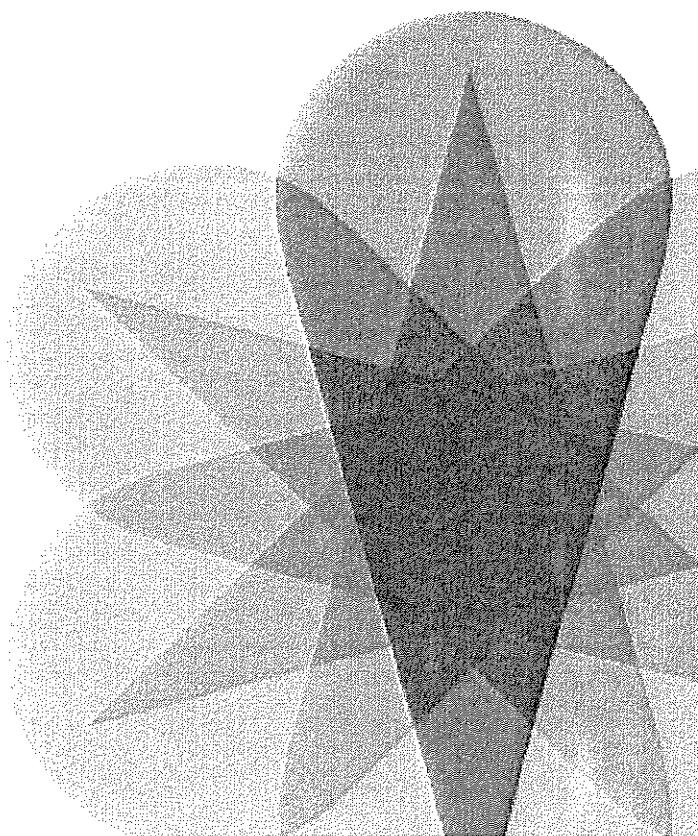
### Lab 3: Provider Bridging

The slide provides the objective for this lab.



# **Junos Service Provider Switching**

## **Chapter 6: Spanning-Tree Protocols**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Explain the purpose of spanning tree protocols
  - Describe the basic operation of STP, RSTP, MSTP, and VSTP
  - Configure and monitor STP, RSTP, MSTP, and VSTP
  - Explain the purpose of BPDU, loop, and root protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | 5-2

### This Chapter Discusses:

- The purpose of spanning tree protocols;
- The basic operation of the Spanning Tree Protocol (STP), the Rapid Spanning Tree Protocol (RSTP), the Multiple Spanning Tree Protocol (MSTP), and the virtual LAN (VLAN) Spanning Tree Protocol (VSTP);
- Configuration and monitoring of STP, RSTP, MSTP, and VSTP; and
- Implementation bridge protocol data unit (BPDU), loop, and root protection.



## Agenda: Spanning Tree Protocols

- Overview of STP
  - Overview of RSTP
  - Overview of MSTP
  - Overview of VSTP
  - Configuring and Monitoring Spanning Tree Protocols
  - Understanding BPDU, Loop, and Root Protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-3

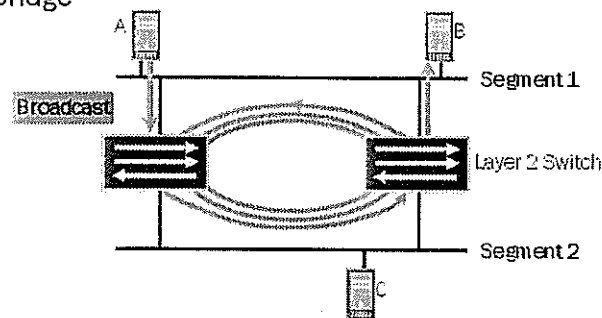
### Overview of STP

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## STP

### ■ STP:

- Defined in the IEEE 802.1D-1998 specification
- Builds loop-free paths in redundant Layer 2 networks
- Automatically rebuilds the tree when the topology changes
  - Configures the port state of every port on every participating bridge



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | E-4

## STP

STP is defined in the Institute of Electrical and Electronics Engineers (IEEE) 802.1D 1998 specification. STP is a simple Layer 2 protocol that prevents loops and calculates the best path through a switched network that contains redundant paths. STP is necessary only when redundant paths exist within a Layer 2 network. STP automatically rebuilds the tree when a topology change occurs.

## Terms and Concepts (1 of 2)

### ■ Key terms and concepts of STP:

- *Bridge ID*: Unique identifier for each switch
- *Root bridge*: Switch with the lowest bridge ID
- *Root port*: The port on each bridge closest to the root bridge
- *Root path cost*: A bridge's calculated cost to get from itself to the root bridge
  - Equal to the received root path cost from configuration BPDUs plus the port cost of the root port on the bridge
- *Port cost*: Every interface on a bridge has an assigned port cost value
  - Used in the calculation of the root path cost for the local bridge
  - Configurable value (1–200000000)
  - The default value is 20000 for 1 Gigabit Ethernet

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | E-5

### STP Terms and Concepts: Part 1

All switches participating in STP have a unique bridge ID. The bridge ID is a combination of the system MAC address and a configurable priority value. The lowest bridge ID determines the *root bridge*.

Once the root bridge is determined, each nonroot switch determines the least-cost path from itself to the root bridge. The port associated with the least-cost path, referred to as the *root path cost*, becomes the *root port* for the switch. Every port on a switch has a configurable *port cost* associated with it. A nonroot switch receives periodic STP BPDUs—described on next slide—that contain a root path cost as determined by the neighboring switch. The local switch adds the received root path cost to each of the port costs for its interfaces. Whichever interface is associated with the lowest value (root path cost + port cost) becomes the root port for the switch.

## Terms and Concepts (2 of 2)

### ▪ Key terms and concepts of STP (contd.):

- *Designated bridge*: A switch representing the LAN segment
- *Port ID*: A unique identifier for each port on each switch
- *Designated port*: The designated bridge's forwarding port on a LAN segment
  - The port used by a designated bridge to send traffic from the direction of the root to the LAN or from the LAN toward the root
- *Bridge protocol data unit*: Packets used to exchange information between switches
  - Configuration BPDU
  - Topology change notification BPDU

## STP Terms and Concepts: Part 2

All switches participating on a common network segment must determine which switch offers the least-cost path from the network segment to the root bridge. The switch with the best path becomes the *designated bridge* for the LAN segment, and the port connecting this switch to the network segment becomes the *designated port* for the LAN segment. If equal-cost paths to the root bridge exist between two or more switches for a given LAN segment, the *bridge ID* acts as a tiebreaker. If the bridge ID is used to help determine the designated bridge, the lowest bridge ID is selected. If two equal-cost paths exist between two ports on a single switch, then *port ID* acts as the tiebreaker (lower is preferable). The designated port transmits BPDUs on the segment.

## Port States

- Each individual port of each bridge can be in one of four states:
  - Blocking
    - The port drops all data packets and listens to BPDUs
    - The port is not used in active topology
  - Listening
    - The port drops all data packets and listens to BPDUs
    - The port is transitioning and will be used in active topology
  - Learning
    - The port drops all data packets and listens to BPDUs
    - The port is transitioning and the switch is learning MAC addresses
  - Forwarding
    - The port receives and forwards data packets and sends and receives BPDUs
    - The port has transitioned and the switch continues to learn MAC addresses

© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

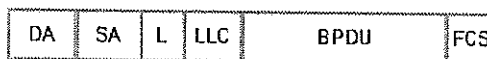
www.juniper.net | 6-7

## Port States

The slide highlights the STP port states along with a brief description of each state. In addition to the states listed on the slide, an interface can have STP administratively disabled (default behavior). An administratively disabled port does not participate in the spanning tree but does flood any BPDUs it receives to other ports associated with the same VLAN. Administratively disabled ports continue to perform basic bridging operations and forward data traffic based on the MAC address table.

## BPDUs—Ethernet Frame Format

### ■ Ethernet frame:



- Source Address—The outgoing port of the originating switch
- Destination Address—The bridge group address (01:30:C2:00:00:00)
- Length
- LLC Header
  - DSAP and SSAP = 0x42 (Bridge Spanning Tree Protocol)

### ■ BPDUs types:

- Configuration BPDUs
  - Used to build the spanning-tree topology
- Topology change notification BPDUs
  - Reports topology changes

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-9

## BPDUs Ethernet Frame

The slide shows the Ethernet frame format of an STP BPDUs. Notice that the Ethernet frame does not contain any 802.1Q-type VLAN tagging. The source address of the frame is the MAC address of the outgoing port of the sending switch. The destination address is the multicast MAC address that is reserved for STP. The frame also contains an LLC header that uses a destination service access point (DSAP) of 0x42, which refers to the bridge STP.

## BPDUs Types

STP uses BPDUs packets to exchange information between switches. Two types of BPDUs exist: configuration BPDUs and topology change notification (TCN) BPDUs. Configuration BPDUs determine the tree topology of a LAN. STP uses the information that the BPDUs provide to elect a root bridge, to identify root ports for each switch, to identify designated ports for each physical LAN segment, and to prune specific redundant links to create a loop-free tree topology. TCN BPDUs report topology changes within a switched network.

## Configuration BPDU Format (1 of 3)

Octets	
Protocol ID	2
Protocol Version	1
BPDU Type	1
Flags	1
Root ID	8
Root Path Cost	4
Bridge ID	8
Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2

### Configuration BPDU fields:

- *Protocol ID*—0 (STP)
- *Protocol Version*—0 (IEEE 802.1D-1998)
- *BPDU Type*—0 (Configuration BPDU)
- *Flags*
  - Topology Change Acknowledgment Flag (Bit 8)
  - Topology Change Flag (Bit 1)
- *Root ID*
  - A unique ID of the bridge that the transmitting bridge believes to be the root
- *Root Path Cost*
  - The sending switch's calculated total cost to get to the root bridge

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net/1.6.9

### Configuration BPDU Format: Part 1

When an STP network is first turned up, all participating bridges send out configuration BPDUs to advertise themselves as candidates for the root bridge. Each bridge uses the received BPDUs to help build the spanning tree and elect the root bridge, root ports, and designated ports for the network. Once the STP network converges and is stable, the root bridge sends a configuration BPDU once every few seconds (the hello time default is 2 seconds).

The following list provides a brief explanation of each of the BPDU fields:

- *Protocol ID*: This value is always 0.
- *Protocol Version*: This value is always 0.
- *BPDU Type*: This field determines which of the two BPDU formats this frame contains—configuration BPDU or TCN BPDU.
- *Flags*: This field is used to handle changes in the active topology; we discuss this field later.
- *Root ID*: This field contains the bridge ID (BID) of the root bridge. After convergence, all configuration BPDUs in the bridged network should contain the same value for this field (for a single VLAN). Some network sniffers break out the two BID subfields: bridge priority and bridge MAC address.
- *Root Path Cost*: This value is the cumulative cost of all links leading to the root bridge.

## Configuration BPDUs Format (2 of 3)

Octets	
Protocol ID	2
Protocol Version	1
BPDUs Type	1
Flags	1
Root ID	8
Root Path Cost	4
Bridge ID	8
Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2

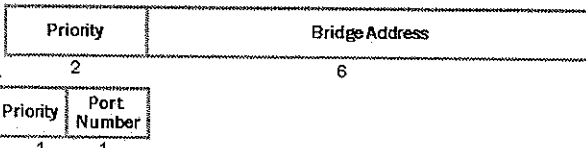
### Configuration BPDUs fields (contd.):

#### • Bridge ID

- Bridge Priority—The priority of becoming the root bridge, the designated bridge, or both (lower is better)
- Bridge Address—The unique MAC address of the bridge itself

#### • Port ID

- Port Priority—Used as the tiebreaker to determine the designated port, the root port, or both for a LAN (lower is better)
- Port Number—The ID of the transmitting port



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 6-10

## Configuration BPDUs Format: Part 2

The following list is a continuation of the explanation of BPDUs fields:

- **BID:** This value is the BID of the bridge that created the current BPDUs. This field is the same for all BPDUs sent by a single switch (for a single VLAN), but it differs between switches. The BID is a combination of the sender bridge's priority to become root or designated bridge and the bridge address (a unique MAC address for the bridge.)
- **Port ID:** This field contains a unique value for every port. This value is a combination of the outbound port's priority and a unique value to represent the port. The default port priority is 128 for every interface on an MX Series 3D Universal Edge Router. The switch automatically generates the port number and you cannot configure it. For example, ge-1/0/0 contains the value 128:41, whereas ge-1/0/1 contains the value 128:42.



## Configuration BPDU Format (3 of 3)

Octets	
Protocol ID	2
Protocol Version	1
BPDU Type	1
Flags	1
Root ID	8
Root Path Cost	4
Bridge ID	8
Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2

### Configuration BPDU fields (contd.):

- **Message Age**—The age of the configuration message
  - The time since the generation of the configuration BPDU
  - Enables a bridge to discard information that exceeds the Max Age
- **Max Age**—A timeout value to remove aging BPDU information (set by the root)
- **Hello Time**—The time interval between the generation of configuration BPDUs by designated ports
- **Forward Delay**—The delay time before transferring the state of a port to forwarding (set by the root)

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-11

## Configuration BPDU Format: Part 3

The following list is a continuation of the explanation of BPDU fields:

- **Message Age:** This field records the time since the root bridge originally generated the information from which the current BPDU is derived.
- **Max Age:** This value is the maximum time that a BPDU is saved. It also influences the bridge table aging timer during the topology change notification process.
- **Hello Time:** This value is the time between periodic configuration BPDUs.
- **Forward Delay:** This value is the time a bridge spends in the listening and learning states. It also influences timers during the topology change notification process.

## TCN BPDU Format

	Octets
Protocol ID	2
Protocol Version	1
BPDU Type	1

### ■ TCN BPDU fields

- Protocol ID—0x0000 (STP)
- Protocol Version—0x00 (IEEE 802.1D-1998)
- BPDU Type—0x80 (TCN BPDU)

© 2010 Juniper Networks, Inc. All rights reserved.

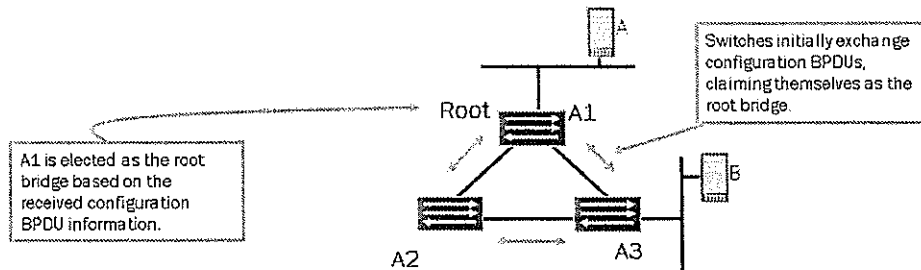
**JUNIPER** Worldwide Education Services

www.juniper.net | 6-12

### TCN BPDU

The slide shows the format of the TCN BPDU. We describe its usage later in this chapter.

## Building a Spanning Tree (1 of 3)



- **Switches exchange configuration BPDUs:**
  - They do not flood—instead each bridge uses information in the received BPDUs to generate its own
- **The root bridge is elected based on BPDU information:**
  - Criterion for election is the bridge ID
    - The election process reviews priority first—lowest priority wins
    - If the priority values are the same, bridge addresses (MAC) are compared—the lowest identifier wins

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-13

### Exchange of BPDUs

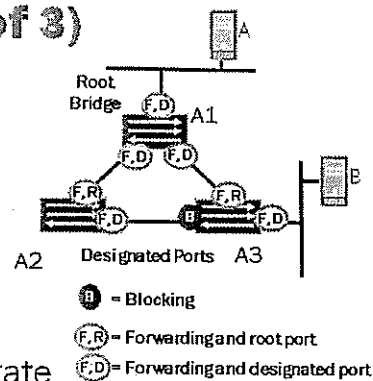
All switches participating in a switched network exchange BPDUs with each other. Through the exchanged BPDUs, neighboring switches become familiar with each other and learn the information necessary to select a root bridge. Each bridge creates its own configuration BPDUs based upon the BPDUs that it receives from neighboring routers. Non-STP bridges simply flood BPDUs as they would any multicast Ethernet frames.

### Root Bridge Election

STP elects the root bridge device based on the BID, which actually consists of two distinct elements: a configurable priority value and a unique device identifier, which is the system MAC address. Each bridge reviews the priority values first and determines the root bridge. If the priority value of one device is lower than the priority value of all other devices, that device receives the root bridge election. If the priority values are equal for all devices, STP evaluates the bridge addresses (MAC), and each bridge elects the device with the lowest MAC address as the root bridge.

## Building a Spanning Tree (2 of 3)

- The least-cost path calculation to the root bridge determines the port role; the port role determines the port state:



- Ports on the root bridge assume the designated port role and forwarding state
- Root ports on switches are placed in the forwarding state
  - The root bridge will have no root ports
- Designated ports on designated bridges are placed in the forwarding state
- All other ports are placed in the blocking state

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-14

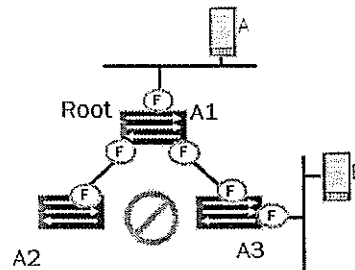
### Port Role and State Determination

Once root bridge election occurs, all nonroot devices perform a least-cost path calculation to the root bridge. The results of these calculations determine the role of the switch ports. The role of the individual switch ports determines the port state.

All switch ports belonging to the root bridge assume the designated port role and forwarding state. Each nonroot switch determines a root port, which is the port closest to the root bridge, based on its least-cost path calculation to the root bridge. Each interface has an associated cost that is based on the configured speed. An interface operating at 10 Mbps assumes a cost of 2,000,000, an interface operating at 100 Mbps assumes a cost of 200,000, an interface operating at 1 Gbps assumes a cost of 20,000, and an interface operating at 10 Gbps assumes a cost of 2000. If a switch has two equal-cost paths to the root bridge, the switch port with the lower port ID is selected as the root port. The root port for each nonroot switch is placed in the forwarding state.

STP selects a designated bridge on each LAN segment. This selection process is also based on the least-cost path calculation from each switch to the root bridge. Once the designated bridge selection occurs, its port, which connects to the LAN segment, is chosen as the designated port. If the designated bridge has multiple ports connected to the LAN segment, the port with the lowest ID participating on that LAN segment is selected as the designated port. All designated ports assume the forwarding state. All ports not selected as a root port or as a designated port assume the blocking state. While in blocked state, the ports do not send any BPDUs. However, they listen for BPDUs.

## Building a Spanning Tree (3 of 3)



- The tree is fully converged
  - All traffic from Host A to Host B flows through the root bridge (Switch A1)

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-15

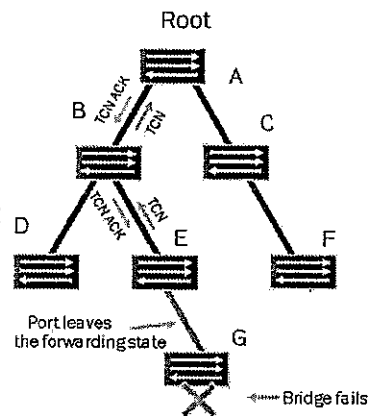
### Full Tree Convergence

Once each bridge determines the role and state for all switch ports, the tree is considered fully converged. The convergence delay can take up to 50 seconds when the default forwarding delay (15 seconds) and max age timer (20 seconds) values are in effect. The formula to calculate the convergence delay for STP is  $2 \times \text{forwarding delay} + \text{maximum age}$ . In the example shown on the slide, all traffic passing between Host A and Host B transits the root bridge (Switch A1).

## Reconvergence Example (1 of 2)

### Steps:

- Bridge G fails
- Bridge E's port leaves the forwarding state
- Bridge E sends a TCN
  - The TCN always travels out the root port: It continues every 2 seconds until the root port from B receives the TCN ACK in the form of a configuration BPDU
- Bridge B send a TCN ACK
- Bridge B sends a TCN out of the root port
- Bridge A sends a TCN ACK



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-16

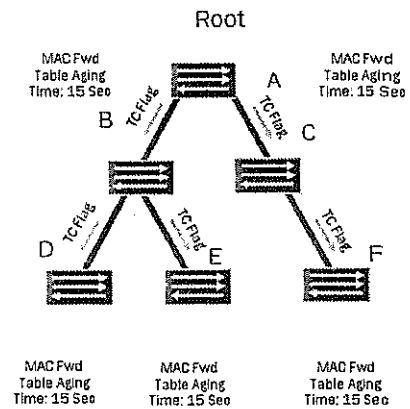
### Reconvergence Example: Part 1

The slide shows the first several steps that occur during a failure and reconvergence scenario.

## Reconvergence Example (2 of 2)

### • Steps (contd.):

7. The root bridge sets the topology change flag and sends an updated configuration BPDU
8. Bridges B and C relay the topology change flag to downstream switches
9. All nonroot bridges change the MAC address forwarding table aging timer to equal the forwarding delay time (default: 15 seconds)



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-17

### Reconvergence Example: Part 2

The slide shows the remainder of the steps involved in a failure and reconvergence scenario. Once the nonroot bridges change their MAC address forwarding table aging timer to the shortened interval and wait that period of time (15 seconds by default), they then delete all entries from the MAC table that were not refreshed within that time frame. All deleted entries must then be learned once again through the normal learning process.

## Agenda: Spanning Tree Protocols

- Overview of STP
- Overview of RSTP
- Overview of MSTP
- Overview of VSTP
- Configuring and Monitoring Spanning Tree Protocols
- Understanding BPDU, Loop, and Root Protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | E-18

### Overview of RSTP

The slide highlights the topic we discuss next.



