SR-OS Fundamentals

Module 1: The High Leverage Network (HLN)
FP3
Service router supported solutions

IPD Development

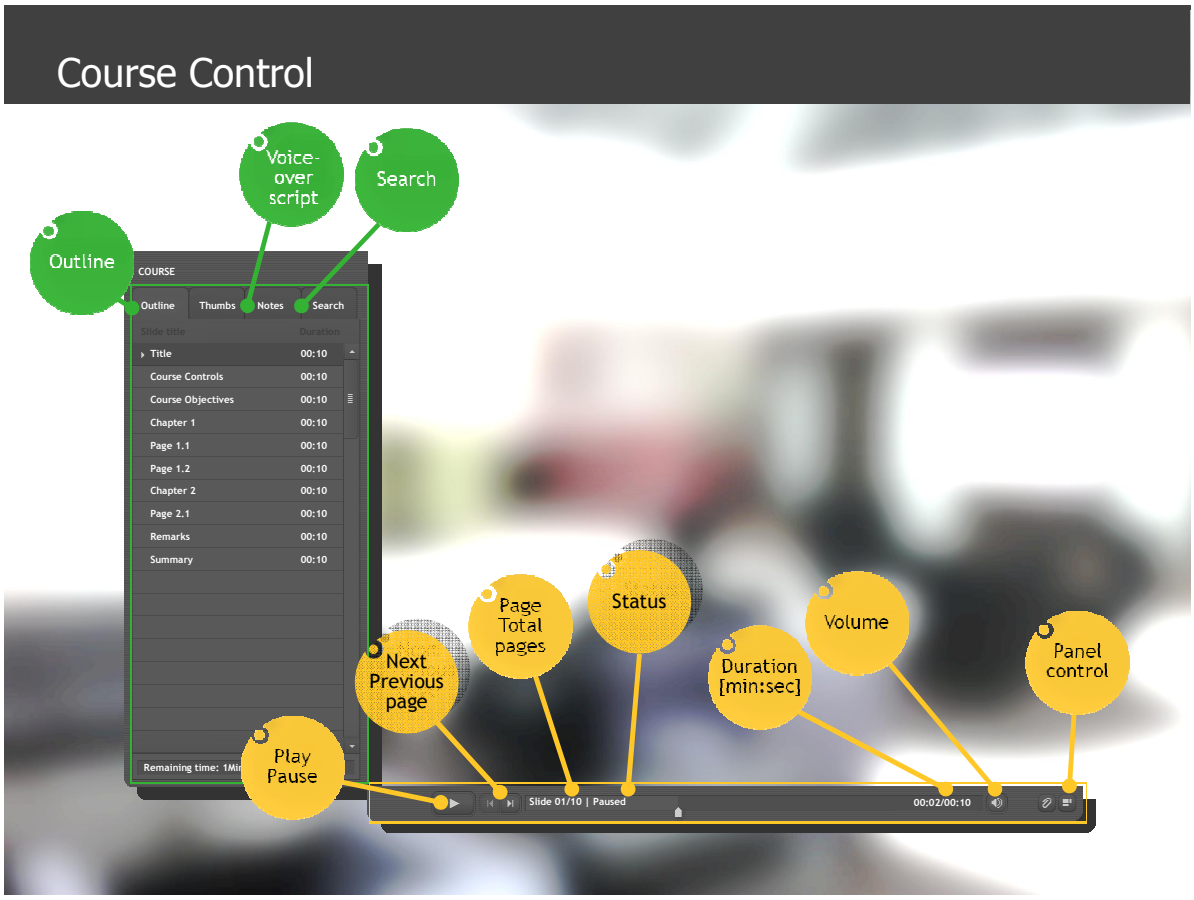
Available as PDF

Welccome to the first module of the SR-OS fundamentals course.

The SR-OS or Service router operating system is an industry leading product platform that offers superior multiservice routing capabilities. The operating system is implemented on the complete range of Alcatel-Lucent IP/MPLS products.

The technical innovations by Bell labs, drive forward the industry-leading capacity, interface density, service-richness and service scale of the market-leading service router portfolio. Leveraging breakthrough FP3 400Gigabit per second technology, Alcatel-Lucent's 7750 SR multi-terabit platform leaps forward into 100Gigabit per second networking. It enables service providers to deal with the dramatic increases in bandwidth that they need to deliver the continually escalating video and data content across business, residential, and mobile service solutions.

Module 1 introduces the High Leverage Network or HLN, the newly developed FP3 chipset and provides an overview of the services supported inside the HLN network.



Before we get started, please have a quick look at the controls of this web-based training:

The side panel allows you to see the outline, display thumbnail images from the pages, view the voice-over script, or search for a topic.

The control bar along the bottom allows you to play or pause, go to next or previous page, indicates current page and total pages and shows the course status. You will also see audio length by minute and seconds, and a volume and mute control. In addition, the panels can be minimized.

Table of Contents

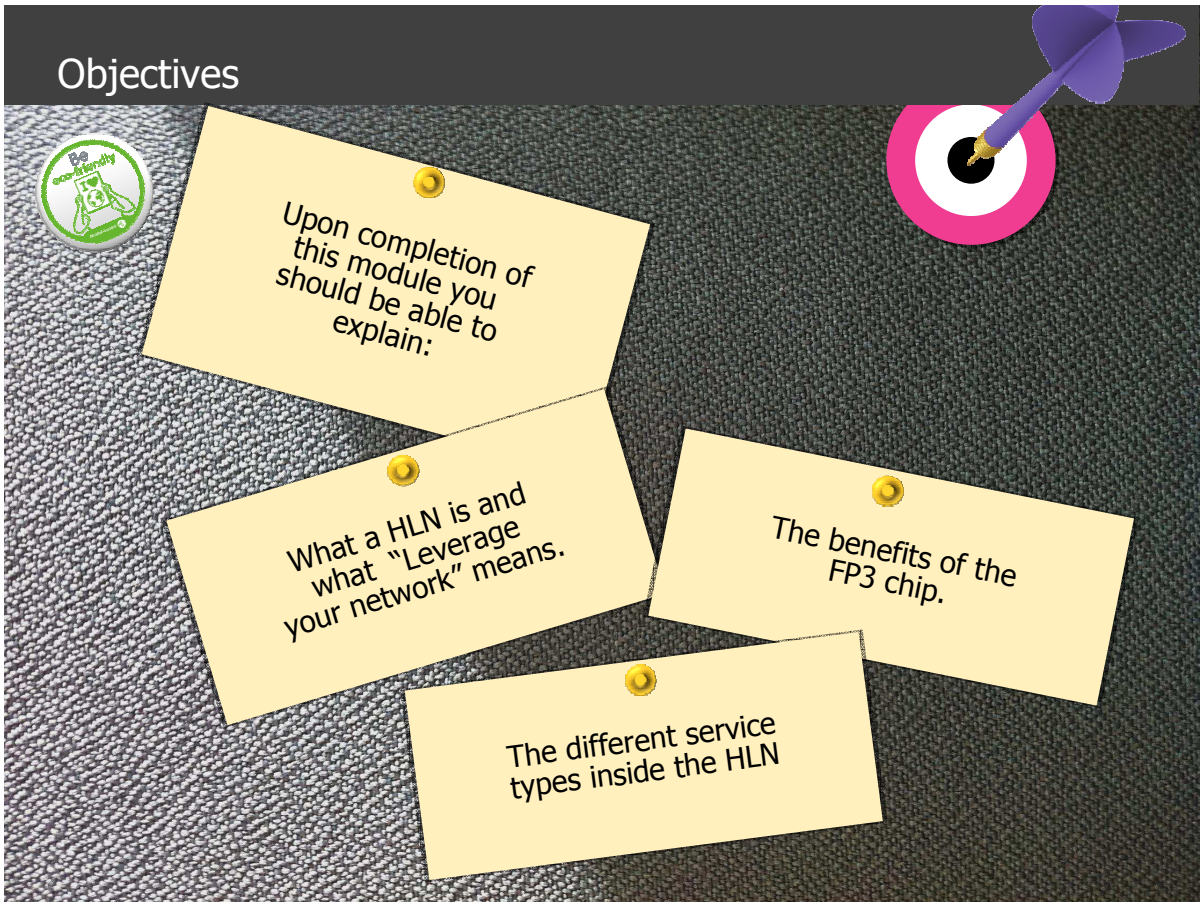
Module 1 is divided into three sections.

Section 1 introduces the High Leverage Network, better know as HLN.

Section 2 introduces the latest FP3 chip or Flexible Path version 3 chip.

Section 3 provides an overview of the service router supported solutions inside the HLN.

Where section 1 is focusing primarily on the need for a high leverage network, section 2 and 3 addresses more the technical aspects of the solutions inside the high leverage network.

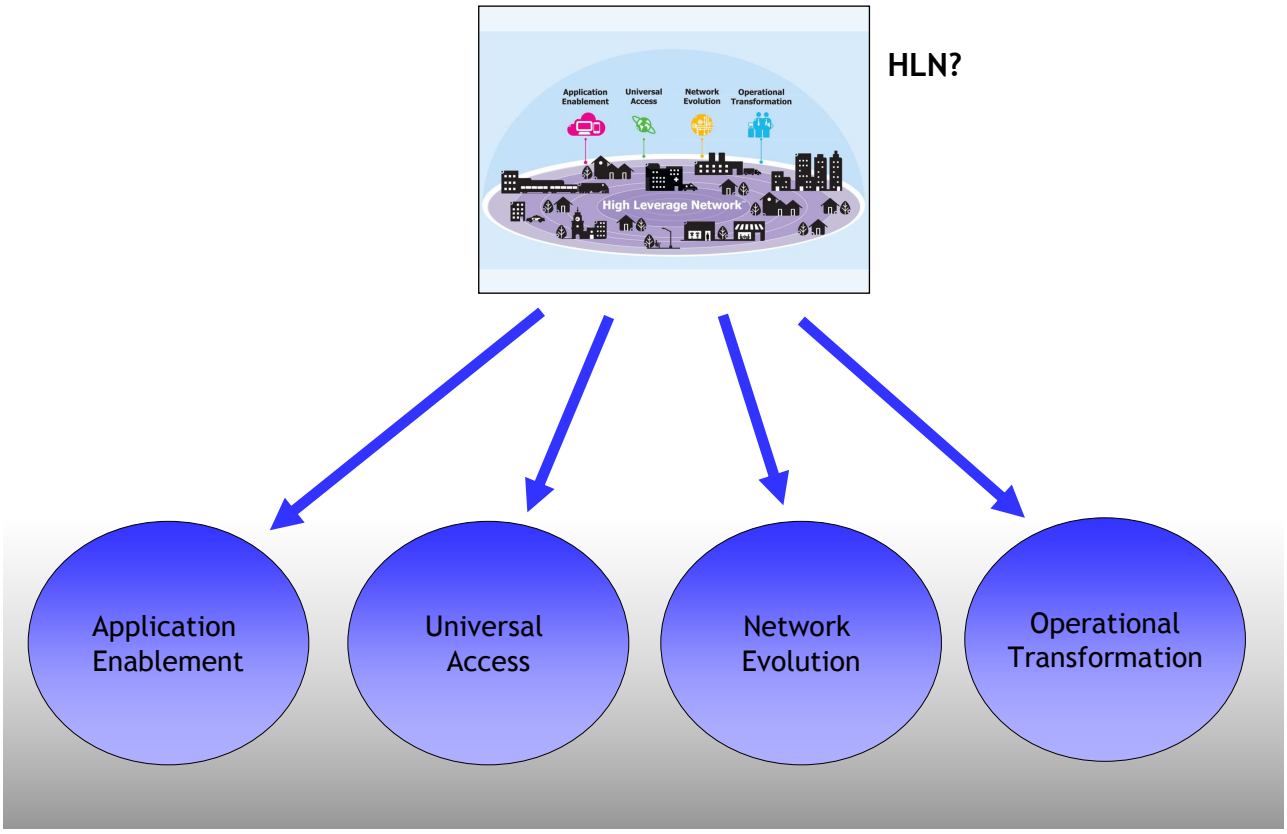


By the end of this module you will be able to explain what is a high leverage network and what does "Leverage your network" mean, the benefits of the FP3 chipset and the different service types inside the HLN.

High Leverage Network (HLN)

Section1: The High Leverage Network.

What is HLN?



Most of today's networks have been converged to an all-IP/MPLS network. In this way, network operators have reduced their cost dramatically. But driven by all types of new video content the network faces an enormous bandwidth explosion. The demand for online content is soaring and end-users have now significantly higher bandwidth requirements and quality of experience expectations. Much of this content, which is preliminary video is free for end users and paid by the advertisers. But for the network operator, that has to expand his network continuously, it means an increase in both the CAPEX and OPEX.

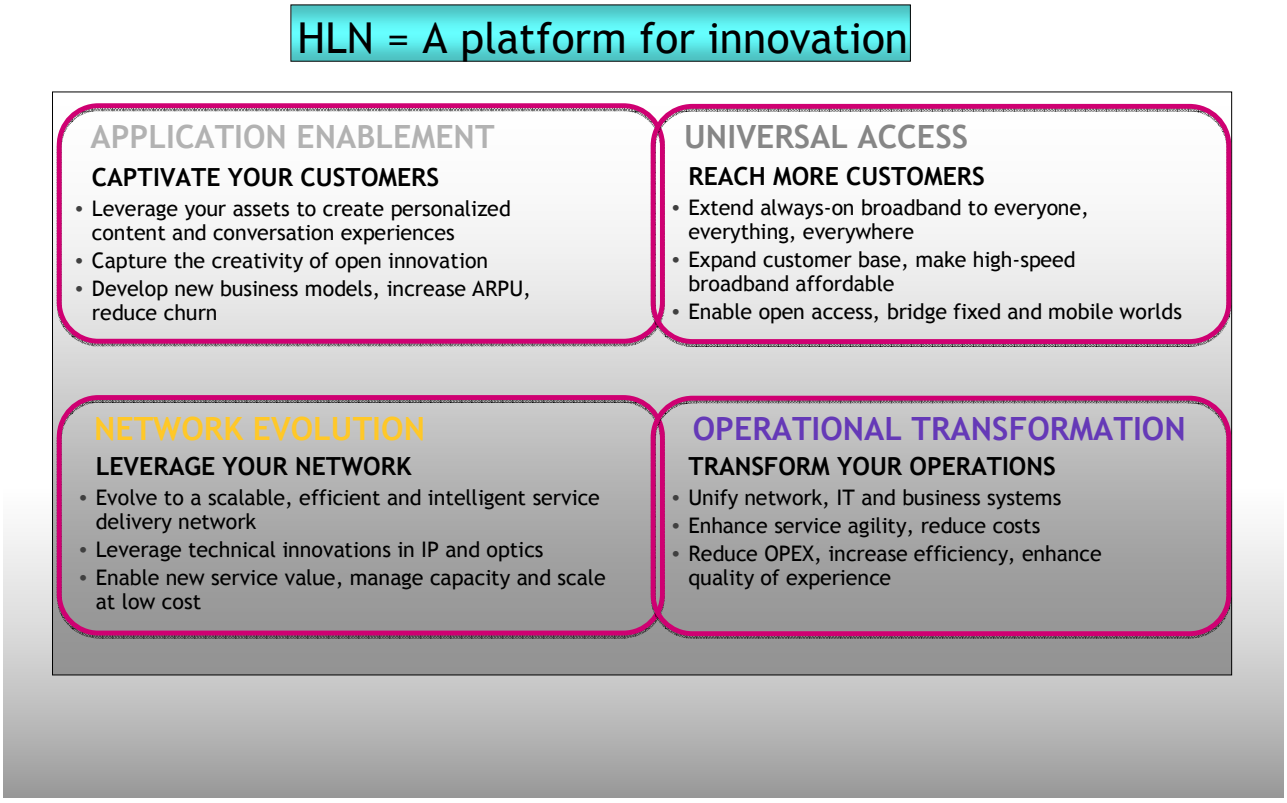
The challenge of the network operator today is to monetize these bandwidth hungry services so that every bit is worth its money.

The answer is building a High Leverage Network.

An High Leverage Network or HLN is a converged, scalable and intelligent all-IP network. It offers distributed service intelligence, broadband access, scalable and efficient IP transport and at the lowest bit cost. But at the same time, it is able to support innovative revenue generating services and business models.

HLN best describes itself as application Enablement, universal Access, network evolution and operational transformation.

What is HLN?



Let's have a closer look at the four pillars of an HLN network.

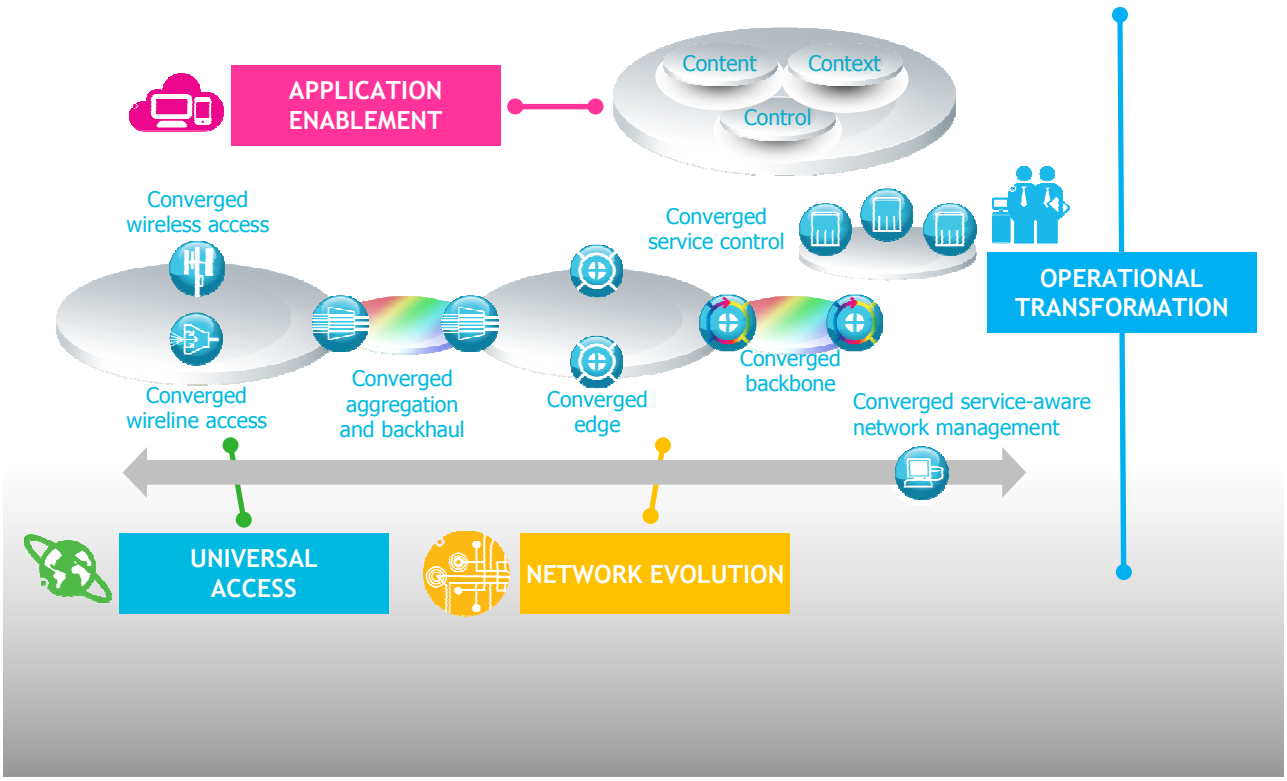
Application Enablement is about delivering more value to end users by delivering new services that enhance the quality of experience. For example, by delivering location based services based on end user preferences.

Network Evolution is about leveraging the intelligence in the network to enhance end user services, increase revenues and capture new business. For example, by delivering enhanced multimedia and multi-screen services. Network Evolution also addresses the data explosion caused by the growth of video and the mobile internet by providing a converged more scalable and efficient all-IP network.

Universal Access is all about bringing broadband to more people and to more devices. The message is about delivering broadband to everyone and everything, wired or wireless, anywhere and on any device. Its about broadband for emerging markets and the developed world, in the most densely populated areas or the most remote.

Operational Transformation is about transforming network, IT and business operations to reduce cost and complexity. It helps service providers to extend the life of network assets and to deliver new services faster while reducing operating costs, making the network more efficient and transforming customers' quality of experience.

The HIGH LEVERAGE NETWORK



For simplicity and for modular evolution, convergence to the single common platform can be split up into seven distinct domains.

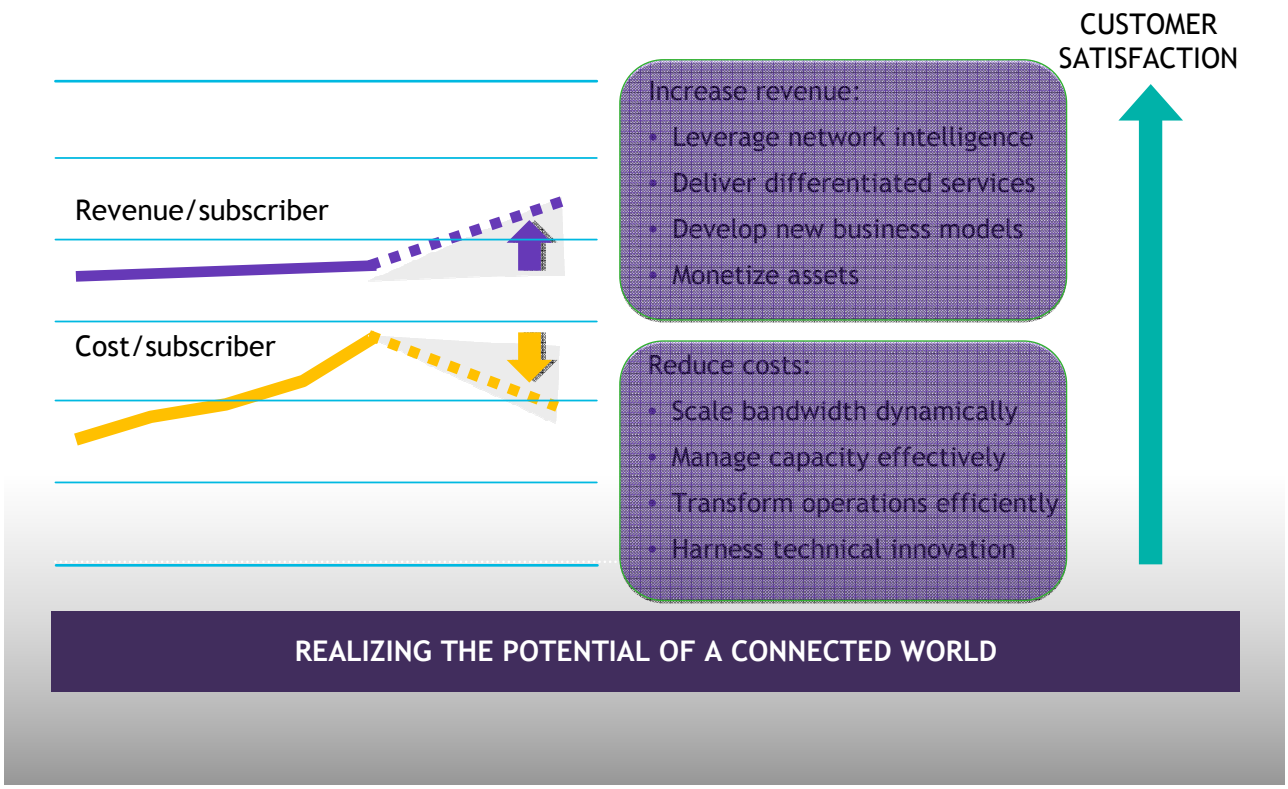
Converged wireless access, converged wireline access, converged aggregation and backhaul, converged Edge, converged backbone, converged service control and converged service-aware network management.

Each of these domains provide a specific set of capabilities and functions.

Let's map the four pillars to the blueprint of the High Leverage Network.

Application Enablement interfaces to the network via its three key enablers; control, content and context

Universal Access and Network Evolution map closely to the network infrastructure while operational transformation spans across the HLN and the three areas of Application Enablement, Universal Access and Network Evolution.



The High Leverage Network addresses two key challenges. How to reduce cost and how to generate new revenues.

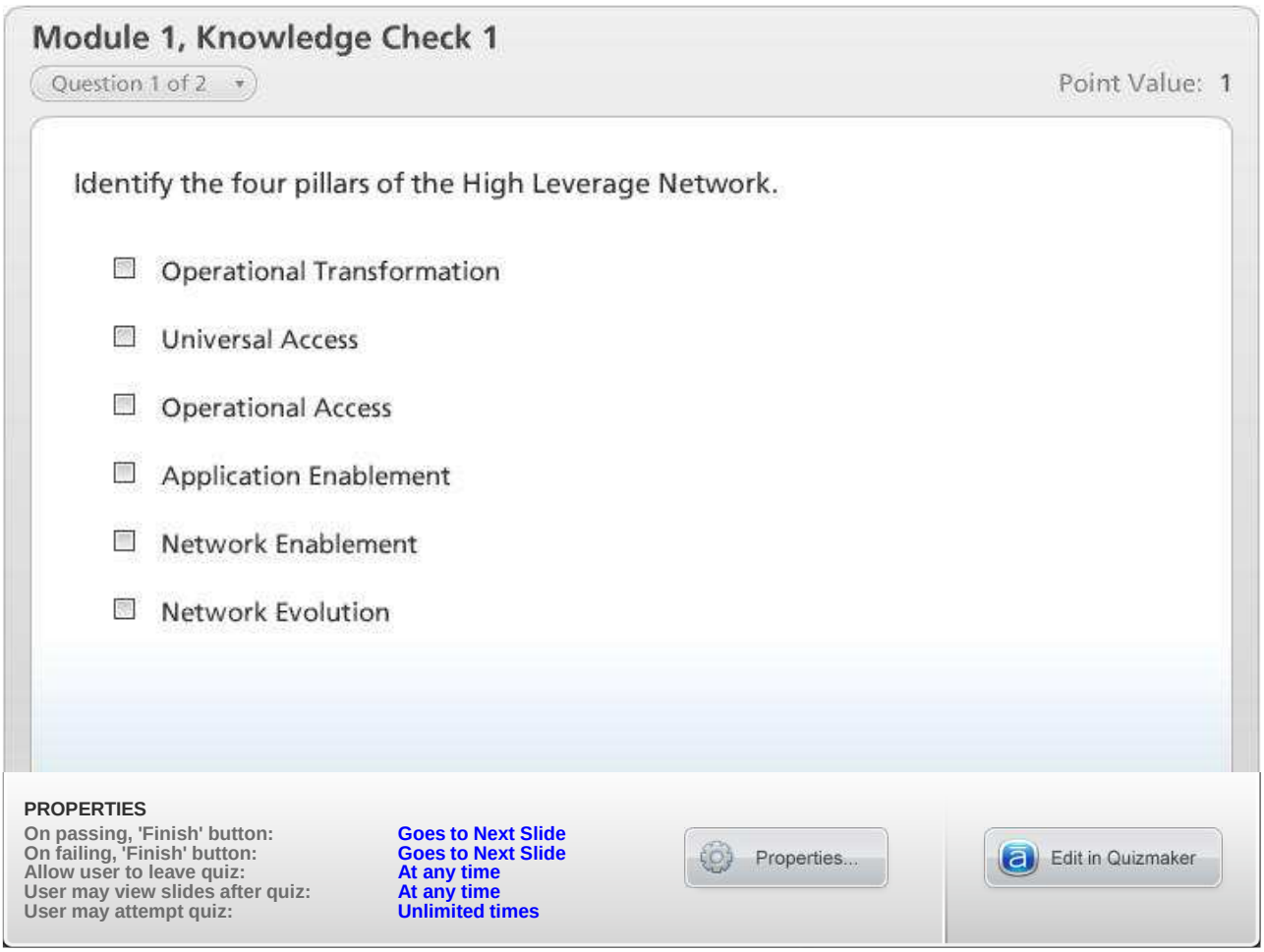
So how does the HLN reduce cost and generate revenues?

Generating new revenues can be achieved by leveraging the intelligence embedded in the network to enhance end user services. For example, by delivering application assured services to enterprises.

Increasing revenue can be enabled by offering differentiated services such as cloud-based services instead of more traditional services.

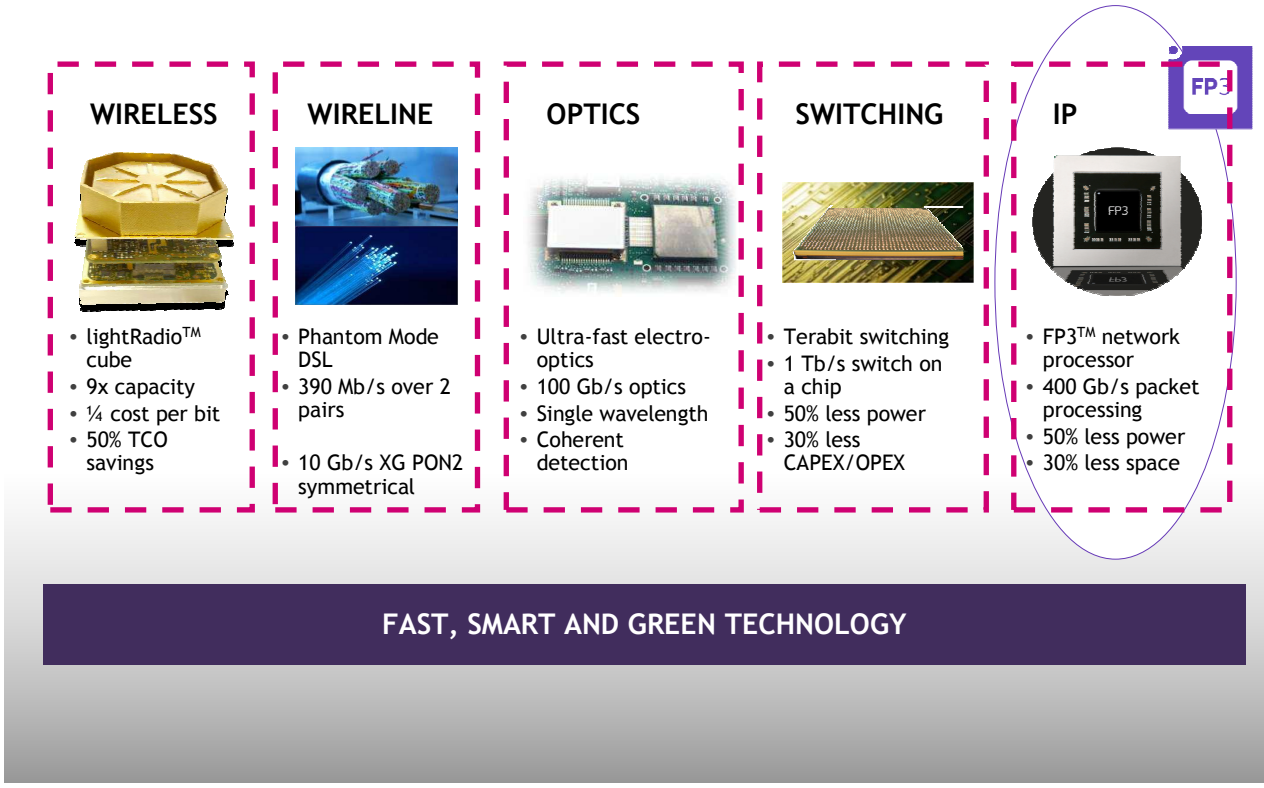
Or by developing new business models by exposing selected network capabilities and assets in a managed and controlled way.

Reducing the cost is achieved by scaling the bandwidth and manage the capacity effectively. And at the same time transforming operation and harnessing technical innovation.



FP3

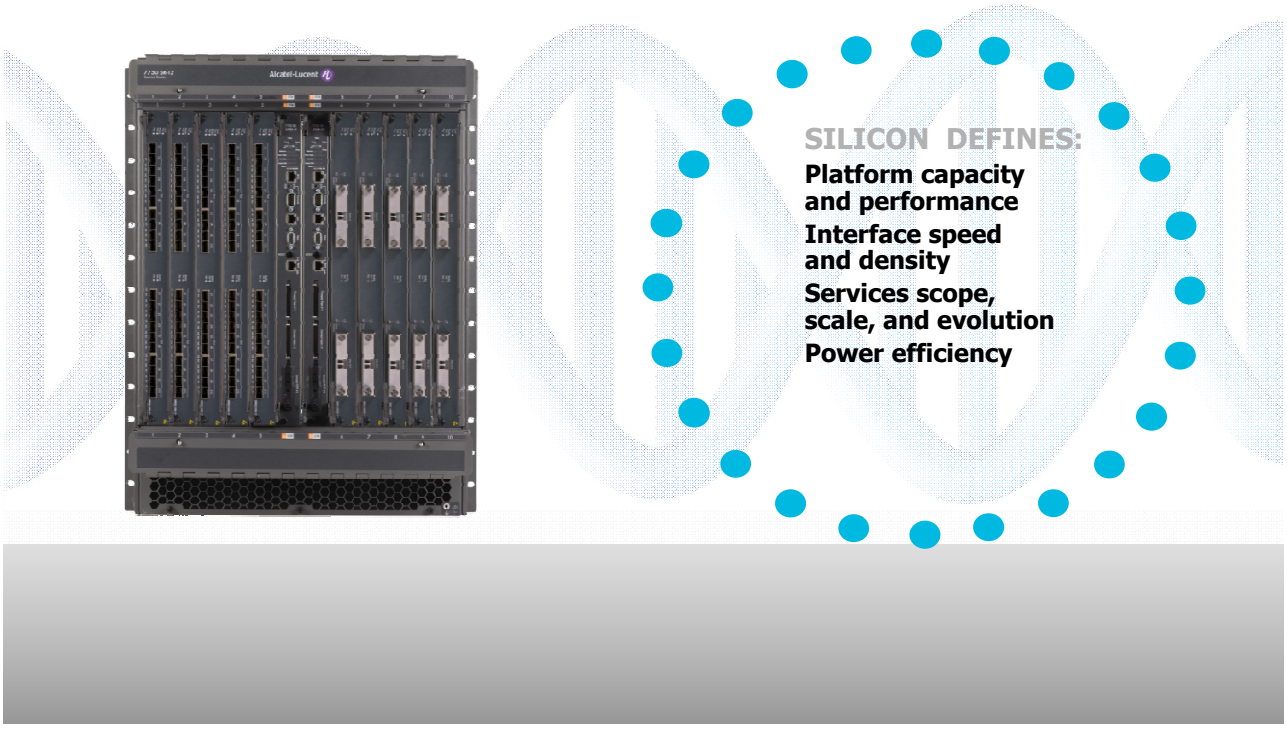
Section2: The FP3 chip.



The world's IP networks are growing faster and faster. On-demand video is driven immense bandwidth growth. Smartphone and tablets have driven the number of connected devices exponentially. There is simply no way to keep up with the growth of the internet without continuous innovation. Innovation happens in all domains of the telecom business. In Wireless and wireline, optics, switching and IP.

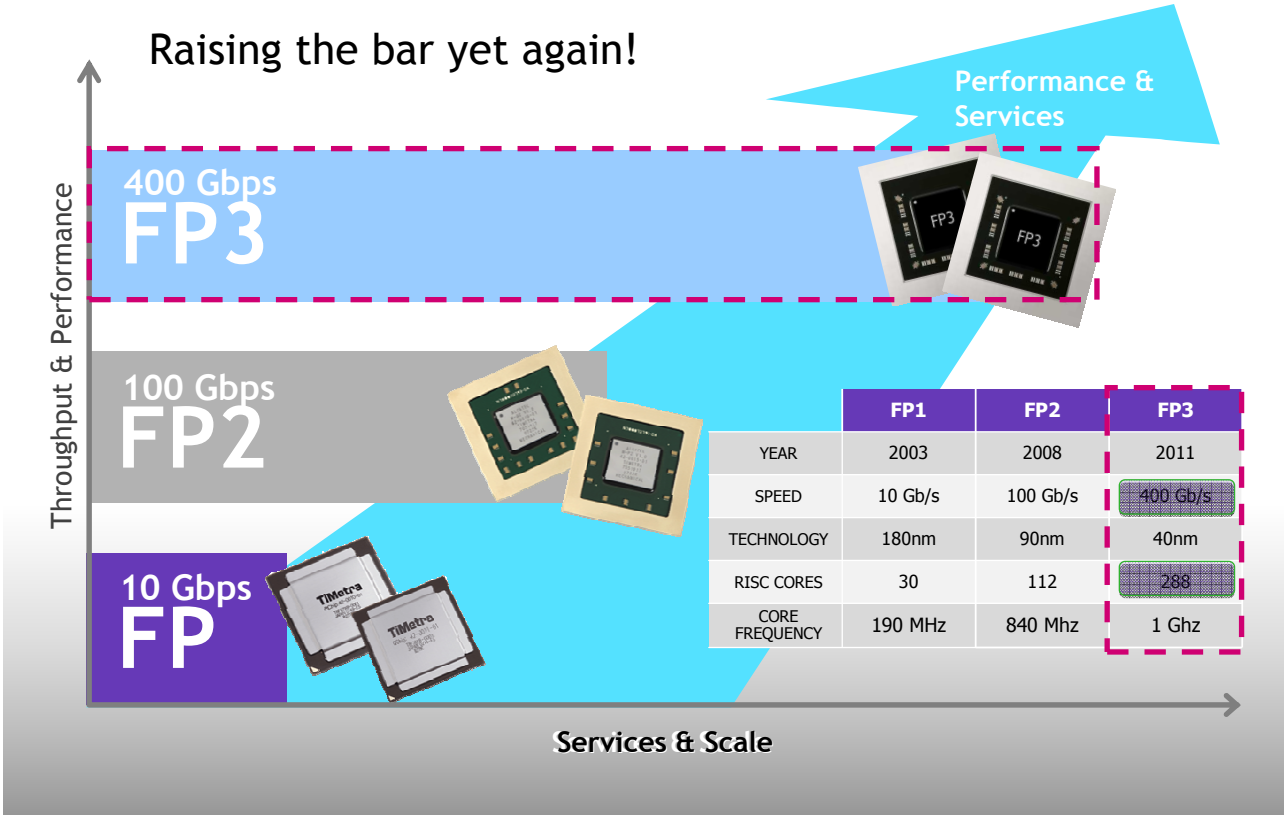
In 2008 Flex Path 2 was announced, a 100 Gig network processor that broke new performance barriers for routing platforms. Today, Alcatel-Lucent has done it again with its new third generation service routing silicon, the flex path 3. The world fastest network processor with less space and less power consumption.

THE IMPORTANCE OF NETWORK PROCESSOR SILICON
THE 'DNA' OF IP ROUTERS



Silicon is the DNA of an IP/MPLS service router. Used in the network processing chipset, silicon determines how the platform behaves, the features it provides and its potential to evolve. With the third generation of in-house designed silicon, Alcatel-Lucent is able to ensure feature consistency from generation to generation, while at the same time massively increasing the performance and service scale of the chipset.

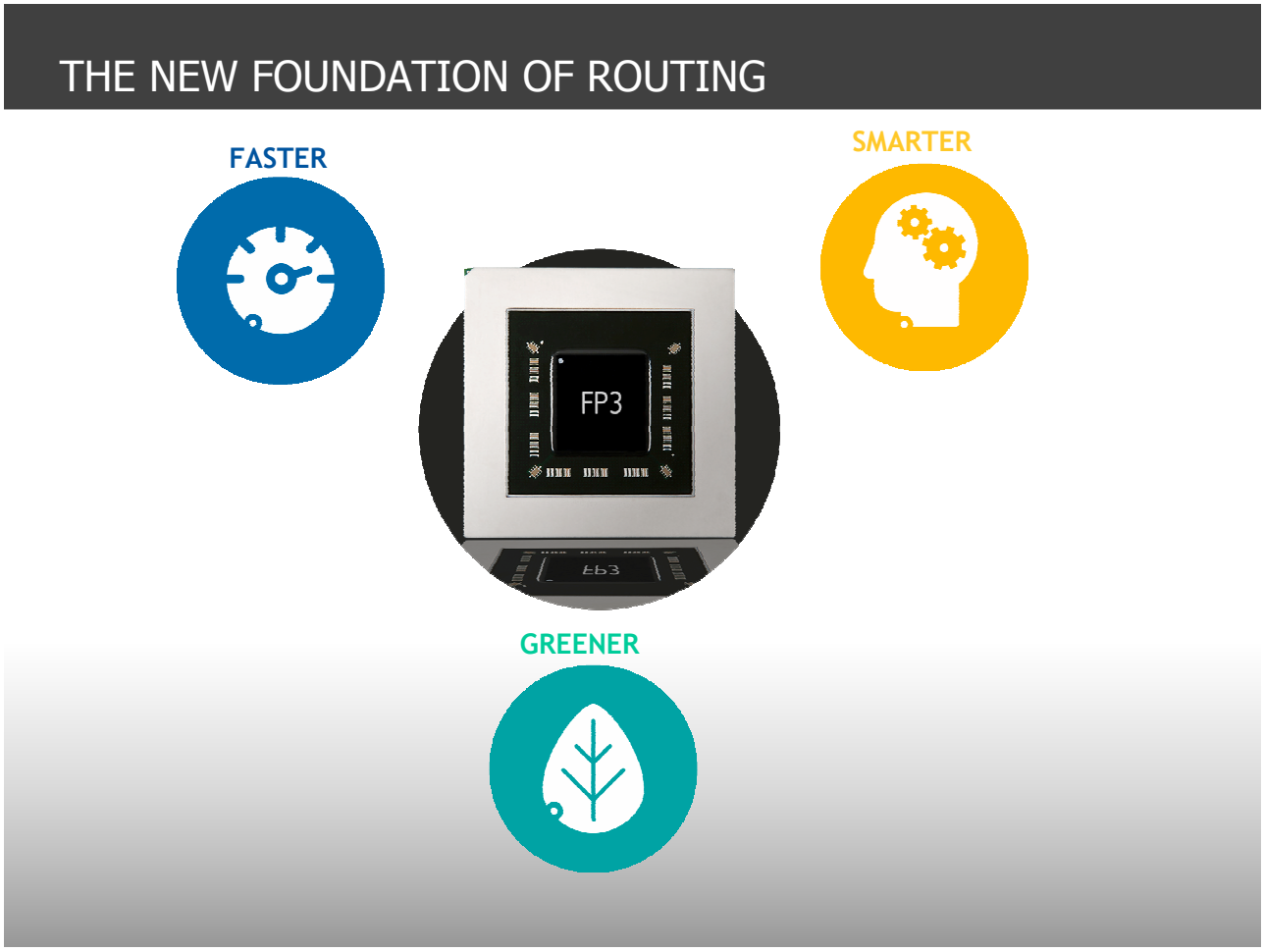
3rd GENERATION OF IN-HOUSE DESIGN



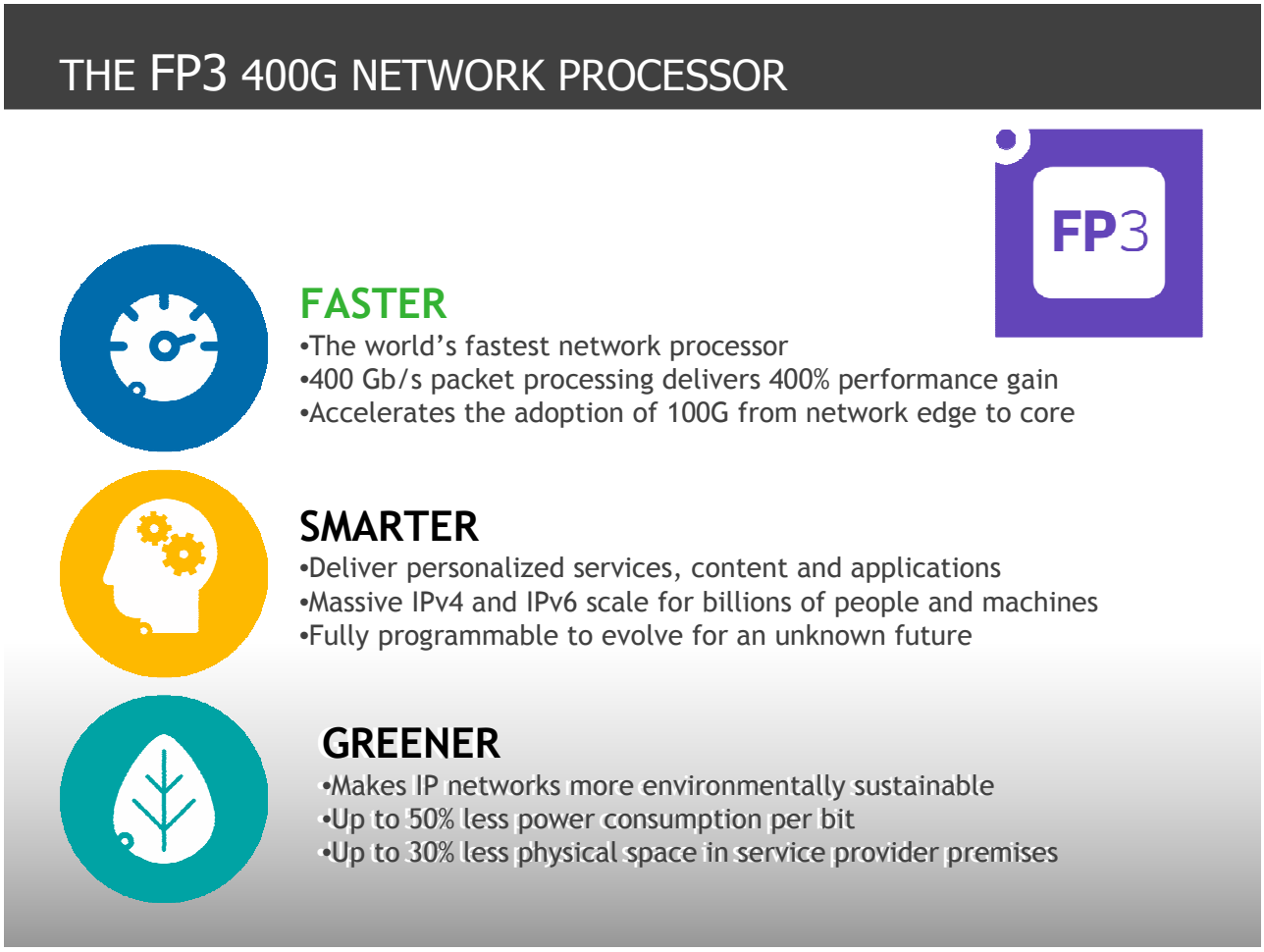
	FP1	FP2	FP3
YEAR	2003	2008	2011
SPEED	10 Gb/s	100 Gb/s	400 Gb/s
TECHNOLOGY	180nm	90nm	40nm
RISC CORES	30	112	288
CORE FREQUENCY	190 MHz	840 Mhz	1 Ghz

Each generation of FP chips has built on the features of the previous generation, providing feature continuity and consistency, while increasing speed and adding functionality. This feature continuity allows the SR portfolio to seamlessly support all generations FP processors while benefiting from the latest performance improvements.

The FP3 is the world fastest network processor with 288 risk processor cores on a single chip. It is able to process packets at 400Gigabits per second line rate. Four times faster than current chipsets in the market. And it still supports deep packet operation which is fundamental for service routers. In fact, FP3 supports more features even though it is four times faster than the FP2.



The FP3 chip is all about being faster, smarter and greener.



The FP3 network processor can quadruple the speed of the most advanced networks available today. It unlocks value for the next generation of online applications, entertainment and communications. It accelerates the adoption of 100G from network edge to core.

Ensuring networks can scale to address the challenge of booming bandwidth demand is not sufficient. Networks must also create value to deliver collaborative multimedia applications, new home entertainment service offerings and cloud-based services. FP3 is capable of delivering personalized services, content and applications with massive IPv4 and IPv6 scale for billions of people and machines. It's a fully programmable chip to be 100% future proof.

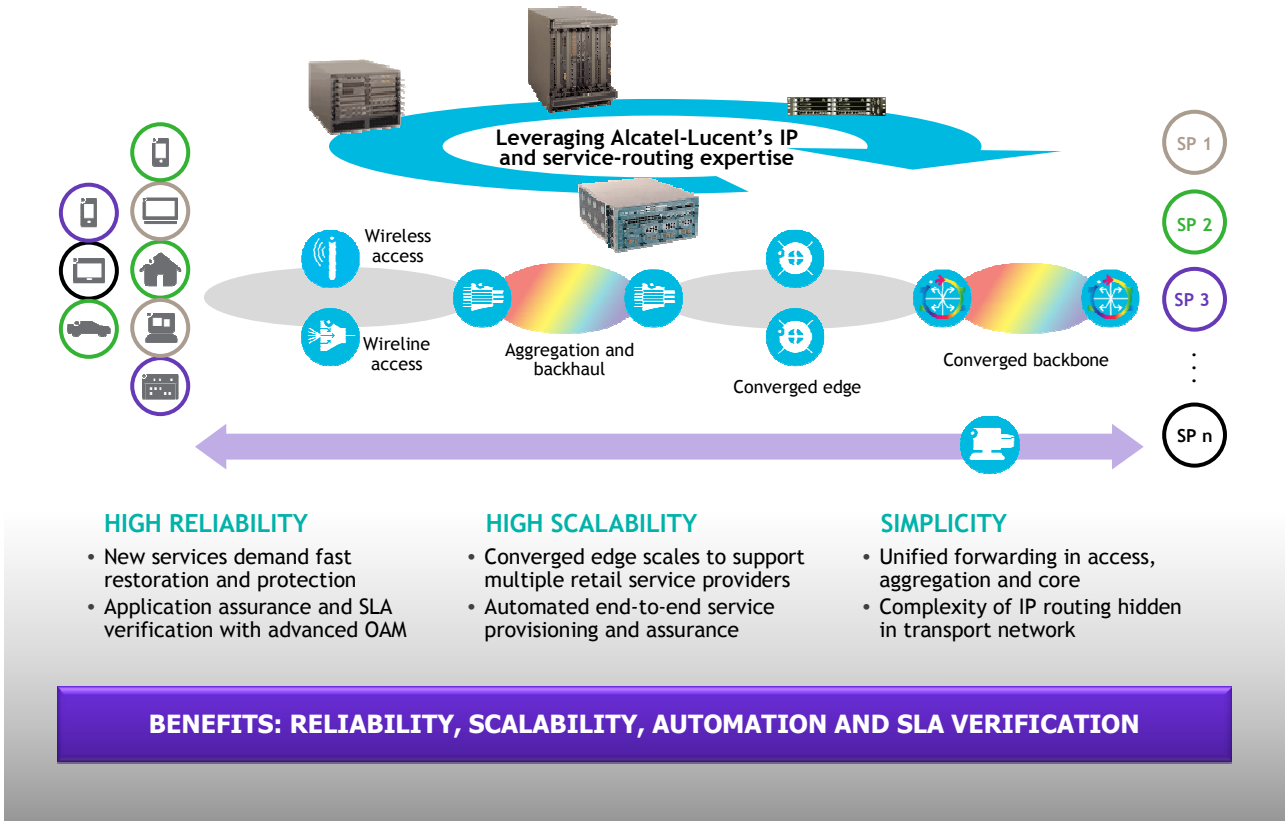
Information and communications industries create 2% of our greenhouse gases. So power efficiency gains in routing equipment are mandatory. The FP3 uses the latest silicon design techniques reducing power consumption by 50% or higher for every packet processed. FP3 uses 30% less physical space in service provider premises.

FP3: EVOLUTION OF SERVICE SCALE

	FP1 Scale	FP2 Scale	FP3 Scale
IP Route Table	1 Million	3 Million	5 Million
L3 Interfaces	5k	20k	32k
VPRN Instances	2k	16k	32k
VPLS Instances	8k	48k	96k
MAC FIB	128k	2M	4M
PW Termination	48k	128k	256k
Queues	32k/IOM	128k/IMM	256k/IMM
Policers	—	1M+	2M

The evolution of the service scale is enormous. The FP3 provides up to 5 million IP routing entries, 32 thousand L3 interfaces and a huge increase in the number of VPRN and VPLS services. 4 million MAC entries can now be stored for all the VPLS instances, 256 thousand pseudowire terminations are possible and up to 256 thousand queues are available in the new IMM input/output card. The number of policers has doubled, from 1 million with the FP2, to 2 million with the FP3 chip.

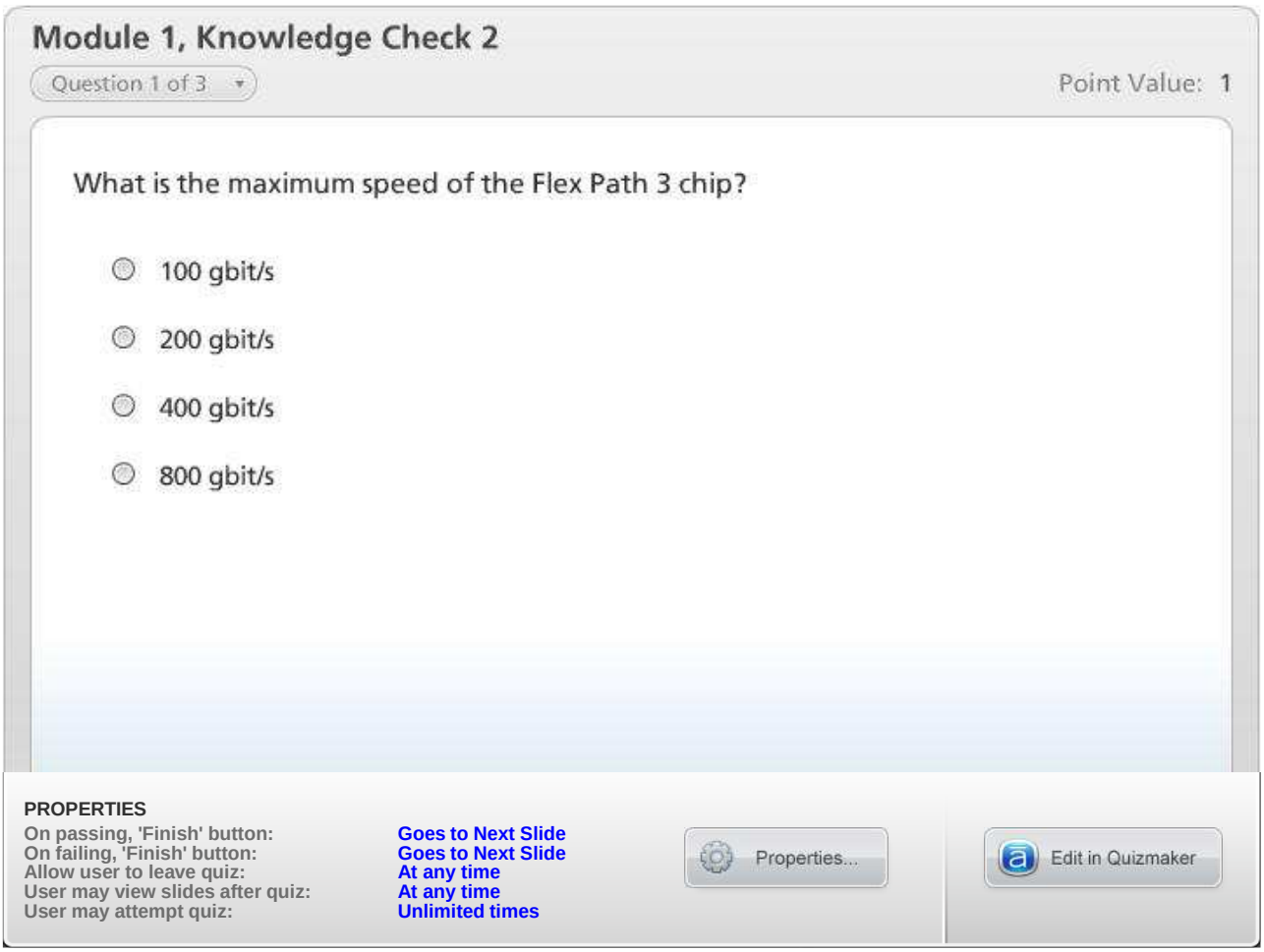
REALIZING UNIVERSAL BROADBAND BY LEVERAGING SERVICE ROUTING AND NETWORK INTELLIGENCE



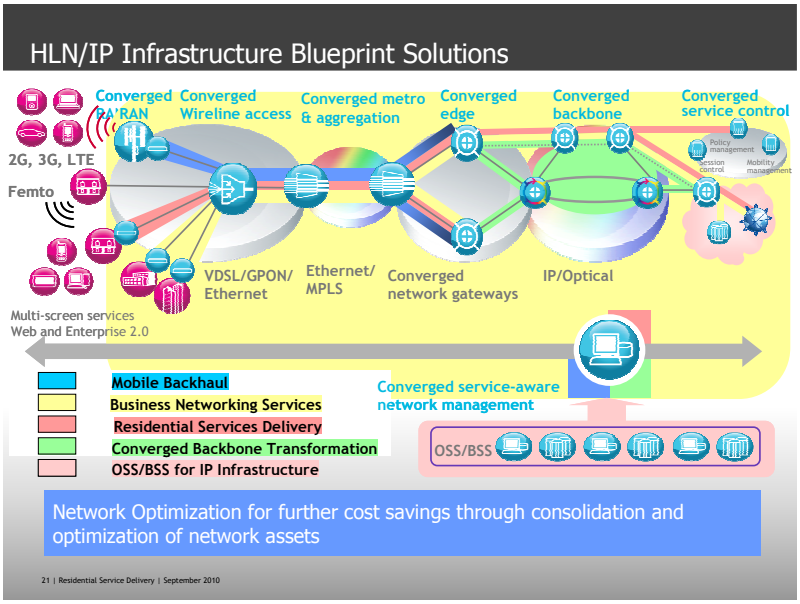
The FP3 plays an important role in the Alcatel-Lucent vision of the High Leverage Network.

The FP3 enables service providers to increase capacity, reduce costs and develop new revenue generating opportunities. The greater capacity of the silicon used in the Alcatel-Lucent portfolio of IP/MPLS routers will allow service providers to move from development to delivery of new services without capacity or reliability concerns.

Realizing universal broadband is achieved by leveraging Alcatel-Lucent's IP and service routing expertise!



Section 3: Service routers supported solutions.



The HLN IP/MPLS infrastructure contains solutions for each type of service application.

Let's have a closer look to the different supported solutions inside a High Leverage Network.

The mobile backhaul solution delivers a carrier grade transport solution from base station to core. Allowing evolution from a circuit based E1/T1 type of backhaul to an all IP network. This packet based solution provides the intelligence to deal with the steadily increasing number of data services. In addition, the solution is easy scalable to meet modern and future mobile bandwidth requirements.

The business networking services solution provides a suite of business networking services across all access technology options. This includes business VPN services with Layer 2 carrier Ethernet VPN's and Layer 3 premium VPN's. These VPN services all provide superior performance, scalability and quality of service.