## Rapid STP

- First defined in IEEE 802.1w and later incorporated into IEEE 802.1D-2004
- Convergence improvements:
  - Point-to-point link designation
    - Allows for rapid recovery from failures because a new root port or designated port can transition to forwarding without waiting for the protocol timers to expire
  - Edge port designation
    - A port that connects to a LAN with no other bridges attached
    - It is always in the forwarding state
  - Direct and indirect link failure and recovery

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-19

### RSTP Defined

RSTP was originally defined in the IEEE 802.1w draft and was later incorporated into the IEEE 802.1D-2004 specification. RSTP introduces a number of improvements to STP while performing the same basic function.

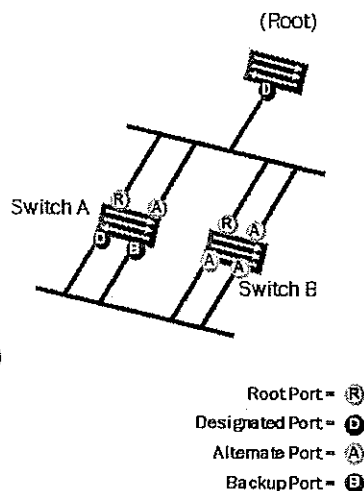
### RSTP Convergence Improvements

RSTP provides better reconvergence time than the original STP. RSTP identifies certain links as point-to-point. When a point-to-point link fails, the alternate link can transition to the forwarding state without waiting for any protocol timers to expire. RSTP provides fast network convergence when a topology change occurs and it greatly decreases the state transition time compared to STP. To aid in the improved convergence, RSTP uses additional features and functionality, such as edge port definitions and rapid direct and indirect link failure detection and recovery. We examine these features in more detail later in this chapter.

## RSTP Port Roles

### ■ RSTP introduces new port roles:

- Alternate port:
  - Provides an alternate path to the root bridge (essentially a backup for the root port of a switch)
  - Blocks traffic while receiving superior BPDUs from a neighboring switch
- Backup port:
  - Provides a redundant path to a segment (on designated switches only)
  - Blocks traffic while a more preferred port functions as the designated port



### ■ RSTP continues to use the root and designated port roles

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-20

## RSTP Introduces New Port Roles

RSTP introduces the alternate and backup port roles. An alternate port is a switch port that has an alternate—generally higher-cost—path to the root bridge. In the event that the root port fails, the alternate port assumes the role of the root port and is placed in the forwarding state. Alternate ports are placed in the discarding state but receive superior BPDUs from neighboring switches. Alternate ports are found on switches participating in a shared LAN segment for which they are not functioning as the designated bridge.

When a designated bridge has multiple ports connected to a shared LAN segment, it selects one of those ports as the designated port. The designated port is typically the port with the lower port ID. RSTP considers all other ports on the designated switch that connects to that same shared LAN segment as backup ports. In the event that the designated port is unable to perform its role, one of the backup ports assumes the designated port role upon successful negotiation and it is placed in the forwarding state.

Backup ports are placed in the discarding state. While in the discarding state, backup ports receive superior BPDUs from the designated port.

## Continued Use of Root and Designated Ports

RSTP continues to use the root and designated port roles. Only ports selected for the root port or designated port role participate in the active topology. We described the purpose of the root port and designated ports previously in this chapter.

## RSTP Port States

- RSTP (802.1D-2004) uses fewer states than STP (802.1D-1998) but has the same functionality

802.1D-1998 STP	802.1D-2004 RSTP	
Disabled	Discarding	Alternate, Backup, and Disabled Ports
Blocking		
Listening		
Learning	Learning	Root, Designated, and Edge Ports
Forwarding	Forwarding	

© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-21

## RSTP Port States

RSTP uses fewer port states than STP. The three possible port states found in RSTP are *discarding*, *learning*, and *forwarding*. Any administratively disabled port excluded from the active topology through configuration, or dynamically excluded from forwarding and learning, is placed in the discarding state. Ports that are actively learning but not currently forwarding are in the learning state, whereas ports that are both learning and forwarding frames simultaneously are in the forwarding state. As the slide indicates, only those ports selected as root ports and designated ports use the forwarding state.

## RST BPDUs

### ■ RST BPDUs:

- Act as keepalives
  - RSTP-designated ports send Configuration BPDUs every hello time (default of 2 seconds)
- Provide faster failure detection
  - If a neighboring bridge receives no BPDU within 3 times the hello interval ( $3 \times 2 = 6$  seconds), connectivity to the neighbor is faulty

© 2007 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-22

## RSTP BPDUs

As previously mentioned, STP uses BPDUs to elect a root bridge, identify root ports for each switch, identify designated ports for each physical LAN segment, prune specific redundant links to create a loop-free tree topology, and report and acknowledge topology changes. RSTP configuration BPDUs also function as keepalives. All RSTP bridges send configuration BPDUs every 2 seconds by default. You can alter this value, if necessary.

By monitoring neighboring switches through the use of BPDUs, RSTP can detect failures of network components much more quickly than STP can. If a neighboring bridge receives no BPDU within three times the hello interval, it assumes connectivity is faulty and updates the tree. By default, it detects failures within 6 seconds when using RSTP, whereas it might take up to 50 seconds when using STP.

On MX Series devices, Ethernet interfaces operating in full-duplex mode are considered point-to-point links. When a failure occurs, a switch port operating as a point-to-point link can become a new root port or designated port and transition to the forwarding state without waiting for the timer to expire. Switch ports operating in half-duplex mode are considered to be shared (or LAN) links and must wait for the timer to expire before transitioning to the forwarding state.

## RST BPD Format

Protocol ID	2
Protocol Version	1
BPD Type	1
Flags	1
Root ID	8
Root Path Cost	4
Bridge ID	8
Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2
Version 1 Length	2

### Small differences from STP BPDUs:

- Protocol Version—0x02 (IEEE 802.1D-2004)
- BPD Type—0x02 (RST BPD)
- Flags
  - Topology Change Acknowledgment Flag (Bit 8)
  - Agreement Flag (Bit 7)
  - Forwarding Flag (Bit 6)
  - Learning Flag (Bit 5)
  - Port Role (Bits 3 and 4)
  - Proposal Flag (Bit 2)
  - Topology Change Flag (Bit 1)
- Version 1 Length—0x0000

© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net 1-8-23

## Configuration BPD Differences

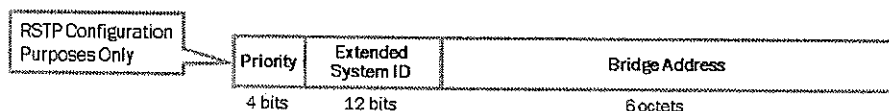
RSTP is backward compatible with STP. If a device configured for RSTP receives STP BPDUs, it reverts to STP. In a pure RSTP environment, a single type of the BPD exists named Rapid Spanning Tree BPD (RST BPD). RST BPDs use a similar format to the STP configuration BPDs. RSTP devices detect the type of BPD by looking at the protocol version and BPD type fields. The BPDs contain several new flags, as shown on the slide. The following is a brief description of the flags:

- TCN Acknowledgment: This flag is used when acknowledging STP TCNs;
- Agreement and Proposal: These flags are used to help quickly transition a new designated port to the forwarding state;
- Forwarding and Learning: These flags are used to advertise the state of the sending port;
- Port Role: This flag specifies the role of the sending port: 0 = Unknown, 1 = Alternate or Backup, 2 = Root, and 3 = Designated; and
- Topology Change: RSTP uses configuration BPDs with this bit set to notify other switches that the topology has changed.

RST BPDs contain a Version 1 Length field that is always set to 0x0000. This field allows for future extensions to RSTP.

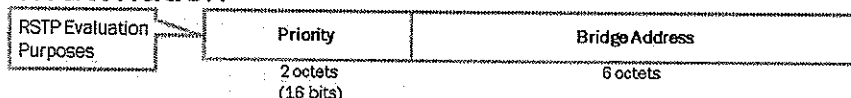
## RSTP Bridge ID

- Bridge priority is configured with a combination of priority and extended system ID
  - Evaluate as a single priority field in the election algorithms
  - MSTP only considers priority as a 4-bit field



```
[edit]
lab@switch# set protocols rstp bridge-priority ?
Possible completions:
<bridge-priority> Priority of the bridge (in increments of 4k - 0,4k,8k,...60k) (0..61440)
[edit]
lab@switch# set protocols rstp extended-system-id ?
Possible completions:
<extended-system-id> Extended system identifier (0..4095)
```

### ▪ Concatenation



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | G-24

## Bridge Priority Configuration

Over time, 16 bits was determined to be too big to represent a bridge's priority for becoming the root or designated bridge. With the advent of MSTP (covered later in this chapter), the older 16-bit priority field was broken into two separate fields: a 4-bit priority field and a 12-bit Extended System ID field. RSTP allows for the configuration of both values. MSTP automatically populates the extended system ID field with a VLAN ID.

## RSTP Bridge Priority Evaluation

Although priority and extended system ID are configured separately, RSTP evaluates a bridge's priority for the root and designated bridge election process by concatenating the two fields together into a single value. That is, it continues to be a 16-bit field during the election process.

## Transitioning to the Forwarding State

### ▪ Original STP (802.1D-1998):

- Takes 30 seconds before the ports start forwarding traffic after port enablement
  - 2x forwarding delay (listening + learning)

### ▪ RSTP (802.1D-2004):

- Uses a proposal-and-agreement handshake on point-to-point links instead of timers
  - Exceptions are alternate ports that immediately transition to root, and edge ports that immediately transition to the forwarding state
  - Nonedge-designated ports transition to the forwarding state once they receive explicit agreement

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-25

## STP Forwarding State Transition

With the original STP, as defined in 802.1D-1998, a port can take more than 30 seconds before it forwards user traffic. As a port is enabled, it must transition through the listening and learning states before graduating to the forwarding state. STP allows two times the forwarding delay (15 seconds by default) for this transition to occur.

## RSTP Forwarding State Transition

RSTP offers considerable improvements when transitioning to the forwarding state. RSTP converges faster because it uses a proposal-and-agreement handshake mechanism on point-to-point links instead of the timer-based process used by STP. On MX Series devices, network ports operating in full-duplex mode are considered point-to-point links, whereas network ports operating in half-duplex mode are considered shared (LAN) links.

Root ports and edge ports transition to the forwarding state immediately without exchanging messages with other switches. Edge ports are ports that have direct connections to end stations. Because these connections cannot create loops, they are placed in the forwarding state without any delay. If a switch port does not receive BPDUs from the connecting device, it automatically assumes the role of an edge port. When a switch receives configuration messages on a switch port that is configured to be an edge port, it immediately changes the port to a normal spanning-tree port (nonedge port).

Nonedge-designated ports transition to the forwarding state only after receipt of an explicit agreement from the attached switch.

## Topology Change Reconvergence

- Topology changes occur only when nonedge ports transition to the forwarding state:
  - Port transitions to the discarding state no longer trigger the STP TCN/TCN Acknowledgment sequence
  - The initiator floods RSTP TCNs (RST BPDU with TCN flag set) out of all designated ports as well as out of the root port
  - Because of the received RSTP TCN, switches flush the majority of MAC addresses in the MAC address forwarding table
    - Switches do not flush MAC addresses learned from edge ports
    - Switches do not flush MAC addresses learned on the port receiving the TCN

©2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-26

### Topology Changes

When using STP, state transitions on any participating switch port cause a topology change to occur. RSTP reduces the number of topology changes and improves overall stability within the network by generating TCNs only when nonedge ports transition to the forwarding state. Nonedge ports are typically defined as ports that interconnect switches. Edge ports are typically defined as ports that connect a switch to end stations.

RSTP also provides improved network stability because it does not generate a TCN when a port transitions to the discarding state. With RSTP, TCNs are not generated when a port is administratively disabled, excluded from the active topology through configuration, or dynamically excluded from forwarding and learning.

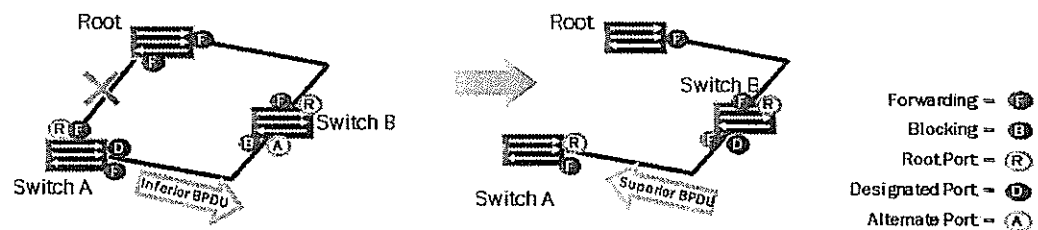
When a TCN is necessary and is generated, the initiating device floods all designated ports as well as the root port. Unlike traditional STP, neighboring switches that are not in the path of the initiator to the root bridge do not need to wait for this information from the root bridge. As the changes propagate throughout the network, the switches flush the majority of the MAC addresses located in their MAC address forwarding tables. The individual switches do not, however, flush MAC addresses learned from their locally configured edge ports.



## Indirect Link Failure

### ■ When an indirect link failure occurs:

- Switch A's root port fails—it assumes it is the new root
- Switch B receives inferior BPDUs from Switch A—it moves the alternate port to the designated port role
- Switch A receives superior BPDUs, knows it is not the root, and designates the port connecting to Switch B as the root port



© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-27

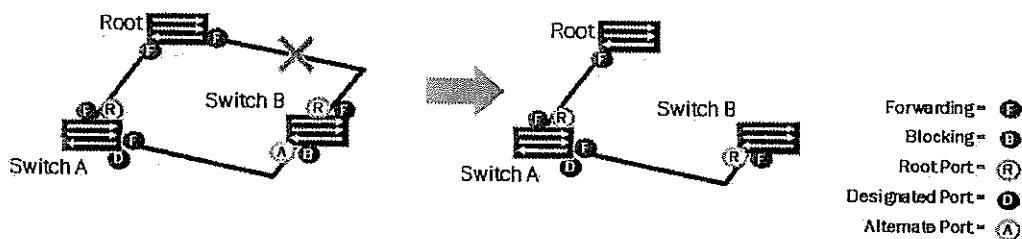
## Indirect Link Failure

RSTP performs rapid recovery for link failures. The slide illustrates a typical scenario for an indirect link failure.

## Direct Link Failure

### ▪ When a direct link failure occurs:

- The alternate port transitions to the forwarding state and assumes the new root port role following the failure of the old root port
- Switch B signals upstream switches to flush their MAC tables by sending RSTP TCNs out of the new root port
  - Upstream switches only flush MAC entries that they learned on active ports that did not receive the RSTP TCNs (except edge ports)



## Direct Link Failure

The slide illustrates a typical scenario in which a direct link failure occurs.

## RSTP Interoperability with STP

### ▪ STP and RSTP interoperability considerations:

- If a switch supports only the 802.1D-1998 STP protocol, it discards any RSTP BPDUs it receives
- If an RSTP-capable switch receives 802.1D-1998 BPDUs, it reverts to 802.1D-1998 STP mode on the receiving interface only
  - Uses STP BPDUs

© 2007 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-29

### Interoperability Considerations

Switches configured for STP and RSTP will interoperate with one another. However, you should keep a few basic considerations in mind. If a switch supports only STP and interconnects with a switch running RSTP, it will discard the RSTP BPDUs. The RSTP-capable switch, upon receiving STP BPDUs, reverts to STP mode, thus allowing interoperability between the two devices.

## Agenda: Spanning Tree Protocols

- Overview of STP
- Overview of RSTP
- Overview of MSTP
- Overview of VSTP
- Configuring and Monitoring Spanning Tree Protocols
- Understanding BPDU, Loop, and Root Protection

© 2010 Juniper Networks, Inc. All rights reserved.

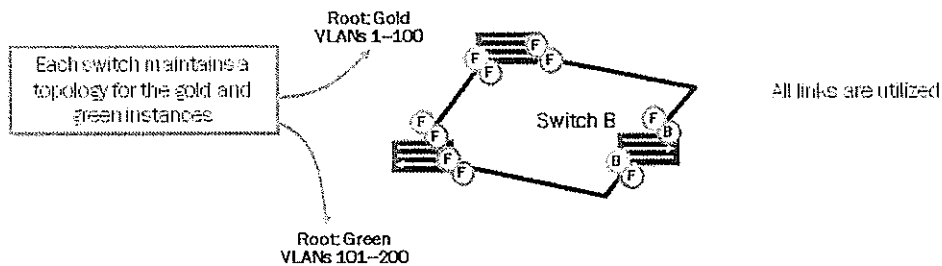
JUNIPER Worldwide Education Services

www.juniper.net | 6-30

### Overview of MSTP

The slide highlights the topic we discuss next.

## Multiple STP



- Originally defined in IEEE 802.1s; later merged into IEEE 802.1Q-2003
- Provides extensions to RSTP:
  - A separate topology tree for each MSTI
  - Resource friendly—maps VLANs to one or more instances; provides for load balancing over available links

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 6-31

### MSTP Defined

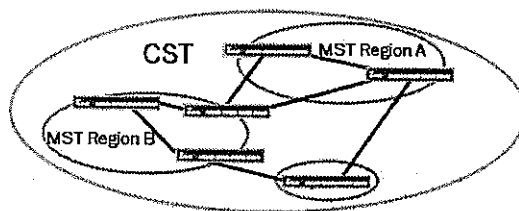
MSTP was originally defined in the IEEE 802.1s draft and later incorporated into the IEEE 802.1Q-2003 specification.

### MSTP Enhancements over RSTP

Although RSTP provides faster convergence than STP, it still does not make good use of all available paths within a redundant Layer 2 network. With RSTP, all traffic from all VLANs follows the same path as determined by the spanning tree; therefore, redundant paths are not utilized. MSTP overcomes this limitation through the use of multiple spanning tree instances (MSTIs). Each MSTI creates a separate topology tree and you can administratively map it to one or more VLANs. Allowing users to administratively map VLANs to MSTIs facilitates better load sharing across redundant links within a Layer 2 switching environment.

## Multiple Spanning Tree Region

- An MST region is a group of switches with the same region name, revision level, and VLAN-to-instance mapping
  - Max of 64 MSTIs per region
  - One regional root bridge per instance
- Backward compatible with STP and RSTP through a CST



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER  
NETWORKS

Worldwide Education Services

www.juniper.net | 6-32

### MST Region

MSTP allows switches to be logically grouped into manageable clusters, known as multiple spanning tree (MST) regions. An MST region is a group of switches that share the same region name, revision level, and VLAN-to-instance mapping parameters.

Each MST region supports up to 64 MSTIs. MSTP greatly reduces the number of BPDUs on a LAN by including the spanning tree information for all MSTIs in a single BPDU. MSTP encodes region information after the standard RSTP BPDU along with individual MSTI messages. The MSTI configuration messages convey spanning tree information for each instance.

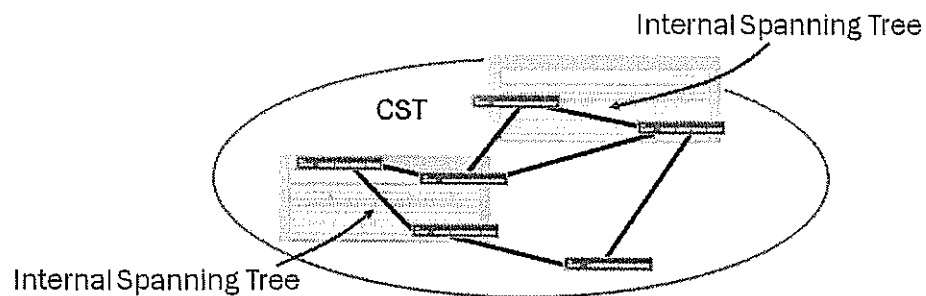
MSTP elects a regional root bridge for each MSTI. The regional root bridge is elected based on the configured bridge priority and calculates the spanning tree within its designated instance.

### MSTP Compatibility with STP and RSTP

Because MSTP encodes region information after the standard RSTP BPDU, a switch running RSTP interprets MSTP BPDUs as RSTP BPDUs. This behavior facilitates full compatibility between devices running MSTP and devices running STP or RSTP. All RSTP switches outside of an MST region view the MST region as a single RSTP switch. The common spanning tree (CST), which interconnects all MST regions as well as STP devices not bound to a particular region, facilitates end-to-end paths within an MSTP environment.

## Common and Internal Spanning Tree

- The CST interconnects MST regions:
  - One root bridge will be elected for the CST
  - Each MST region appears as a virtual bridge
    - Internal spanning tree extends CST into regions



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

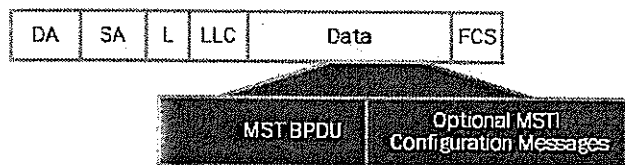
www.juniper.net | 6-33

### Common and Internal Spanning Tree

All MSTP environments contain a CST, which is used to interconnect individual MST regions and independent STP devices. All bridges in the CST elect a single root bridge. The root bridge is responsible for the path calculation for the CST. As illustrated on the slide, bridges outside of the MST region treat each MST region as a virtual bridge, regardless of the actual number of devices participating in each MST region.

The common and internal spanning tree (CIST) is a single topology that connects all switches (RSTP and MSTP devices) through an active topology. The CIST includes a single spanning tree as calculated by RSTP together with the logical continuation of connectivity through MST regions. MSTP calculates the CIST and the CIST ensures connectivity between LANs and devices within a bridged network.