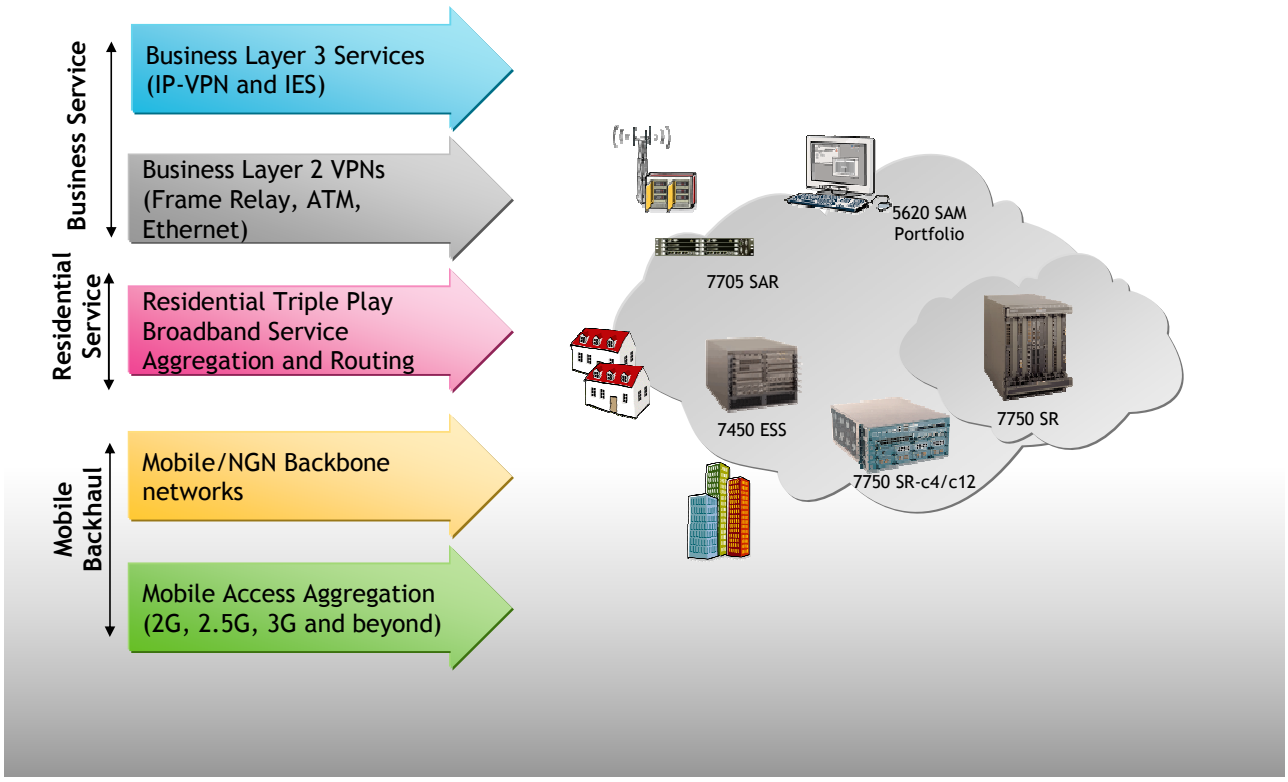
The residential services delivery solution addresses the boom and demand more multi-media service delivery. It provides the higher bandwidth required per subscriber by distributing network intelligence and subscriber management to the edge domain. The solution offers exposure to embedded quality of service and application awareness as well as smooth migration to IPv6.

With the converged backbone solution Alcatel-Lucent offers the first industry comprehensive solution to optimize core transport across the full range of traffic mixes. The Alcatel-Lucent approach leverage the companies leadership in the optics and IP domains to offer an integrated solution across the data plane, control plane and management plane.

Network evolution towards IP creates new operations needs such as the end to end management of services, quality of service and service level agreements. To allow for this end to end management of this IP infrastructure, Alcatel-Lucent provides the IP infrastructure support solution. This solution is the OSS/BSS system solution for the high leverage network. It is integrated with the converged service network management domain over an open standard based architecture.

In conclusion. One converged network supports multiple service types. Further cost saving is achieved through the consolidation and optimization of network assets.

Different service types inside the HLN Network



The high leverage network may be purchased as an end to end solution or maybe viewed as an architectural framework to help network providers evolve in their own individual modular way.

This architectural framework is relevant to all service providers. From the small service providers providing business services over a fixed network to the mobile service provide as well to the huge multi service, multi national residential service provider.

These service providers will evolve towards a coherent vision of a converged, scalable and intelligent network.

If we take a closer look to the business service provider, L2 and/or L3 point to point or point to multi point services are the most popular offered services. L2 services is mainly focussed around Ethernet, but also legacy service like frame relay, ATM, TDM are HDLC are supported.

The Residential services are mainly focussed on aggregating and routing of a large residential customer base.

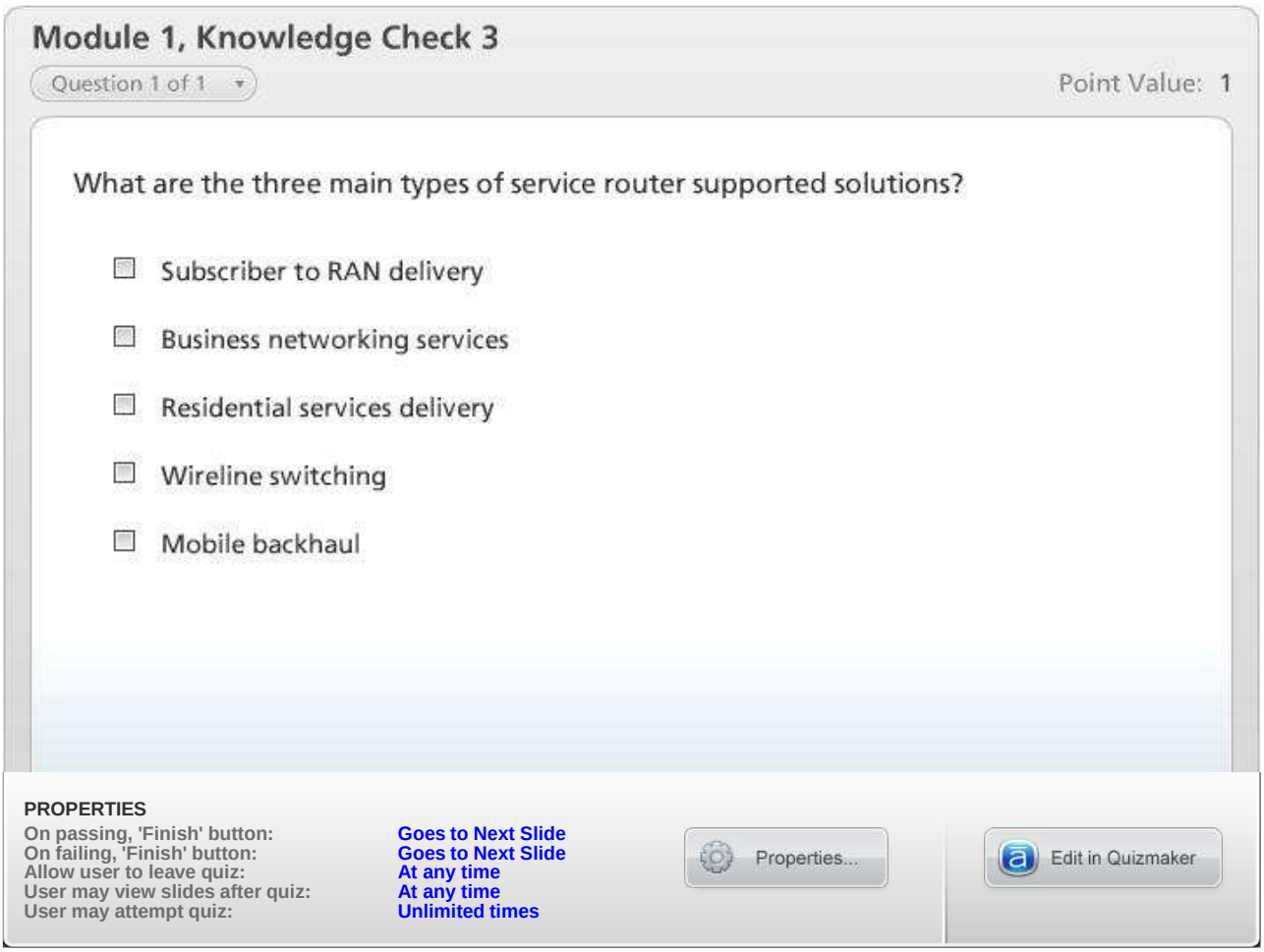
The strength of the mobile access aggregation relies on the mix of different generations of mobile infrastructures and therefore access technologies.

Other SR-OS supported solutions

- Other important supported solutions
  - IP Core
  - Lightradio
  - IPv6 and IPv4 to IPv6 transition
  - Enabling the cloud
  - Video delivery
  - Ipsec
  - Application Assurance
  - Strategic industries

Energy | Transportation | Public sector | Enterprises | Defense & Security

Traditional telecom operators are typically active in the residential, mobile and/or business markets. But there are a number of important SR-OS supported solutions and products like the new products for the IP core and the lightradio solution, the transition solutions to go from IPv4 to IPv6, the cloud based services, video delivery options, IP security, application assurance and the whole strategic industry market solutions. The following modules are referring to these other solutions without going in depth. There are separate courses describing each of the solutions.

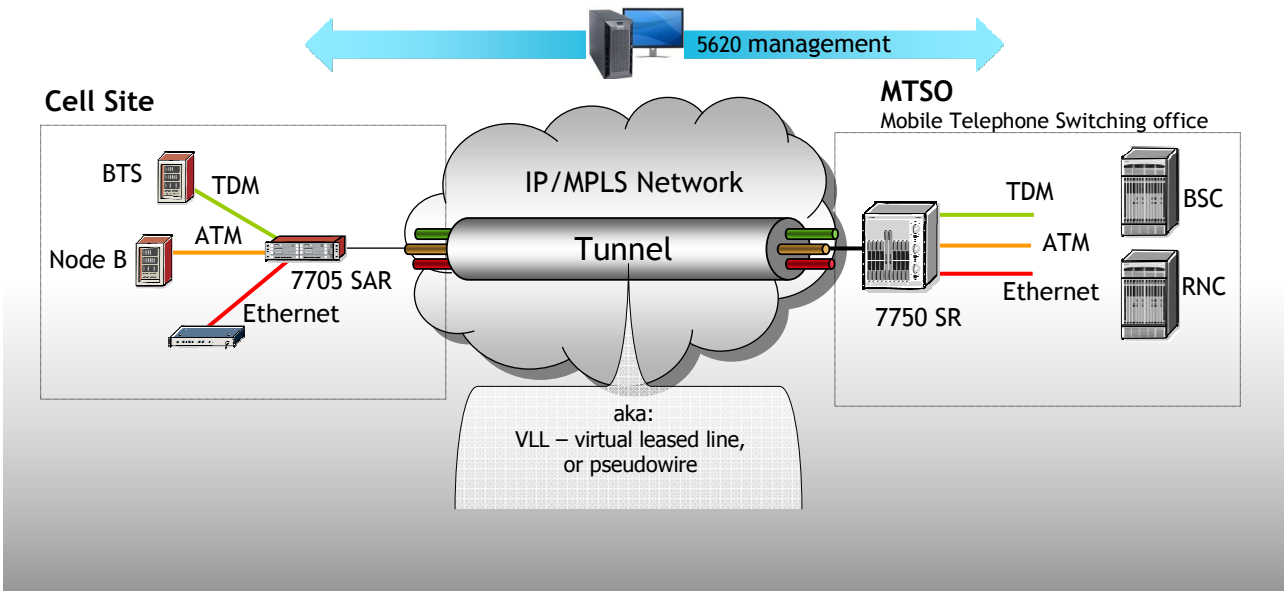


3.1 Mobile Backhaul/LTE

Let's have a look at each of the three supported solutions inside the high leverage network starting with the mobile backhaul and LTE solution.

Mobile backhaul

- Transport all type of traffic from Mobile Cell
  - 2G traffic in a TDM VLL
  - 3G traffic in an ATM VLL
  - Additional traffic in an Ethernet VLL



Mobile backhaul or aggregation of the traffic from different cell sites towards a central location in the network faces two main challenges.

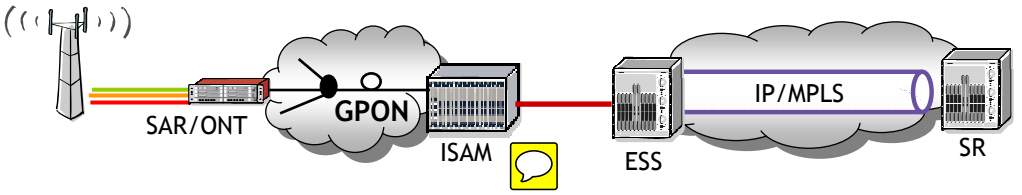
Connecting all types of traffic in a cost effective way from different generations of mobile technologies is the first important challenge. One common device should be capable of connecting these different technologies and provide one uplink towards the network.

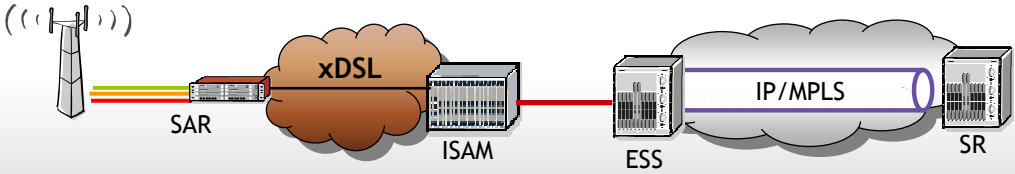
The second challenge is to have a unified network which is both cost effective, scalable and highly resilient. The answer: a fully managed IP/MPLS infrastructure.

The tunnel or virtual leased line created over the IP/MPLS network that connect the cell site with the MTSO or Mobile Telephone Switching office is also called a pseudowire or VLL. The MPLS pseudowire was developed by the PWE3 working group. The pseudowire edge to edge emulation working group or PWE3 was established to develop an architecture for service provider edge-to-edge pseudowires or virtual leased lines.

In telecommunications, a pseudowire is an emulation of a native service over a packet switched network. The native service may be ATM, frame relay, Ethernet, low-rate TDM, or SONET/SDH, while the packet switched network is IP/MPLS.

Different mobile backhaul solutions

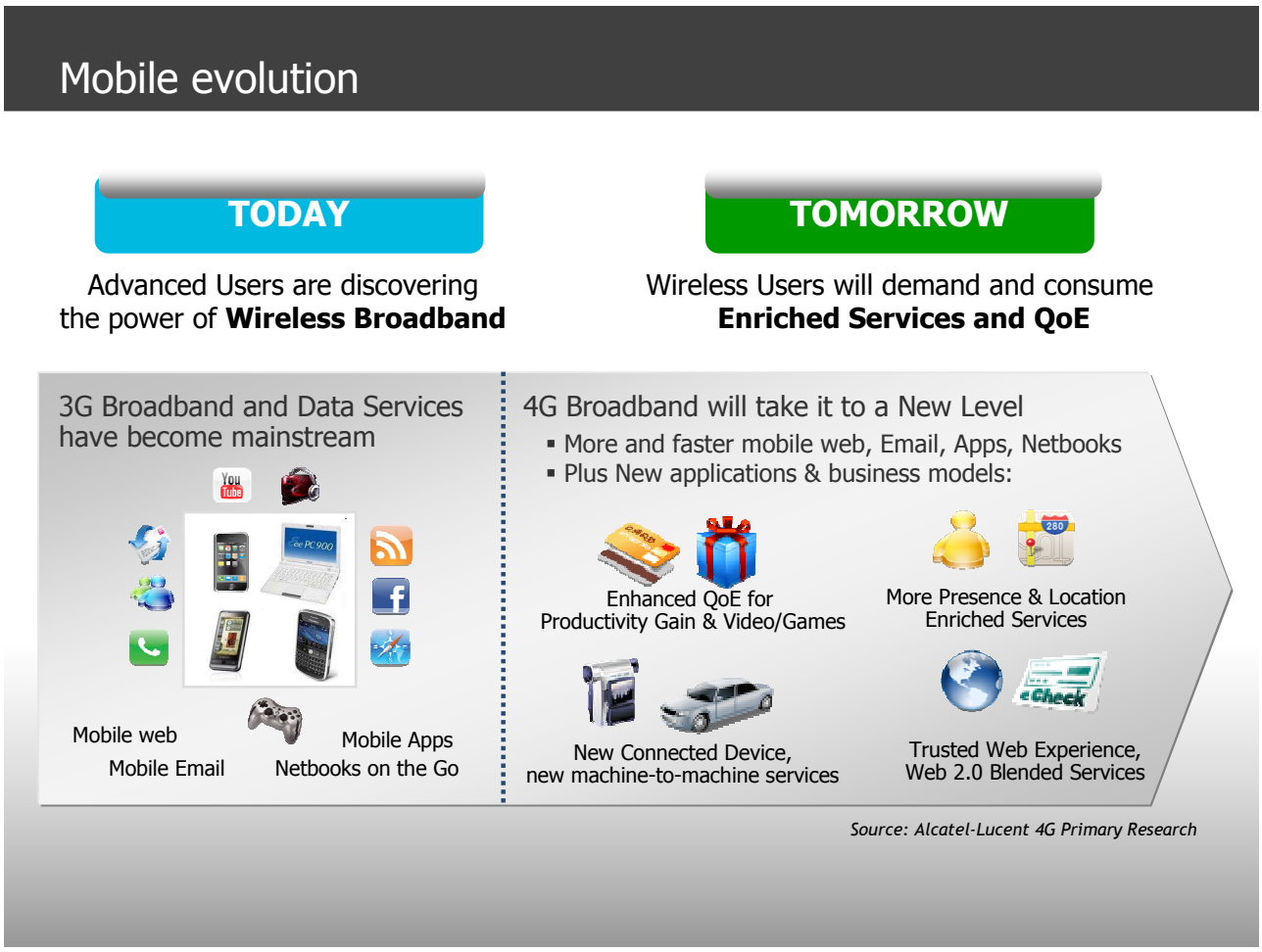




Depending of the existing infrastructure, different alternative options of mobile backhaul solutions might exist. Two solutions are given as an example. The first solution uses the Gigabit passive optical network as transport network. The mobile backhaul device at the cell site serves as an ONT or optical network termination. This ONT can be one of the many ONT's connected the all passive network. The ISAM on the picture is the access aggregator and provides an uplink towards the IP/MPLS aggregation network.

As seen from the second alternative solution, the GPON network may also be an xDSL network allowing all mobile traffic to be aggregated over any flavor of xDSL technology.

The SAR, ISAM, ESS and SR are carrier grade Alcatel-Lucent products which will be explained in separated modules of this course.



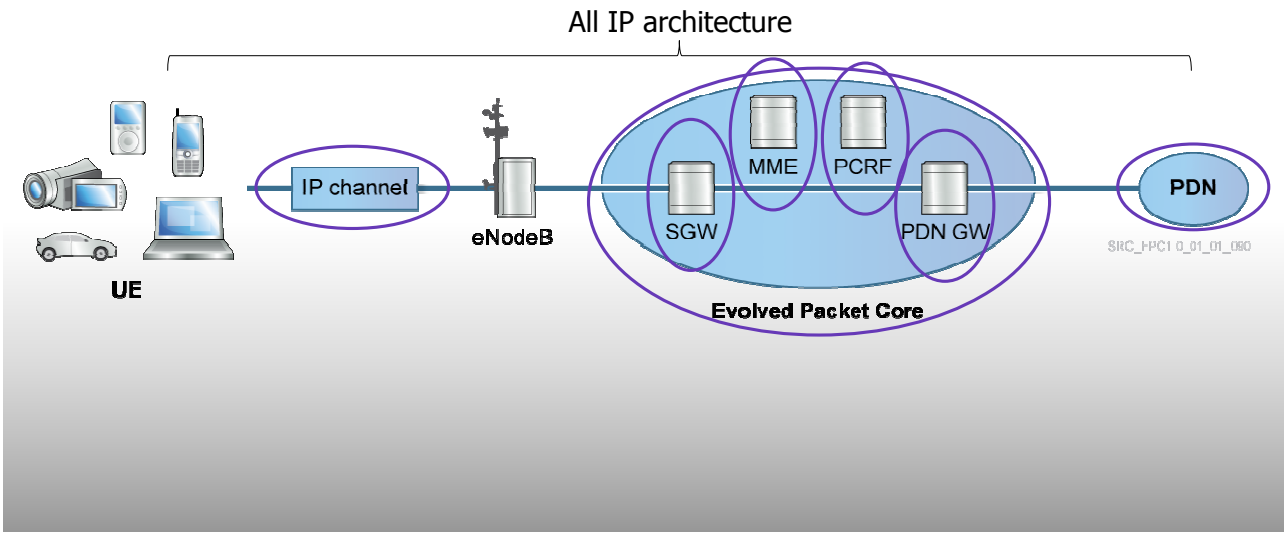
Today, most of the mobile customers have access to the third generation or 3G of wireless broadband. The fourth generation of 4G will take broadband to the next level.

The emerging next-generation 4G offers faster wireless broadband in addition to new services and business models with an end-to-end quality of service. Amongst those are interactive gaming, web 2.0 and new machine-to-machine services.

Machine-to-machine refers to machines or devices that can communicate with other machines or devices, on either wired or wireless networks. For the last few years, SMS was the approach used for machine-to-machine for mobile-based communications. M2M services such as home security systems, automatic meter readers for the smart grid and basic vehicle telematics have low bandwidth requirements and hence are acceptable 2G or 3G applications. 4G allows mobile operators to offer new M2M services that require low latency and high throughput such as fleet management, wireless healthcare, security and video surveillance. M2M will provide a multitude of wireless applications for consumers and businesses in various industries including energy, health care, and consumer products.

LTE Network – a closer look

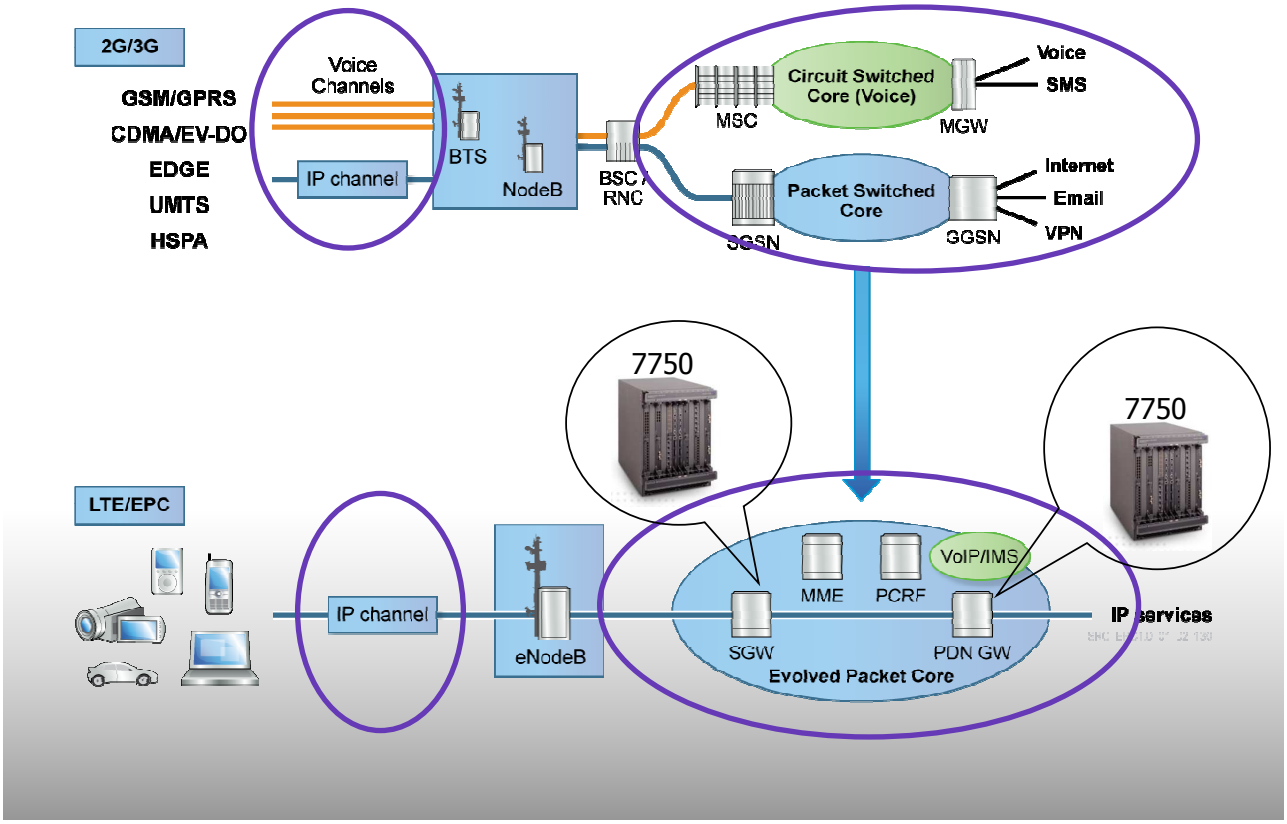
- An end-to-end all-IP architecture
  - Evolved NodeB (eNodeB) for the radio access
  - Evolved Packet Core (EPC) for the core network



An Long Term Evolution or LTE 4G network offers an end-to-end all-IP architecture. One IP channel connects the UE or user equipment to the Evolved NodeB.

The Evolved Packet Core or EPC is a packet switched core that consists of four elements: the Serving Gateway or SGW, the Mobility Management Entity or MME, the Packet Data Network Gateway or PDN gateway and the Policy and Charging Rules Function or PCRF. The EPC provides the user with an IP connectivity to the Packet Data Network .

New **All-IP** Simplified Network Architecture



The new EPC introduces a single IP packet core for both voice and data whereas these were separated networks in the past. The aim is to have one IP core that can offer all services and can be used by all access networks including LTE, UMTS and CDMA.

On the radio side, the separate voice channel and IP channel become a single IP channel in LTE.

On the network core side, the separate circuit switched network that is used for voice and the packet switched network that is used for data are converged into a single all-IP packet routing core. This convergence creates operational savings for the network operators as they move to support a single physical network as opposed to two separate ones.

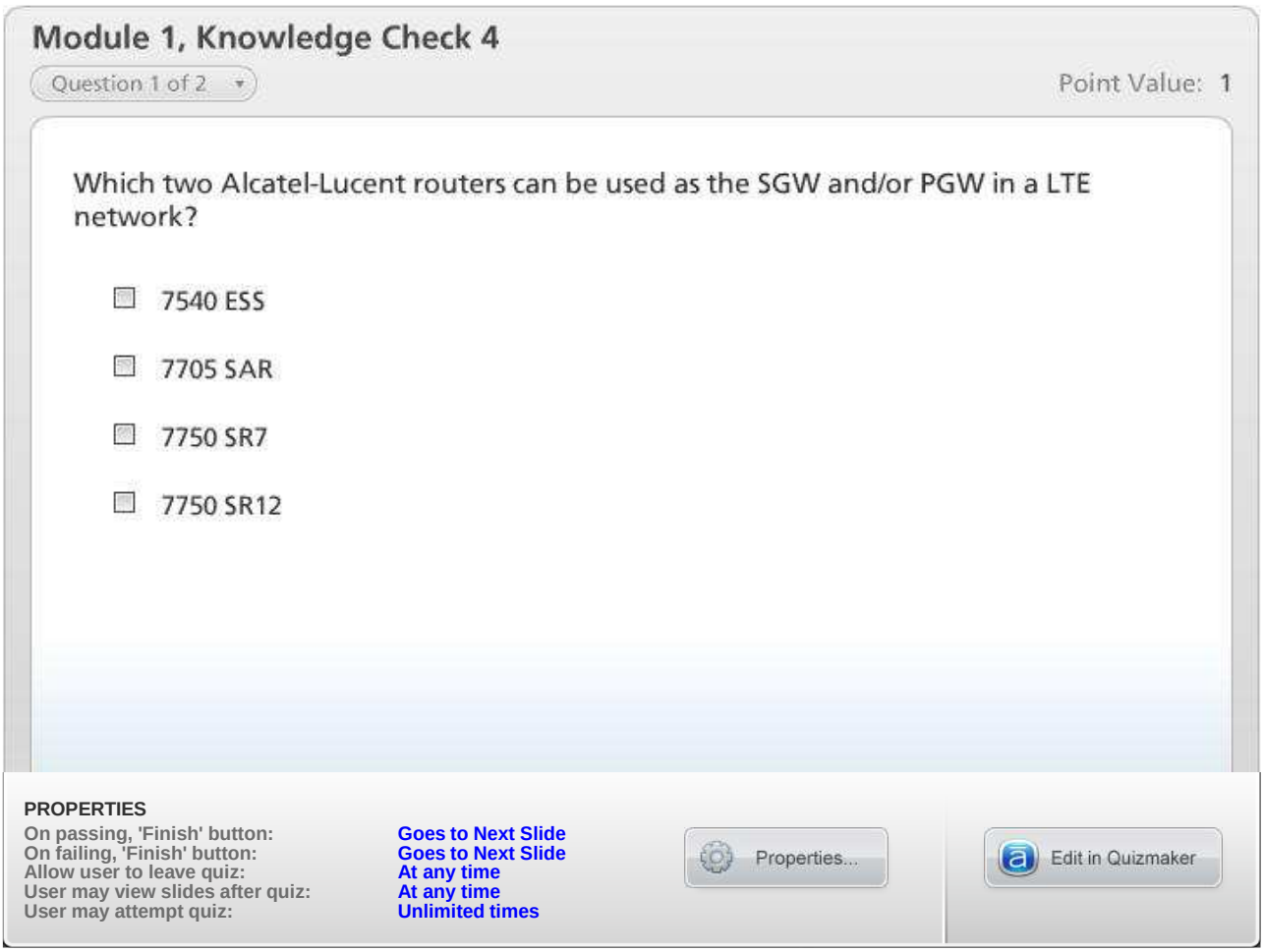
The same products that are used for business and residential customers are used as the Serving Gateway (SGW) and the Packet Data Network Gateway (PGW). The Alcatel-Lucent 7750 service router (7750 SR) can be equipped as a SGW and/or PGW.

Alcatel-Lucent EPC mobile gateways

Making an S- and/or P gateway from a service router used to deliver business and residential services is simple. The MG-ISM or mobile gateway integrated service module is the only added hardware required to enable the S/P gateway function.

Each MG-ISM enables flexible configuration for either the S or P gateway function.

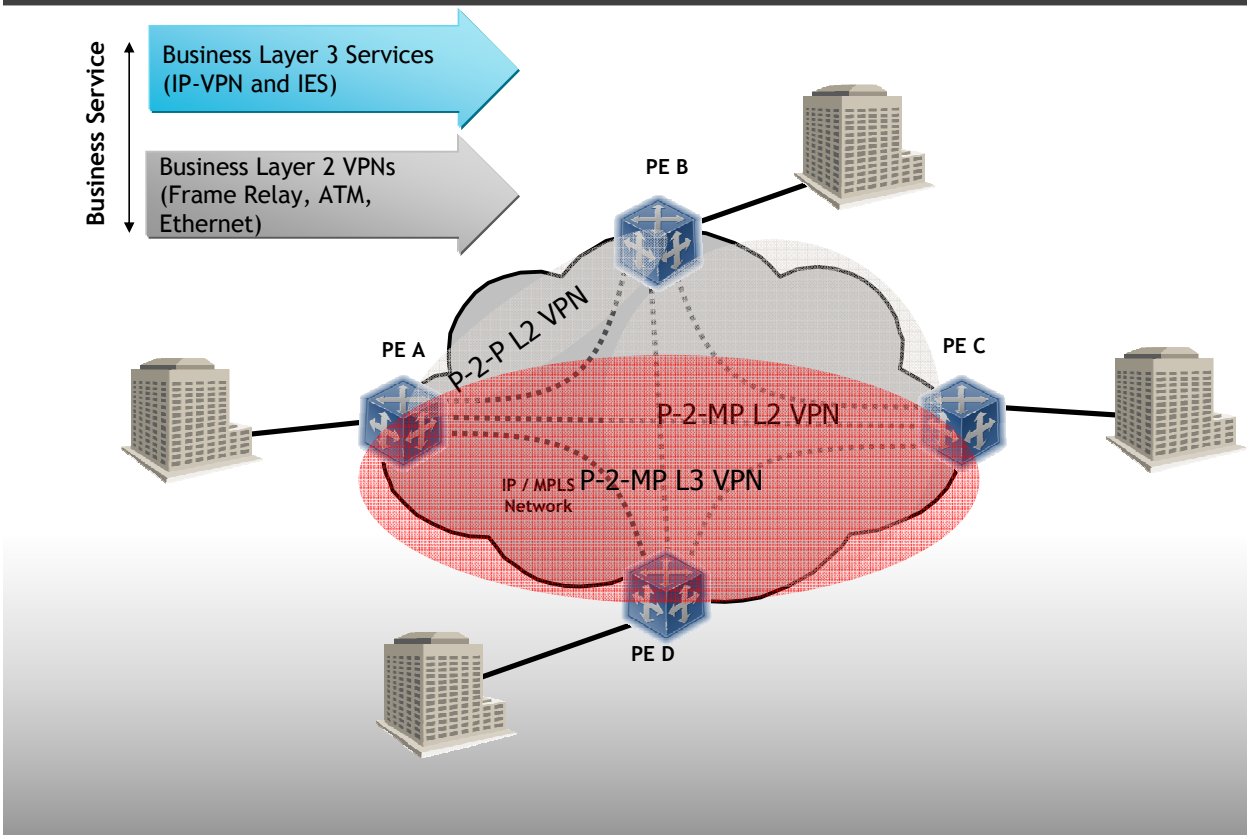
The MG-ISM includes a dedicated processing engine including wire-rate deep packet inspection.



3.2 Business network service

Let's take a closer look at the business network services provided inside the High Leverage Network.

Business network services



Business services can mainly be divided into two categories. The layer 2 and the layer 3 VPN's.

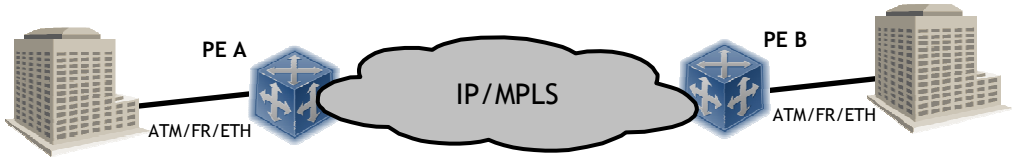
Each of the two categories can be both offered as a point to point service or point to multi point service.

An enterprise wants to connect their two premises in a point to point way. The connection at either side could be any available L2 access technology. The customer data packets are transported over an MPLS pseudowire between provider edge router A and provider edge router B. There is no interference with other customers traffic not belonging to the VPN and the customer does not make use of the base routing and forwarding tables of the service provider.

More than two premises can be connected to the same Ethernet VPN of virtual private LAN service. This is a typical Ethernet point to multi point service on a service router that uses MAC forwarding tables on the provider edge routers. Forwarding is based on a MAC addresses look-up on each PE router.

The service provider can also provide L3 services towards its customer in a point to point or point to multipoint way. The forwarding of data packets is now based on the routing table on each PE service router where each customer has his dedicated routing table of VRF.

L2 Business network Services => options



PE A \ PE B	ATM	Frame Relay	Ethernet
ATM	Apipe	Apipe	Ipipe Epipe
Frame Relay	Apipe	Fpipe	Ipipe Epipe
Ethernet	Ipipe Epipe	Ipipe Epipe	Epipe

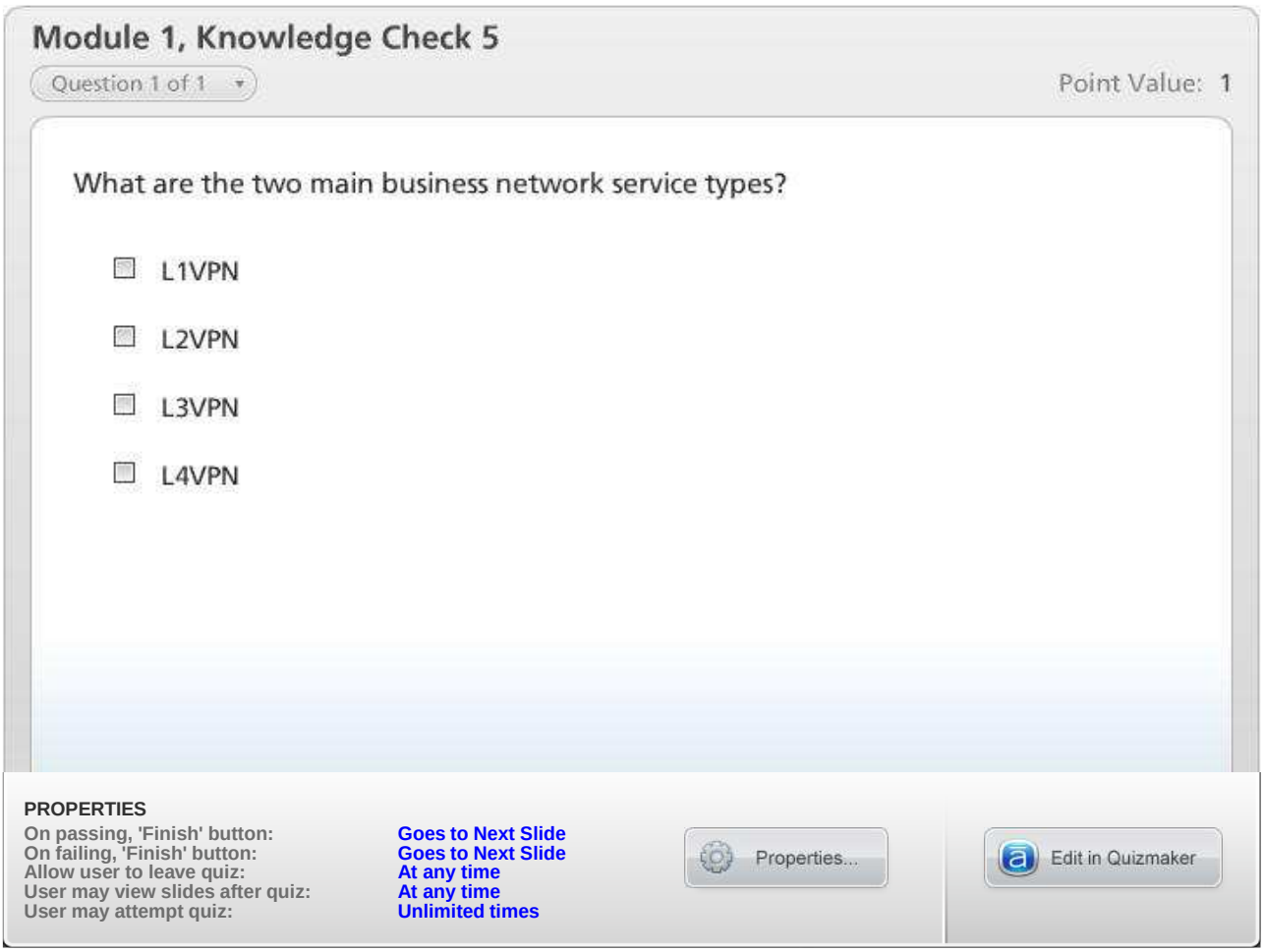
Let's have look to the different options of connecting business point to point L2 VPN's together. The access technology on one side does not have be the same on the other side of the point to point link. This is called the interworking function.

The table lists the various interworking scenarios possible with ATM, Frame Relay and Ethernet ports at either end of an MPLS pseudowire.

This result in a pseudowire service called a "pipe" service. Different options of pipe services are available on a modern service router.

For example, when IP traffic, encapsulated inside an Ethernet frame, needs to be transported by a bridge/router over its ATM port to a bridge/router with an Ethernet port, bridged encapsulation is used. An E-pipe is required to transport this type of traffic. More details are given in the service module of this course.

Routed encapsulation is used when native IP traffic is transported across a bridge/router with an ATM port to a bridge/router with an Ethernet port. An Ipipe is required to transport this type of traffic. For further details, please refer to the Ipipe section in this module.

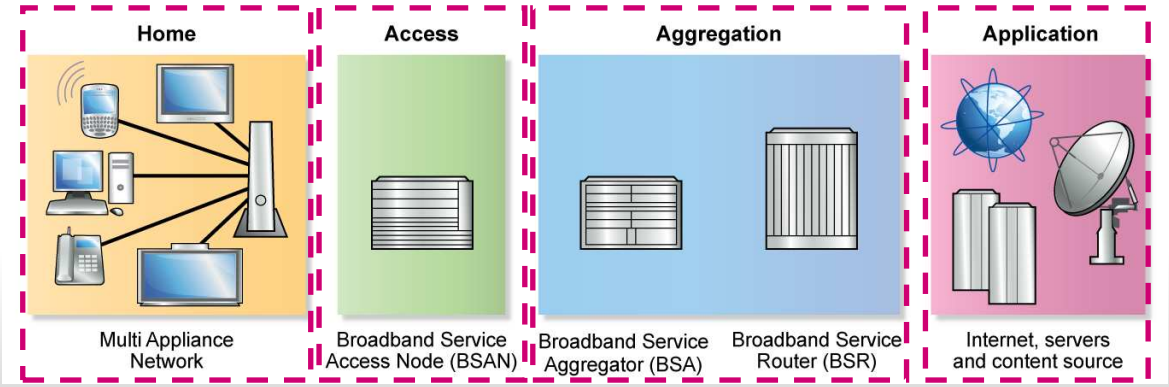


3.3 Residential service delivery

Let's take a look at the residential service delivery network solution.

The residential service delivery network solution

Alcatel-Lucent meets the demands of delivering broadband to each household in an intelligent way.

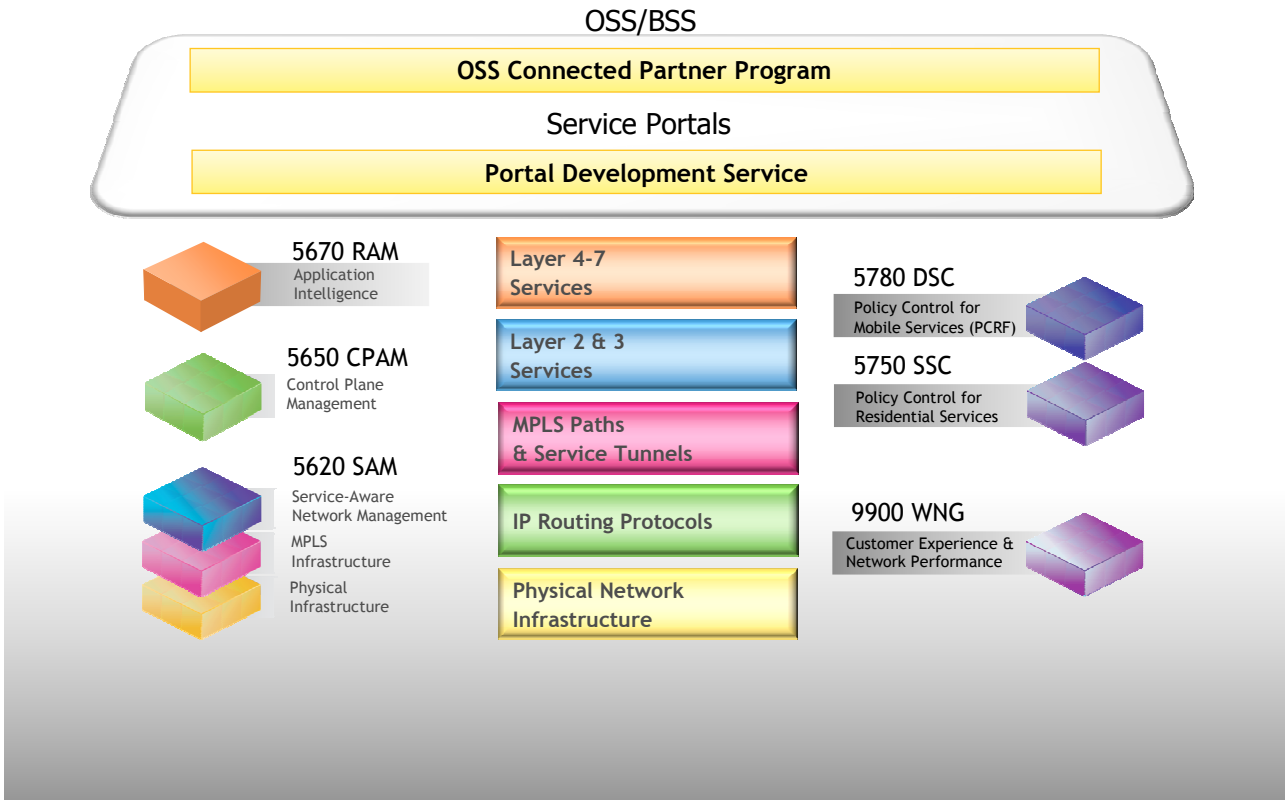


The Alcatel-Lucent residential service delivery network inside the high leverage network is divided in four major parts. The home network, the access network, the aggregation network and the application network. It gives access to individual households, independently if they are connected via xDSL, fiber or any wireless device. The intelligence is built in all the parts of the residential service delivery network. A mechanism which is called enhanced subscriber management or better known as ESM allows to install the subscriber and SLA profiles for each customer individual. This network is capable of terminating the PPP over Ethernet or IP over Ethernet access methods. Based on this, access is granted for each subscriber with his allowed credentials. Each customer or residential user will get it's queue buffer space and allowed bandwidth. IP addresses are given out of the embedded databases on the service routers or are coming from a central DHCP server.

3.4 OSS/BSS system

The last section of this module addresses the unified management OSS/BSS systems.

IP/MPLS Management Architecture & Platforms



The OSS or operations support system and BSS or business support system most frequently describes "network systems" dealing with the supporting processes such as maintaining network inventory, provisioning services, configuring network components, and managing faults. The complementary term business support systems or BSS is a newer term and typically refers to "business systems" dealing with customers, supporting processes such as taking orders, processing bills, and collecting payments. The two systems together are often abbreviated as OSS/BSS.

Alcatel-Lucent has developed a one common framework for its supporting systems. This includes a base graphical management system for setting up services, MPLS tunnels and the whole physical infrastructure. This system , the 5620 SAM, also serves as an alarm management system. On top of the 5620 SAM there is also a control plane assurance management system that visualize the routing and MPLS tunnels inside the network. If the service provider want to see the rapport of the applications running over the network, the 5670 RAM needs to be installed.

Customer experience and policy management and enforcement is done by the 5780 DSC for the mobile services or 5750 SSC of the residential services. Whereas the 9900 WNG is targeted to measure network performance for the RAN or radio access network.

WRAP-UP

Wrap-up: The SR-OS or Service Router Operating System is an industry leading product platform that plays a crucial role in the high leverage network. One of the key technology innovations is the introduction of the new Flex Path 3 chip. This chip is all about being faster, smarter and greener. The SR-OS products are also continiuously evolving and are found in different types of networks. In the residential, business and mobile networks, which are typically owned by the traditional service providers. But also in the strategic industry markets like energy, transportation, public sector, enterprises, defense and security markets as seen in the next modules.

Module Summary

- A High Leverage Network, or HLN, is a converged, scalable and intelligent all-IP network.
- HLN is best described as Application Enablement, Universal Access, Network Evolution and Operational Transformation.
- HLN monetizes the bandwidth hungry services of today.
- The FP3 plays an important role in the Alcatel-Lucent vision of the High Leverage Network. It enables service providers to increase capacity, reduce costs and develop new revenue generating opportunities.
- The new Flex Path 3 is the world fastest network processor using less physical space and consumes less power.
- The FP3 chip is all about being faster, smarter and greener.
- The three main types of service routers supported solutions are mobile backhaul solution, business networking services solution, and the residential services delivery solution.
- Alcatel-Lucent has developed a single common framework for its supporting systems. This includes a base graphical management system for setting up services, MPLS tunnels and the entire physical infrastructure.

What have we learned in this module? A High Leverage Network, or HLN, is a converged, scalable and intelligent all- IP network which is best described as Application Enablement, Universal Access, Network Evolution and Operational Transformation. It's also about getting money out the bandwidth-hungry services of today. One of the key technical innovations that makes this high leverage network happen is the introduction of the new Flex Path 3 Processor, the world fastest network processor using less physical space and consumes less power than processors Installed in the field today. It enables service providers to increase capacity, reduce costs and develop new revenue generating opportunities. The three main types of service routers supported solutions you will find in a high leverage network are mobile backhaul solution, business networking services solution, and the residential services delivery solution. The strategic industry Market which equally important for Alcatel-Lucent is explained in the next module. To perform the graphical management for setting up services, MPLS tunnels and the entire physical infrastructure, a single common framework is developed for its supporting systems.

End of Module 1

Alcatel·Lucent

This completes module 1.

SR-OS Fundamentals
Module 2: IP/MPLS SR-OS product family

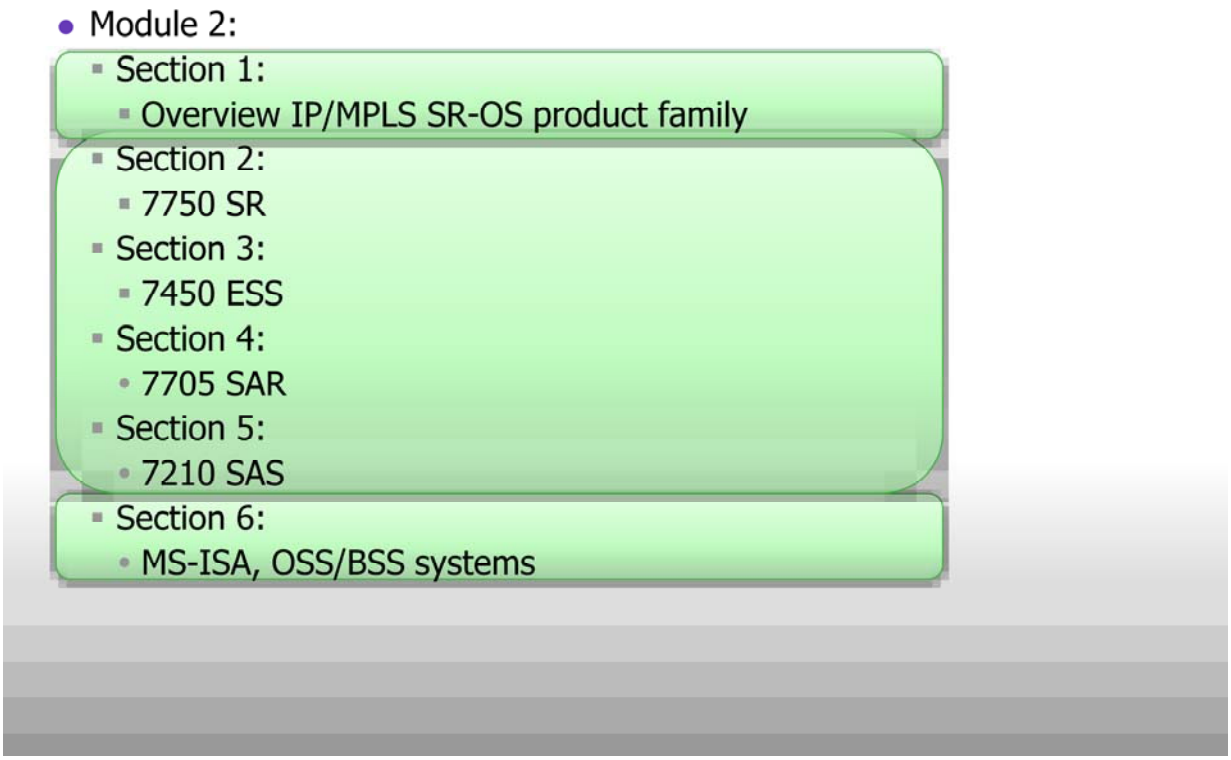
IPD Development

Available as PDF

Welcome to the second module of the SR-OS fundamentals course.

Module 2 provides an overview of the complete range of IP/MPLS products available within Alcatel-Lucent.

Agenda

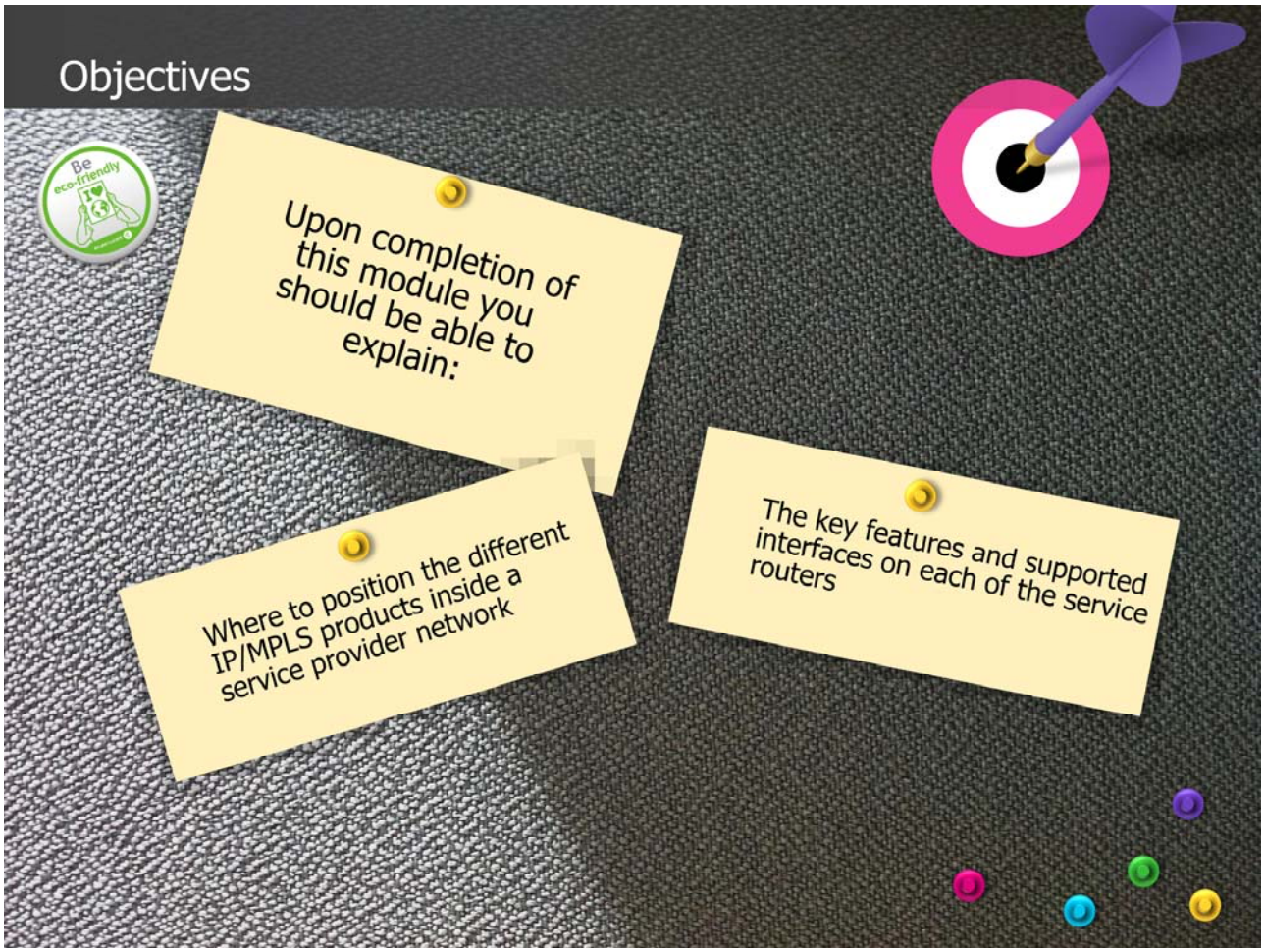


Module 2 is divided into six sections.

Section 1 gives a brief overview of Alcatel-Lucent IP/MPLS products.

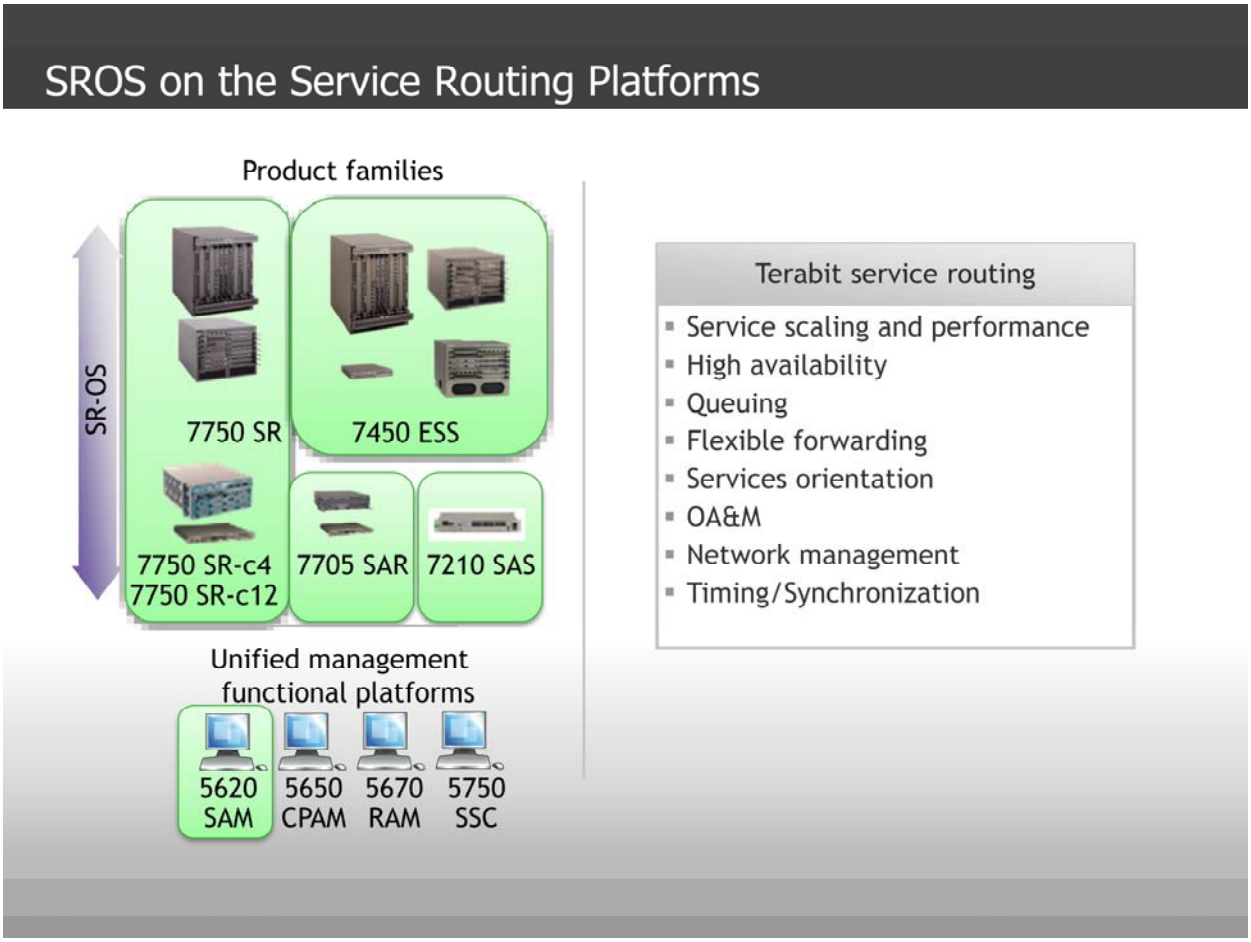
Section 2 till section 5 describes the major product families and explains how they are positioned inside the service provider network.

At the end of this module, Section 6 provides an overview of all supporting and add-on products to the suite of the IP/MPLS base products.



By the end of this module you will be able to explain where to position the different IP/MPLS products inside a service provider network and you are able to explain the key features and supported interfaces on the service routers.