## MST BPDU Format (1 of 3)



### ■ Ethernet frame:

- Source Address—The outgoing port of the originating switch
- Destination Address—The bridge group address (01:30:C2:00:00:00)
- Length
- LLC Header
  - DSAP and SSAP = 0x42 (bridge STP)

### MST BPDU Format: Part 1

The slide shows that MSTP uses the same Ethernet frame as STP and RSTP. However, the BPDU information in the data field is different. The next few slides discuss the MST BPDU information and the optional MSTI configuration messages.



## MST BPDUs Format (2 of 3)

	Octets
Protocol ID	2
Protocol Version	1
BPDUs Type	1
CIST Flags	1
CIST Root ID	8
CIST External Path Cost	4
CIST Regional Root ID	8
CIST Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2
Version 1 Length = 0	2
...	

■ The MST BPDUs fields and format listed on this slide are what allow MSTP to be compatible with RSTP and STP BPDUs

- Switches that are external to an MSTP region use only this information in their spanning-tree calculation
  - This information is used to build the CST
  - Essentially, RSTP is used to interconnect MST regions or RSTP-only bridges

More on the next slide

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

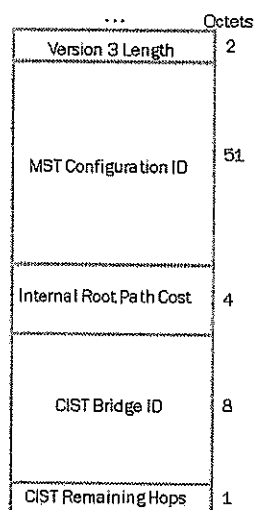
Worldwide Education Services

www.juniper.net | 6-35

### MST BPDUs Format: Part 2

The first 13 fields in the MST BPDUs contain similar information to what you would find in an RSTP BPDUs. In fact, an RSTP-speaking switch evaluates these fields in the same manner as it would any other RSTP PDU. To the outside world (other MSTP regions or standalone RSTP speakers), these fields are a representation of the virtual bridge that is an individual MSTP region. This information is used to build the CST.

## MST BPDU Format (3 of 3)



■ The MST BPDU fields listed on this slide, combined with others, allow each MST region to build an internal spanning tree:

- The CST between regions, combined with the internal spanning trees built within regions, result in a single CIST between all bridges
- By default, all traffic on all VLANs within a region will follow the internal spanning tree
  - MSTI configuration allows for traffic to follow a different path than the internal spanning tree

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 6-38

### MST BPDU Format: Part 3

Each MSTP region builds a spanning tree for the region, referred to as an internal spanning tree, based upon the BPDU fields on this slide as well as some of the fields on the previous slide (CIST Port ID, CIST Regional Root ID, and so forth). For a switch to participate in a region's internal spanning tree and use the information in this portion of the BPDU, it must be configured with the same configuration ID. Therefore, all switches in the same region must be configured with the same configuration ID. This approach to configuration ensures that when MSTP switches outside of the local MSTP region receive MSTP BPDUs, those switches will evaluate only the CST-related information (previous slide). Once the internal spanning tree is built, by default, all traffic on all VLANs will follow it.

## MSTI Configuration Messages

MSTI Flags	1
MSTI Root ID	8
MSTI Root Path Cost	4
MSTI Bridge Priority	2
MSTI Port Priority	2
MSTI Remaining Hops	2

■ MSTI configuration messages allow for more spanning trees to be built within a region:

- Each switch participating in the MSTI will go through the process of electing a root bridge, root ports, designated ports, and so forth for the MSTI
  - The local configuration on a switch determines which VLAN IDs belong to which MSTI
  - The configuration of each switch in a region should use the same VLAN ID to MSTI mappings

## MSTI Configuration Methods

Without the use of MSTI configuration methods, traffic for all VLANs within a region flows along the path of the internal spanning tree. To override this behavior and allow some VLANs to take one path through the region and let others take other paths (64 paths are possible for each region), you must configure MSTIs as part of the router MSTI configuration. The information carried in the MSTI configuration messages allows each switch to elect root bridges, root ports, designated ports, designated bridges, and so forth for each MSTI. Each MSTI will have one or more VLANs associated with them. One VLAN cannot be in more than one MSTI. Notice that the MSTI messages do not carry VLAN ID information. The VLAN-to-MSTI mappings are configured locally on each switch and each switch configuration should use the same mappings.

## Agenda: Spanning Tree Protocols

- Overview of STP
- Overview of RSTP
- Overview of MSTP
- Overview of VSTP
- Configuring and Monitoring Spanning Tree Protocols
- Understanding BPDU, Loop, and Root Protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

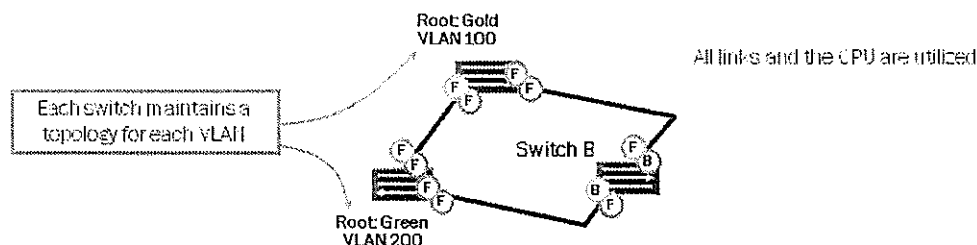
Worldwide Education Services

www.juniper.net | 6-38

### Overview of VSTP

The slide highlights the topic we discuss next.

## VLAN STP



- VSTP maintains a separate spanning-tree instance for each VLAN:
  - Different VLANs can use different spanning-tree topologies
  - Supports up to 4094 different spanning-tree topologies
  - Compatible with Per-VLAN Spanning Tree Plus and Rapid-PVST+ protocols
  - As you add VLANs, they consume more CPU resources

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-39

## VSTP

VSTP allows for spanning trees to be calculated for each VLAN. VSTP supports up to 4094 separate paths through the network. VSTP is a nonstandard protocol, yet it is compatible with Cisco's Per-VLAN Spanning Tree Plus (PVST+) and Rapid Per-VLAN Spanning Tree Plus (Rapid-PVST+) protocols. As you add more VLANs to the network, they consume more CPU resources. For example, imagine a network that is configured for 4000 VLANs. If VSTP is in use, each switch must participate in the election of 4000 root bridges, 4000 root ports, and so forth.

## VSTP Versus RSTP

- VSTP is most similar to RSTP:
  - VSTP uses the same terminology as RSTP and all of the terms have the same meaning
    - Root bridge
    - Designated bridge
    - Root port
    - Designated port
    - Bridge ID
    - Root path cost
    - Port cost
    - Port ID
  - Allows forcing the version to STP

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-40

## VSTP Versus RSTP

VSTP uses the same terminology as the other spanning tree protocols. It is most similar to RSTP. VSTP also provides for the ability to force the version to STP.

## VSTP BPDUs—Ethernet Frame Format

### ▪ Ethernet frame:

DA	SA	VLAN TAG	L	LLC	SNAP	BPDU	FCS
----	----	-------------	---	-----	------	------	-----

- Source Address—The outgoing port of the originating switch
- Destination Address—The bridge group address (01:00:0C:CC:CC:CD)
- IEEE 802.1Q tag
- Length
- LLC Header
  - DSAP and SSAP = 0xAA (Snap Header)
- SNAP Header
  - Protocol ID = 0x010B (Shared Spanning Tree or PVST+)
- BPDU format is the same as RSTP format with an added type, length, and value that advertises the same VLAN ID found in the VLAN tag

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net J-641

### VSTP Frame Format

The slide shows the format of the VSTP BPDU. Notice that it uses a special destination MAC address and it is also carried in an IEEE 802.1Q tag.

## Spanning Tree Protocols Summary

### ■ Spanning tree protocols summary:

- STP (802.1D-1998) is used in Layer 2 networks to prevent logical loops
  - Automated—user selects root switch and STP does the rest
  - STP is slow to converge and can be difficult to troubleshoot
- RSTP (802.1D-2004) reduces link-convergence time to subseconds on point-to-point links
- STP and RSTP support a single STP instance
  - Lacks load-balancing mechanism: creates underutilized links
- MSTP (802.1Q-2005) supports up to 64 instances
  - Overcomes the shortcomings of a single spanning tree
- VSTP (proprietary) supports up to 4094 instances
  - Compatible with proprietary protocols from other vendors

© 2012 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-42

### Spanning Tree Protocols Summary

The slide provides a quick overview along with the highlights of STP, RSTP, MSTP, and VSTP.



## Agenda: Spanning Tree Protocols

- Overview of STP
- Overview of RSTP
- Overview of MSTP
- Overview of VSTP
- Configuring and Monitoring Spanning Tree Protocols
- Understanding BPDU, Loop, and Root Protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-43

### Configuring and Monitoring Spanning Tree Protocols

The slide highlights the topic we discuss next.

## Configuring STP

```
[edit protocols rstp]
user@switch# set ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except  Don't inherit configuration data from these groups
backup-bridge-priority Priority of the bridge (in increments of 4k - 4k,8k,..60k) (4096..61440)
bpdu-block-on-edge     Block BPDU on all interfaces configured as edge (BPDU Protect)
bpdu-destination-mac-address Destination MAC address in the spanning tree BPDUs
bridge-priority         Priority of the bridge (in increments of 4k - 0,4k,8k,..60k) (0..61440)
disable                Disable STP
extended-system-id     Extended system identifier (0..4095)
force-version         Force protocol version
forward-delay          Time spent in listening or learning state (4..30 seconds)
hello-time             Time interval between configuration BPDUs (1..10 seconds)
> interface            Interface options
max-age               Maximum age of received protocol bpdu (6..40 seconds)
priority-hold-time     Hold time before switching to primary priority when core domain becomes up (1..255 seconds)
> system-id           System ID to IP mapping
> traceoptions         Tracing options for debugging protocol operation
vpls-flush-on-topology-change Enable VPLS MAC flush on root protected CE interface receiving topology change
```

```
[edit protocols rstp]
user@switch# show
bridge-priority 32k;
max-age 20;
hello-time 2;
forward-delay 15;
interface ge-1/0/0;
force-version stp;
```

Configuration example illustrates default STP settings

You must enable STP on at least one interface

### Configuring STP

The slide shows some STP configuration options along with a basic STP configuration. MX Series devices use a version of STP based on IEEE 802.1D-2004, with a forced protocol version of 0, running RSTP in STP mode. Because of this implementation, you can define RSTP configuration options, such as `hello-time`, under the `[edit protocols rstp]` configuration hierarchy. To specify running RSTP in STP mode, you simply need to specify **force-version stp**.

## Configuring RSTP

```
[edit protocols rstp]
user@switch# show
bridge-priority 32k;
max-age 20;
hello-time 2;
forward-delay 15;
interface ge-1/0/1 {
  priority 128;
  mode point-to-point;
}
interface ge-1/0/3 {
  cost 20000;
  mode shared;
}
interface ge-1/3/6 {
  edge;
}
extended-system-id 0;
```

Diagram illustrating the configuration of RSTP settings and their default values:

- Default RSTP settings:** bridge-priority 32k; max-age 20; hello-time 2; forward-delay 15;
- Default port priority value (used to determine the designated port used by the designated bridge):** priority 128;
- Default interface mode for interfaces operating in full-duplex mode:** mode point-to-point;
- Default cost value for interfaces operating at 1 Gbps:** cost 20000;
- Default interface mode for interfaces operating in half-duplex mode:** mode shared;
- Default value for interfaces that do not connect to STP-enabled devices:** edge;
- Default value for the Extended System ID:** extended-system-id 0;

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-45

### Configuring RSTP

The sample RSTP configuration provided on the slide shows the typical configuration structure along with various settings.

## Monitoring STP and RSTP (1 of 2)

```
User@switch> show spanning-tree ?
```

Possible completions:

bridge	Show STP bridge parameters
interface	Show STP interface parameters
mstp	Show Multiple Spanning Tree Protocol information
statistics	Show STP statistics

```
User@switch> show spanning-tree bridge
```

STP bridge parameters

Routing instance name

Context ID

Enabled protocol

Root ID

Root cost

Root port

Hello time

Maximum age

Forward delay

Message age

Number of topology changes

Time since last topology change

Local parameters

Bridge ID

Extended system ID

: GLOBAL

: 0

: RSTP

: 32768.00:21:59:01:97:d0

: 20000

: ge-1/3/6

: 2 seconds

: 20 seconds

: 15 seconds

: 1

: 7

: 4170 seconds

: 32768.00:22:83:75:cf:d0

: 0

Root Bridge's ID

Cumulative Cost  
to Root Bridge

Root Port

Local Device's Bridge ID

### Monitoring Spanning Tree Operation: Part 1

This slide and the next illustrate some common operational mode commands used to monitor the operation of STP and RSTP.

## Monitoring STP and RSTP (2 of 2)

```
user@switch> show spanning-tree interface
```

Spanning tree interface parameters for instance 0

Interface	Port ID	Designated port ID	Designated bridge ID	Port Cost	State	Role
ge-1/3/8	128:1	128:1	32768.00228375cfd0	20000	FWD	DESG
ge-1/3/9	128:2	128:2	32768.00228375cfd0	20000	FWD	DESG
ge-1/3/6	128:6	128:77	32768.0021590197d0	20000	FWD	ROOT
ge-1/3/7	128:7	128:78	32768.0021590197d0	20000	BLK	ALT

```
user@switch> show spanning-tree statistics interface
```

Interface	BPDUs sent	BPDUs received	Next BPDU transmission
ge-1/3/8	5915	1418	1
ge-1/3/9	5179	2041	1
ge-1/3/6	4103	4267	0
ge-1/3/7	3979	4393	0

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-47

### Monitoring Spanning Tree Operation: Part 2

The slide shows typical output for the **show spanning-tree interface** and **show spanning-tree statistics interface** commands.

## Configuring MSTP

```
[edit protocols mstp]
```

```
user@switch# show
```

```
configuration-name region1;
```

```
revision-level 1;
```

```
interface ge-1/0/0;
```

```
interface ge-1/0/2;
```

```
interface ge-1/0/4;
```

```
interface ge-1/0/5;
```

```
interface ge-1/0/6;
```

```
interface ge-1/0/7;
```

```
msti 1 {
```

```
    bridge-priority 4k;
```

```
    vlan 100-199;
```

```
}
```

```
msti 2 {
```

```
    bridge-priority 8k;
```

```
    vlan 200-299;
```

```
}
```

User-defined configuration-name and revision-level (must match on all switches within the same region)

MSTP instances defined with individual bridge-priority values and VLAN ranges

### Configuring MSTP

The sample MSTP configuration provided on the slide shows the typical configuration structure along with various settings.

## Monitoring MSTP (1 of 3)

```
user@switch> show spanning-tree ?
```

Possible completions:

bridge	Show STP bridge parameters
interface	Show STP interface parameters
mstp	Show Multiple Spanning Tree Protocol information
statistics	Show STP statistics

```
user@switch> show spanning-tree mstp configuration
```

MSTP configuration information

```
Context identifier      : 0
Region name            : region1
Revision               : 1
Configuration digest    : 0xb0c8664d7263bcd9aea9f884439230e4
```

Values must match for all switches within a common MST region

MSTI	Member VLANs
0	0-99, 300-4094
1	100-199
2	200-299

Configuration digest is determined by the contents of MSTI to VID table

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 6-49

### Monitoring MSTP Operation: Part 1

This slide and the next two slides illustrate some common operational mode commands used to monitor MSTP. This slide highlights the **show spanning-tree mstp configuration** command, which you can use to verify MSTP configuration parameters including region, revision, and assigned MSTI parameters.

## Monitoring MSTP (2 of 3)

```
user@switch> show spanning-tree interface
```

Spanning tree interface parameters for **instance 0**

Interfaces and associated details are listed by instance

Interface	Port ID	Designated port ID	Designated bridge ID	Port Cost	State	Role
...						
ge-1/0/7	128:48	128:58	32768.0021590197d2	20000	BLK	ALT

Spanning tree interface parameters for **instance 1**

Interface	Port ID	Designated port ID	Designated bridge ID	Port Cost	State	Role
...						
ge-1/0/7	128:48	128:48	4097.0021590197d3	20000	FWD	DESG

Spanning tree interface parameters for **instance 2**

Interface	Port ID	Designated port ID	Designated bridge ID	Port Cost	State	Role
...						
ge-1/0/7	128:48	128:48	8194.0021590197d3	20000	FWD	DESG

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 6-50

### Monitoring MSTP Operation: Part 2

The slide highlights the use of the **show spanning-tree interface** command, which you use to verify the MSTP interface status and role assignment along with various other details.



## Monitoring MSTP (3 of 3)

```
user@switch> show spanning-tree bridge
STP bridge parameters
Routing instance name      : GLOBAL
Context ID                 : 0
Enabled protocol           : MSTP
```

```
STP bridge parameters for CIST
Root ID                    : 32768.00:21:59:01:97:d1
Root cost                  : 0
Root port                  : ge-1/0/4
CIST regional root         : 32768.00:21:59:01:97:d1
...
```

```
STP bridge parameters for MSTI 1
MSTI regional root        : 4097.00:21:59:01:97:d3
Hello time                 : 2 seconds
Maximum age                : 20 seconds
Forward delay              : 15 seconds
Number of topology changes : 2
Time since last topology change : 841 seconds
Local parameters
  Bridge ID                : 4097.00:21:59:01:97:d3
...
```

STP details are listed by instance

© 2004 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-51

### Monitoring MSTP Operation: Part 3

The slide highlights the **show spanning-tree bridge** command, which you use to display STP bridge parameters for the CIST and individual MSTIs.

## Configuring VSTP

```
[edit protocols vstp]
```

```
user@switch# show
```

```
interface ge-1/0/1;
```

```
interface ge-1/0/3;
```

```
interface ge-1/3/6;
```

```
interface ge-1/3/7;
```

```
vlan 100 {
```

```
    bridge-priority 60k;
```

```
    interface ge-1/0/1;
```

```
    interface ge-1/0/3;
```

```
    interface ge-1/3/6;
```

```
    interface ge-1/3/7;
```

```
}
```

```
vlan 200 {
```

```
    bridge-priority 8k;
```

```
    interface ge-1/0/1;
```

```
    interface ge-1/0/3;
```

```
    interface ge-1/3/6;
```

```
    interface ge-1/3/7;
```

```
}}
```

Specify interfaces that will run VSTP. Each must also be listed under a particular VLAN.

VLANs defined with individual bridge-priority values and interfaces

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 6-52

## Configuring VSTP

The sample VSTP configuration provided on the slide shows the typical configuration structure along with various settings.

## Monitoring VSTP (1 of 2)

```
user@switch> show spanning-tree interface
```

Interfaces and associated details are listed by VLAN ID

Spanning tree interface parameters for **VLAN 100**

Interface	Port ID	Designated port ID	Designated bridge ID	Port Cost	State	Role
ge-1/0/1	128:42	128:42	8292.002159ab8fd0	20000	FWD	ROOT
ge-1/0/3	128:44	128:44	8292.002159ab8fd0	20000	BLK	ALT
ge-1/3/6	128:77	128:46	8292.002159ab8fd0	20000	BLK	ALT
ge-1/3/7	128:78	128:46	8292.002159ab8fd0	20000	BLK	ALT

Spanning tree interface parameters for **VLAN 200**

Interface	Port ID	Designated port ID	Designated bridge ID	Port Cost	State	Role
ge-1/0/1	128:42	128:42	8392.0021590197d0	20000	BLK	DESG
ge-1/0/3	128:44	128:44	8392.0021590197d0	20000	BLK	DESG
ge-1/3/6	128:77	128:42	8392.0021590197d0	20000	BLK	BKUP
ge-1/3/7	128:78	128:42	8392.0021590197d0	20000	BLK	BKUP

© 2008 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-53

### Monitoring VSTP: Part 1

The slide highlights the use of the **show spanning-tree interface** command, which you use to verify the VSTP interface status and role assignment along with various other details.

## Monitoring VSTP (2 of 2)

```
user@switch> show spanning-tree bridge
```

```
STP bridge parameters
```

```
Routing instance name      : GLOBAL
Enabled protocol           : RSTP
```

```
STP bridge parameters for VLAN 100
```

```
Root ID                    : 8292.00:21:59:ab:8f:d0
Root cost                   : 20000
Root port                  : ge-1/0/1
Hello time                  : 2 seconds
Maximum age                 : 20 seconds
Forward delay               : 15 seconds
Message age                 : 1
Number of topology changes : 3
Time since last topology change : 141 seconds
Local parameters
  Bridge ID                 : 61540.00:21:59:01:97:d0
  Extended system ID        : 100
```

```
STP bridge parameters for VLAN 200
```

```
Root ID                    : 8392.00:21:59:01:97:d0
Hello time                  : 2 seconds
Maximum age                 : 20 seconds
Forward delay               : 15 seconds
...
```

STP details are  
listed by VLAN ID

### Monitoring VSTP: Part 2

The slide highlights the **show spanning-tree bridge** command, which you use to display STP bridge parameters for the individual VLANs.

## Agenda: Spanning Tree Protocols

- Overview of STP
- Overview of RSTP
- Overview of MSTP
- Overview of VSTP
- Configuring and Monitoring Spanning Tree Protocols
- Understanding BPDU, Loop, and Root Protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-55

### Understanding BPDU, Loop, and Root Protection

The slide highlights the topic we discuss next.