Section1: Overview Of the IP/MPLS SR-OS product family



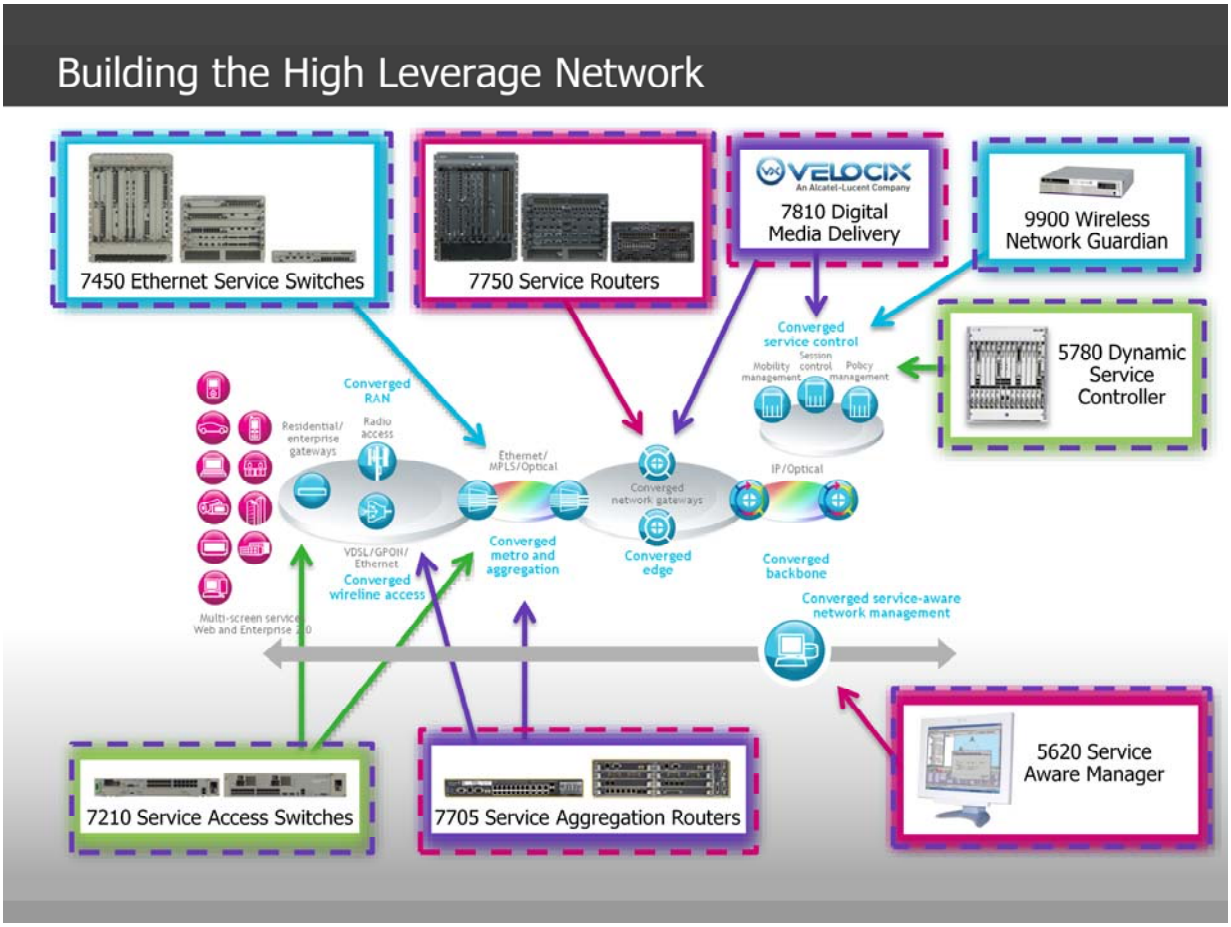
The four major SR-OS product families are the 7750 service router, the 7450 Ethernet service switch, the 7705 service access router and the 7210 service access switch.

Each of the families are designed to deliver flexible and scalable networking solutions. Utilizing customized interface options, field replaceable components and operational software. The portfolio includes service routers and Ethernet service switches with a broad range of capacities. From 90 Gigabits per second for the 7750 SR-c4 up to 400 Gigabits per second for the 7750 SR-12.The complete portfolio is managed by the 5620 Service Aware Manager or SAM

The superior functions and capabilities that differentiate our service routing platforms from a typical edge and Internet router are the enormous service scale and performance, the high availability features like non- stop routing and non-stop service, the huge queue buffer space available and the deep packet inspection that allows a flexible way of forwarding.

The SR-OS products have a standard set of operation and maintenance tools allowing to troubleshoot on MPLS, IP and service level.

Sync-E and 1588v2 are included as the standard for packet synchronization.



Where do we position all those products in the high leverage network? The 7750 service routers are typical core routers which you will find in the converged network backbone and edge. Closer towards the customers, the converged aggregation part, the 7450 Ethernet service switches are positioned. Enterprises that require a couple of Gig interfaces with a couple of Gig interfaces as uplink, or directly connected the IP/MPLS network might want to install the 7210 service access switches. In the same area of the HLN network, the converged access domain, the 7705 service aggregation routers are positioned to transport and aggregate mobile backhaul traffic.

The positioning of products is not fixed by rules, meaning that it can vary depending on the specific customer requirements, supported interface type, cost, numbers of interfaces and future growth capabilities.

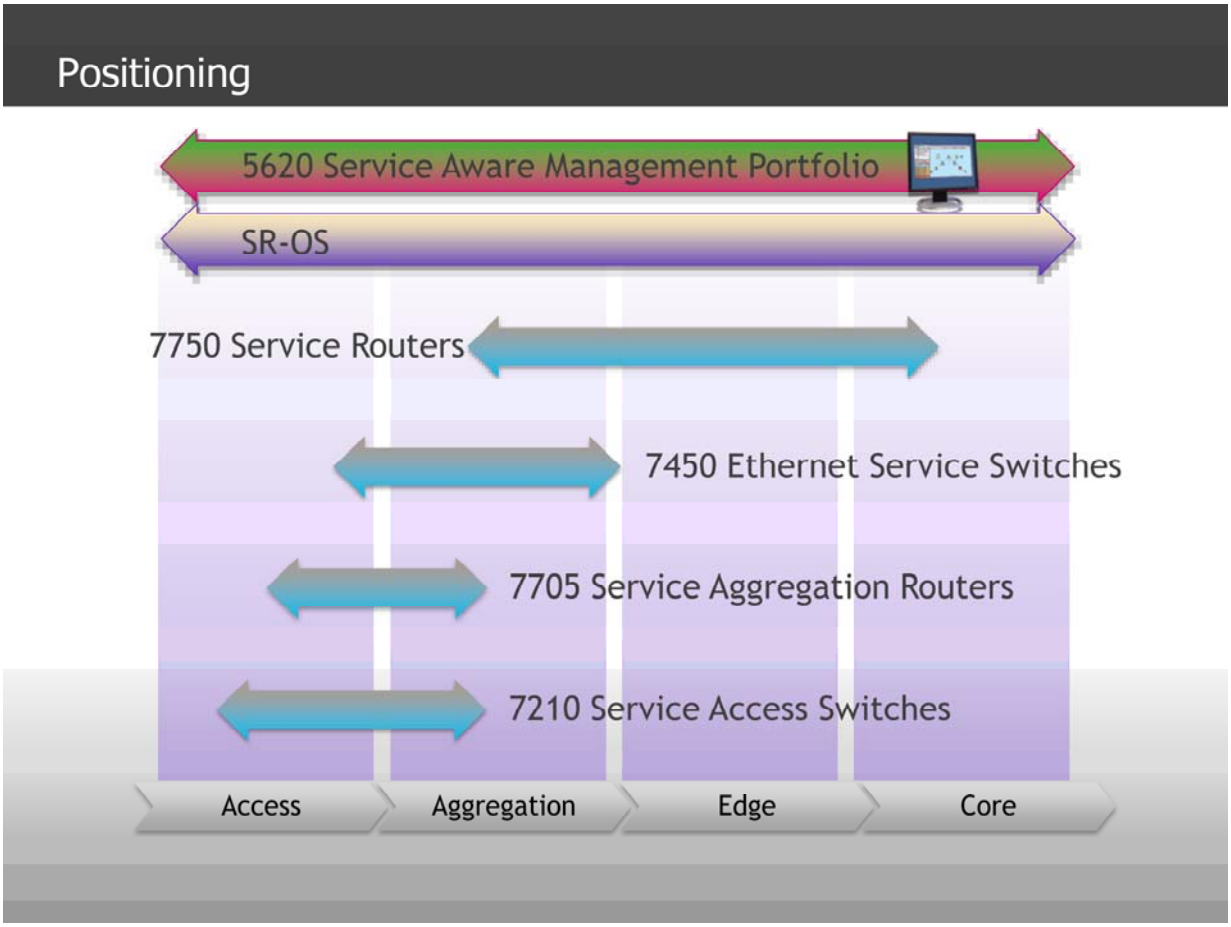
The base network management platform, the 5620 Service Aware Manager, is not limited to one domain in the HLN network as it manages the complete range of Alcatel-Lucent's IP/MPLS products.

In the converged service control area, three main products are positioned. The 7810 digital media delivery, the 9900 wireless network guardian and the 5780 dynamic service controller.

The 7810 digital media delivery platform is specifically designed for network service providers to help them transform into entertainment providers and capitalize on growing consumer demand for rich digital media with a smarter, content-savvy network infrastructure.

The Alcatel-Lucent 9900 Wireless Network Guardian or WNG provides powerful capabilities for wireless data service providers to accurately design, engineer, optimize, manage, and price their networks.

The Alcatel-Lucent 5780 Dynamic Services Controller or DSC enables delivery of personalized services that monetize the service provider's network and is also pivotal in optimizing these services to get the most value out of the network.



Let us try to position the different service routers on a scale from access to core. After a first look, it seems there are overlaps. However, product positioning does also depend on the required type of interfaces, the access technologies, cost and future growth capacity.

The 7210 SAS and 7705 SAR are typically found on the access and aggregation side. Closer to the edge, you will find the 7450 ESS products and in the core, the 7750 SR is positioned.

Let's have a closer look at the 7750 service router family.

The 7750 SR product family

Available in five chassis types, the 7750 SR is a scalable platform that provides cost-effective solutions to cover from the smallest to the largest network locations.

The Alcatel-Lucent 7750 SR multiservice edge routers are designed to provide high-performance and high-availability routing with service-aware operations, administration, management, and provisioning. Leveraging the Alcatel-Lucent 400 Gbit/s FP3 silicon technology, the 7750 SR delivers exceptional performance and scale to a wide range of IP services, with service intelligence to further drive operational efficiency.

The 7750 SR offers an advanced and comprehensive feature set, and can serve as a Broadband Network Gateway for residential services, as a Multiservice Edge for Carrier Ethernet and IP VPN business services, as the aggregation router in mobile backhaul applications, or as a mobile packet core for second generation, third generation and long-term evolution wireless networks.

7750 SR capacity and components at a glance

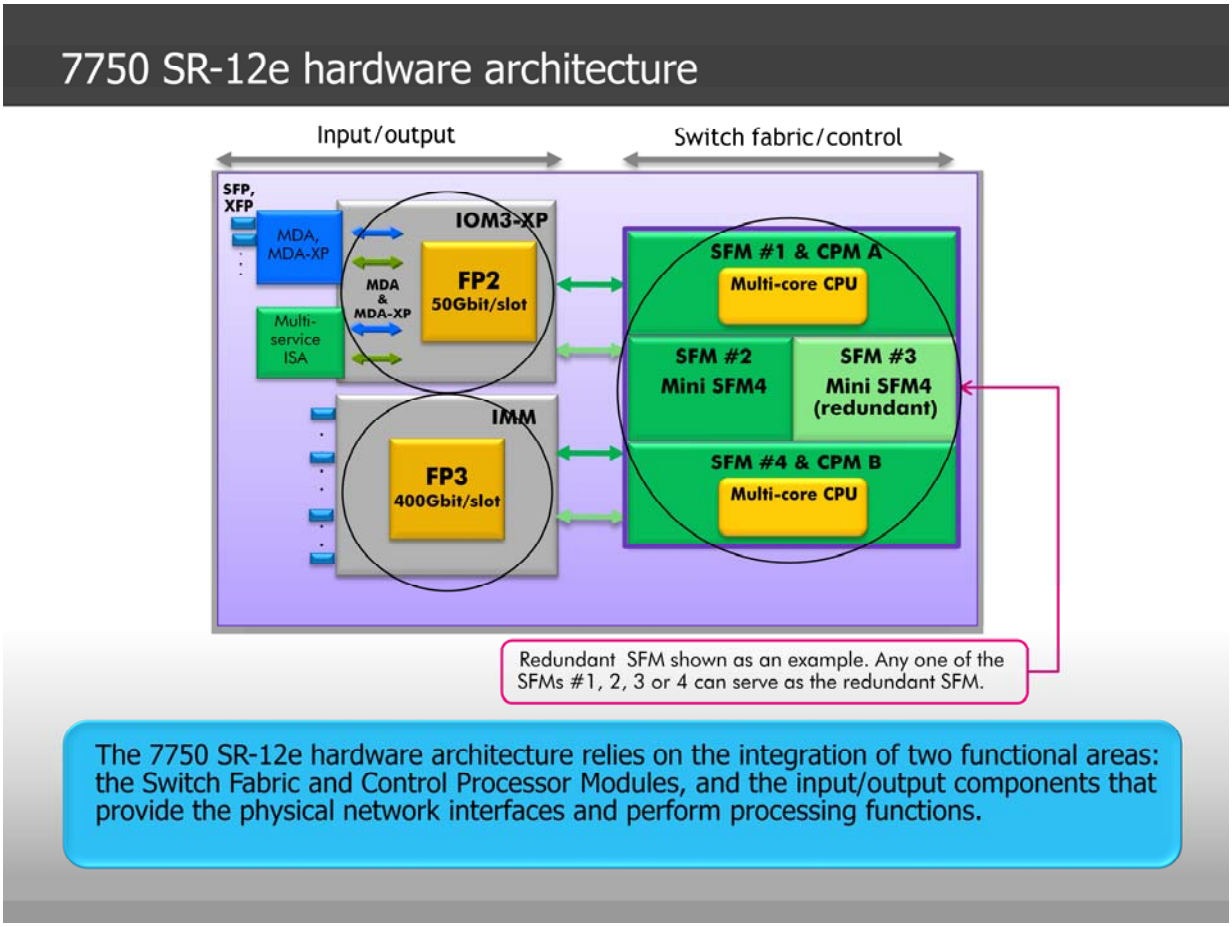
7750 SR model	Capacity	Slots	Components	Slot capacity
7750 SR-12e	7.2 Tbit/s	12	2 SF/CPMs 2 half-slot mini-SFMs 9 IOMs or IMMs	400 Gbit/s and more
7750 SR-12	2 Tbit/s	12	2 SF/CPMs and 10 IOMs or IMMs	Up to 100 Gbit/s
7750 SR-7	1 Tbit/s	7	2 SF/CPMs 5 IOMs or IMMs	Up to 100 Gbit/s
7750 SR-c12	90 Gbit/s	12 compact (3 full)	2 CFMs Interface blocks containing MDAs and CMAs	Up to 10 Gbit/s
7750 SR-c4	90 Gbit/s	4 compact	Integrated CCM/CFM Separate interface block containing CMAs	Up to 10 Gbit/s

This table summarizes the capacity and components of the 7750 SR models that are currently available. The differences in capacity indicate that the 7750 SR-12e, 7750 SR-12 and SR-7 are suitable for large core applications, while the 7750 SR-c12 and SR-c4 are appropriate for small Point-of-Presence deployments.

The following slides will describe in greater detail all the 7750 SR components: Switch Fabric and Control Processing Modules, mini Switch Fabric Modules, Input Output Modules, Integrated Media Modules, Control and Forwarding Modules, Chassis Control Module, Media Dependent Adapters, and Compact Media Adapters.

7750 SR-12e architecture and hardware overview

Section 3.1 describes the 7750 SR-12e architecture and hardware.



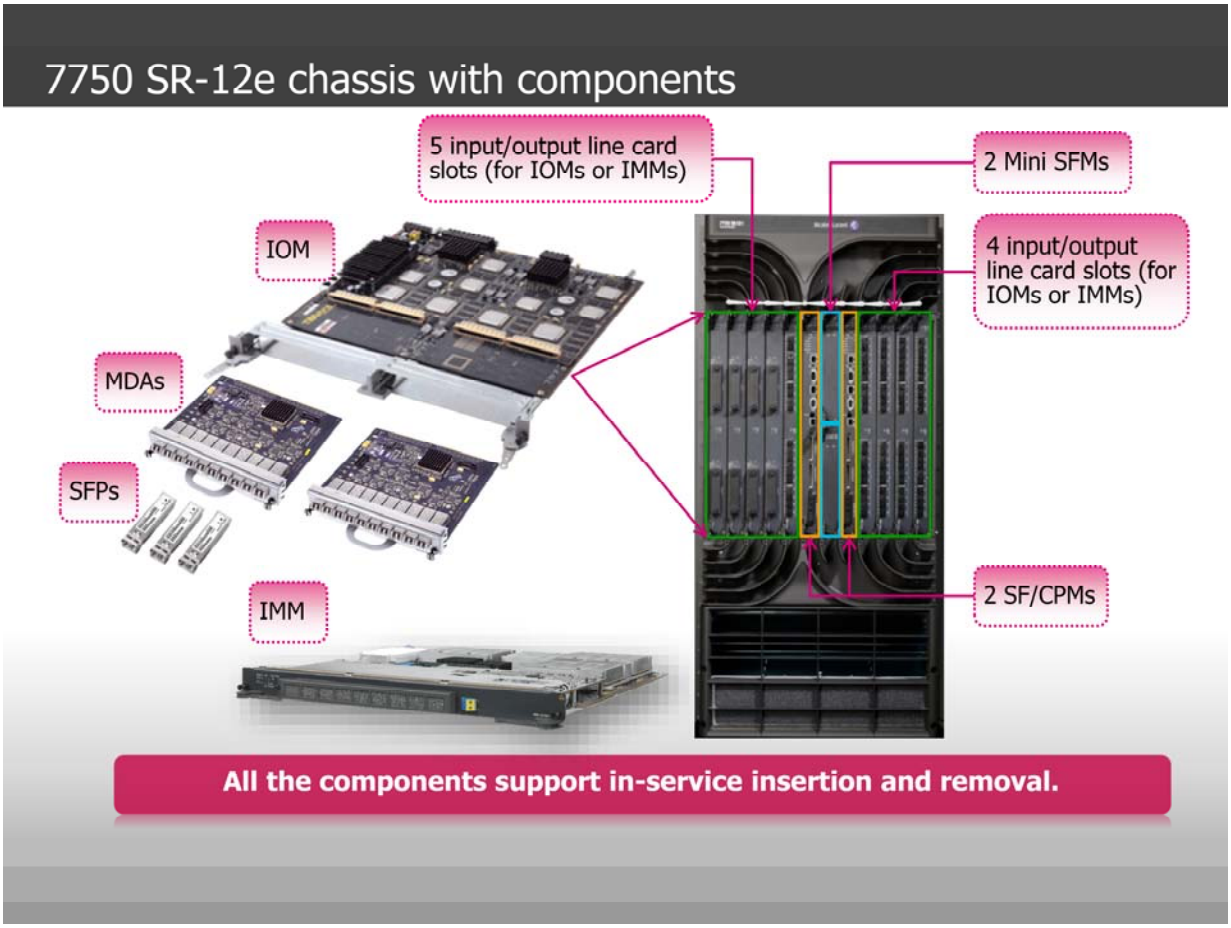
The 7750 SR-12e is a new, high-end addition to the 7750 SR product family, and provides high-scale capacity for future expansion The 7750 SR-12e has been designed to deliver differentiated, high-performance, and high-availability services. It supports specialized service-aware application processing, advanced quality of service, and a wide range of Ethernet and multi-service interfaces and protocols. The 7750 SR-12e provides industry-leading scale and intelligence to deliver residential, business, and wireless broadband IP services on a converged edge routing platform.

A fully equipped 7750 SR-12e integrated chassis contains two Switch Fabric and Control Processor Modules and two Mini Switch Fabric Modules and provides a fully redundant fabric.

The SF/CPM is a combined control processing (CPM) and switch fabric (SFM) module. The control processing function operates in a 1+1 active/standby redundancy model when two SF/CPM cards provide a fully redundant and hot synchronized control plane. Each SF/CPM integrates the central processing unit and switch fabric, controls the routing and switching functions, and provides the management and console interfaces, as well as the interface for the external synchronization signal.

The Mini Switch Fabric Modules are required (along with the switch fabric function of the SF/CPM card) to provide a fully redundant fabric for the 7750 SR-12e platform.

The input output section can contain Input Output Modules or Integrated Media Modules, or both.

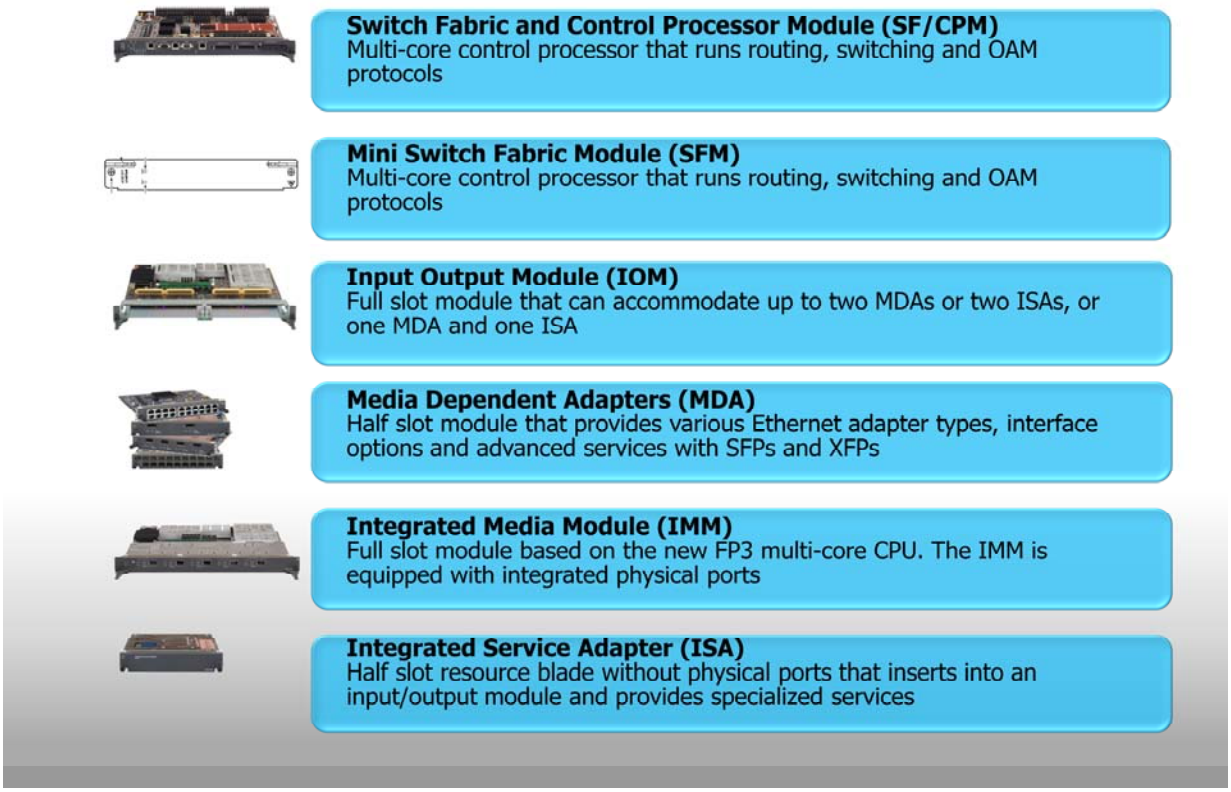


The 7750 SR-12e has a vertical slot layout with two slots for Switch Fabric and Control Processor Modules, two half-slots for mini SFMs, and nine slots for input output modules or integrated media modules.

Each IOM card can accommodate Media Dependent Adapters, which provide physical network interfaces through the Small Form-factor Pluggable units. The IOM can also accommodate Integrated Service Adapters, which provide specialized application processing and buffering.

The IMM cards integrate the processing and physical interfaces on a single board. The 7750 SR-12e introduces support for a series of IMMs based on the new FP3 multi-core CPU.

7750 SR-12e components at a glance



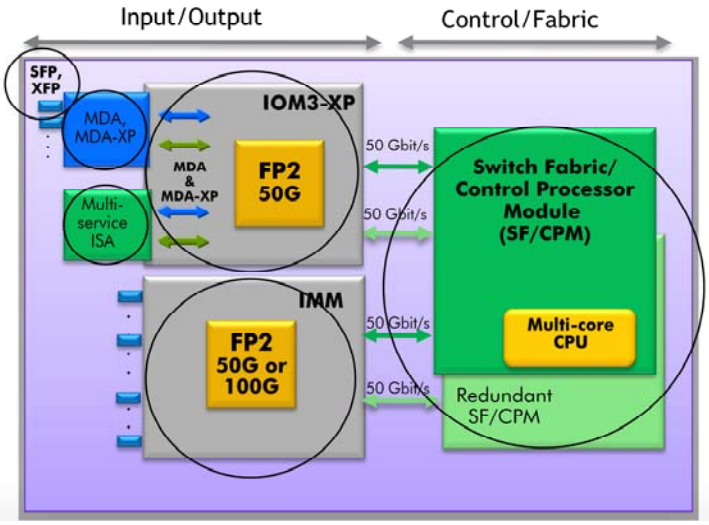
This slide provides brief descriptions of the 7750 SR-12e components. You will find out more information about each component later in the course.

With the exception of the mini SFM, the 7750 SR-12e chassis accommodates the same component types as the 7750 SR-12 and SR-7. However, the characteristics of the components supported by the 7750 SR-12e can be different. For details, see the Release Notice document that applies to your equipment.

7750 SR-12 and SR-7 architecture and hardware overview

Section 3.2 describes the 7750 SR-12 and SR-7 architecture and hardware.

7750 SR-12 and SR-7 hardware architecture

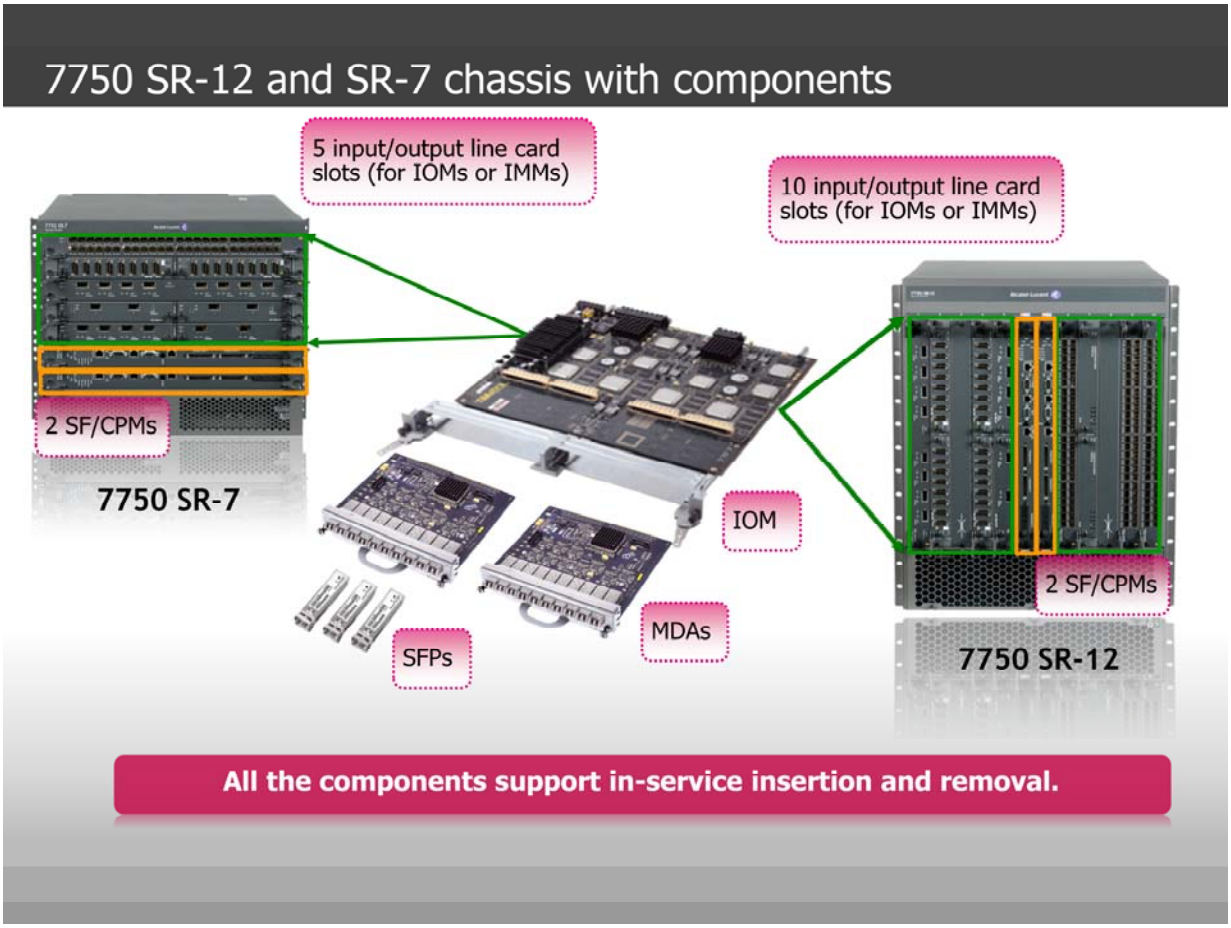


The 7750 SR-12 and SR-7 hardware architecture relies on the integration of two functional areas: the Switch Fabric and Control Processor Modules, and the input/output components that provide the physical network interfaces and perform processing functions.

Each 7750 SR-12 and SR-7 unit contains two Switch Fabric and Control Processor Modules. Each SF/CPM card integrates the central processing unit and switch fabric, controls the routing and switching functions, and provides the management and console interfaces, as well as the interface for the external synchronization signal. The two SF/CPM cards share the traffic load and provide system redundancy.

The input output section can contain Input Output Modules or Integrated Media Modules, or both. Each IOM card can accommodate Media Dependent Adapters, which provide physical network interfaces through the Small Form-factor Pluggable units. The IOM can also accommodate Integrated Service Adapters, which provide specialized application processing and buffering.

The IMM cards integrate the processing and physical interfaces on a single board.



The 7750 SR-7 has a horizontal slot layout with two slots for Switch Fabric and Control Processor Modules and five slots for input output cards. The 7750 SR-12 chassis has a vertical layout with two slots for Switch Fabric and Control Processor Modules and 10 slots for input output cards. The greater capacity makes the 7750 SR-12 router the perfect choice for high-end multi-service edge or core applications. The 7750 SR-7 model is better suited for mid-range multi-service deployments.

7750 SR-12 and SR-7 components at a glance

Switch Fabric and Control Processor Module (SF/CPM)
Multi-core control processor that runs routing, switching and OAM protocols

Input Output Module
Full slot module that can accommodate up to two MDAs or two ISAs, or one MDA and one ISA

Media Dependent Adapters
Half slot module that provides various Ethernet adapter types, interface options and advanced services with SFPs and XFPs

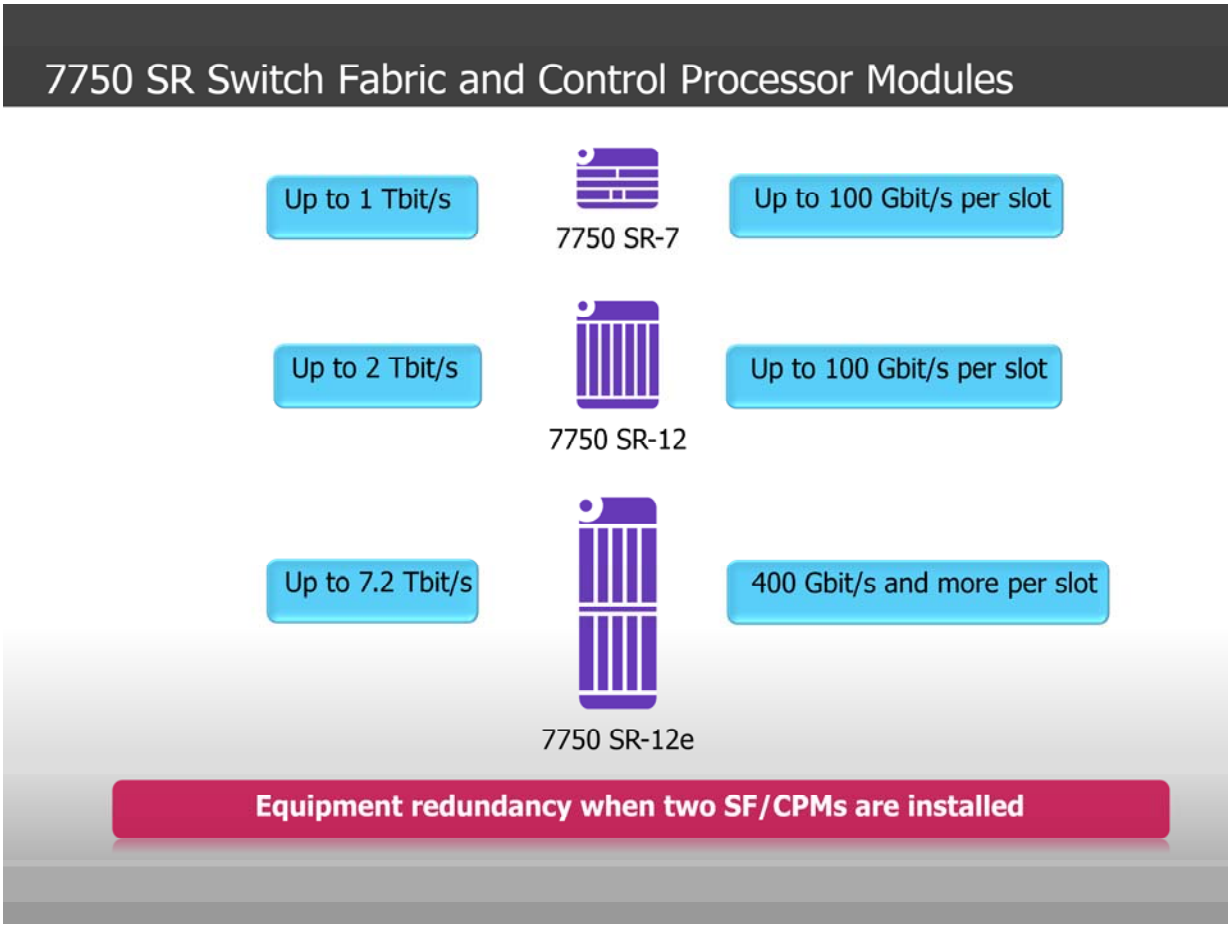
Integrated Media Module
Full slot module with FP2 network processor complex equipped with integrated physical ports

Integrated Service Adapter
Half slot resource blade without physical ports that inserts into an input/output module and provides specialized services

This slide provides brief descriptions of the 7750 SR-12 and SR-7 components. You will find out more about each component later in the course.

7750 SR-12e, SR-12 and SR-7 components

Section 3.3 describes the 7750 SR-12e, SR-12 and SR-7 components.



The Switch Fabric and Control Processor Module integrates the switch fabric and control functions into one module. Each SF/CPM has three slots for compact flash cards: cf1, cf2, and cf3. A new SR system is supplied with a compact flash card that contains the files required to start the system; this is typically inserted in the cf3 slot. You can use the cards in the cf1 and cf2 slots to store debugging and accounting logs.

An 7750 SR-7, SR-12 or SR-12e system requires one SF/CPM to function. When two SF/CPMs are installed, the switch fabrics are active on both cards and operate in a load-sharing arrangement, doubling the throughput of the chassis switch fabric.

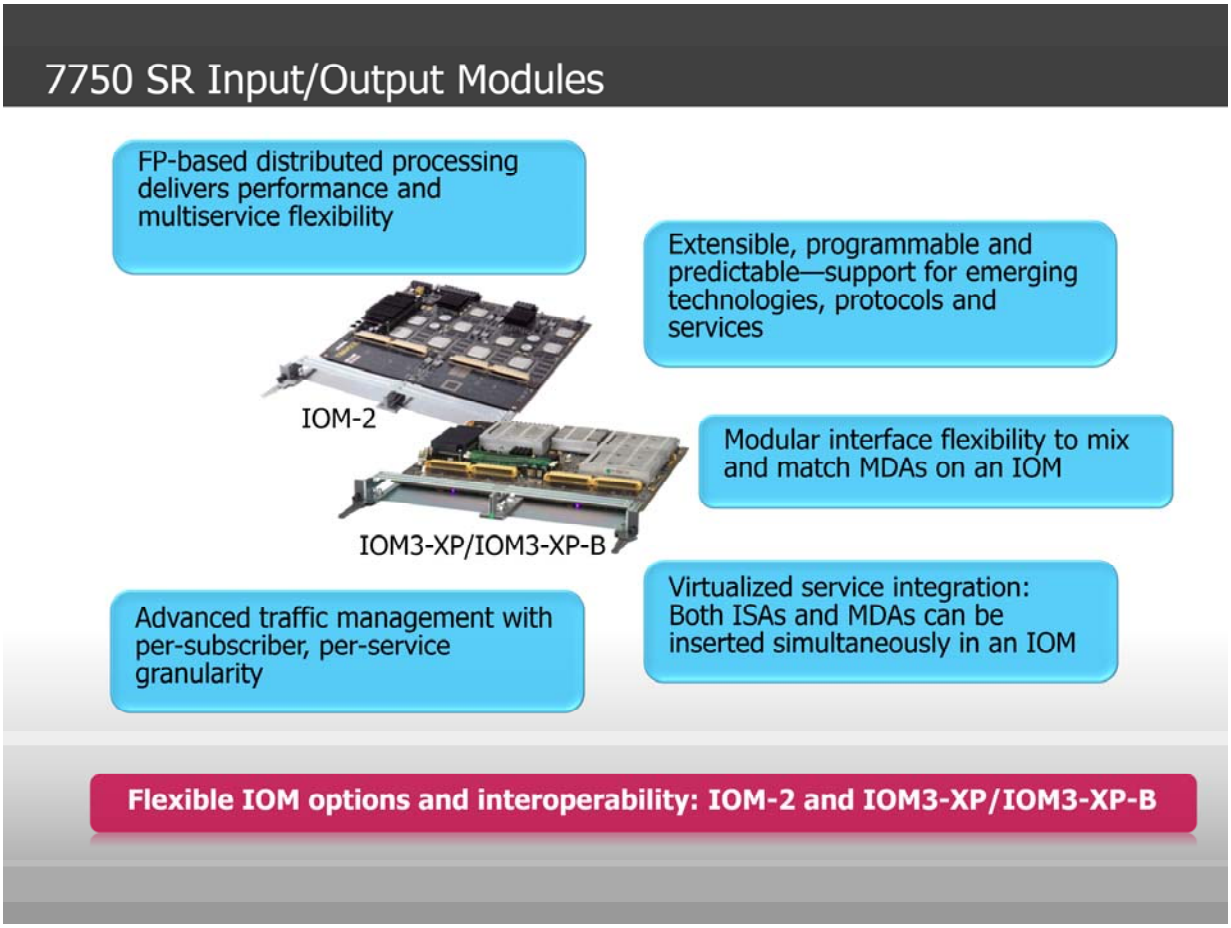
The performance of the 7750 SR-12e is even more stunning due to the presence of the mini SFMs, which add more capacity to the switching fabric. The next slide describes the mini SFM.

7750 SR-12e Mini Switch Fabric Module

The Mini SFM module is a smaller form-factor card compared to the SF/CPM module, containing only the SFM unit, controlling the switching functions for the 7750 SR-12e system.

The Mini SFM connects directly to the router backplane and carries traffic between line cards. The backplane provides high-speed access to the modules and chassis components.

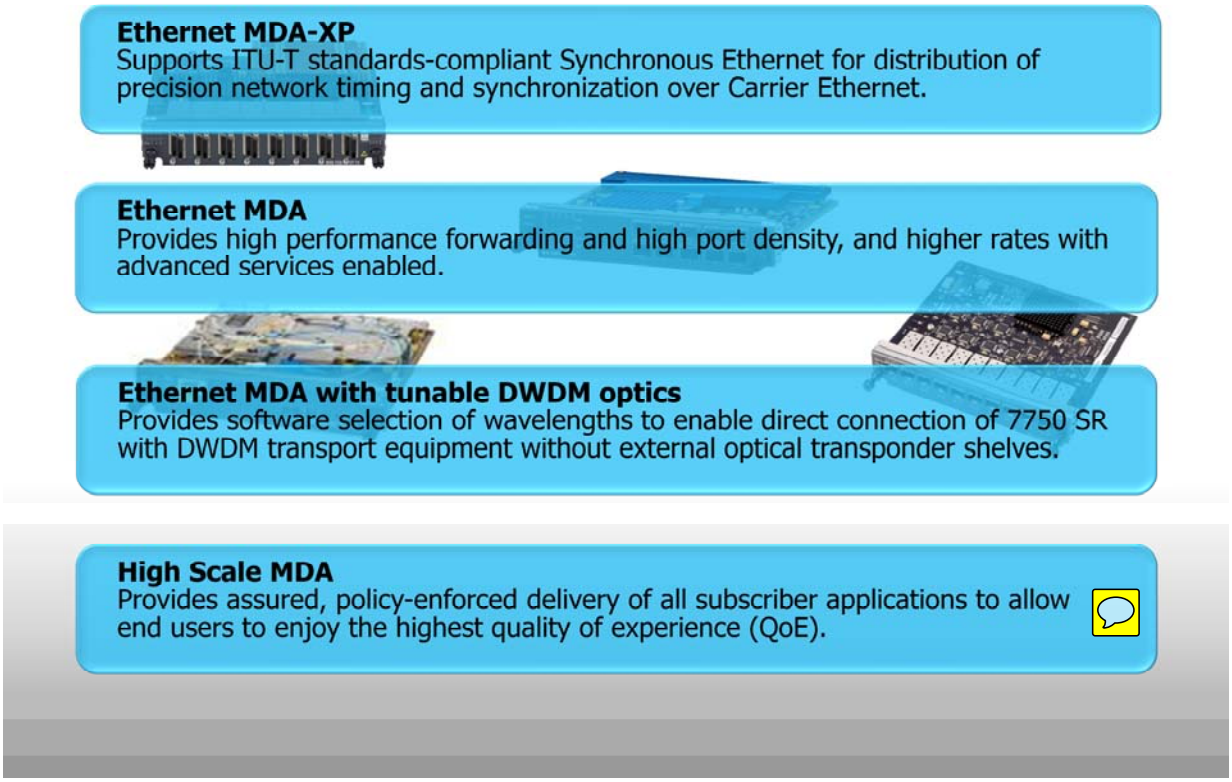
Two Mini Switch Fabric Modules are required (along with the switch fabric function of the SF/CPM card) to provide a fully redundant fabric for the 7750 SR-12e platform. The switch fabric in the 7750 SR-12e operates in a 3+1 redundancy scheme when two of the fabric elements are present on each of the SF/CPMs, and the other two are present on the Mini SFM4 modules. A configuration with two SF/CPM modules and two Mini-SFM4 modules delivers a fully-redundant performance of 200 Gbit/s per slot in full-duplex mode.



The Input Output Module is a full-slot module that inserts into a 7750 SR-12 or SR-7 chassis slot. Each IOM supports up to two Media Dependent Adapters, or up to two Integrated Service Adapters, or a combination of both. The IOM is the distributed forwarding and packet services engine, while the Media Dependent Adapter provides the physical network interfaces, and the Integrated Service Adapters provide processing resources for integrated advanced services.

The 7750 SR IOM supports Layer 2 and Layer 3 IPv4 and IPv5 routing services, IP/MPLS, Ethernet over MPLS, IP VPN and VPLS.

7750 SR Media Dependent Adapters—part 1 of 3

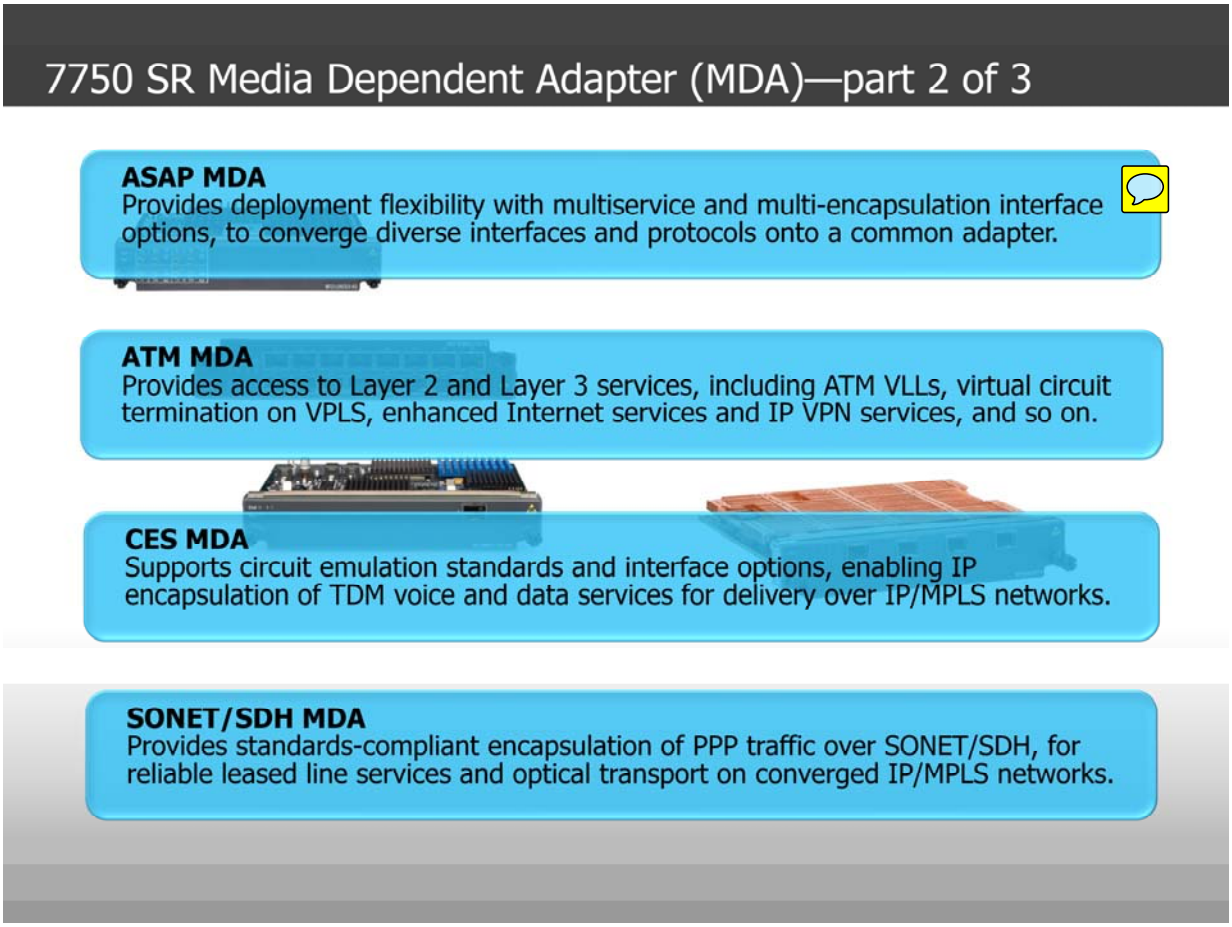


The 7750 SR supports a wide range of Media Dependent Adapter types, interface options, and advanced service delivery capabilities. As a result, the 7750 SR provides the flexibility, performance and scalability to meet the full range of service routing and service requirements.

The Ethernet MDAs support Carrier Ethernet applications and services, including mobile core and backhaul aggregation, Layer 2 and Layer 3 business virtual private networks and broadband residential services and content delivery networks, as well as traditional IPv4 and routing applications.

The Ethernet MDA-XP, Ethernet MDA, Ethernet MDA with tunable Dense Wavelength Division Multiplexing optics, and High Scale MDA deliver high-density, high-performance Carrier Ethernet for business, mobile and residential applications.

All Ethernet MDA types support a wide range of SFP and XFP optical modules, and include variants with tunable optics to interface with DWDM transport equipment.



The 7750 SR supports the following multiservice MDAs: Any Service Any Port MDA, Asynchronous Transfer Mode MDAs, Circuit Emulation Services MDA, and SONET/SDH MDA. These MDAs are available in multiple interface options to support next-generation applications, including Broadband Network Gateways for residential services, multiservice edge for Layer 2 and Layer 3 business VPN services. The Multiservice MDA interface options also support mobile backhaul and mobile packet core for second generation, third generation and fourth generation long-term evolution wireless services.

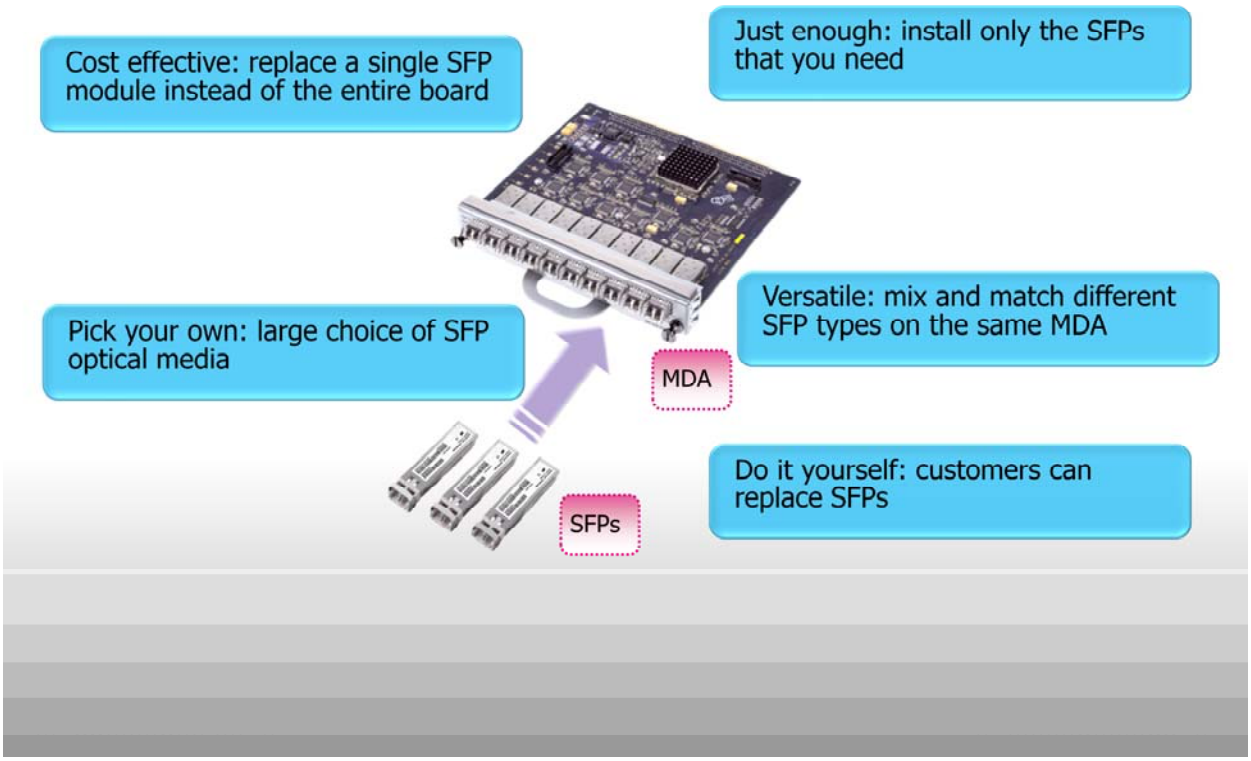
7750 SR Media Dependent Adapter (MDA)—part 3 of 3

VSM-XP
Enables the creation of new services by internally interconnecting services. Multiple VSM-XPs can be installed in a chassis and logically grouped, increasing the overall interconnect capacity through load sharing across VSM-XPs in the group.

Each VSM-XP occupies half a slot in an IOM and can interconnect up to 25 Gbit/s (half-duplex) of services when installed in an IOM3-XP/IOM3-XP-B and 10 Gbit/s when installed in an IOM2 or IOM1.

The 7750 SR Versatile Services Module XP provides high-performance, inter-service connectivity within a single chassis with advanced Quality of Service features. You can use the VSM XP to interconnect pairs of services internally and to create new service offerings. For example, you can interconnect a Layer 2 service and a Layer 3 service to create a "routed" Layer 2 service.

7750 SR Small Form-Factor Pluggable transceivers

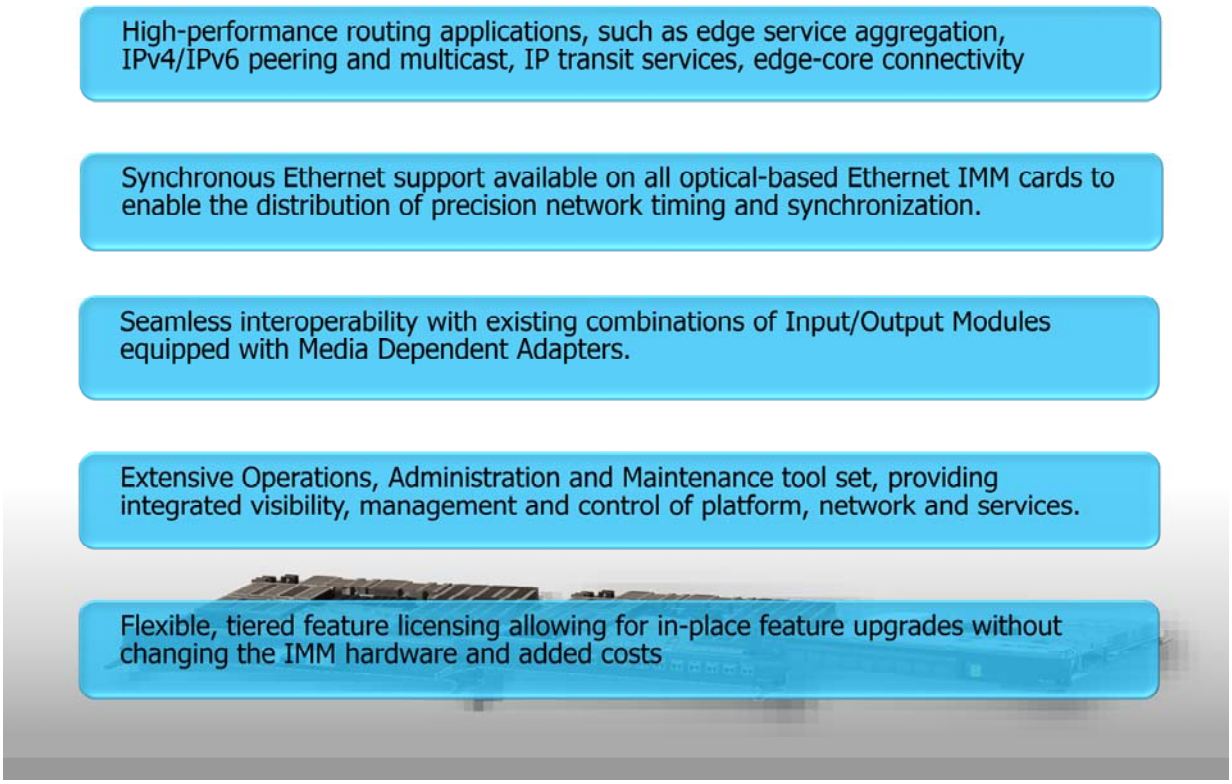


The SFP transceivers are available in a variety of formats. Alcatel-Lucent recommends you use the SFPs that have been tested, verified and validated for the 7750 SR. Alcatel-Lucent programs its SFPs with a type and a part number. As a result, the operator can easily determine the SFP type and replacement part number when troubleshooting.

The use of SFPs allows you to populate the MDAs with the required features on a per-port basis. The combination of MDAs and SFPs means that carriers can defer much of the capital expense of each point-of-presence customer interface until a firm order for service is received.

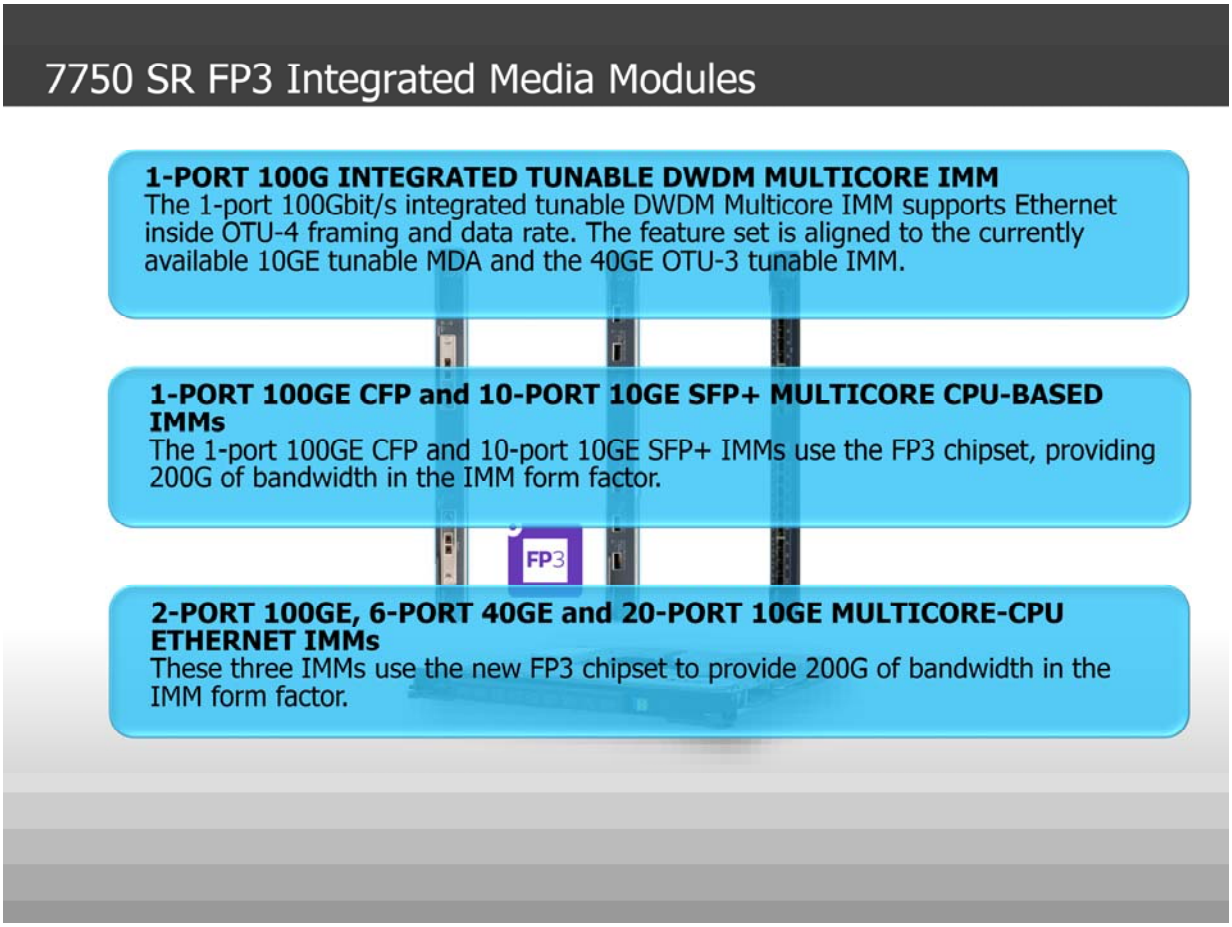
To differentiate them from the Gigabit Ethernet transceivers, the 10-Gigabit Ethernet transceivers are commonly referred to as XFPs.