## Purpose for BPDU Protection

### ■ Problem:

- Bridge applications running on PCs or “personal” switches can generate BPDUs
- STP, RSTP, or MSTP running on a switch could detect those BPDUs and trigger spanning-tree miscalculations leading to network outages

### ■ Solution:

- Enable BPDU protection on Layer 2 interfaces connected to user devices or on interfaces on which no BPDUs are expected
- If a protected interface receives a BPDU, the bridge disables the interface and stops forwarding frames by transitioning to a blocking state

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-56

### Purpose for BPDU Protection—Problem

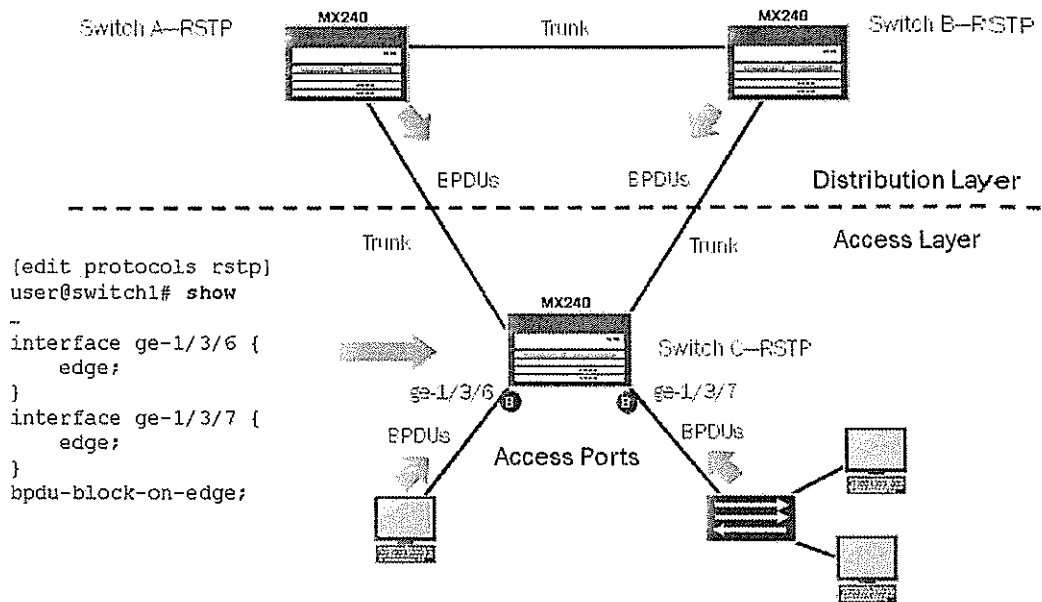
As mentioned previously in this chapter, the purpose for STP, RSTP, and MSTP is to prevent Layer 2 loops in the network. The exchange of BPDUs achieves a loop-free Layer 2 network. A user bridge application running on a PC or a user’s “personal” switch for connecting multiple end devices can also generate BPDUs. If these BPDUs are picked up by STP, RSTP, or MSTP applications running on the switch, they can trigger spanning-tree miscalculations, which in turn could cause loops, leading to network outages.

### Purpose for BPDU Protection—Solution

You can enable BPDU protection on switch interfaces on which no BPDUs are expected. If a protected interface receives BPDUs, the switch disables the interface and stops forwarding frames by transitioning to a blocking state.

You can configure BPDU protection on a switch with a spanning tree as well as on a switch that is not running STP.

## Configuring BPDU Protection (1 of 2)



© 2014 Juniper Networks, Inc. All rights reserved.

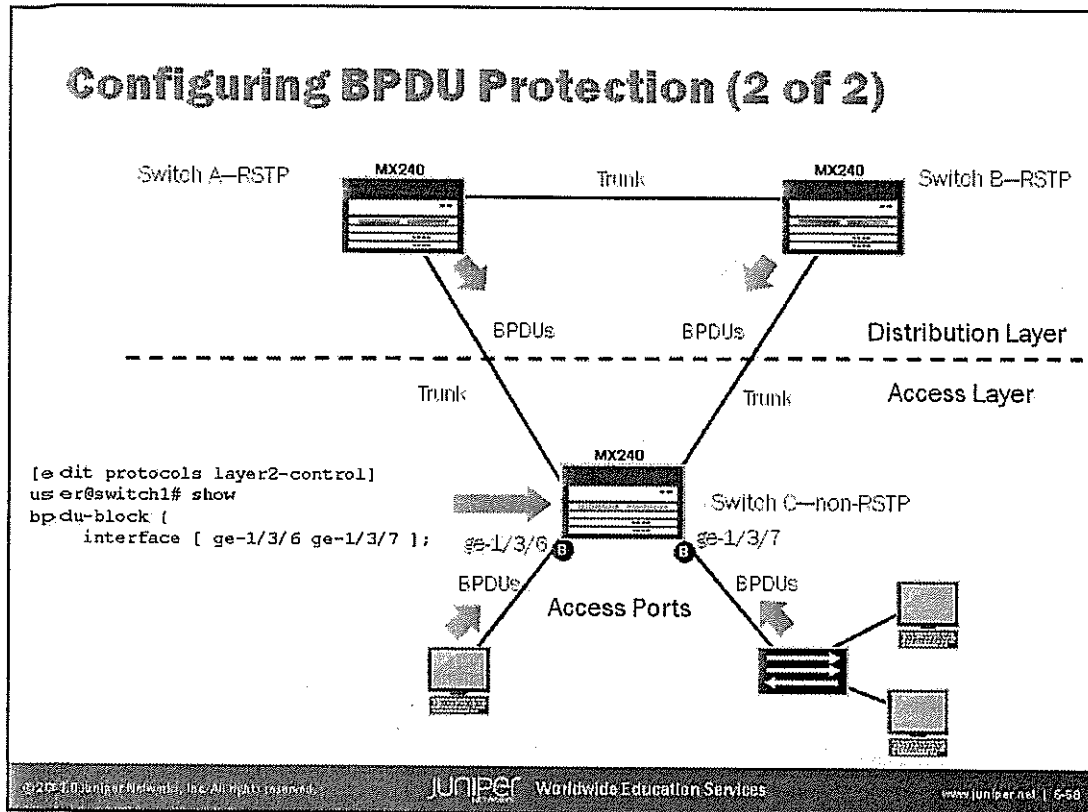
JUNIPER

Worldwide Education Services

www.juniper.net | 6-57

### Configuring BPDU Protection: Part 1

Consider a network in which two switches, Switch A and Switch B, are at the Distribution Layer. A third switch, Switch C, is at the Access Layer. Switch C connects to a PC and a user's unauthorized switch by means of the access ports, ge-1/3/6 and ge-1/3/7, respectively. You have configured Switch A, Switch B, and Switch C to use STP. You should configure BPDU protection on Switch C access ports. The slide illustrates the required configuration. When BPDU-protected interfaces receive BPDUs, the interfaces transition to a blocking state and stop forwarding frames.



### Configuring BPDUs Protection: Part 2

Now consider a network in which two switches, Switch A and Switch B, are at the Distribution Layer. A third switch, Switch C, is at the Access Layer. Switch C connects to a PC and a user's unauthorized switch by means of the access ports—ge-1/3/6 and ge-1/3/7, respectively. Switch C is not configured to participate in STP. However, you should protect its access ports—ge-1/3/6 and ge-1/3/7. The slide illustrates the required configuration. When BPDUs-protected interfaces receive BPDUs, the interfaces transition to a blocking state and stop forwarding frames.



## Verifying BPDU Protection Functionality

- To verify BPDU protection functionality:
  - Use the **show spanning-tree interface** command before and after enabling the BPDU protection feature on an STP-running switch
  - Use the **show 12-learning interface** command on a non-STP switch
  - Watch for state changes, role changes, or both in the output:
    - FWD state transitions to BLK
    - DESG role transitions to DIS (loop inconsistent)
    - unblocked transitions to blocked
- To unblock the interface:
  - Use the **clear error bpdu interface** operational mode command

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-59

### Verifying of BPDU Protection Functionality

To confirm that the configuration is working properly on the STP-running switch, use the **show spanning-tree interface** operational mode command. To confirm that the configuration is working properly on the switch that is not running STP, you should observe the interfaces using the **show 12-learning interface** operational mode command.

These commands provide the information on the state and role changes on the protected interfaces. For example, if the PCs send BPDUs and the protected interfaces receive them, the interfaces transition to the DIS role. The BPDU inconsistent state changes the interface state to blocking (BLK), preventing it from forwarding traffic.

To unblock the interfaces, you must use the **clear error bpdu interface** operational mode command.

## The Purpose of Loop Protection

### ■ The problem:

- Switch hardware and configuration errors could cause an STP loop
- A nondesignated port might stop receiving superior BPDUs from the designated bridge, causing an interface to go into the forwarding state (causing a loop)

### ■ The solution:

- Enable loop protection on all root and alternate ports
- Once enabled, selected ports do not interpret the lack of BPDUs as a false positive for making the interface a designated port
  - Ports that detect the loss of BPDUs transition to the “loop inconsistent state” (essentially the same as blocking)
- The interface recovers and transitions back to the blocking state when it receives a BPDU

© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-60

### The Purpose of Loop Protection—Problem

Although the purpose of STP, RSTP, and MSTP is to provide Layer 2 loop prevention, switch hardware or software errors could result in an erroneous interface-state transition from the blocking state to the forwarding state. Such behavior could lead to Layer 2 loops and consequent network outages.

### The Purpose of Loop Protection—Solution

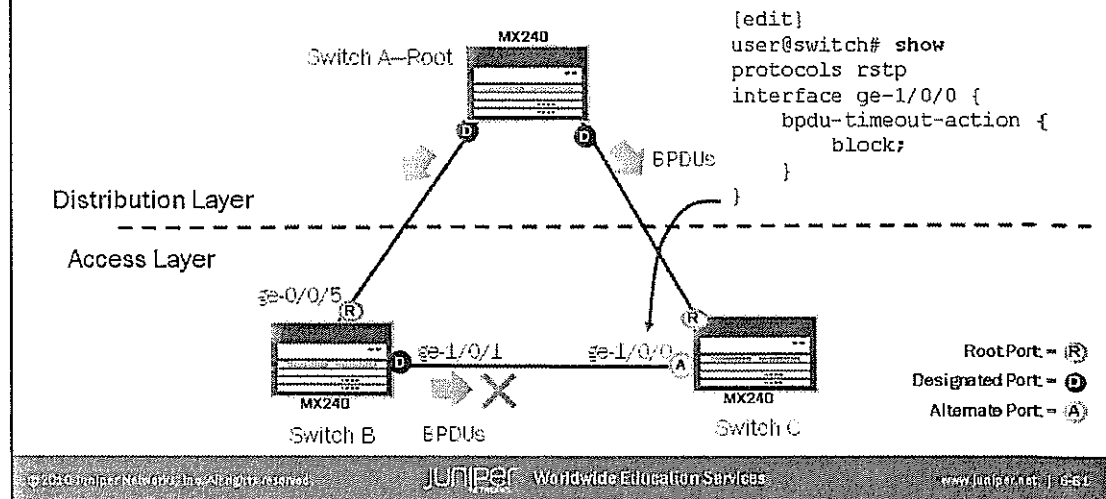
When loop protection is enabled, the spanning-tree topology detects root ports and alternate ports, and ensures that both are receiving BPDUs. If a loop-protection-enabled interface stops receiving BPDUs from its designated port, it reacts as it would react to a problem with the physical connection on the interface. It does not transition the interface to a forwarding state. Instead, it transitions the interface to a loop-inconsistent state. The interface recovers and then it transitions back to the spanning-tree blocking state when it receives a BPDU.

We recommend that if you enable loop protection, you should enable it on all switch interfaces that have a chance of becoming root or designated ports. Loop protection is most effective when it is enabled on all switches within a network.

You can configure an interface for either loop protection or root protection, but not for both.

## Configuring Loop Protection

- Stop hardware problems from causing loops in the spanning-tree topology



## Configuring Loop Protection

Now consider a network in which one switch, Switch A, is at the Distribution Layer, and two switches, Switch B and Switch C, are at the Access Layer. Being in an alternate state, interface ge-1/0/0 of Switch C is blocking traffic between Switch B and Switch C; therefore, the traffic forwards through Switch A from Switch C. BPDUs are traveling from the root bridge on Switch A to both its interfaces. Switch C, during normal operation, receives BPDUs from Switch B. In the example on the slide, assume that a hardware problem exists on Switch B's ge-1/0/1 interface such that both Switch B and Switch C believe the interface is up, but Switch C cannot receive BPDUs from Switch B. Without loop protection, Switch C might place its ge-1/0/0 interface into the forwarding state, causing a loop. The slide illustrates the required configuration, and shows how to configure loop protection on interface ge-1/0/0 to prevent it from transitioning from a blocking state to a forwarding state and causing a loop in the spanning-tree topology.

You can configure an interface for either loop protection or root protection, but not both.

## Verifying Loop Protection Functionality

- To verify loop protection functionality:
  - Use the **show spanning-tree interface** command before and after enabling the loop protection feature
  - Watch for state changes, role changes, or both in the interface output
    - BLK state remains BLK
    - ALT role transitions to DIS (loop inconsistent)
  - The interface recovers and transitions back to its original state when it receives BPDUs

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-62

### Verification of Loop Protection Functionality

To confirm that the configuration is working properly on the STP-running switch, use the **show spanning-tree interface** operational mode command prior to configuring loop protection. This command provides information for the interface's spanning-tree state, which should be blocking (BLK).

Once BPDUs stop arriving at the protected interface, the loop protection is triggered on that interface. You can use the **show spanning-tree interface** command to observe the state of the interface. This command displays the loop-inconsistent state for the protected interface, which prevents the interface from transitioning to the forwarding state. The interface recovers and transitions back to its original state when it receives BPDUs.

## The Purpose of Root Protection

### ■ The problem:

- Bridge applications running on PCs can generate BPDUs and interfere with root port election
- Erroneous root port election on a switch

### ■ The solution:

- Enable root protection on the switch interfaces that should not receive superior BPDUs from the root bridge and should not be elected as the root port
- The interfaces become designated ports
- Once a superior BPDU arrives on a port with root protection enabled, the port transitions to an inconsistency state, blocking the interface
- The interface recovers and transitions back to the forwarding state when it stops receiving superior BPDUs

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-63

### The Purpose of Root Protection—Problem

Although the purpose of STP, RSTP, and MSTP is to provide Layer 2 loop prevention, a root port elected through the spanning-tree algorithm has the possibility of having been wrongly elected. In addition, user bridge applications running on PCs can generate BPDUs, interfering with root port election.

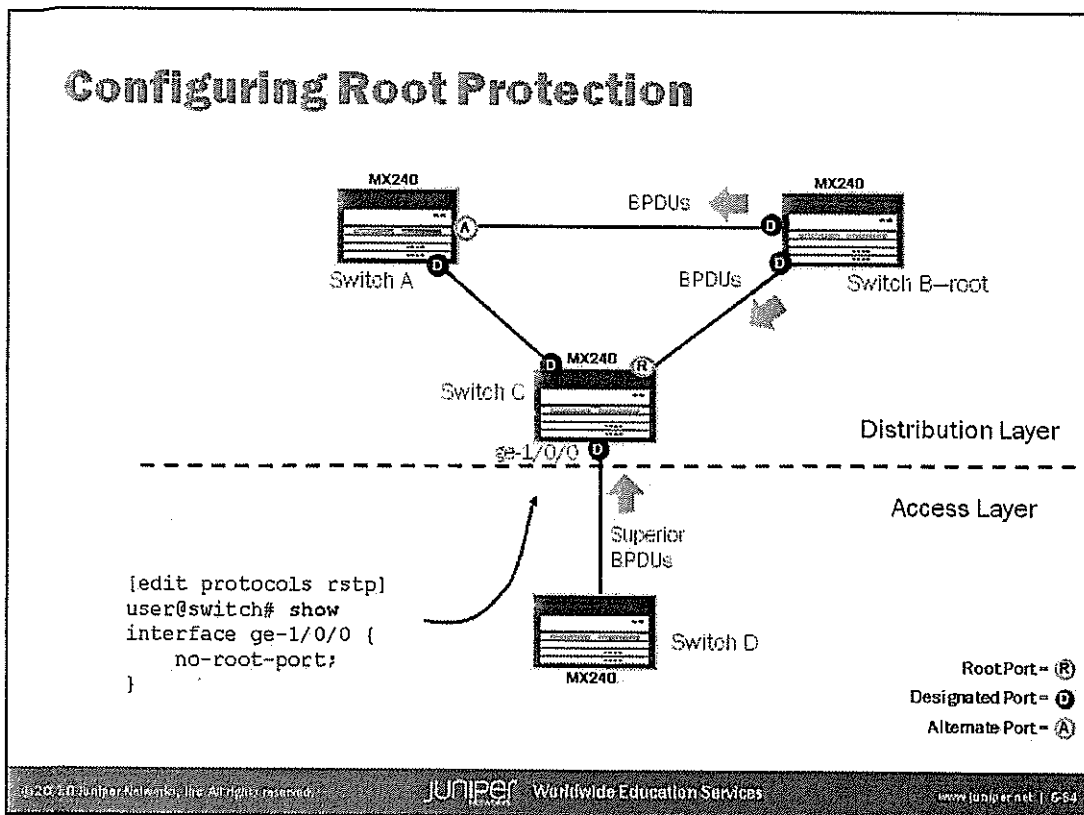
### The Purpose of Root Protection—Solution

Enable root protection on interfaces that should not receive superior BPDUs and should not be elected as the root port. These interfaces become designated ports. If the bridge receives superior BPDUs on a port that has root protection enabled, that port transitions to an inconsistency state, blocking the interface. This blocking prevents a bridge that should not be the root bridge from being elected.

After the bridge stops receiving superior BPDUs on the interface with root protection, the interface returns to a listening state, followed by a learning state, and ultimately returns to a forwarding state. Recovery back to the forwarding state is automatic.

When root protection is enabled on an interface, it is enabled for all the STP instances on that interface. The interface is blocked only for instances for which it receives superior BPDUs. Otherwise, it participates in the spanning-tree topology.

You can configure an interface for either loop protection or root protection, but not for both.



### Configuring Root Protection

Now consider a network in which three switches, Switch A, Switch B, and Switch C, are at the Distribution Layer, and one switch, Switch D, is at the Access Layer.

Interface ge-1/0/0 of Switch C is configured with root protection. If Switch D sends superior BPDUs, they trigger root protection on interface ge-1/0/0, blocking it. The slide illustrates the required configuration.

You can configure an interface for either loop protection or root protection, but not both.

## Verifying Root Protection Functionality

- To verify root protection functionality:
  - Use the **show spanning-tree interface** command before and after you enable the root protection feature
  - Receipt of superior BPDUs on the watched interface triggers root protection
    - FWD state changes to BLK
    - DESG role transitions to DIS (loop inconsistent)
  - The interface recovers and transitions back to its original state when it no longer receives superior BPDUs

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-65

### Verification of Root Protection Functionality

To confirm that the configuration is working properly on the STP-running switch, use the **show spanning-tree interface** operational mode command prior to configuring loop protection. This command provides information for the interface's spanning-tree state.

Once you configure root protection on an interface and that interface starts receiving superior BPDUs, root protection is triggered. You can use the **show spanning-tree interface** command to observe the state of the impacted interface. This command displays the loop-inconsistent state for the protected interface, which prevents the interface from becoming a candidate for the root port. When the root bridge no longer receives superior BPDUs from the interface, the interface recovers and transitions back to a forwarding state. Recovery is automatic.

## Summary

### ■ In this chapter, we:

- Explained the purpose of STPs
- Described the basic operation of STP, RSTP, MSTP, and VSTP
- Configured and monitored STP, RSTP, MSTP, and VSTP
- Explained the purpose of BPDU, loop, and root protection

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-66

### This Chapter Discussed:

- The purpose of STPs;
- The basic operation of STP, RSTP, MSTP, and VSTP;
- Configuration and monitoring of STP, RSTP, MSTP, and VSTP; and
- Implementation of BPDU, loop, and root protection.



## Review Questions

1. What is the purpose of STP?
2. Describe the operation of the STP port states.
3. Describe how to build a spanning tree.
4. How are STP, RSTP, MSTP, and VSTP different?

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 6-67

## Review Questions

- 1.
- 2.
- 3.
- 4.

## Lab 4: MSTP

- Perform configuration and verification steps typically associated with MSTP.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | C-68

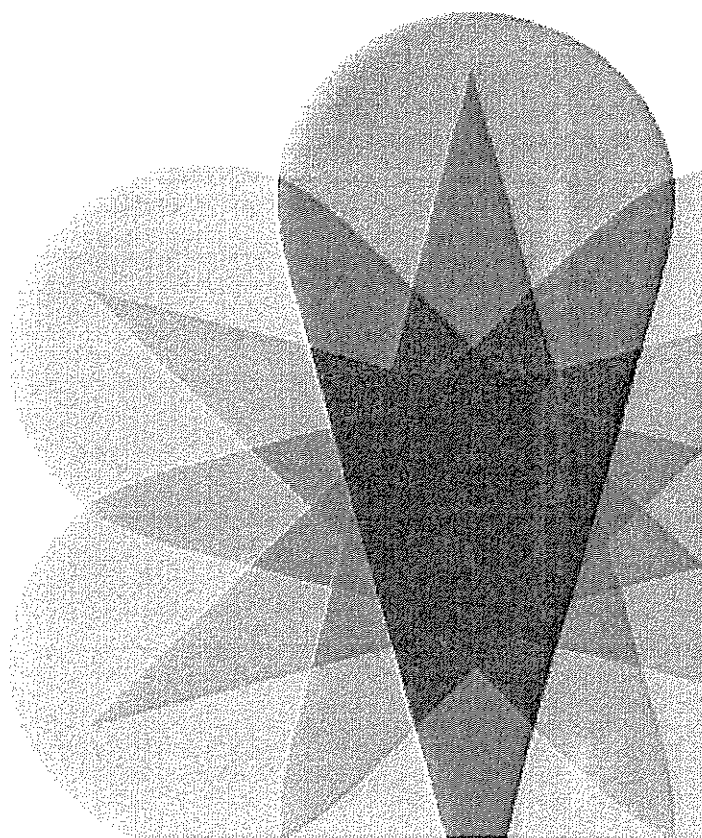
## Lab 4: MSTP

The slide provides the objective for this lab.