7750 SR Integrated Media Modules



The IMM is a full-slot module with an FP2 or FP3 network processor complex, equipped with integrated physical ports. An IMM combines the functionality of an IOM board equipped with MDAs on a single interface module that you can install in the 7750 SR. The IMMs provide high-density Gigabit Ethernet and 10 Gigabit Ethernet, high-speed 40-Gb/s and 100-Gb/s interfaces, and high-performance IP/MPLS routing and services. The IMMs support advanced traffic management with Hierarchical Quality of Service, a full range of Layer 2 and Layer 3 routing capabilities, Layer 2 and Layer 3 VPN services and residential services.

The IMM is also available in variants with tunable DWDM optics with Fixed Optic—LC connector interfaces. Refer to the product datasheet for details about the available IMMs for each chassis type.



The most recent SR-OS release introduces support for FP3-based IMMs on 7750 SR-7, SR-12, and SR-12e chassis equipped with the latest SF/CPM.

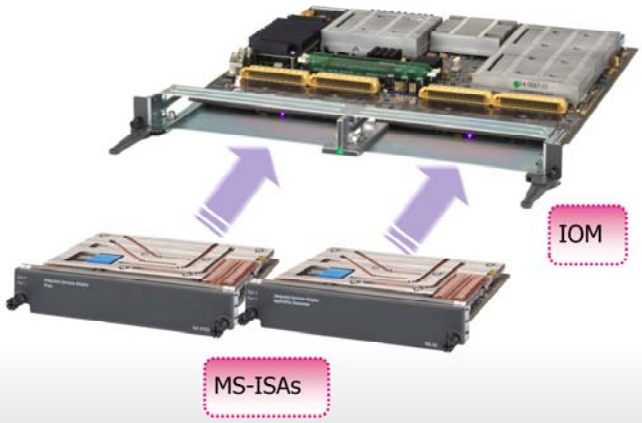
• The one-port 100 Gigabit per second integrated tunable D-W-D-M multicore IMM extends optical reach in long-haul applications without requiring optical signal amplification or dispersion.

• The one-port 100 Gigabit Ethernet C-F-P and ten-port 10 Gigabit Ethernet SFP+ multicore IMMs provide flexible interface options for MDA-like capabilities.

• Each of the two-port 100 Gigabit Ethernet, six-port 40 Gigabit Ethernet and twenty-port 10 Gigabit Ethernet multicore IMMs support 200 Gbit per second throughput when two SF/CPM-4 cards are installed and operational.

7750 SR Integrated Service Adapter—description

The Multiservice Integrated Service Adapter is a half-slot resource blade that inserts into a 7750 SR Input/Output Module



A single Multiservice Integrated Service Adapter is equivalent to multiple external servers, offering scalability and resiliency advantages while reducing costs.

The Alcatel-Lucent Multiservice Integrated Service Adapter delivers advanced business and residential services on the 7750 SR, eliminating the need for expensive external platforms to support these services. The MS-ISA virtualizes the services and makes them available to all the ports across the 7750 SR chassis. The MS-ISA reduces the need for standalone network elements, the space and cabling requirements, the power consumption, the topology churn and the network latency.

7750 SR Integrated Service Adapter—services

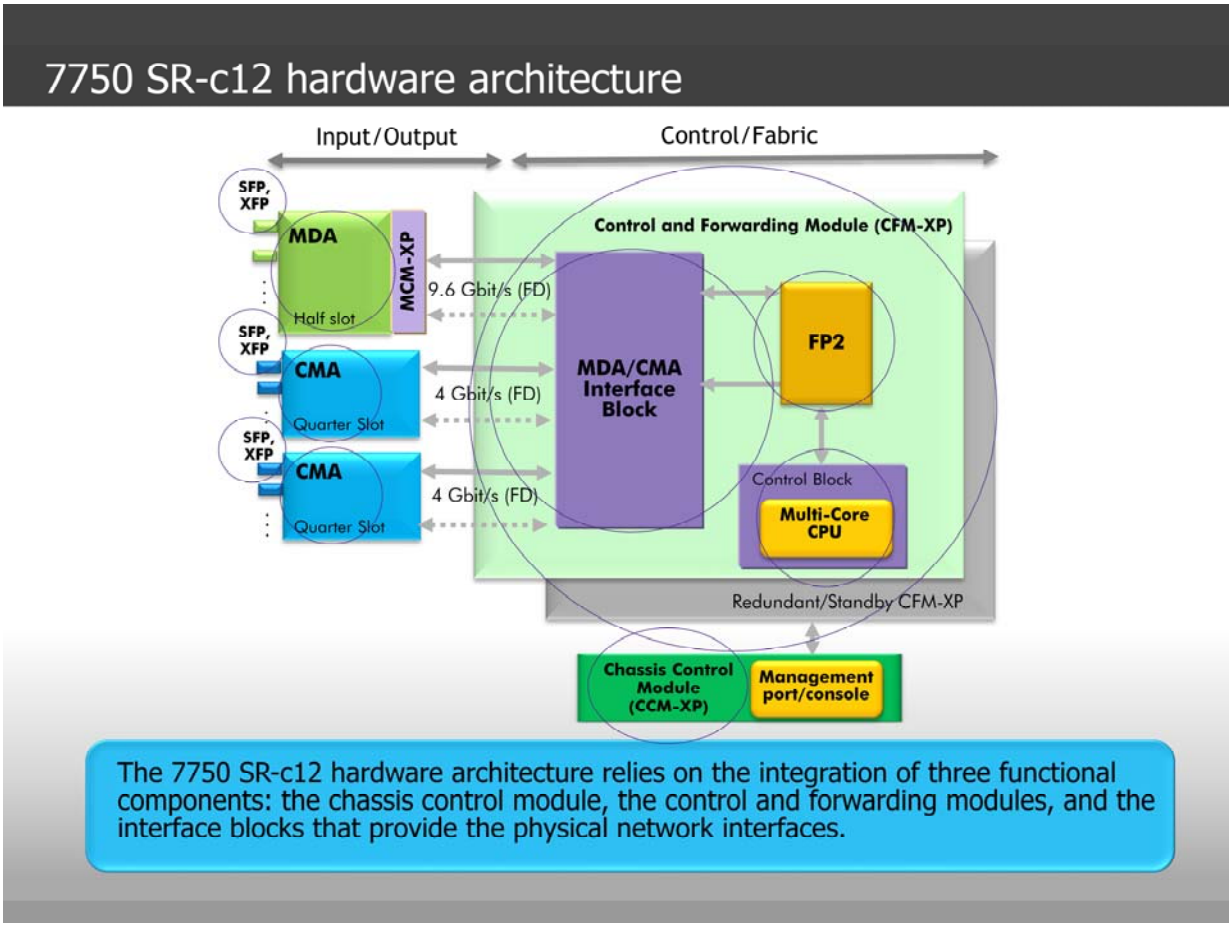
- Application Assurance—including identification and reporting
- L2TP Network Server (LNS) with support for IPv4 and IPv6 features
- Advanced video services—superior Internet Protocol Television QoE
- Network Address Translation (NAT) services—high transaction rates

- MS-ISA Threat Management System (TMS) DDOS Mitigation
- Dual-Stack Lite Address Family Transition Router (AFTR) services
- IPsec services—including Layer 3 VPN connectivity
- WLAN Gateway services with tunneled traffic aggregation

The Alcatel-Lucent MS-ISAs provide purpose-built, extended functionality, and enable deeper levels of integrated services. For details about all supported services and ordering information, refer to the 7750 SR MS-ISA datasheet.

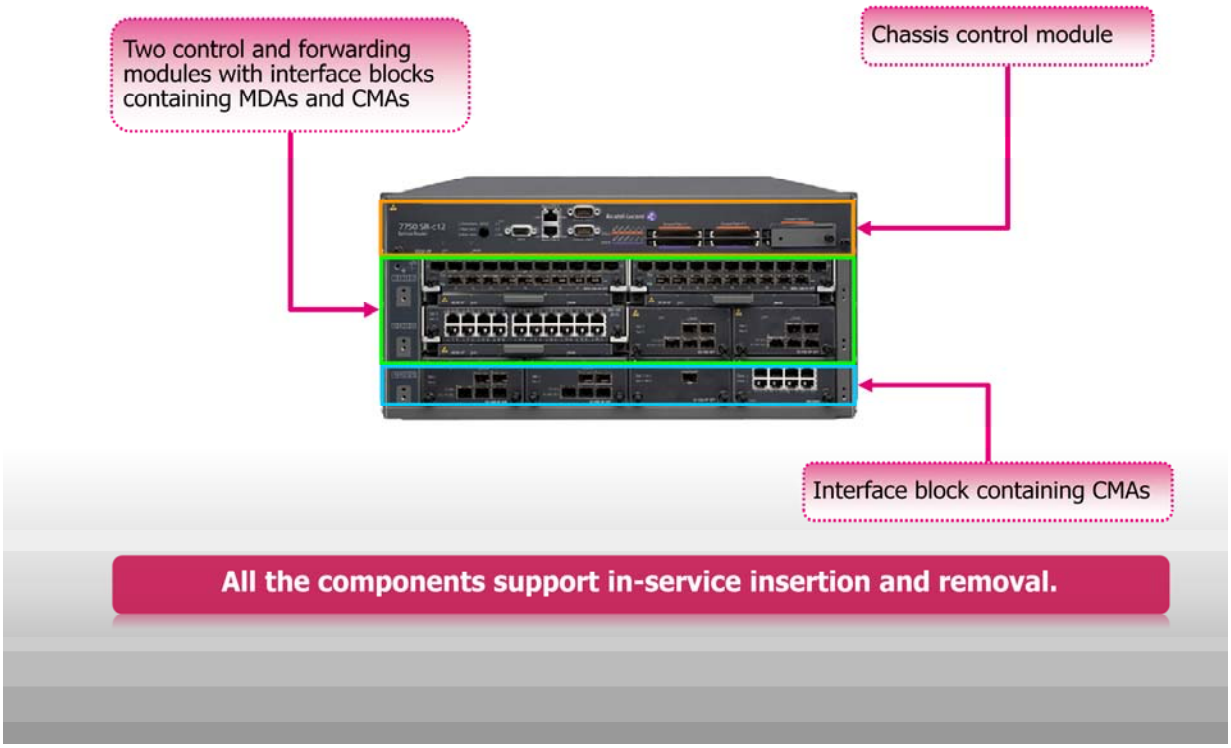
7750 SR-c12 and SR-c4 architecture and hardware overview

Section 3.4 describes the 7750 SR-c12 and SR-c4 architecture and hardware.



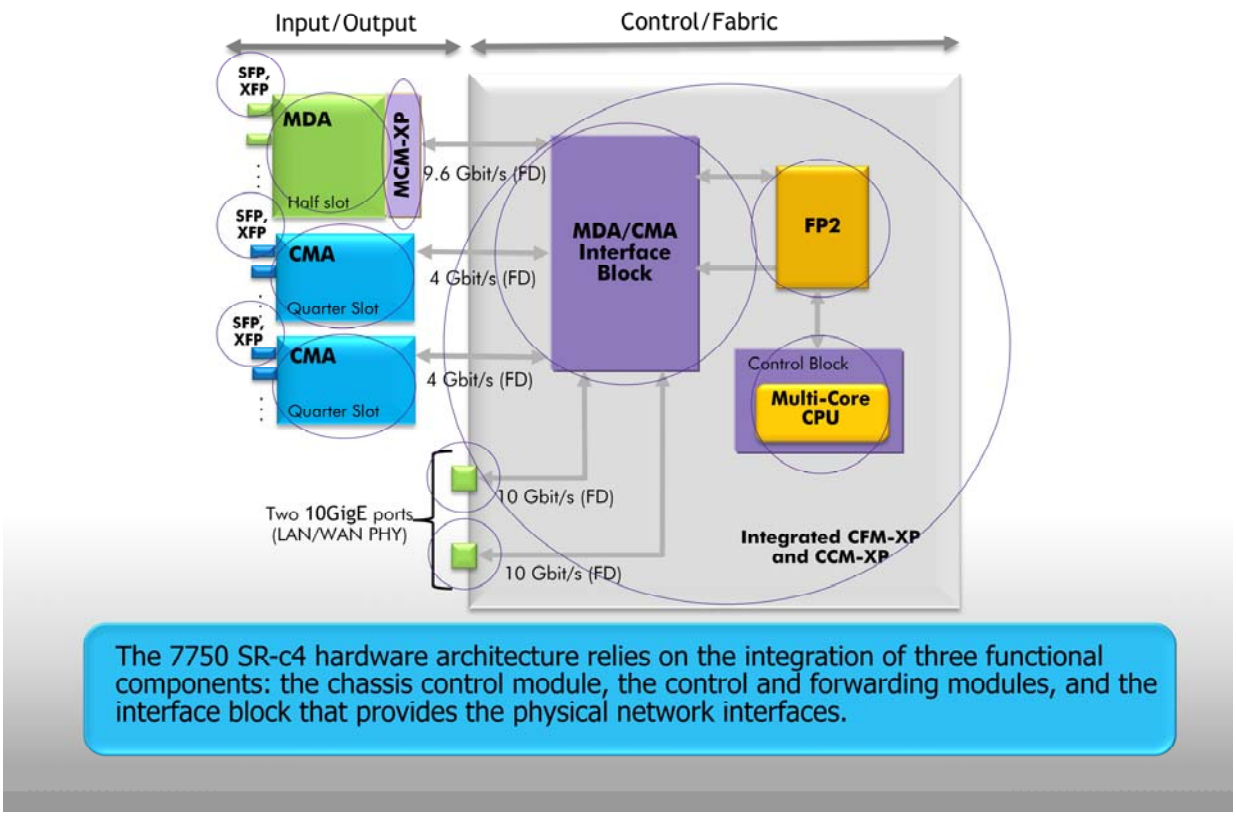
Each 7750 SR-c12 unit contains a Chassis Control Module XP, which provides control ports, connectors, and three compact flash card slots. The CCM XP is connected to two Control and Forwarding Modules XP for active and standby redundancy. Each CFM XP card integrates an FP2 unit for packet processing, traffic management and advanced Quality of Service, and a 10-core CPU for system control, centralized protocol processing, and management. The CFM XP cards are connected to interface blocks, which accommodate Media Dependent Adapters or Compact Media Adapters, or a combination of both. The MDAs and CMAs provide the physical interface connectivity. An MDA carrier module is necessary to adapt an MDA for use in the SR-c12.

7750 SR-c12 with components



The 7750 SR-c12 has a horizontal slot layout with the top slot dedicated to the chassis control module, two middle slots for the control and forwarding modules with interface blocks, and a slot for a third interface block. The chassis control module provides control ports, connectors, and three compact flash card slots; redundant systems have two sets of three compact flash card slots, for a total of six slots.

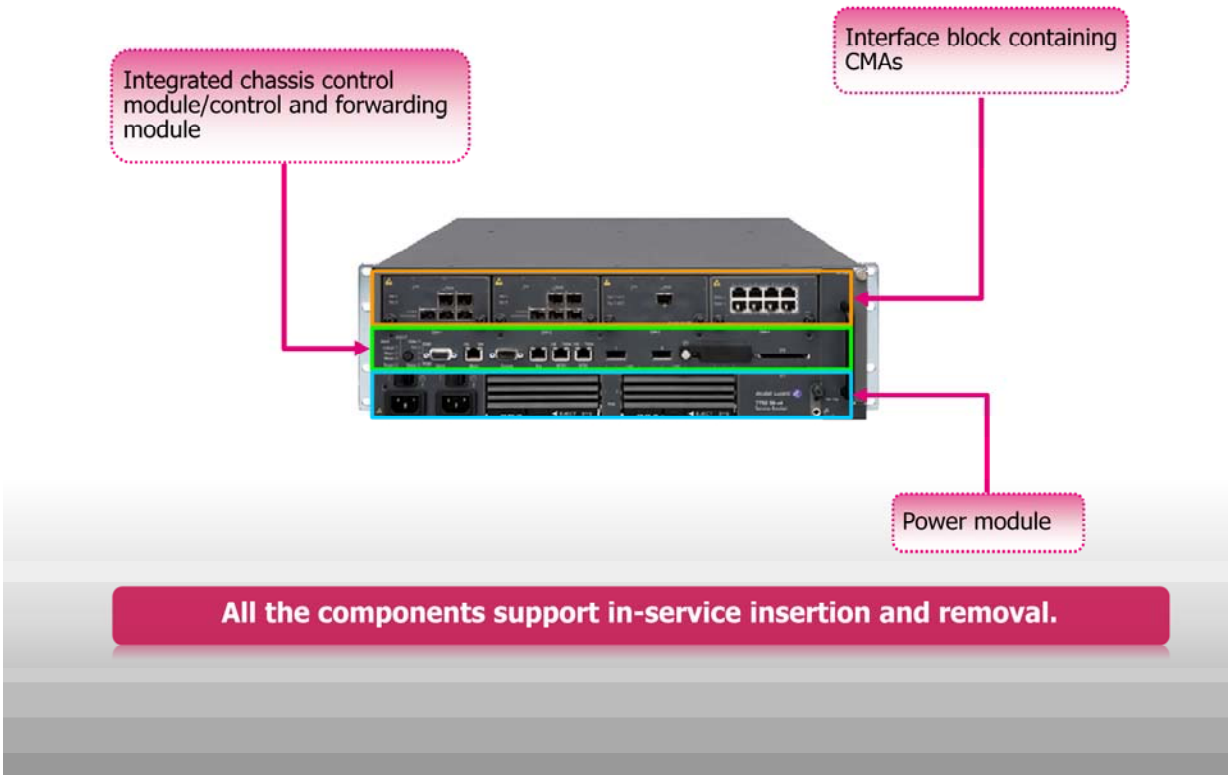
7750 SR-c4 hardware architecture



Each 7750 SR-c4 unit contains a CCM XP, which provides control ports, connectors, and compact flash card slots, integrated with a CFM XP. The CFM contains an FP2 unit for packet processing, traffic management and advanced Quality of Service, and a 10-core CPU for system control, centralized protocol processing and management. The integrated CCM XP CFM XP module is connected to an interface block, which can accommodate MDAs or CMAs, or a combination of both. The MDAs and CMAs provide physical interface connectivity for SFP or XFP units. An MDA carrier module is necessary to adapt an MDA for use in the SR-c4.

The 7750 SR-c4 also provides two built-in 10-Gigabit Ethernet physical ports.

7750 SR-c4 with components



The 7750 SR-c4 has a horizontal slot layout with an integrated CCM and CFM module. The integrated module provides control ports, connectors, and three compact flash card slots, and connects to an interface block, which can accommodate MDAs or CMAs, or a combination of both. In this slide, the interface block contains four CMAs. The bottom unit of the chassis is the power module.

7750 SR-c12 and SR-c4 components

Section 3.5 describes the 7750 SR-c12 and SR-c4 components.

7750 SR-c12 and SR-c4 components at a glance

Chassis Control Module-XP (CCM-XP)
The CCM-XP provides the console and management interface, as well as alarm information for the CFM and power supplies.

Control and forwarding module (CFM-XP)
The CFM-XP controls the routing and switching functions for the entire system.

Compact Media Adapters (CMA)
Quarter-slot card providing physical layer termination—supported options: Ethernet, POS, ATM, ASAP, TDM and CES

Media Dependent Adapters
Half slot module providing various Ethernet adapter types, interface options and advanced services with SFPs and XFPs

MDA Carrier Module-XP (MCM-XP)
Half-slot module required to adapt an MDA for use in the SR-c12 and the SR-c4—support for subset of SR-12 and SR-7 MDAs

The CCM is located at the front of the chassis, and has three slots for compact flash memory cards that store system boot images, software images, and configuration files and logs.

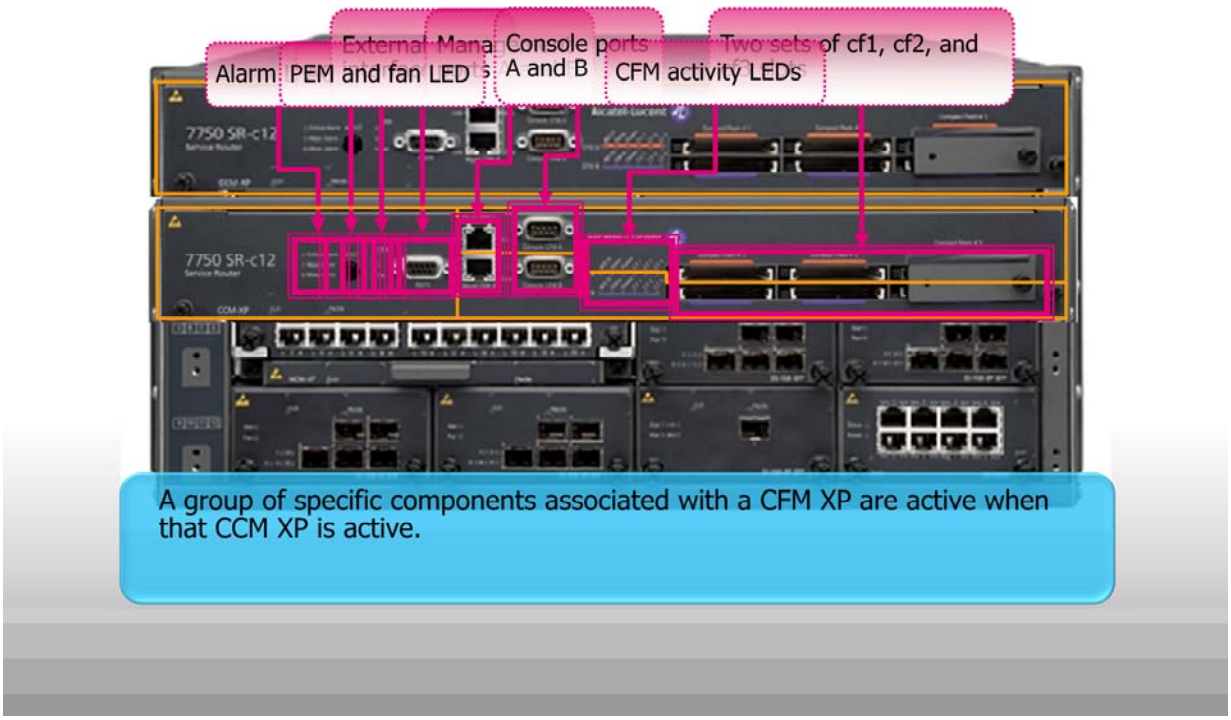
The CFM units are located at the back of the chassis, and carry traffic between the MDAs or the CMAs. The switch fabric portion of a CFM receives and directs traffic to the appropriate destinations according to the routing information.

The CMAs support lower-speed, lower-density applications for maximum interface flexibility.

A 7750 SR-c12 or SR-c4 system supports up to two MDAs of types ASAP, ATM or CES .

Each MDA requires an MDA Carrier Module to function in a 7750 SR-c12 or SR-c4 system.

7750 SR-c12 Chassis Control Module XP



The 7750 SR-c12 has a dedicated Chassis Control Module XP, connected to redundant Control Forwarding Module XPs.

The CCM includes indicators and interfaces that serve the entire chassis, and two duplicated groups of components: one for each CFM.

The CCM-specific elements are: alarm indicators an audible alarm cutoff lamp test button, a power entry module indicator and a fan operation indicator, as well as an alarm port.

The CFM-specific components are: management ports console ports activity LEDs and compact flash card slots.

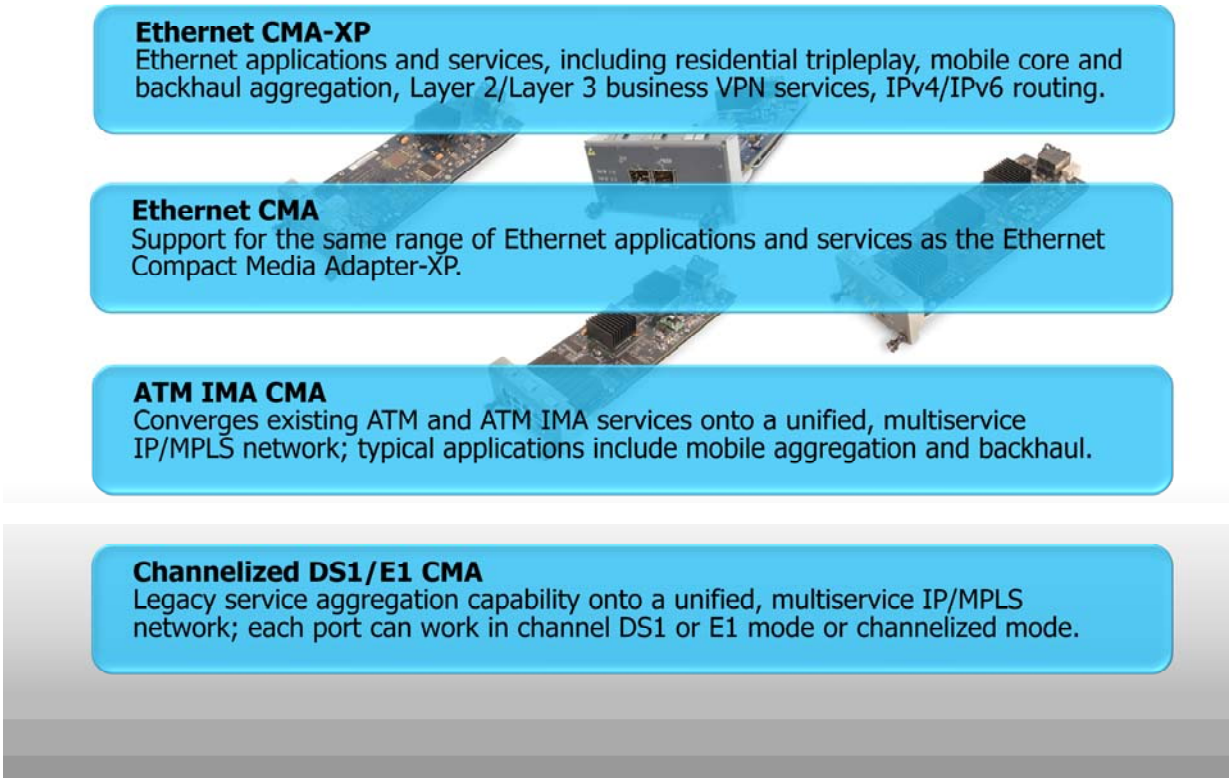
A group of specific components associated with a CFM are active when that CCM is active.

7750 SR-c4 Chassis Control Module XP

The 7750 SR-c4 has an integrated Chassis Control Module and Control Forwarding Module.

The integrated module includes the following components: alarm indicators an audible alarm cutoff lamp test button, a power entry module and fan operation indicators, an alarm port a management port a console port an auxiliary port two BITS ports two 10 Gigabit Ethernet ports, and three compact flash card slots.

7750 SR-c12 and SR-c4 Compact Media Adapters—part 1 of 2

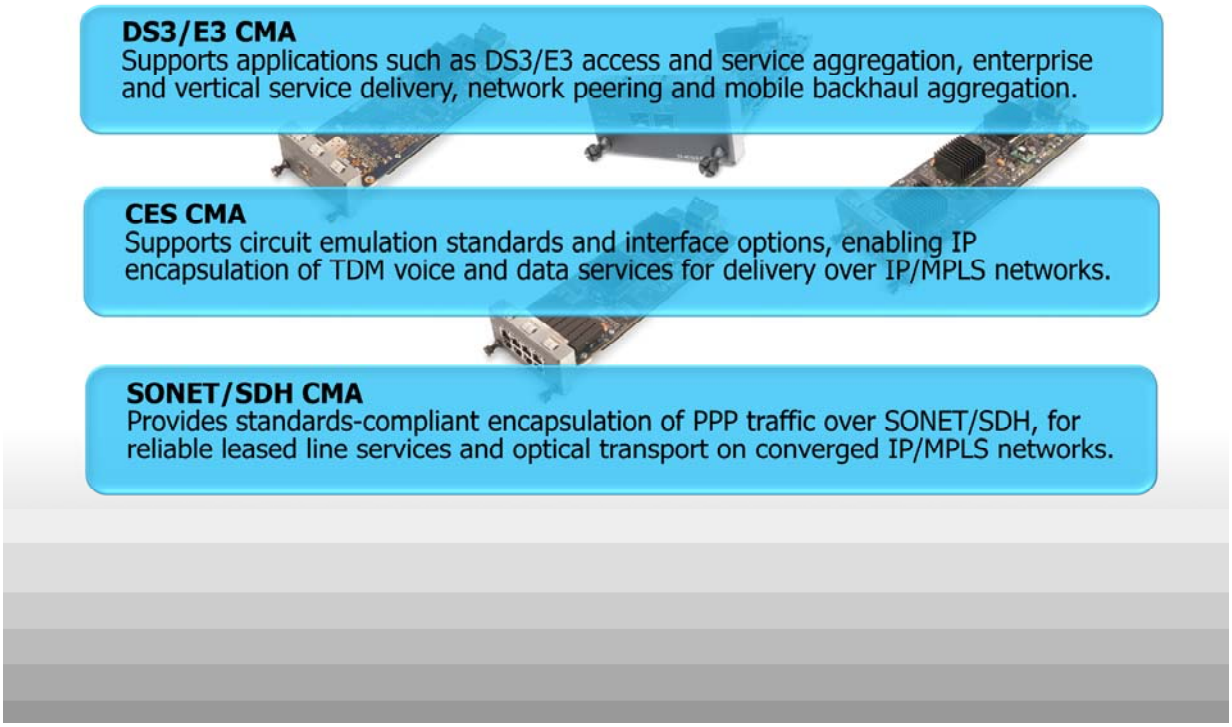


The CMA and CMA XP modules provide service-rich interface options for the 7750 SR-c12 and 7750 SR-c4. The 7750 SR-c12 can accommodate up to eight CMAs in addition to two MDAs and the 7750 SR-c4 can accommodate up to four CMAs.