# **Junos Service Provider Switching**

## **Chapter 7: Ethernet OAM**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Explain typical Operation, Administration, and Maintenance features
  - Describe the basic operation of link fault management
  - Describe the basic operation of connectivity fault management
  - Configure and monitor Ethernet OAM

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.junipernet | 7-2

### This Chapter Discusses:

- Typical Operation, Administration, and Maintenance (OAM) features;
- The basic operation of link fault management (LFM);
- The basic operation of connectivity fault management (CFM); and
- Configuration and monitoring of Ethernet OAM.

## Agenda: Ethernet OAM

### → OAM Overview

- LFM
- CFM
- Configuring and Monitoring Ethernet OAM

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

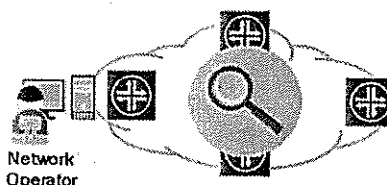
www.juniper.net/173

### OAM Overview

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## What Is OAM?

- OAM is a set of functions that allows network operators to monitor the health of the network:
  - Determines the location of faulty links or faulty conditions
  - Measures performance of the network
  - Allows for diagnosis testing (loopback and so forth)
- Measurements that can be taken:
  - Availability
  - Frame delay
  - Frame delay variation
  - Frame loss



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-4

### OAM Monitors the Health of a Network

For years, SONET and Asynchronous Transfer Mode (ATM) devices have been able to perform the functions of OAM. Only recently have standards been introduced to bring these same types of features to Ethernet. The purpose of OAM is to help network operators determine the location of faulty conditions, measure the performance of the network, and allow for diagnosis testing. One key to Ethernet's success in becoming *carrier-class* is support for OAM features.

### OAM Measurements

A device supporting OAM should determine the following measurements:

- *Availability*: The ratio of uptime over total time the measure takes;
- *Frame delay*: The time required to transmit a frame from one device to another;
- *Frame delay variation*: The variation in frame delay measurements between consecutive test frames; and
- *Frame loss*: The number of frames lost over time.

## Ethernet OAM Standards

- IEEE, MEF, and ITU have developed complimentary standards to allow for Ethernet OAM:
  - IEEE 802.3-2008, clause 57 (was 802.3ah)
    - LFM—detecting faults on a single link in an Ethernet network
    - Known as Ethernet in the First Mile OAM
  - IEEE 802.1ag and ITU-T Y.1731
    - CFM—detecting faults along an entire path of an Ethernet network
  - MEF 17
    - Provides the requirements that OAM mechanisms must satisfy
    - Provides a framework to discuss and implement those mechanisms

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net/17-5

## Ethernet OAM Standards

The Institute of Electrical and Electronics Engineers (IEEE), Metro Ethernet Forum (MEF), and International Telecommunication Union Telecommunication Standardization (ITU-T) organizations have developed complementary standards to allow for Ethernet OAM. IEEE 802.3-2008, clause 57, defines a method of OAM to monitor link performance, detect faults, and perform loopback testing over a single link. It is usually used on the customer's access link to the provider. The standard is referred to as Ethernet in the First Mile OAM (EFM OAM). The IEEE 802.1ag standard specifies the requirements to detect faults along an end-to-end path of an Ethernet network using CFM. CFM provides for fault monitoring, path discovery, fault isolation, and frame delay measurement (ITU-T Y.1731). The MEF 17 technical specification defines the requirements that must be satisfied by both an equipment vendor and service provider in the area of fault management, performance monitoring, autodiscovery, and intraprovider and interprovider service OAM.

## OAM Basics (1 of 4)

- The primary purpose of OAM is to detect network defects:
  - Defect—some network function is not working as expected
  - Failure—defects over some time that can cause a network function to stop
  - Alarm—an indication that something has failed
- Continuity check messages:
  - Unidirectional messages
    - Sent at regular intervals by one endpoint
    - If the remote end does not receive the message within a certain interval, a fault is detected, potentially causing an alarm

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 26

### OAM Basics

In general, OAM has the main purpose of detecting network defects. A defect is a network function that is not working as expected. If a defect continues to occur over time, a device considers the recurring defect a failure. The device signals the failure as an alarm. An alarm is a notification that alerts a human and potentially other devices that something has gone wrong in the network.

### Continuity Check Messages

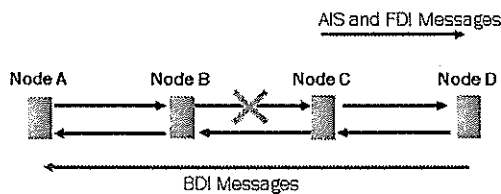
A common feature of OAM is the usage of continuity check (CC) messages. These messages are unidirectional (no acknowledgments), and they are sent between devices. These messages notify a remote device that the local device is still reachable along the path of the CC message. If for some reason a failure occurs that prevents the CC messages from being delivered, the remote device might consider that network path down and generate an alarm.



## OAM Basics (2 of 4)

### ■ Indications:

- Alarm indication signal and Forward Defect Indicator
  - Notify downstream network nodes when a failure or defect occurs
- Backward Defect Indicator
  - Notifies upstream network nodes when a failure occurs in the reverse direction



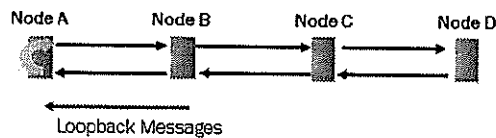
### Indications

Another feature of OAM protocols is the use of indicators to signal failures in the network. The diagram on the slide shows several devices that are interconnected by some network media (SONET for example). The diagram also shows that the network media has a transmit path and a receive path. The example shows that a failure has occurred on Node B's transmit path and Node C's receive path on the link between Node B and Node C. OAM capabilities allow the devices in the data path to *indicate* that a failure has occurred. Node C uses an alarm indication signal (AIS) and Forward Defect Indicators (FDIs) to inform downstream devices (Node D) that a failure has occurred along the upstream path. Once Node D receives the AIS and FDI messages, it might send a Backward Defect Indicator (BDI) along its transmit path to inform nodes in the reverse direction of the failure (Node A and Node B) that a problem exists downstream from those devices.

## OAM Basics (3 of 4)

### ▪ Loopback messages:

- Help narrow the scope of the issue if a problem exists on a network with multiple nodes
- Allow for detection of a defect between nodes
- Comprise two different types:
  - Nonintrusive loopback messages—do not cause disruption to service (like the ping facility)
  - Intrusive loopback messages—signal a remote node to go into a special test mode (normal transit traffic cannot flow)



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-8

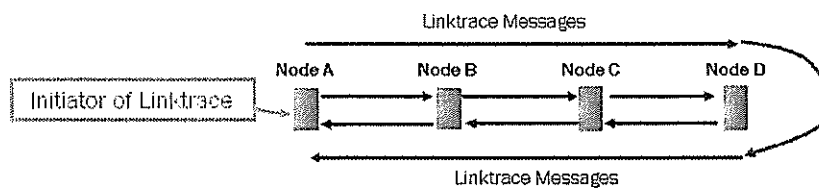
### Loopback Messages

Loopback messages are a feature of OAM that help an administrator to find faults in the network. Two types of loopback messages exist and both types are initiated by an administrator of a device. Non-intrusive loopback messages (like ping for IP) allow an administrator to direct a device to send a loopback message downstream to another device with the expectation that the device will respond with a loopback response message. If the response is not received, the administrator might need to perform further testing. Another type of loopback message is an intrusive type. This type of message also is initiated by the administrator of a device, but it signals a downstream device to place a loop on its interface. Usually, when the remote device's interface is in a loop, it cannot pass normal traffic. Instead of receiving any traffic arriving on the receive path of the remote device's looped interface, the remote device loops the traffic around and sends it back out its transmit path. By setting a loop on a remote interface, an administrator can send traffic toward the remote device, and if it returns to the initiating device, then the administrator can eliminate the data path from this list of potentially bad data paths.

## OAM Basics (4 of 4)

### ▪ Linktrace messages:

- Bidirectional continuity check
  - Similar to traceroute for IP
- Identifies nodes along the path of the messages



### Linktrace Messages

Linktrace messages are a feature of OAM that is similar to the nonintrusive loopback messages that we discussed previously. The difference is that each device along the path of the linktrace message will identify itself by sending a response to the initiator of the linktrace message. After gathering the responses from all of the devices along the path of the message, the administrator knows the identity of all of the devices along the data path.

## Agenda: Ethernet OAM

- OAM Overview
- LFM
- CFM
- Configuring and Monitoring Ethernet OAM

© 2010 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

www.juniper.net | 7-10

### LFM

The slide highlights the topic we discuss next.

## LFM Capabilities

### ▪ LFM is limited to a single Ethernet link:

- Remote failure indication
- Remote loopback
- Link monitoring
  - Event notification
  - Device polling
- OAM capability discovery
- No AIS

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

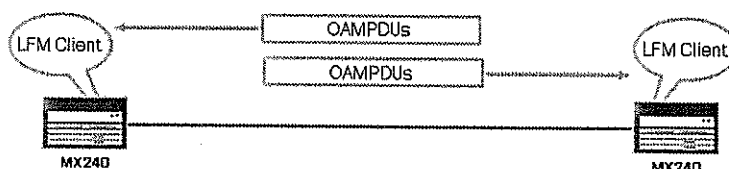
www.juniper.net | 7-11

## LFM Capabilities

LFM is defined in IEEE 802.3, Clause 57. It specifies a method of OAM to be used on a single link. Usually, LFM is used on user-to-network interface (UNI) links between a customer and the service provider. On the single Ethernet link, LFM can provide for remote failure indication, remote loopback (intrusive), link monitoring, event notification, and LFM capability discovery. Because LFM is used on a single link, no AIS capabilities are available.

## LFM Clients

- LFM clients communicate at the Ethernet layer:
  - No IP addressing is necessary
  - Clients exchange OAM protocol data units
    - OAMPDUs are sent with a source address of the outgoing port and a destination address of 01-80-c2-00-00-02 (they are never flooded)



- LFM clients can be either active or passive:
  - Passive clients cannot initiate the discovery process or loopback control messages
  - At least one client must be in active mode (both can be)

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 7-12

### LFM Clients

A switch must have LFM client capabilities to support LFM. LFM clients exchange OAM protocol data units (OAMPDUs) that are addressed to 01-80-c2-00-00-02—a multicast MAC address. These messages are sent only across a single link and never flooded. To determine whether an LFM client is on the opposite side of an Ethernet link, an LFM client will go through a discovery process. The discovery process is where LFM clients discover their peers and determine each other's supported capabilities.

### Active and Passive

A client can be either active or passive (both clients can also be active). Only active LFM clients can initiate the discovery process. Also, only active clients can send loopback messages.

## OAM PDU Codes and Flags

Destination MAC Address	Source MAC Address	Type/ Length	Sub Type OAM	Flags	Code	Data (TLVs)	Frame Check Sequence
-------------------------------	--------------------------	-----------------	--------------------	-------	------	----------------	----------------------------

- The type of carried information depends on the code and the flags:

- Codes

- 0x00 Information
- 0x01 Event notification
- 0x02-0x03 Variable request and response (polling for MIBs)
- 0x04 Loopback control

- Flags

- Bit 0: Link Fault
- Bit 1: Dying Gasp
- Bit 2: Critical Event
- Bit 3-4: Used during the discovery process

## OAM PDUs

Several different types of OAM PDUs exist. We discuss each over the next few slides. This slide shows the format of an OAM PDU. The type of OAM PDU is determined by the code and the flag settings. In general, four types of OAM PDUs exist: information, event notification, variable request and response, and loopback control. Except in the case of the discovery process, the flags are used as BDIs, as described in the previous slides, which notify the remote LFM client of a failure. The following events will result in the setting of flags:

- *Link Fault*: Signal loss is detected on the receive path.
- *Dying Gasp*: An external failure condition occurred. A power failure is a good example of a Dying Gasp event.
- *Critical Event*: An unspecified failure event. By configuring an *action-profile* (described on the following slides), an administrator can specify which events can cause the local switch to send OAM PDUs with the Critical Event bit set.

## OAMPDU Types (1 of 2)

### ■ Information OAMPDU:

- Discovery
  - OAM clients discover each other and exchange capabilities (remote loopback, and so on)
- Heartbeat
  - At least one OAMPDU must be sent each second
  - An empty information OAMPDU is sent if the client has no information to transmit
- Critical Events
  - Link Fault: The Physical Layer determined a fault occurred in the receive direction of the client
  - Dying Gasp: An unrecoverable local failure condition occurred
  - Critical Event: An unspecified critical event occurred

### Information OAMPDUs

Information OAMPDUs serve several purposes. They are used during the discovery process to discover neighboring clients and exchange capabilities. They also are used to perform continuity checks between clients. When the clients have no information to transmit, they exchange empty information OAMPDUs on a regular, configured interval. In this case, they provide a heartbeat between the two clients. If the peer stops receiving so many messages, it determines that a failure occurred on the link between the two peers. Also, the information OAMPDU signals Critical Events, as described previously.



## OAMPDU Types (2 of 2)

### ■ Event notification OAMPDU:

- Signals various link events and statistics
  - Errored Frame Event
  - Errored Frame Period Event
  - Errored Frame Seconds Summary Event
  - Errored Symbol Period Event

### ■ Loopback control OAMPDU

- Signals a remote peer to set or unset a looped interface

### ■ Variable request and response OAMPDU

- Optional OAMPDU
- Allows peers to exchange IEEE 802.3 Clause 30 MIB variables

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-15

### Event Notification OAMPDUs

Event notification OAMPDUs are used as BDIs. That is, they inform the upstream client that errors have occurred on the local receive path. The slide lists the four different types of event notifications.

### Loopback Control OAMPDU

The loopback control OAMPDU allows an LFM client to direct the remote clients to set or unset a loop on its interface.

### Variable Request and Response OAMPDUs

Variable request and response OAMPDUs are not currently supported in the Junos operating system. They are used to allow for one client to gather information about the remote clients by polling it for IEEE 802.3, clause 30 MIBs (read only).

## Reacting to Events

- Reaction to events is locally configurable:
  - Except for Critical Events
    - Interface goes to link-down state automatically
  - Possible actions
    - Syslog
    - Link down
    - Begin sending OAMPDUs with a Critical Event bit set

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 7-16

### Reacting to Events

When an LFM client receives a Critical Event, it automatically places the interface in a down state (it removes routes and causes spanning-tree recalculation). It does, however, continue to monitor the interface for LFM messages in the event that the link becomes stable again.

For other types of events, the action an LFM client performs is configurable. To specify the action to be performed for a particular event, you must create an `action-profile` and apply it to an interface. Actions you can configure are generation of a syslog message, placement of the interface in a down state, or the sending of an OAMPDU to the remote peers with the Critical Event bit set.

## Agenda: Ethernet OAM

- OAM Overview
- LFM
- CFM
- Configuring and Monitoring Ethernet OAM

© 2013 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

www.juniper.net | 7-17

### CFM

The slide highlights the topic we discuss next.

## CFM Features

### ■ Features:

- Fault monitoring using continuity check
  - Neighbor discovery and health check protocol
- Path discovery and fault verification using linktrace
  - Similar to IP traceroute
- Fault isolation using a loopback protocol
  - Similar to IP ping
- Frame delay measurement
  - Defined in Y.1731 using vendor specific TLVs

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

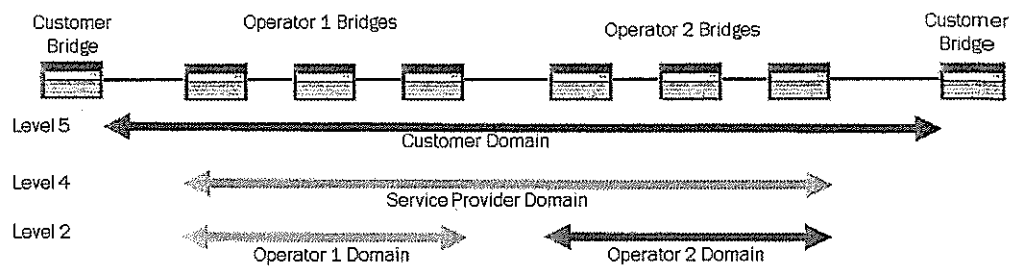
www.juniper.net | 7-10

### CFM Features

The slide lists the CFM features that an MX Series 3D Universal Edge Router supports. Other features defined in Y.1731 are not yet supported (such as AIS).

## Maintenance Domains

- An end-to-end network is broken up into maintenance domains:
  - Each maintenance domain is assigned a level (0 to 7)
    - Assignment of the level is based on hierarchy—the outermost domain must be higher than the innermost domain



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

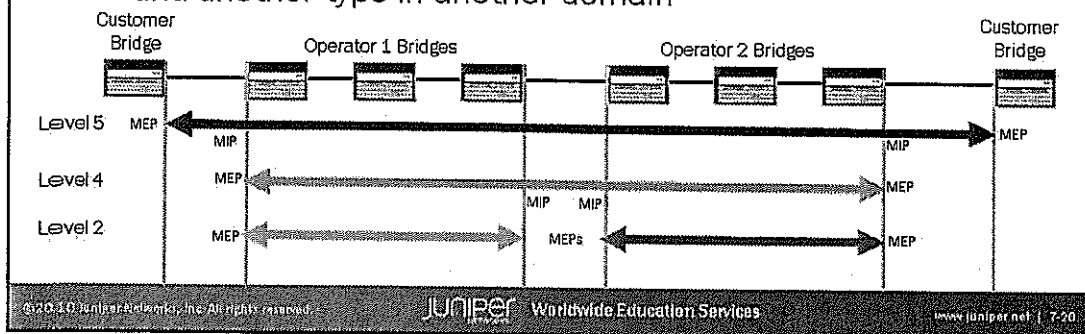
www.juniper.net | 7-19

## Maintenance Domains

CFM operates in a layered environment. As you will see on the next few slides, this layering allows end-to-end OAM functionality without having to expose all of the details of the service provider network to the customer. Each layer of the CFM network is assigned a maintenance domain ID and a level. A level can be in the range of 0 to 7. Level 5 through Level 7 are reserved for customers, Level 3 and Level 4 are reserved for providers, and Level 0 through Level 2 are reserved for operators. An operator level maintenance domain represents a subset of the provider network. Besides hiding the details of the network from the upper-level maintenance domains, this layering allows for quicker fault detection in the network because a fault detected in the Operator 1 maintenance domain actually eliminates the need to troubleshoot devices in the Operator 2 maintenance domain.

## Maintenance Point

- A maintenance point is a port on a bridge:
  - Three types
    - Maintenance association endpoint—at the edge of a domain
    - Maintenance association intermediate point—internal to a domain
    - Transparent point—does not respond to CFM messages
  - A maintenance point can be a certain type in one domain and another type in another domain



### Maintenance Points

A maintenance point is a port or interface on a switch. Three possible types of maintenance points might be present in a maintenance domain. A maintenance domain has at least two maintenance endpoints (MEPs). MEPs are interfaces found at the edge of the maintenance domain. A MEP forms a relationship with a single MEP or several MEPs that are in the same maintenance domain, and level, and that protect the same Ethernet virtual circuit (EVC) (also called a maintenance association). A MEP forms a relationship with a single MEP when protecting an Ethernet Line (E-Line) EVC and forms relationships with multiple MEPs when protecting an Ethernet LAN (E-LAN) EVC. Among other things, MEPs exchange CC messages with each other to ensure that the path between them is up and available.

Another type of maintenance point is a maintenance intermediate point (MIP). MIPs are completely optional. MIPs are used to expose some of the network at a lower maintenance domain level to an upper level. For example, consider the diagram where MIP functionality was configured on the Level 4 MEPs. A linktrace from the customer bridge on the left side of the diagram at Level 5 shows three hops to the customer bridge on the right side. The two MIPs that were configured at Level 4 and the final MEP at Level 5 respond to the linktrace message. MIPs respond only to CFM messages that were received from a MEP at one higher level than their own. The final type of maintenance point is a transparent point. A transparent point is not configured for CFM messages and simply forwards them as regular data traffic.

## Maintenance Point Roles

- Each maintenance point has a role to perform:
  - The goal is to maintain the integrity of each domain while still providing each domain with enough information to isolate faults

Tasks	MEP	MIP	Transparent
Initiate CFM messages	Yes	No	No
Respond to loopback and linktrace messages	Yes	Yes	No
Track CC messages	Yes	Yes	No

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

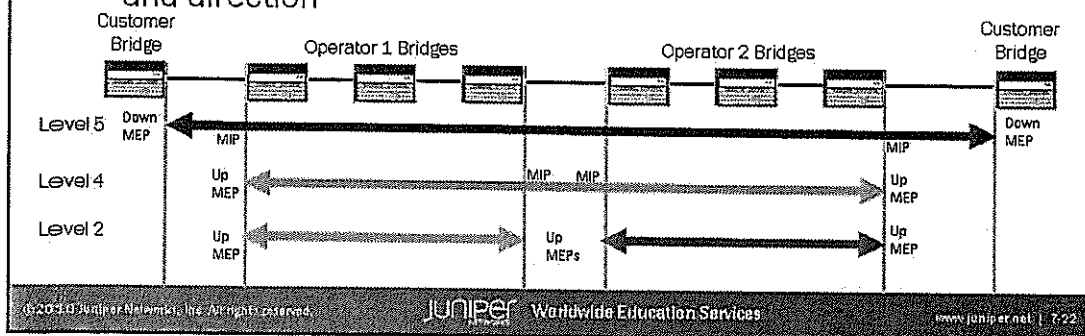
www.juniper.net | 7-21

## Maintenance Point Roles

The slide shows the roles that each type of maintenance point plays in a CFM network.

## MEP

- A MEP forms a neighbor relationship with other MEPs in the same domain with the exchange of CC messages
  - Two types of MEPs:
    - Down MEP—A MEP (interface) that faces a neighboring down MEP
    - Up MEP—A MEP (interface) that faces away from a neighboring up MEP
  - To become neighbors, two MEPs must be configured with the same maintenance domain, maintenance association, level, and direction

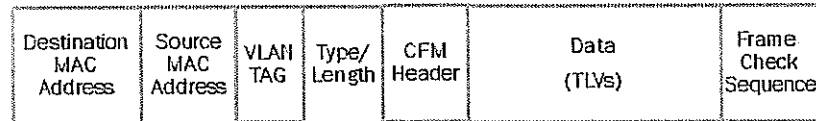


### MEP-to-MEP Relationship

Once a MEP is configured, it attempts to form a neighbor relationship with other MEPs that have are similarly configured. The relationship establishes by means of the exchange of CC messages. Each MEP is configured with a MEP ID (a number). The MEP ID must be unique among all MEPs in the network. Each MEP also is configured with a direction—either up or down. A down MEP expects to find neighboring MEPs downstream. An up MEP expects to find neighboring MEPs upstream. To become neighbors, two MEPs must be configured with the same maintenance domain, maintenance association, level, and direction. This data is carried in each CFM message.



## CFM Messages



Destination Address = 01-80-C2-00-00-3y



- The Opcode specifies the type of message:
  - As defined in IEEE 802.1ag
    - Continuity check
    - Loopback reply
    - Loopback message
    - Linktrace reply
    - Linktrace message
  - ITU-T Y.1731 specifies others (AIS, RDI, APS, and so on)

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

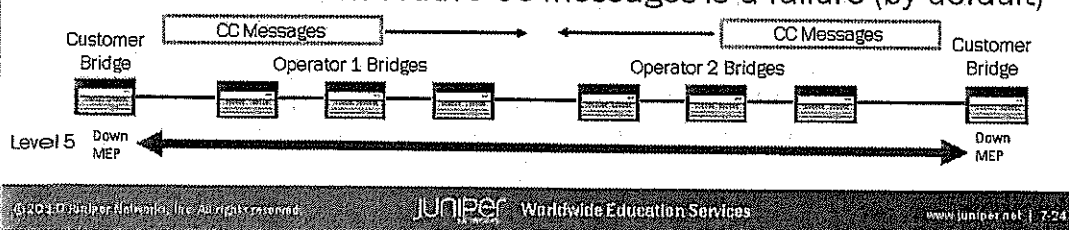
www.juniper.net | 7-23

## CFM Messages

The slide shows the format of a CFM message. IEEE 802.1ag defines five types of messages, as shown on the slide. CC and linktrace messages are sent to a multicast destination address. The same specification has reserved a set of type, length, and values (TLVs) to allow for the future expansion of the CFM protocol. The ITU-T uses some of these extended TLVs to allow for Ethernet frame delay measurement, remote defect indications, and more. Note that the last four bits of the destination address represent the level of the sending MEP as well as the type of message. If y equal 0-7, then the message is a CC message destined to the appropriate level. If y = 8-F, then the message is a linktrace message destined for Levels 0-7, respectively. CFM messages are also encapsulated with the virtual LAN (VLAN) tag of the EVC that is being protected.

## Continuity Check

- A MEP sends CC messages at regular intervals:
  - The interval is configurable to 100 ms, 10 ms, 1 s, 1 m, or 10 m
  - Contains several values
    - Maintenance domain ID
    - Level
    - Maintenance association ID
    - MEP ID (unique among MEPs)
  - CC messages are multicast
    - MEPs only act upon them in the same level
  - Loss of three consecutive CC messages is a failure (by default)



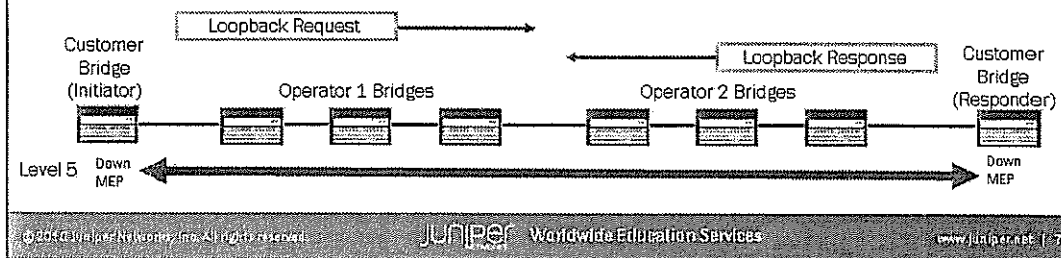
### CC Messages

The slide shows the details of CC messages. The remote MEP considers the loss of three CC messages a failure by default. This loss threshold is configurable.

## Loopback Protocol

### ■ Loopback protocol:

- Loopback initiator sends a loopback request
  - Sent to a specific MAC address
- Loopback responder sends a loopback response message
- A lack of received loopback response message by the initiator allows an administrator to determine whether a problem exists in the network



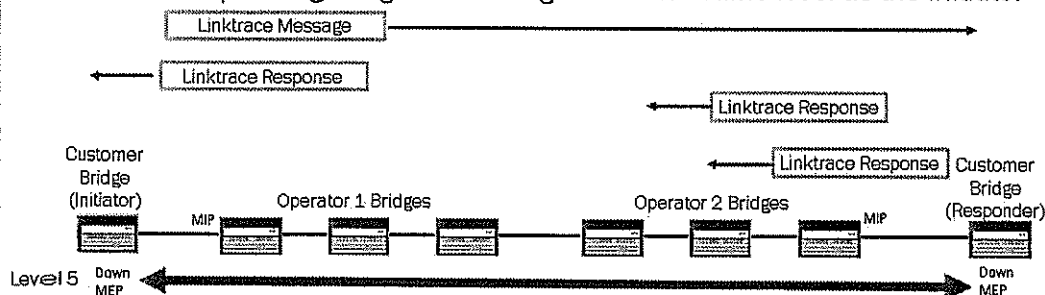
### Loopback Protocol

The slide shows the details of the loopback protocol. The loopback protocol is similar to ping in IP.

## Linktrace Protocol

### ■ The administrator initiates the linktrace protocol:

- The linktrace initiator sends a linktrace message
  - Sent to a specific MAC address
- Each of the maintenance points along the path forwards the original linktrace message to the destination MAC address and also sends a linktrace reply listing their own MAC addresses
  - Responding bridges are configured at the same level as the initiator



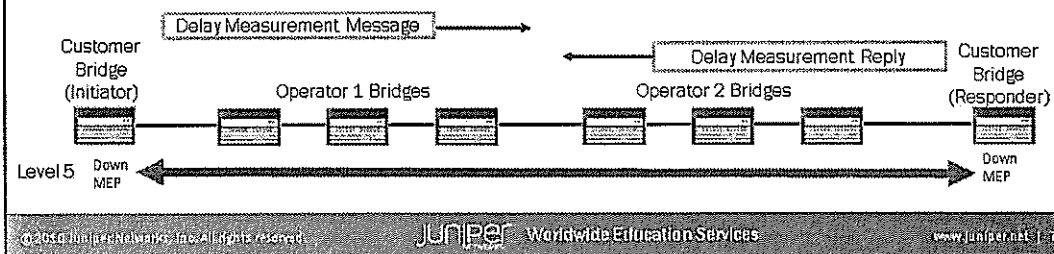
## Linktrace Protocol

The slide shows the details of the linktrace protocol. The linktrace protocol is similar to the traceroute function in IP.

## Frame Delay Measurement

- The administrator initiates frame delay measurements:

- The initiator sends the delay measurement message
- The delay measurement responder sends a delay measurement reply message
- The initiator calculates the two-way delay
  - $(\text{Time reply received}) - (\text{Time message sent}) = \text{Delay}$



### Frame Delay Measurement

Two types of delay tests exist: one-way and two-way. An MX Series router uses hardware-assisted timestamping. When an administrator initiates a one-way frame delay test, a delay measurement message is sent to the remote MEP. The delay measurement message contains a timestamp. The remote MEP then calculates the delay from the time the frame was sent to the time it arrived. For the measurement to be accurate, both devices must have their clocks synchronized. A two-way test does not require the two devices to have their clocks synchronized. The slide shows the details of a two-way frame delay test.

## Agenda: Ethernet OAM

- OAM Overview
- LFM
- CFM
- Configuring and Monitoring Ethernet OAM

© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

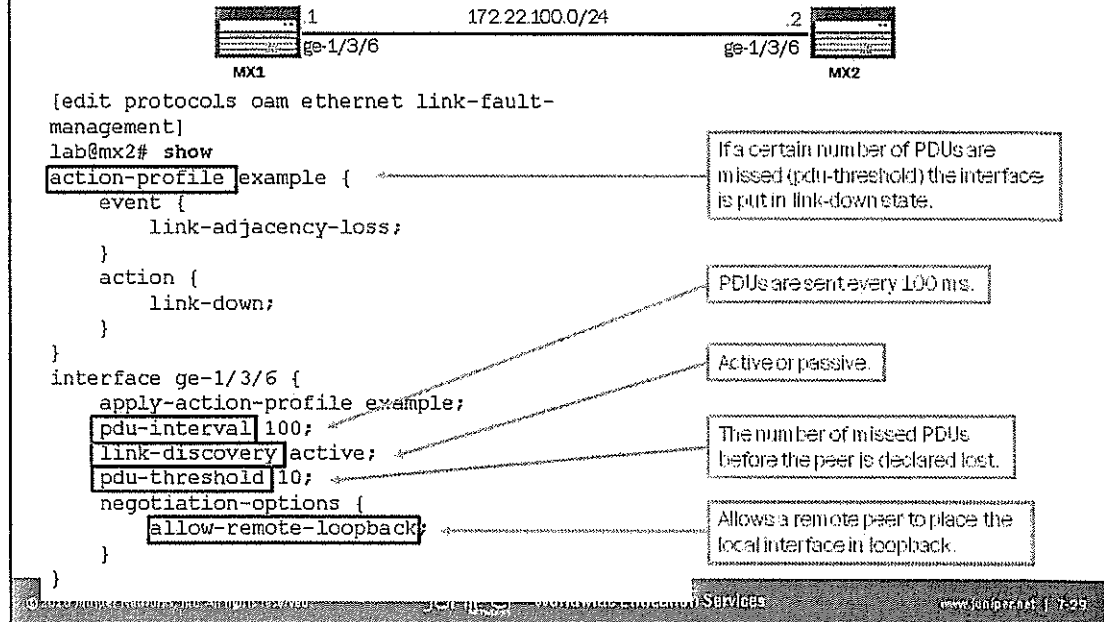
[www.juniper.net](http://www.juniper.net) | 7-28

### Configuring and Monitoring Ethernet OAM

The slide highlights the topic we discuss next.

## LFM Configuration

### ■ LFM settings



### LFM Settings

The slide shows some of the typical LFM configuration settings.

## Action Profile

- Use action profiles to specify how a switch should react to certain events:
  - Critical Events cause the interface to go into link-down state automatically

```
[edit protocols oam ethernet link-fault-management]
lab@mx1# set action-profile example event ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except   Don't inherit configuration data from these groups
link-adjacency-loss     Loss of adjacency with OAM peer
> link-event-rate       Upper layer indication on protocol down
protocol-down           Upper layer indication on protocol down

lab@mx1# set action-profile example action ?
Possible completions:
+ apply-groups          Groups from which to inherit configuration data
+ apply-groups-except   Don't inherit configuration data from these groups
link-down              Mark the interface down for transit traffic
send-critical-event     Start sending OAM PDUs with critical event bit set
syslog                 Generate syslog message
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-30

## Action Profiles

An action profile allows you to configure how the switch should react to certain events. You configure a profile for the following events:

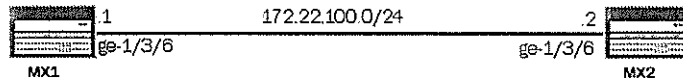
1. `link-adjacency-loss`: Occurs when CC messages are no longer being received from the remote peer.
2. `link-event-rate`: Allows you to specify a rate of receiving different types of event messages that cause an action to take place.
3. `protocol-down`: Allows the MEP to monitor when maintenance associations at higher levels go down.

The slide shows the actions you can take when the events specified in the action profile occur.



## LFM Status

- Display the status of LFM with the `show oam ethernet link-fault-management` command



```

lab@mx1> show oam ethernet link-fault-management
Interface: ge-1/3/6
Status: Running, Discovery state: Send Any
Peer address: 00:22:83:75:cc:8a
Flags: Remote-Stable Remote-State-Valid Local-Stable 0x50
Remote entity information:
  Remote MUX action: forwarding, Remote parser action: forwarding
  Discovery mode: active, Unidirectional mode: unsupported
  Remote loopback mode: supported, Link events: supported
  Variable requests: unsupported
  
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

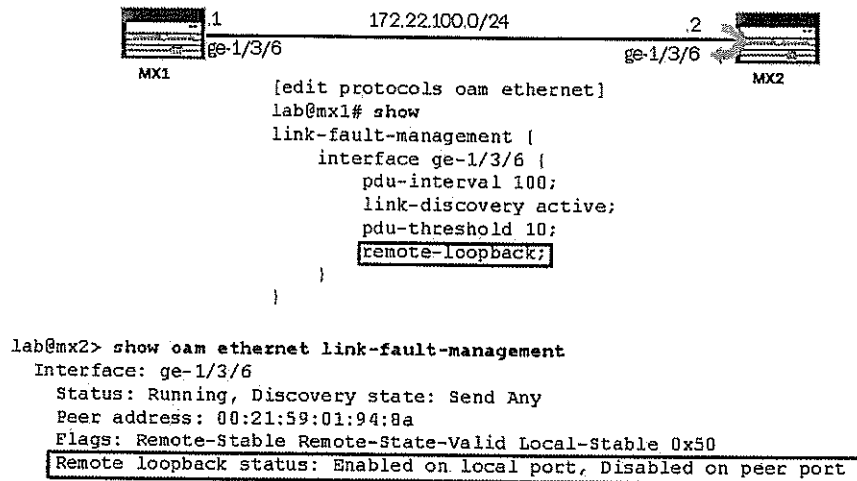
www.juniper.net | 7-31

## LFM Status

Use the `show oam ethernet link-fault-management` command to determine the status of the interfaces running LFM. The output shows the peer's MAC address, the state of the relationship with that peer, and the capabilities of the peer.

## Remote Loop

- Set a loop on the remote peer by adding the `remote-loopback` statement to the configuration



© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 7-32

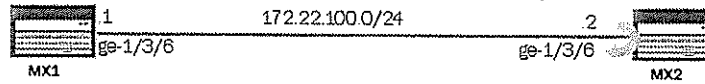
### Setting a Remote Loop

In the example on the slide, MX1 is configured to send a loopback message to MX2 so that MX2 externally loops its ge-1/3/6 interface. Once MX1's configuration is committed, the loopback message travels to MX2. Looking at MX2's command output, you can see that the local interface was placed in a loop.

## Testing the Looped Circuit

### ■ Send pings from MX1 to MX2:

- Success appears at TTL expirations
- A static Address Resolution Protocol entry is necessary



```
lab@mx1> show interfaces ge-1/3/6 | match hardware
Current address: 00:21:59:01:94:8a, Hardware address: 00:21:59:01:94:8a

[edit interfaces ge-1/3/6 unit 0 family inet]
lab@mx1# set address 172.22.100.1/24 arp 172.22.100.2 mac 00:21:59:01:94:8a

lab@mx1> ping 172.22.100.2
PING 172.22.100.2 (172.22.100.2): 56 data bytes
36 bytes from 172.22.100.1: Time to live exceeded
Vr HL TOS Len ID Flg off TTL Pro cks Src Dst
 4 5 00 0054 f7e7 0 0000 01 01 a191 172.22.100.1 172.22.100.2

36 bytes from 172.22.100.1: Time to live exceeded
Vr HL TOS Len ID Flg off TTL Pro cks Src Dst
 4 5 00 0054 f7eb 0 0000 01 01 a18d 172.22.100.1 172.22.100.2
--- 172.22.100.2 ping statistics ---
2 packets transmitted, 0 packets received, 100% packet loss
```

© 2009 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-33

## Testing the Looped Circuit

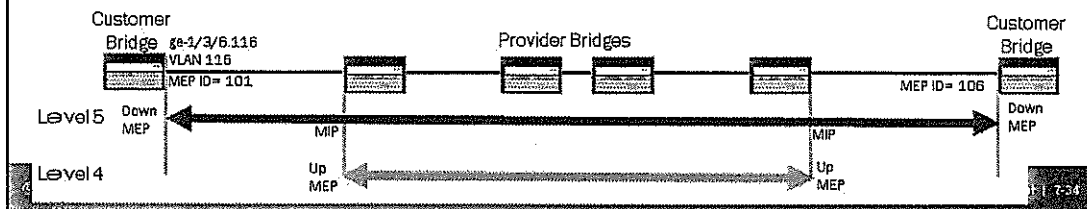
The slide shows the steps necessary to generate ping traffic across the looped circuit. A successful test of the link comes in the form of TTL expiration messages. The time-to-live (TTL) expirations show that the pings were successfully transmitted, received, and looped through the network until the TTL finally expired on the Internet Control Message Protocol (ICMP) packets.

## CFM Configuration—Customer Bridge

- Remote MEP must be configured similarly to the local MEP:

- Remote MEP must have the same maintenance domain, maintenance association, interval, and level

```
[edit protocols oam ethernet connectivity-fault-
management]
lab@switch1# show
action-profile evc1-profile {
    event {
        adjacency-loss;
    }
    action {
        interface-down;
    }
}
maintenance-domain customer {
    level 5;
    maintenance-association evc1 {
        continuity-check {
            interval 100ms;
        }
    }
    mep 101 {
        interface ge-1/3/6.116 vlan 116;
        direction down;
        auto-discovery;
        remote-mep 106 {
            action-profile evc1-profile;
        }
    }
}
```



### Down MEP Configuration

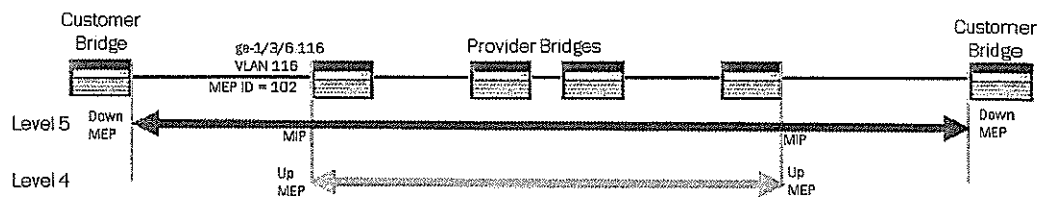
Any given MEP must be configured with a named maintenance domain, a level, a named maintenance association, a unique MEP ID, a direction, and a remote MIP ID (autodiscovery can be used as well). Also, a CC message interval must be specified to begin the neighbor discovery process and monitoring of the end-to-end EVC. The slide also shows how to apply an action profile to a remote MEP. You cannot apply an action profile when using autodiscovery.

## CFM Configuration—Provider Bridge

- You configure a provider edge bridge as a MEP and also as a MIP:

- It acts as a MIP only for level 5 (level 4 + 1)

```
[edit protocols oam ethernet connectivity-fault-management]
lab@switch2# show
maintenance-domain provider {
  level 4;
  maintenance-association evc1 {
    continuity-check {
      interval 100ms;
    }
    mip-half-function default;
    mep 102 {
      interface ge-1/3/6.116 vlan 116;
      direction up;
      auto-discovery;
    }
  }
}
```



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.junipernet | 7-35

### Up MEP Configuration

The slide shows a typical configuration of an up MEP. To allow for an up MEP to also act as a MIP for a higher level, simply add the **mip-half-function default** statement to the configuration.

## Displaying Status of CFM

- Show information related to CFM

```
Lab@switch1> show oam ethernet connectivity-fault-management ?
Possible completions:
  delay-statistics      Show Ethernet OAM maintenance endpoint delay statistics information
  forwarding-state      Show Ethernet OAM forwarding state for received packets
  interfaces            Show Ethernet OAM information for interface
  mep-database          Show Ethernet OAM maintenance endpoint database information
  mep-statistics        Show Ethernet OAM maintenance endpoint statistics information
  mip                  Display MIP information
  path-database         Display the linktrace path-database for a remote host
  policer              Show Ethernet OAM policer information
```

### Status of CFM

The slide shows all of the possible troubleshooting commands that you can use for CFM.

## Verifying Continuity

- View the status of the switch's MEP-neighbor relationships

```
lab@switch2> show oam ethernet connectivity-fault-management interfaces ge-1/3/6.116 vlan 116
```

Interface	Link	Status	Level	MEP Identifier	Neighbors
ge-1/3/6.116	Up	Active	4	102	1

```
lab@switch2> show oam ethernet connectivity-fault-management interfaces ge-1/3/6.116 vlan 116 extensive
```

Interface name: ge-1/3/6.116, vlan 116, Interface status: Active, Link status: Up

Maintenance domain name: provider, Format: string, Level: 4

Maintenance association name: evc1, Format: string

Continuity-check status: enabled, Interval: 100ms, Loss-threshold: 3 frames

Interface status TLV: none, Port status TLV: none

MEP identifier: 102, Direction: up, MAC address: 00:22:83:75:cc:8a

MEP status: running

Defects:

Remote MEP not receiving CCM	: no
Erroneous CCM received	: no
Cross-connect CCM received	: no
RDI sent by some MEP	: no
Some remote MEP's MAC in error state	: no

Statistics:

Remote MEP count: 1

Identifier	MAC address	State	Interface
105	00:21:59:ad:f4:8a	ok	ge-1/3/5.116

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-37

## CC Status

Use the **show oam ethernet connectivity-fault-management interface** command to determine the status of a MEP's relationship with a remote MEP.

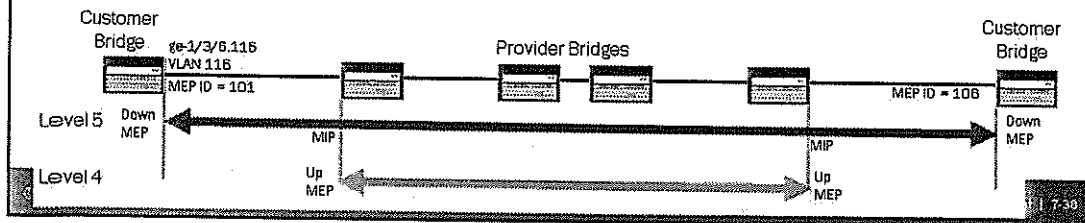
## CFM Loopback

- Use the `ping ethernet` command to initiate a CFM loopback test

- Specify the remote MEP by MEP ID or MAC address

```
lab@switch1> ping ethernet maintenance-domain customer maintenance-association evcl mep 106
PING to 00:22:83:30:fc:8a, Interface ge-1/3/6.116
68 bytes from 00:22:83:30:fc:8a: lbm_seq=0
68 bytes from 00:22:83:30:fc:8a: lbm_seq=1
68 bytes from 00:22:83:30:fc:8a: lbm_seq=2
68 bytes from 00:22:83:30:fc:8a: lbm_seq=3
--- ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss

lab@switch1> ping ethernet maintenance-domain customer maintenance-association evcl 00:22:83:30:fc:8a
PING to 00:22:83:30:fc:8a, Interface ge-1/3/6.116
68 bytes from 00:22:83:30:fc:8a: lbm_seq=4
68 bytes from 00:22:83:30:fc:8a: lbm_seq=5
68 bytes from 00:22:83:30:fc:8a: lbm_seq=6
68 bytes from 00:22:83:30:fc:8a: lbm_seq=7
--- ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
```



## CFM Loopback

The slide shows how to perform a loopback test with the `ping ethernet` command.

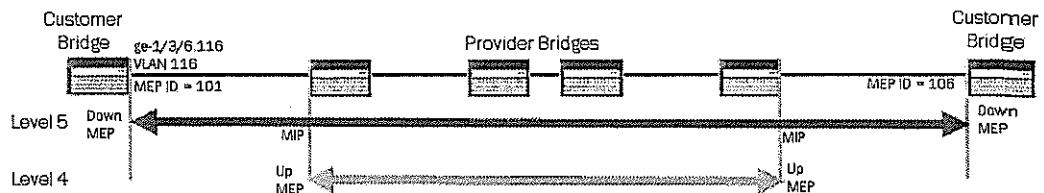


## CFM Linktrace

- Use the `traceroute ethernet` command to initiate a CFM linktrace test
  - Intermediary MIPs respond along with the remote MEP

```
lab@switch1> traceroute ethernet maintenance-domain customer maintenance-association evcl mep 106
Linktrace to 00:22:83:30:fc:8a, Interface : ge-1/3/6.116
Maintenance Domain: customer, Level: 5
Maintenance Association: evcl, Local Mep: 101
Transaction Identifier: 1
```

Hop	TTL	Source MAC address	Next-hop MAC address
1	63	00:22:83:75:cc:8a	00:22:83:75:cc:89
2	62	00:21:59:ad:f4:89	00:21:59:ad:f4:8a
3	61	00:22:83:30:fc:8a	00:00:00:00:00:00



© 2012 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-39

### Linktrace

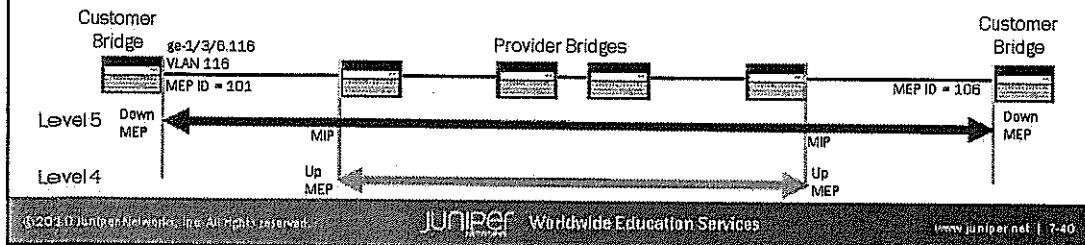
To perform a CFM linktrace to a remote MEP, issue the `traceroute ethernet` command. Note that any MIPs configured at Level 4 also respond to the linktrace message initiated by a Level 5 MEP.

## CFM Frame Delay Measurement

- Use the `monitor ethernet delay-measurement` command to initiate a CFM frame delay test

```
lab@switch1> monitor ethernet delay-measurement maintenance-domain customer maintenance-association
evl1 mep 106 two-way
Two-way ETH-DH request to 00:22:83:30:fc:8a, Interface ge-1/3/6.116
DHR received from 00:22:83:30:fc:8a Delay: 232 usec Delay variation: 0 usec
DHR received from 00:22:83:30:fc:8a Delay: 201 usec Delay variation: 31 usec
DHR received from 00:22:83:30:fc:8a Delay: 187 usec Delay variation: 14 usec
DHR received from 00:22:83:30:fc:8a Delay: 180 usec Delay variation: 7 usec
DHR received from 00:22:83:30:fc:8a Delay: 197 usec Delay variation: 17 usec
DHR received from 00:22:83:30:fc:8a Delay: 178 usec Delay variation: 19 usec
DHR received from 00:22:83:30:fc:8a Delay: 199 usec Delay variation: 21 usec
DHR received from 00:22:83:30:fc:8a Delay: 163 usec Delay variation: 36 usec
DHR received from 00:22:83:30:fc:8a Delay: 193 usec Delay variation: 30 usec
DHR received from 00:22:83:30:fc:8a Delay: 179 usec Delay variation: 14 usec

--- Delay measurement statistics ---
Packets transmitted: 10, Valid packets received: 10
Average delay: 190 usec, Average delay variation: 18 usec
Best case delay: 163 usec, Worst case delay: 232 usec
```



### Frame Delay Measurement

The slide shows how to use the `monitor ethernet delay-measurement` command to test the frame delay between MEPs.

## Review Frame Delay Statistics

- The switch maintains a record of previous frame delay measurements

```
lab@switch1> show oam ethernet connectivity-fault-management mep-statistics maintenance-  
domain customer maintenance-association evc1
```

```
...
```

```
Delay measurement statistics:
```

Index	One-way delay (usec)	Two-way delay (usec)
1		191
2		189
3		182
4		185
5		173
6		180
7		189
8		184
9		182
10		205
11		232
12		201
13		187
14		180
15		197
16		178
17		199
18		163
19		193
20		179

```
Average two-way delay      : 187 usec  
Average two-way delay variation: 15 usec  
Best case two-way delay    : 160 usec  
Worst case two-way delay   : 232 usec
```

## Save Frame Delay Measurements

To go back and look at historical frame delay measurements, issue the **show oam ethernet connectivity-fault-management mep-statistics** command.

## Summary

- In this chapter, we:
  - Explained typical OAM features
  - Described the basic operation of LFM
  - Described the basic operation of CFM
  - Configured and monitored Ethernet OAM

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 7-42

### This Chapter Discussed:

- Typical OAM features;
- The basic operation of LFM;
- The basic operation of CFM; and
- Configuration and monitoring of Ethernet OAM.

## Review Questions

1. Which Ethernet OAM protocol allows for setting a loop on a remote switch's interface?
2. What is the difference between an up MEP and a down MEP?
3. What must be true for a MIP to respond to a linktrace message?

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-43

## Review Questions

- 1.
- 2.
- 3.

## Lab 5: Ethernet OAM

- Configure LFM.
- Configure CFM.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 7-44

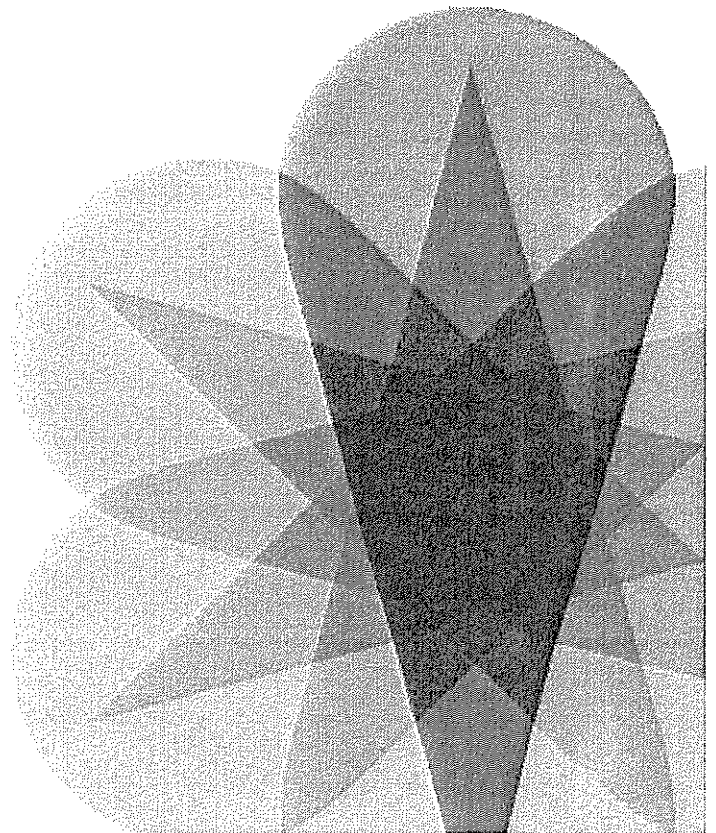
### Lab 5: Ethernet OAM

The slide provides the objectives for this lab.



# **Junos Service Provider Switching**

## **Chapter 8: Ethernet Ring Protection**



## Chapter Objectives

- After successfully completing this chapter, you will be able to:
  - Explain Ethernet Ring Protection
  - Configure and monitor Ethernet Ring Protection

© 2010 Juniper Networks, Inc. All rights reserved.



Worldwide Education Services

[www.juniper.net](http://www.juniper.net) | B-2

### This Chapter Discusses:

- Ethernet Ring Protection (ERP); and
- Configuration and monitoring of ERP.



## Agenda: Ethernet Ring Protection

→ ERP Overview

▪ Configuring and Monitoring ERP

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

[www.juniper.net/t/8-3](http://www.juniper.net/t/8-3)

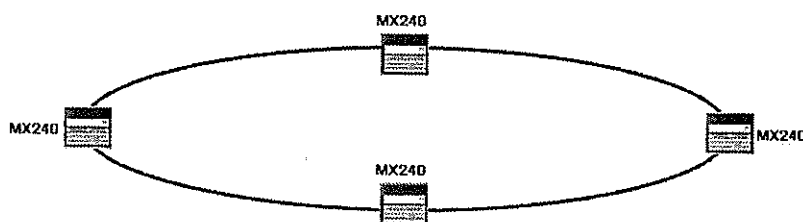
### ERP Overview

The slide lists the topics we cover in this chapter. We discuss the highlighted topic first.

## ERP

### ■ ERP is defined in ITU-T G.8032:

- Designed to provide sub-50 ms, loop-free protection to an Ethernet network
- Ethernet network must be in a ring topology
- Because of the faster failover times, ERP can replace spanning tree protocols on the ring
- Works best in conjunction with connectivity fault management—especially on copper links



© 2005 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

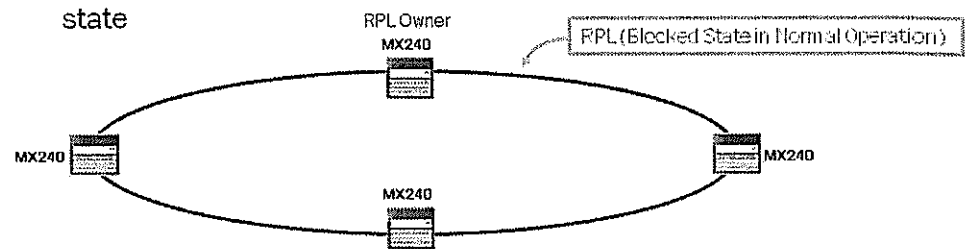
www.juniper.net | B-4

## ERP

Defined in the International Telecommunication Union Telecommunication Standardization (ITU-T) G.8032 recommendation, ERP provides highly reliable, stable, and loop-free protection for Ethernet ring topologies. ERP is a solution for an Ethernet ring where each ring node (switch) connects to two adjacent nodes, participating in the same ring, using two independent links. The minimum number of nodes on a ring is two. Because ERP can provide sub-50 ms, loop-free protection for a ring topology, it can viably replace any spanning tree protocol on the ring. Using an Ethernet fiber ring of less than 1200 km and less than 16 nodes, the switch completion time at the time of failure should be less than 50 ms. Copper links can also be used, but we recommend that you use connectivity fault management (CFM) to help detect failures between nodes.

## Ring Protection Link

- A single link acts as the RPL for the ring:
  - The RPL-owner node controls the RPL
  - During normal operation, the RPL-owner node places the RPL in a blocked state to prevent a loop in the ring topology
  - When a link failure occurs on the ring, the RPL-owner node places the RPL in a forwarding state
    - When the failed link is repaired, the Junos operating system acts in a revertive manner and the RPL owner places the RPL in a blocking state



© 2003 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

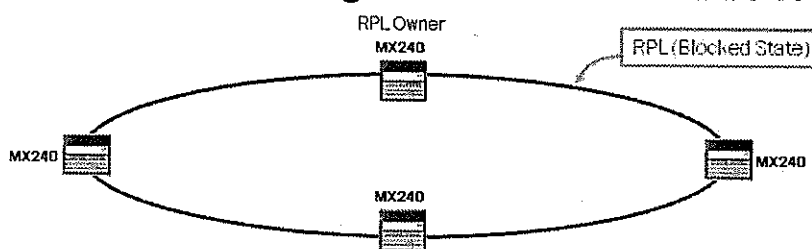
www.juniper.net | B-5

## Ring Protection Link

To protect the Ethernet ring, a single link between two nodes acts as the ring protection link (RPL) on the ring. One of the adjacent nodes, which is referred to as the RPL owner, controls the state of the RPL. During normal operation with no failures (idle state), the RPL owner places the RPL in the blocking state, which results in a loop-free topology. If a link failure occurs somewhere on the ring, the RPL owner places the RPL in a forwarding state until the failed link is repaired. Once the failed link is repaired, the Junos operating system acts in a revertive manner, returning the RPL to the blocking state.

## Node Types

- **RPL-owner node:**
  - Controls the state of the RPL
  - Initiates Ring Automatic Protection Switching messages
- **Normal node:**
  - All other nodes on the ring with no special role
  - Configured to listen to and forward APS messages
  - Generate R-APS messages when a local link failure occurs



© 2013 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-6

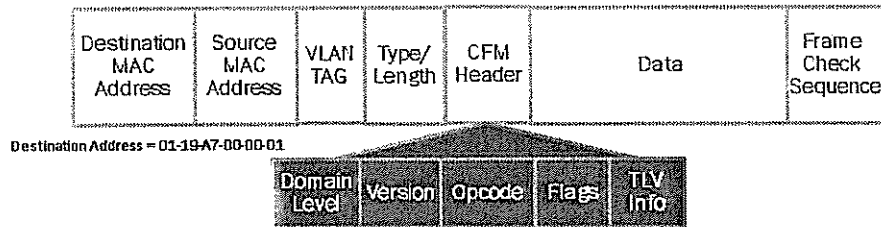
### RPL-Owner Node

The RPL owner controls the state of the RPL. During the idle state, it is the only node that sends periodic Ring Automatic Protection Switching (R-APS) messages to notify the other nodes about the state of the RPL. The next few slides discuss the details of the Automatic Protection Switching (APS) protocol and R-APS messages.

### Normal Node

A normal node is any other node on the ring besides the RPL owner. It listens to and forwards R-APS messages. Also, if a local ring link failure occurs, a normal node signals all other nodes that the failure has occurred using R-APS messages.

## APS Protocol



- APS coordinates the protection actions over the ring:
  - Requires a dedicated channel (a VLAN) to deliver R-APS messages between nodes:
    - A single VLAN is chosen to send and receive R-APS messages; however, all VLANs on the trunk are affected by the APS algorithm
  - Uses CFM frame format
    - Opcode = 40 (R-APS)
    - Flags = 0

© 2015 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-7

## APS

To coordinate the effort of protecting the Ethernet ring, each node participates in the APS. Each of the two ports on each node must be configured for a dedicated channel—a virtual LAN (VLAN) or a bridge domain—to communicate using the APS protocol. Although the APS protocol uses a single VLAN to communicate, the changes in the forwarding state of interfaces that occur as a result of the exchange of R-APS messages affect the entire port of a node (all VLANs). ITU-T G.8032 specifies the use of the CFM frame format as described in the Operation, Administration, and Maintenance (OAM) chapter of this course. To allow differentiation between an R-APS message from a CFM message, an R-APS message uses a destination address of 01-19-A7-00-00-01, as well as an opcode of 40.

## APS Frame Data Fields

Destination MAC Address	Source MAC Address	VLAN TAG	Type/ Length	CFM Header	Data	Frame Check Sequence
-------------------------------	--------------------------	-------------	-----------------	---------------	------	----------------------------

Destination Address = 01-19-A7-00-00-01

Request/ State	Reserved 1	RPL Blocked	Do Not Flush	Status Reserved	Node ID	Reserved 2
-------------------	------------	----------------	-----------------	--------------------	------------	------------

### ■ R-APS-specific data fields include the following fields:

- Request/State (4 bits)
  - Request/State = 1011 represents signal fail
  - Request/State = 0000 represents no request
- RPL Blocked (1 bit)
  - RPL Blocked = 1 represents the RPL is blocked
  - RPL Blocked = 0 represents the RPL is unblocked
- Do not flush (1 bit)
- Node ID (MAC address of the node, informational only)

© 2014 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-8

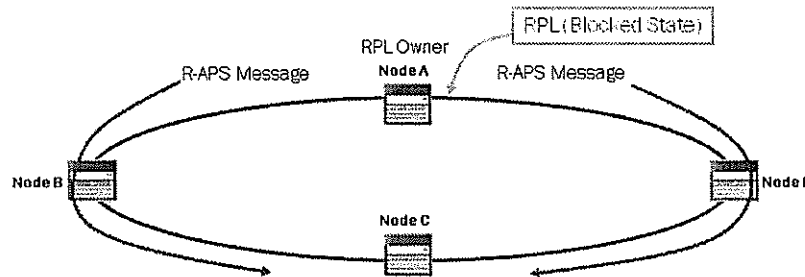
## R-APS Data Fields

Currently, APS has no specified type, length, and values (TLVs). The slide shows the data fields found in an R-APS message. The following list describes each data field:

- *Request/State* (4 bits): Currently only two values are defined. A value of 0000 is used when a node wants to signal that it detects no failure on the ring (No request). A value of 1011 is used when a node wants to signal that an interface has failed (Signal Fail state).
- *Reserved 1* (4 bits): This value is always 0000. This field is reserved for future use.
- *RPL Blocked* (1 bit): Usage for this field is shown on the slide. Only the RPL owner can signal RPL Blocked.
- *Status Reserved* (6 bits): This value is always 000000. This field is reserved for future use.
- *Node ID* (6 octets): This field is a MAC address unique to the ring node.
- *Reserved 2* (24 octets): This value is all zeros. This field is reserved for future use.

## APS—Idle State (Normal Operation)

- RPL-owner node places RPL in the blocked state:
  - RPL owner (Node A) sends R-APS messages out of all ports every 5 seconds
    - Request/State = No request
    - Do not flush = 0 (Flush)
    - RPL Blocked = 1 (RPL is blocked)
  - All other switches place ring ports in the unblocked state



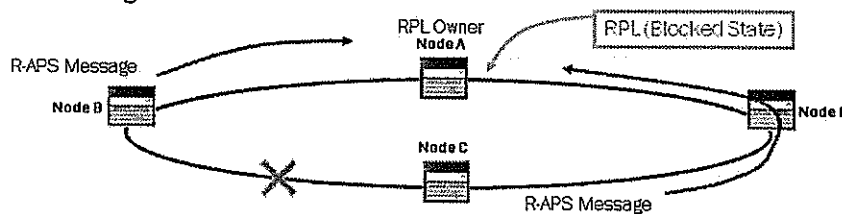
### Idle State

When no failures occur on the Ethernet ring, all nodes are in the idle state. During the idle state, the RPL owner places the RPL in a blocking state. Also, the RPL owner sends periodic (every 5 seconds) R-APS messages that signal that no failure is present on the ring (Request/State = no request), that all switches should flush their MAC tables (Do not flush = 0), and that the RPL is currently blocked (RPL Blocked = 1). All other switches flush their MAC tables once (on the first received R-APS message) while unblocking both of their ring ports.

## APS—Signal Failure (1 of 2)

### ▪ Signal failure:

- Occurs when a failure is detected on an unblocked ring link (CFM failure detection, and so on)
- Node B and Node C:
  - Wait for hold interval to expire (default 0)
  - Switch from idle state to protection state
  - Block failed ports and flush MAC tables
  - Send three R-APS messages in the first 10 ms followed by one every 5 seconds (Request/State = signal fail: Do not flush = 0) until signal failed condition clears



### Signal Failure: Part 1

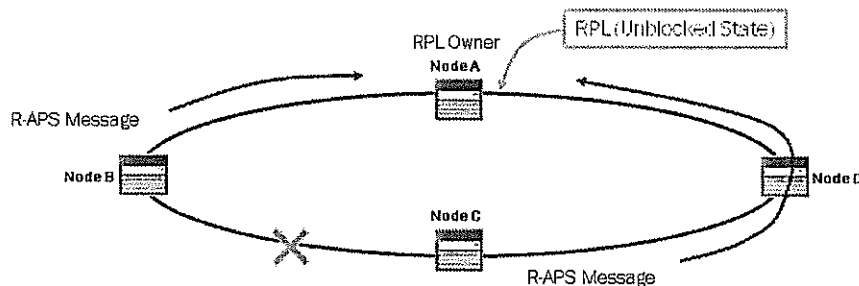
A signal failure occurs when a node detects a failure on a ring port. In the example, Node B and Node C detect a failure on the link between them. The Junos OS does not currently support hold interval. In other words, Node B and Node C react immediately to the failed link. The nodes switch from the idle state to the protection state, block the failed ports, flush their MAC table, and signal to all the other nodes that a signal failure has occurred using R-APS messages. The R-APS messages tell the other nodes that a failure has occurred (Request/State = signal fail) and that the nodes should flush their MAC tables (Do not flush = 0). Node B and Node C continually send R-APS messages every 5 seconds until the signal failure condition clears.



## APS—Signal Failure (2 of 2)

### ▪ Signal failure (contd.):

- All switches except Node B and Node C:
  - Switch from the idle state to the protection state
  - Flush MAC tables and stop sending R-APS messages
- RPL owner (Node A):
  - Unblocks RPL
  - Listens for subsequent R-APS messages from Node C and Node D (subsequent signal fail R-APS messages do not re-trigger flushes)



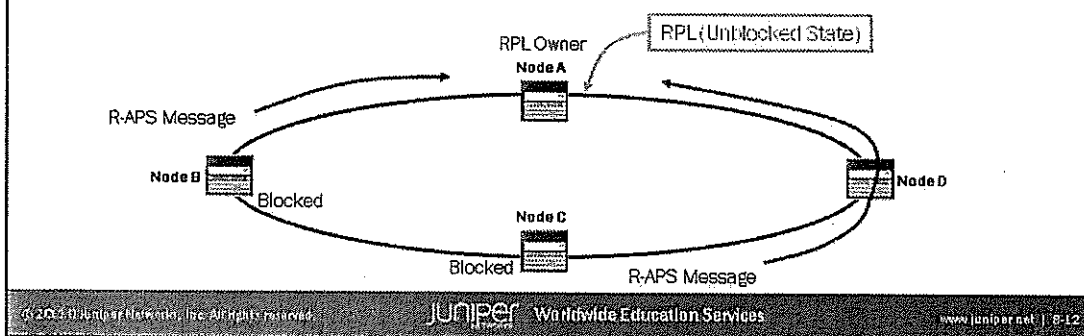
### Signal Failure: Part 2

Upon receiving the signal fail R-APS messages from Node B and Node C, all other nodes (including the RPL owner) switch to the protection state, flush their MAC tables, and stop sending R-APS messages. The RPL owner unblocks the RPL and listens for subsequent R-APS message from Node A and Node B.

## APS—Restoration (1 of 2)

### ▪ Signal fail condition clears on the link between Node B and Node C

- Node B and Node C:
  - Continue to block the previously failed link
  - Send no request R-APS messages (Request/State = no request: Do not flush = 1) and continue until they receive an R-APS message from Node A

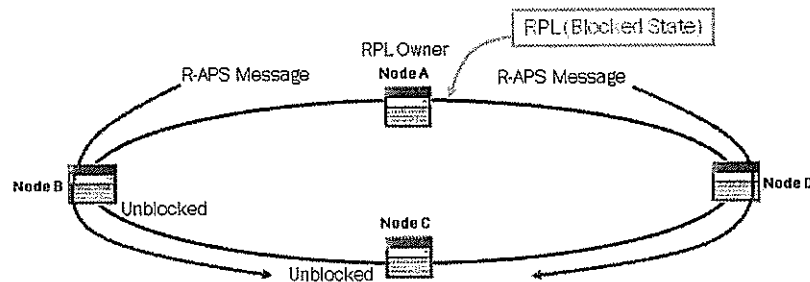


### Restoration of a Failed Link: Part 1

When the failure is repaired between Node B and Node C, they begin sending out new R-APS messages. The R-APS messages tell the other nodes that the failure (Request/State = no request) is no longer present and that they should not flush their MAC tables (Do not flush = 1). Node B and Node C keep the previously failed ports in the blocked state (preventing a loop) until they receive R-APS messages from Node A as described in the following slide.

## APS—Restoration (2 of 2)

- Signal fail condition clears on the link between Node B and Node C (contd.)
  - Node A (after receiving no request R-APS messages):
    - Waits for the restore timer to expire (default is 5 minutes)
    - Blocks RPL and transmits a R-APS message (Request/State = no request; RPL Blocked = 1, Do not flush = 0)
    - All other switches flush their MAC tables and unblock any blocked ring ports
    - All switches change from the protection state to the idle state



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | 8-13

### Restoration of Failed Link: Part 2

Upon receiving the no request R-APS messages from Node B and Node C, Node A starts a restore timer. The default is 5 minutes. You can configure the restore timer in 1-minute steps between 5 and 12 minutes. Once the restore timer expires, Node A blocks the RPL and transmits RPS messages that signal to the other nodes that no failure is present on the ring (Request/State = no request), that the RPL has been blocked (RPL Blocked = 1), and that the other nodes should flush their MAC tables (Do not flush = 0). Once they receive the R-APS messages from Node A, the other nodes flush their MAC tables and unblock any ring ports that had been blocked. At this point, all switches will be in the idle state.

## **Agenda: Ethernet Ring Protection**

- ERP Overview
- Configuring and Monitoring ERP

© 2014 Juniper Networks, Inc. All rights reserved.

**JUNIPER** Worldwide Education Services

www.juniper.net | 8-14

### **Configuring and Monitoring ERP**

The slide highlights the topic we discuss next.

## ERP Configuration

### ■ Configuration options

```

protection-group {
  guard-interval number;
  hold-interval number;
  restore-interval number;
  ethernet-ring ring-name {
    east-interface {
      ring-protection-link-end;
      control-channel channel-name {
        vlan number;
      }
    }
    guard-interval number;
    hold-interval number;
    node-id mac-address;
    restore-interval number;
    ring-protection-link-owner;
    west-interface {
      control-channel channel-name {
        vlan number;
      }
    }
  }
}

```

### ERP Configuration Options

The slide shows all of the options available when configuring ERP. You must configure an `east-interface` and a `west-interface`. You need not configure the two interfaces in any specific order. You can specify global or ring-specific versions of the three intervals (timers) for ERP:

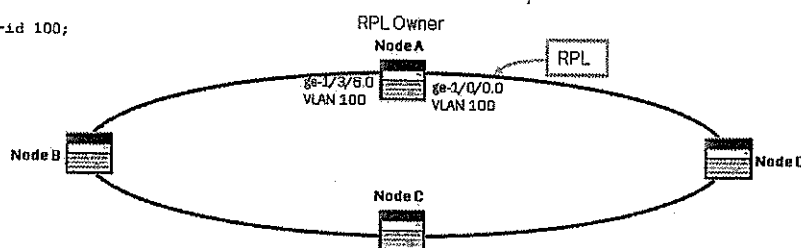
- `guard-interval` (disabled by default): Configurable in 10 ms intervals from 10 ms to 2000 ms. It is used to prevent a node from receiving outdated R-APS messages. Once an R-APS message is received, the guard timer starts. Any R-APS messages that arrive before the expiration of the guard timer drop.
- `hold-interval`: We described this interval on the previous slides.
- `restore-interval`: We described this interval on the previous slides.

## ERP Configuration—RPL Owner

### ■ Node A configuration

```
[edit]
lab@nodeA# show interfaces
ge-1/0/0 {
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list 100;
    }
  }
}
--
ge-1/3/6 {
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list 100;
    }
  }
}
--
[edit]
lab@nodeA# show bridge-domains
bd {
  vlan-id 100;
}
```

```
[edit]
lab@nodeA# show protocols protection-
group
ethernet-ring pg100 {
  ring-protection-link-owner;
  east-interface {
    control-channel {
      ge-1/3/6.0;
      vlan 100;
    }
  }
  west-interface {
    control-channel {
      ge-1/0/0.0;
      vlan 100;
    }
  }
  ring-protection-link-end;
}
```



© 2011 Juniper Networks, Inc. All rights reserved.

JUNIPER

Worldwide Education Services

www.juniper.net | E-16

### RPL Owner Configuration

The slide shows a typical configuration for the RPL owner node. First, you must configure the two interfaces that participate in the Ethernet ring for the APS channel (VLAN and bridge domain). In this case, VLAN 100 is used as the communication channel between nodes. Configure ERP under [edit protocols protection-group]. The following are a few things to note about the ERP configuration for the RPL owner:

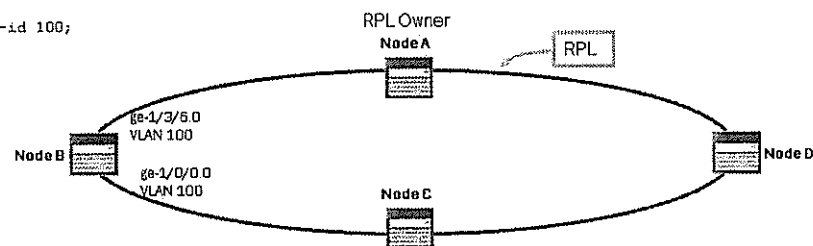
- You must configure the RPL owner node specifically as the ring-protection-link-owner;
- The interfaces are interchangeable with regard to selecting them to act as the west-interface and east-interface as long as you specify one of them as being the ring-protection-link-end; and
- For trunk-mode interfaces, you must also specify the VLAN.

## ERP Configuration—Normal Node

### ■ Node B configuration

```
[edit]
lab@nodeB# show interfaces
ge-1/0/0 {
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list 100;
    }
  }
}
ge-1/3/6 {
  unit 0 {
    family bridge {
      interface-mode trunk;
      vlan-id-list 100;
    }
  }
}
[edit]
lab@nodeB# show bridge-domains
bd {
  vlan-id 100;
}
```

```
[edit]
lab@nodeB# show protocols protection-
group
ethernet-ring pg100 {
  east-interface {
    control-channel {
      ge-1/0/0.0;
      vlan 100;
    }
  }
  west-interface {
    control-channel {
      ge-1/3/6.0;
      vlan 100;
    }
  }
}
```



© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-17

### Normal Node Configuration

The slide shows a typical configuration for a normal node.

## ERP Monitoring

- Display the status of ERP

```
lab@nodeA> show protection-group ethernet-ring ?
```

Possible completions:

aps	Show RAPS PDU information for ethernet ring
interface	Show interface information for ethernet ring
node-state	Show RAPS state machine information for ethernet ring
statistics	Show statistics for ethernet ring

### ERP Status

The slide shows all of the possible commands to monitor ERP. We discuss each one on the next few slides.



## R-APS Information

- Display the status of R-APS protocol data units

```
lab@nodeA> show protection-group ethernet-ring aps
Ethernet Ring Name Request/state No Flush Ring Protection Link Blocked Originator
Remote Node ID
pg100 NR No Yes Yes

lab@nodeA> show protection-group ethernet-ring aps detail
Ethernet-Ring name : pg100
Request/State : NR
No Flush Flag : No
Ring Protection Link blocked : Yes
Originator : Yes
```

## R-APS Information

The command on the slide shows the details of the R-APS messages to which the local node is currently listening or which it is forwarding. Based on the output, you can tell that the local node (Node A) is the RPL owner because the R-APS message originates from it and it is advertising that the RPL is currently blocked.

## ERP Interface Status

### ■ Display the status of ring ports

```
lab@nodeA> show protection-group ethernet-ring interface
```

```
Ethernet ring port parameters for protection group pg100
```

Interface	Control Channel	Forward State	Ring Protection Link End	Signal Failure	Admin State
ge-1/0/0	ge-1/0/0.0	discarding	Yes	Clear	IFF ready
ge-1/3/6	ge-1/3/6.0	forwarding	No	Clear	IFF ready

```
lab@nodeA> show protection-group ethernet-ring interface detail
```

```
Ethernet ring port parameters for protection group pg100
```

```
Interface name           : ge-1/0/0
Control channel name     : ge-1/0/0.0
Ring Protection Link End : Yes
Signal Failure           : Clear
Forward State            : discarding
Interface Admin State    : IFF ready
```

```
Interface name           : ge-1/3/6
Control channel name     : ge-1/3/6.0
Ring Protection Link End : No
Signal Failure           : Clear
Forward State            : forwarding
Interface Admin State    : IFF ready
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | B-20

### Interface Status

The command on the slide shows the state of the local node interfaces in relation to ERP. Note that the Admin State shows that it is IFF ready. This state means that the Ethernet flow forwarding function (the control channel) is available to forward R-APS traffic.

## ERP Node State

- Display the status of the local node

```
lab@nodeA> show protection-group ethernet-ring node-state
Ethernet ring  APS State  Event      Ring Protection Link Owner  Restore Timer  Guard Timer  Operation state
pg100          idle      NR-RB      Yes                          disabled       disabled     operational

lab@nodeA> show protection-group ethernet-ring node-state detail
Ethernet-Ring name      : pg100
APS State                : idle
Event                   : NR-RB
Ring Protection Link Owner : Yes
Restore Timer            : disabled
Guard Timer              : disabled
Operation state          : operational
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-21

### Local Node Details

The command on the slide shows the APS State of the local node, as well as some of the locally configured timer values.

## ERP Statistics

- Display statistics of the local node

```
lab@nodeA> show protection-group ethernet-ring statistics
```

```
Ethernet Ring statistics for PG pg100
RAPS event sent           : 2
RAPS event received       : 2
Local SF happened:        : 0
Remote SF happened:       : 1
NR event happened:        : 1
NR-RB event happened:     : 2
```

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-22

## APS Statistics

The command on the slide shows the quantities of specific events that have occurred. You can reset these values to 0 by issuing the **clear protection-group ethernet-ring statistics group-name name** command.

## Summary

- In this chapter, we:
  - Explained ERP
  - Configured and monitored ERP

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-23

### This Chapter Discussed:

- The fundamentals of ERP; and
- Configuration and monitoring of ERP.

## Review Questions

1. When copper links exist between ring nodes, which other protocol should you use in conjunction with ERP?
2. What can cause the RPL owner to stop sending periodic R-APS messages?
3. Once a signal fail condition clears, how long does the RPL owner wait until it begins sending periodic R-APS messages? What is the name of the timer?

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-24

## Review Questions

- 1.
- 2.
- 3.

## Lab 6: Ethernet Ring Protection

- Configure ERP.
- Monitor ERP.

© 2010 Juniper Networks, Inc. All rights reserved.

JUNIPER Worldwide Education Services

www.juniper.net | 8-25

### Lab 6: Ethernet Ring Protection

The slide provides the objectives for this lab.





# Appendix A: Acronym List

---

AIS	alarm indication signal
APS	Automatic Protection Switching
ATM	Asynchronous Transfer Mode
BCB	backbone core bridge
BDI	Backward Defect Indicator
BEB	backbone edge bridge
BID	Bridge ID
BPDU	bridge protocol data unit
B-TAG	Backbone VLAN tag
BUM	broadcast, unicast with unknown destination, and multicast
B-VLAN	Backbone VLAN
CBP	Customer Backbone Port
CC	continuity check
CFI	Canonical Format Indicator
CFM	connectivity fault management
CIST	common and internal spanning tree
CLI	command-line interface
CoS	class of service
CSMA/CD	carrier-sense multiple access with collision detection
CST	common spanning tree
C-TAG	customer VLAN tag
C-VLAN	customer VLAN
DEI	Drop Eligibility Indicator
DPC	Dense Port Concentrator
DSAP	destination service access point
E-LAN	Ethernet LAN
E-Line	Ethernet Line
ERP	Ethernet Ring Protection
EVC	Ethernet virtual connection
FDI	Forward Defect Indicator
GRE	generic routing encapsulation
GUI	graphical user interface
ICMP	Internet Control Message Protocol
IEEE	Institute of Electrical and Electronics Engineers
IP-IP	IP over IP
IRB	integrated routing and bridging
I-SID	Backbone Service Instance ID
I-TAG	Backbone Service Instance tag
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union Telecommunication Standardization
JNTCP	Juniper Networks Technical Certification Program
LFM	link fault management
MAC	media access control
MAN	Metropolitan Area Network
MEF	Metro Ethernet Forum
MEP	maintenance association end point
MIP	maintenance association intermediate point
MISTP	Multiple Instance STP
MSO	multiple service operator

MST	multiple spanning tree
MSTI	multiple spanning tree instance
MSTP	Multiple Spanning Tree Protocol
OAM	Operation, Administration, and Maintenance
OAMPDU	OAM protocol data unit
OSI	Open Systems Interconnection
PBB	provider backbone bridge
PBBN	provider backbone bridged network
PBN	provider bridged network
PEB	Power and Ethernet Board
PFE	Packet Forwarding Engine
PIM	Physical Interface Module
PIP	Provider Instance Port
PVC	permanent virtual circuit
PVST+	Per-VLAN Spanning Tree Plus
Rapid-PVST+	Rapid Per-VLAN Spanning Tree Plus
R-APS	Ring Automatic Protection Switching
RE	Routing Engine
RPL	ring protection link
RST BPDU	Rapid Spanning Tree BPDU
RSTP	Rapid Spanning Tree Protocol
SLA	service-level agreement
S-TAG	service VLAN tag
STP	Spanning Tree Protocol
S-VLAN	service VLAN
TCN	topology change notification
TDM	time-division multiplexing
TLV	type/length/value
TPID	Tag Protocol Identifier
TTL	time to live
UCA	User Customer Address
UNI	user-to-network interface
VLAN	virtual LAN
VPLS	virtual private LAN service
VSTP	VLAN Spanning Tree Protocol

# Appendix B: Answer Key

---

## Chapter 1: Course Introduction

This chapter contains no review questions.

## Chapter 2: Carrier Ethernet

1.

Service providers can offer more than one service to an individual customer over a single access port with carrier Ethernet, and enterprise customers need only hire Ethernet experts to manage the entire network.

2.

The three prominent Ethernet standards organizations are the MEF, the IEEE, and the ITU.

3.

An MX Series device can provide provider bridging, provider backbone bridging, and OAM.

## Chapter 3: Ethernet Switching and Virtual LANs

1.

A bridge domain allows you to specify which VLANs will be used for Layer 2 switching.

2.

A bridge generally forwards multicast frames out of every interface except for the one from which they were received.

3.

A IRB interface eliminates the need for an external router to route between VLANs. It acts as an IP gateway for the hosts attached to a VLAN.

4.

The match condition used in a Layer 2 firewall filter to match on 802.1p priority bits is `learn-vlan-1p-priority`.

## Chapter 4: Virtual Switches

1.

For multiple routers, you can configure virtual-router routing instances. For multiple switches, you can configure virtual-switch routing instances.

2.

You must list the interface at the `[edit routing-instances vs1]` level of the hierarchy to ensure that it appears as part of the `vs1` virtual switch.

3.

By default, you can find the routes associated with IRB interfaces in the `inet.0` routing table.

## Chapter 5: Provider Bridging

1.

The service provider and potentially thousands of customers must share a limited number of VLAN IDs when a service provider uses IEEE 802.1Q VLANs to provide LAN service. Also, each service provider switch must learn the MAC addresses of its customers.

2.

Three VLAN tag operations that a switch can perform on a frame are pop, push, and swap.

3.

The I-component translates S-VLAN tags to I-SIDs.

## Chapter 6: Spanning-Tree Protocols

1.

STP is a simple Layer 2 protocol that prevents loops and calculates the best path through a switched network that contains redundant paths. STP automatically rebuilds the tree when a topology change occurs.

2.

*Blocking* drops all data packets and receives BPDUs. *Listening* drops all data packets and listens to BPDUs. *Learning* does not forward data traffic but builds the MAC address table. *Forwarding* forwards all data traffic and transmits and receives BPDUs. *Disabled* does not participate in STP (administratively disabled).

3.

The basic steps involved in building a spanning tree are that switches exchange BPDUs, each individual bridge elects a single root bridge based on the received BPDUs, and the bridges determine the role and state of individual ports, at which time the tree is considered fully converged.

4.

RSTP improves link-convergence time significantly over STP. MSTP supports up to 64 regions and allows load balancing over redundant links—STP and RSTP do not. VSTP allows for per-VLAN spanning trees.

## Chapter 7: Ethernet OAM

1.

LFM allows for setting a remote loop.

2.

A down MEP expects to find neighboring MEPs downstream. An up MEP expects to find neighboring MEPs upstream.

3.

The MIP must be configured at one level below the MEP that initiated the linktrace message.

## Chapter 8: Ethernet Ring Protection

1.

You should use CFM with ERP for faster protection times.

2.

An RPL owner stops sending its own R-APS messages when it receives an R-APS message for another node that specifies signal failure.

3.

The RPL owner waits until the Restore Timer has expired. The default is 5 minutes.