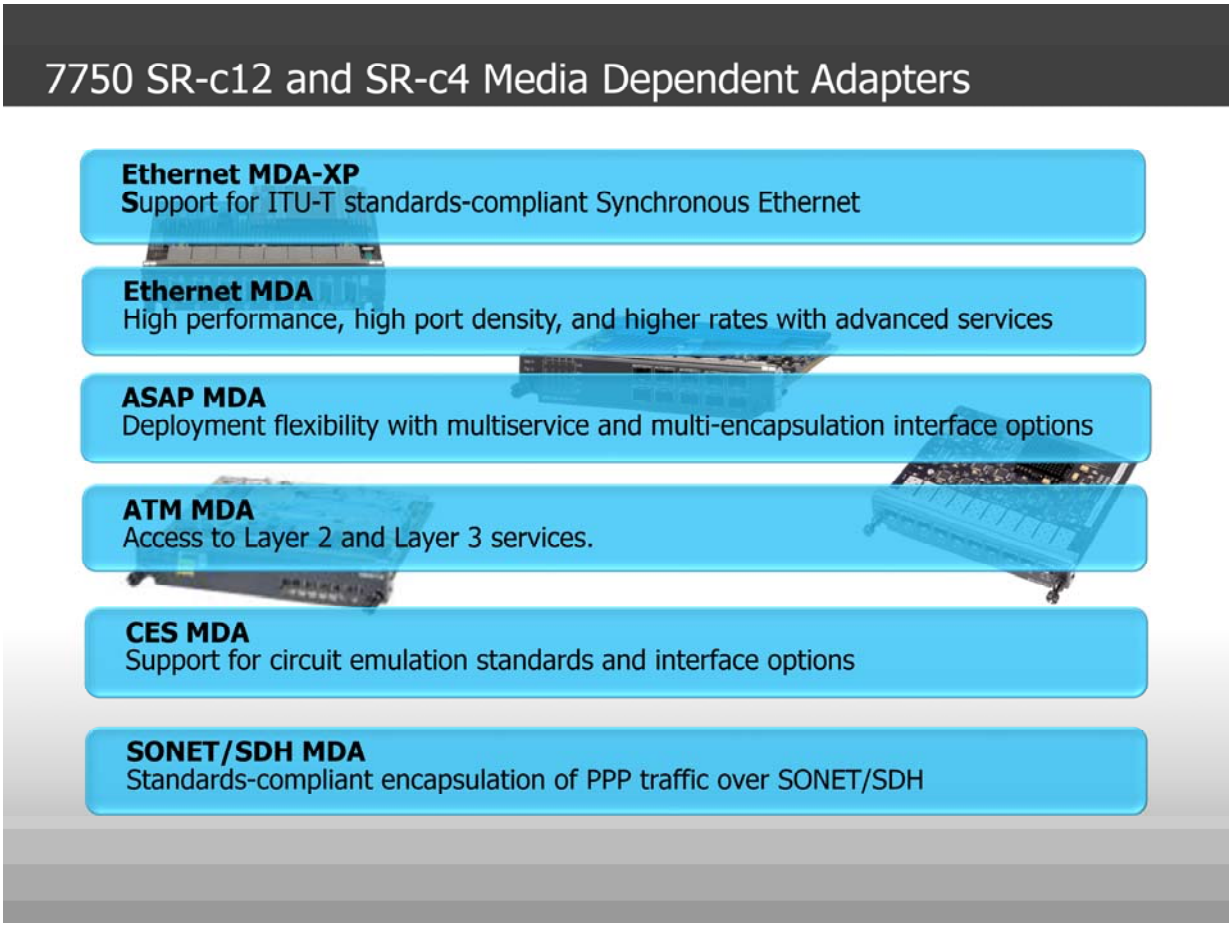
The CMA XP modules are available for Ethernet applications, and the CMA modules are offered for Ethernet, ATM, SONET/SDH, CES, DS1/E1 and DS3/E3 services. You can mix CMA-XP and CMA modules on the same chassis for flexible configuration options.

The CMA XP and CMA modules that accept SFP transceivers support a wide range of pluggable optics with Digital Diagnostic Monitoring for improved installation, activation and troubleshooting capabilities.

7750 SR-c12 and SR-c4 Compact Media Adapters—part 2 of 2

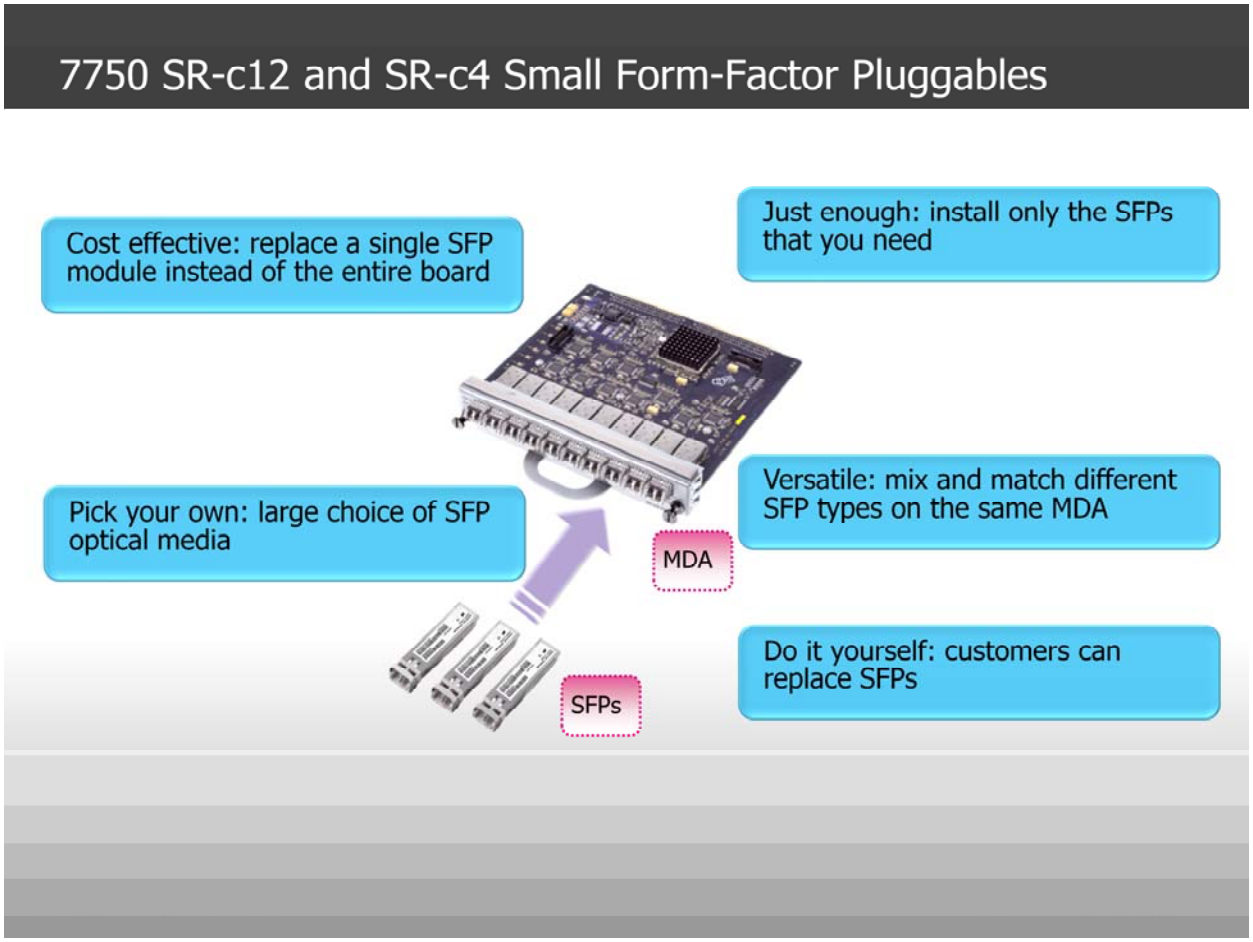


For details about all the features and applications that the CMA and CMA XP modules support and for ordering information, refer to the product datasheet.



The 7750 SR-c12 and SR-c4 support most of the MDA types that are supported by the 7750 SR-12 and SR-7 systems, with the following exceptions: Ethernet MDA with tunable DWDM optics, High Scale MDA, and Versatile Services Module XP.

The Ethernet MDA types support a wide range of SFP and XFP optical modules. Some SFPs and XFPs provide tunable optics to interface with DWDM transport equipment.



The 7750 SR-c12 and SR-c4 support most of the SFP and XFP transceivers that are supported by the 7750 SR-12 and SR-7 systems, with the exceptions of the units that apply to the unsupported MDAs.

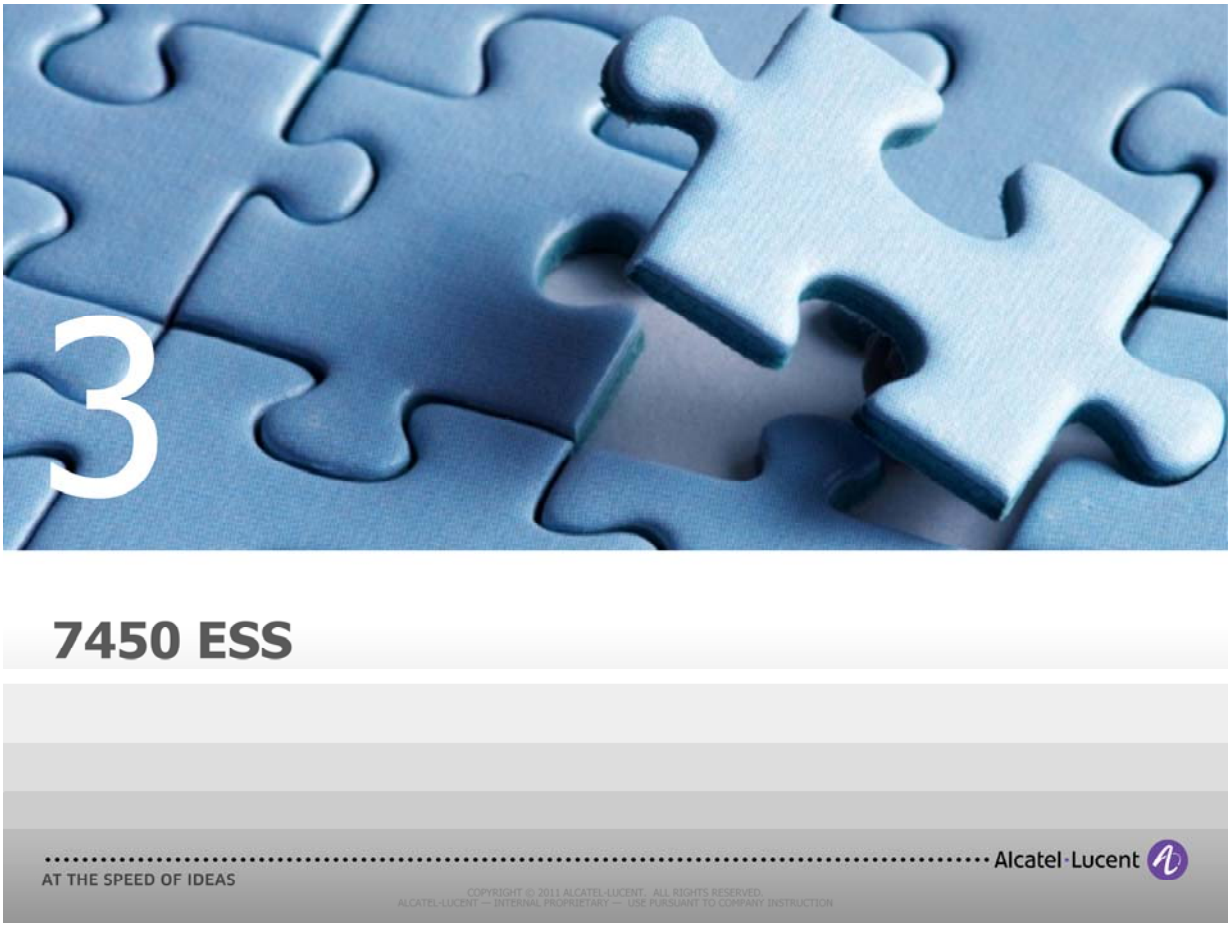
7750 SR-c12 and SR-c4 Integrated Service Adapter

The MS-ISA is a half-slot resource blade that is equivalent to multiple external servers, offering scalability and resiliency advantages while reducing costs.

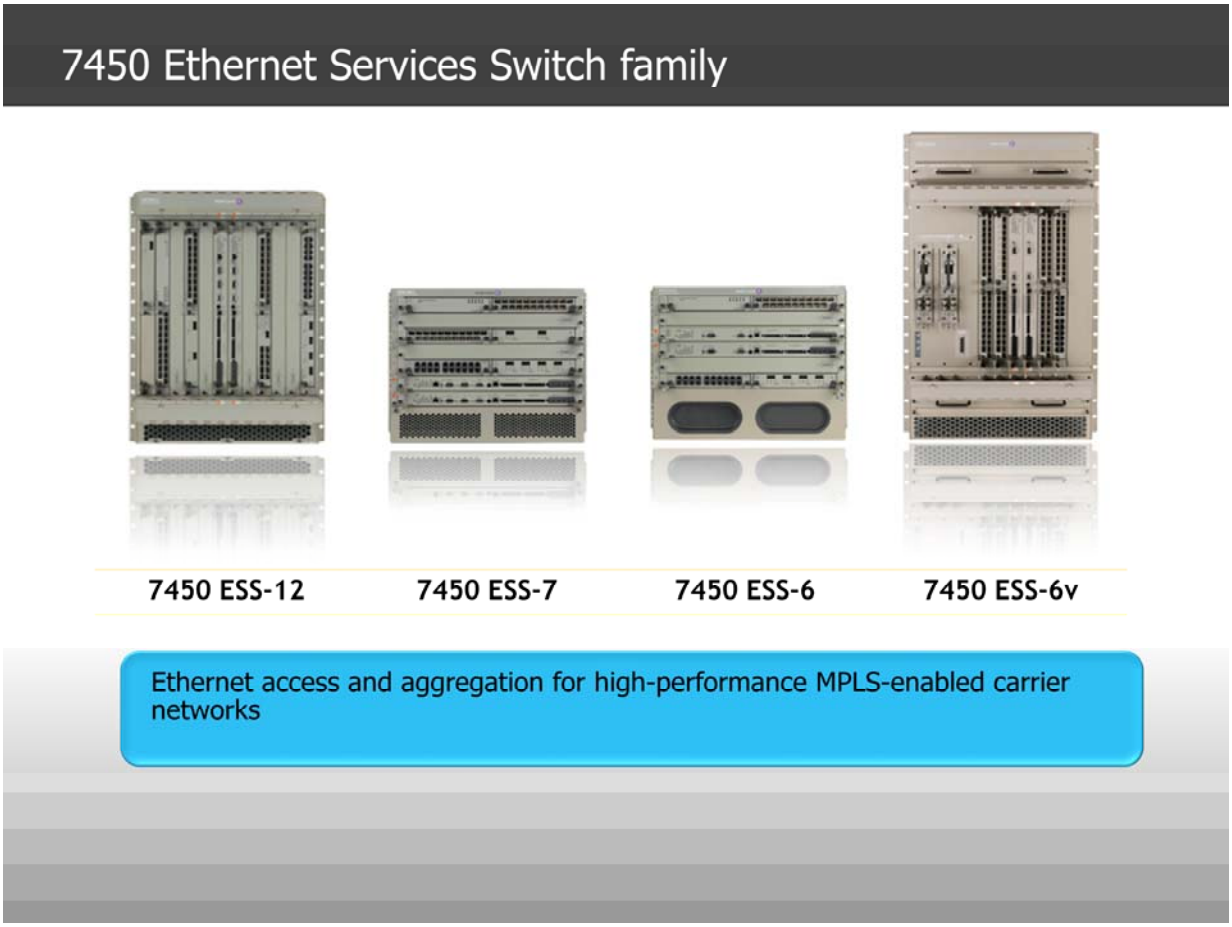
The 7750 SR-c12 supports up to two MS-ISAs in each chassis, for up to 20 Gbit/s of wire-speed processing and application recognition.

The 7750 SR-c4 supports one MS-ISA in each chassis, for up to 10 Gbit/s of wire-speed processing and application recognition.

The Multiservice Integrated Service Adapter provides extended functionality and enables deeper levels of integrated services. The 7750 SR-c12 supports the Application Assurance and IPsec services MS-ISAs, and the 7750 SR-c4 supports the Application Assurance MS-ISA.



Section 3 introduces the 7450 Ethernet Service Switch or ESS.



The 7450 ESS provides Ethernet access and aggregation services over high-performance MPLS-enabled carrier networks of various sizes. The 7450 ESS-12 offers the high capacity required to cover an international or national reach, while the 7450 ESS-7, ESS-6 and ESS-6v models can be deployed in metro or national reach networks.

A 7450 ESS system consists of a physical chassis with slots that house cards. The cards provide the intelligence of the system and the physical interfaces with the network.

7450 ESS capacity and components at a glance

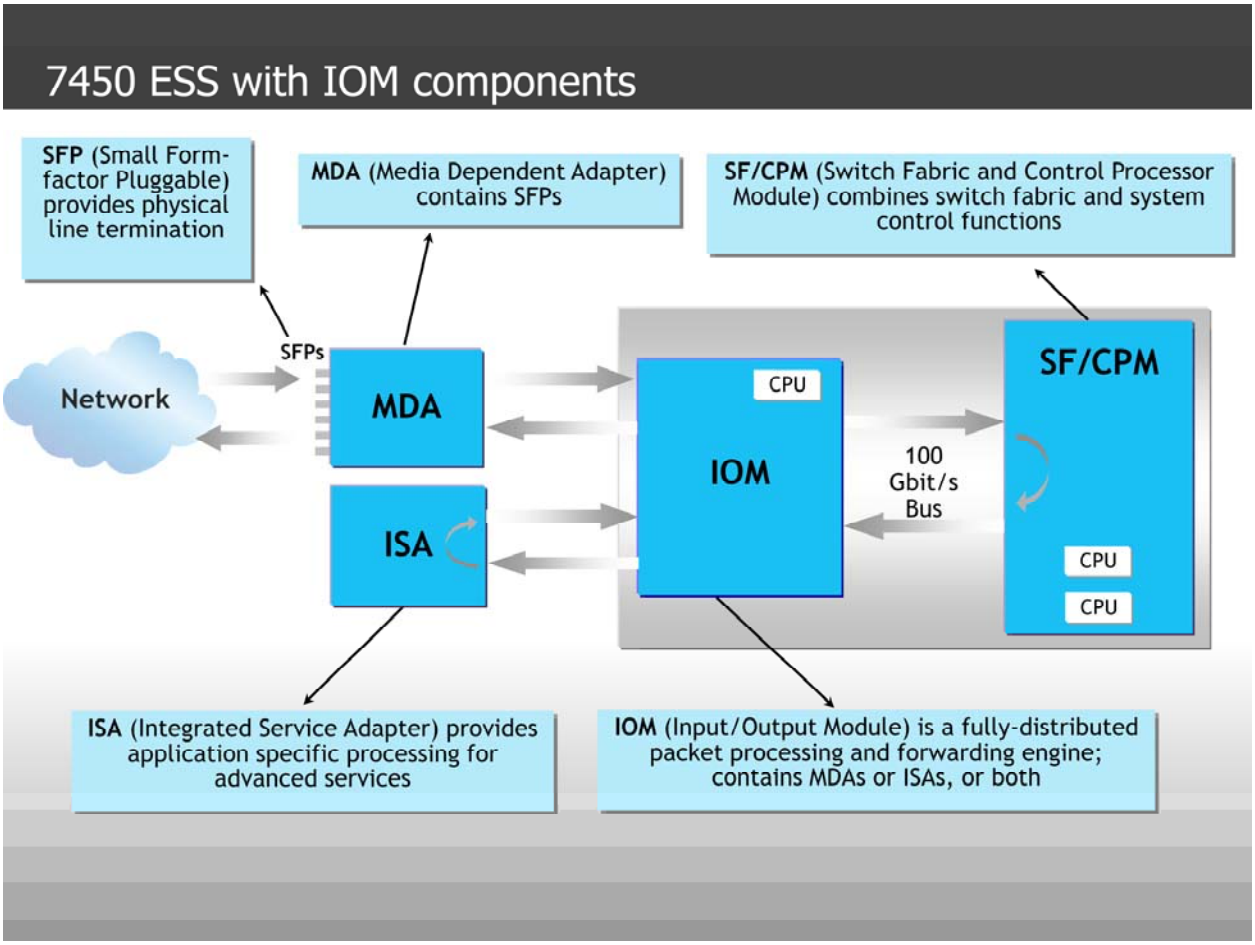
7450 ESS model	Capacity	Slots	Components	Slot capacity
7450 ESS-12	2 Tbit/s	12	2 SF/CPMs and 10 IOMs or IMMs	Up to 100 Gbit/s
7450 ESS-7	1 Tbit/s	7	2 SF/CPMs and 5 IOMs or IMMs	Up to 100 Gb/s
7450 ESS-6	320 Gbit/s	6	2 SF/CPMs and 4 IOMs	Up to 40 Gbit/s
7450 ESS-6v	320 Gbit/s	6	2 SF/CPMs and 4 IOMs (vertical configuration)	Up to 40 Gbit/s

The table on this slide summarizes the capacity and components of the 7450 ESS models.

There is no capacity difference between the 7450 ESS-6 and the 7450 ESS-6v models. The 7450 ESS-6v has a vertical slot configuration.

7450 ESS architecture and hardware overview

Section 4.1 describes the 7450 ESS architecture and hardware components.



This diagram illustrates the structure of the 7450 ESS with Input Output Modules.

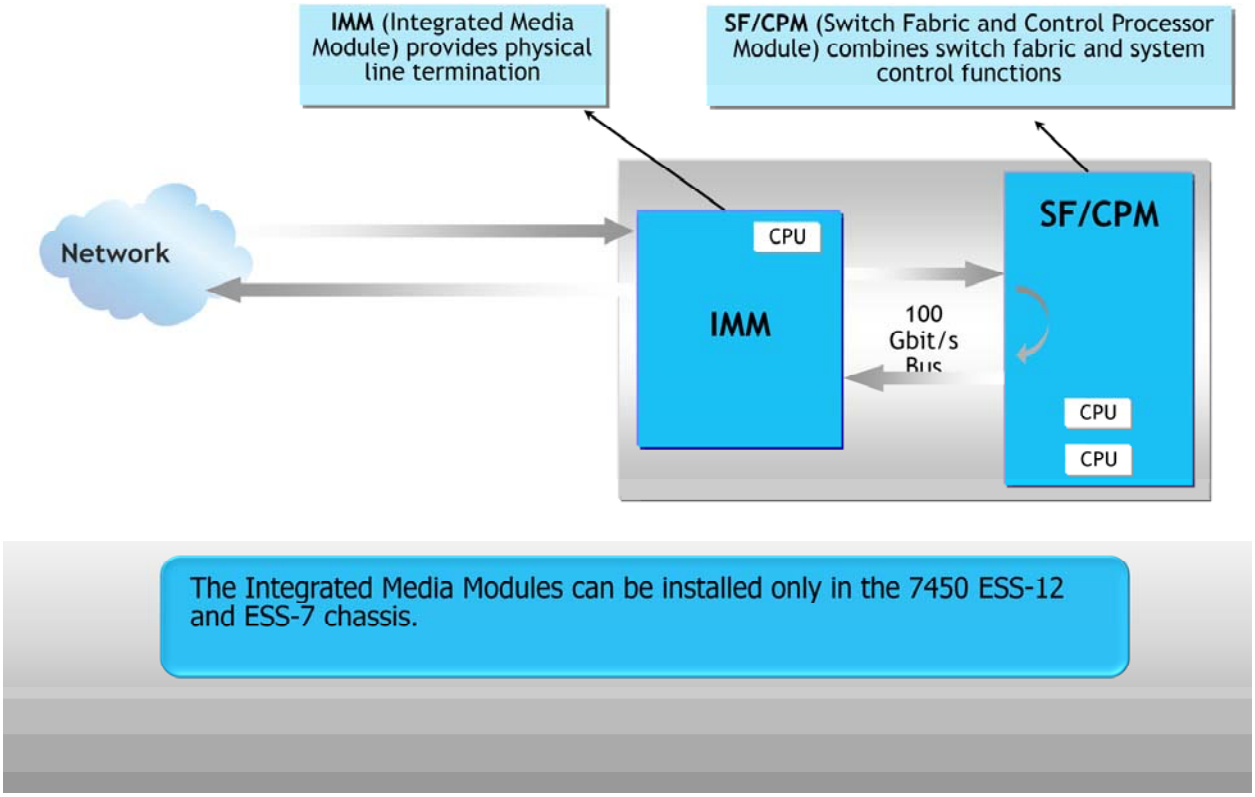
The Switch Fabric and Control Processor Module controls the routing and switching functions for the entire system and provides the management and console interfaces, as well as an interface for the external synchronization signal.

The IOM is a fully-distributed packet processing and forwarding engine, which receives and processes packets to accomplish switching and routing tasks. The IOM provides connectivity to the switch fabric module to move each packet from the ingress interface to the egress interface.

The Media Dependent Adapter contains Small Form-factor Pluggable units that provide the physical network interfaces. The 7450 ESS supports the mixing and matching of different SFP types on the same MDA.

The Multiservice Integrated Service Adapter provides application-specific processing for advanced services such as Application Assurance and Video Processing. The MS-ISA module can be plugged into the IOM cards like the MDAs.

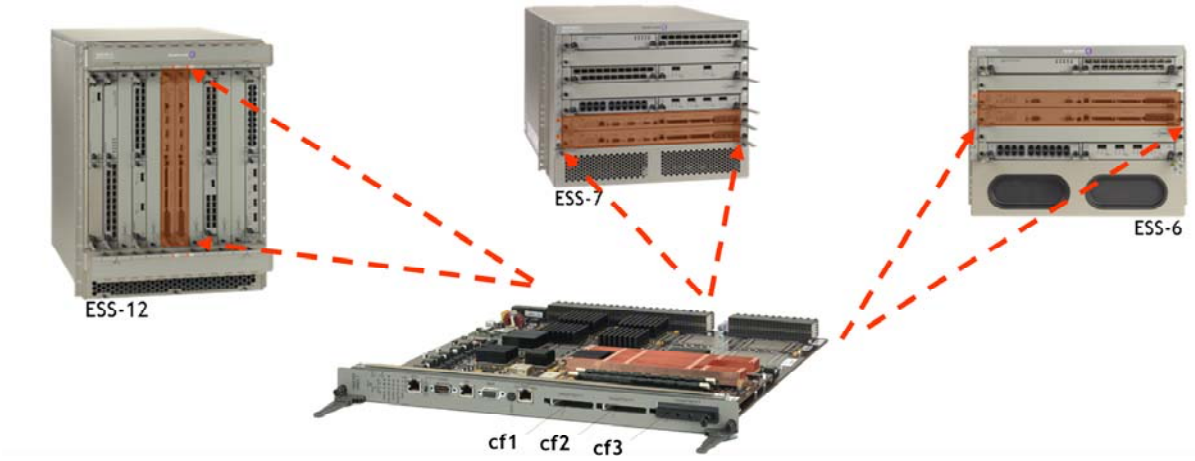
7450 ESS with IMM components



This diagram illustrates the structure of the 7450 ESS with Integrated Media Modules.

The IMM integrates IOM, MDA and SFP functionality onto a single full-slot, high-density, high-capacity interface module.

7450 ESS—SF/CPM description

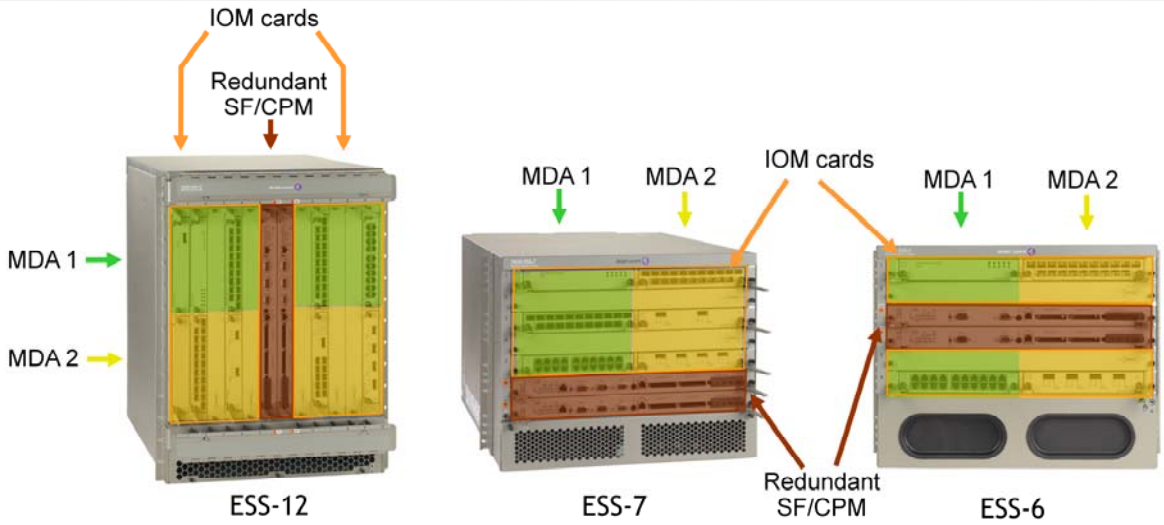


The two load-sharing, redundant Switch Fabric and Control Processor Modules, located in slots A and B, provide full redundancy.

Each SF CPM contains two central processors. The SF CPM supports hitless switchover using non-stop routing and non-stop service capabilities, as well as routing, switching and Operations, Administration, and Maintenance protocols.

Each SF CPM has three slots for compact flash cards: cf1, cf2, and cf3. Each new system is supplied with a compact flash card that contains the files required to start the system; this is typically inserted in the cf3 slot. You can use the cards in the cf1 and cf2 slots to store debugging and accounting logs.

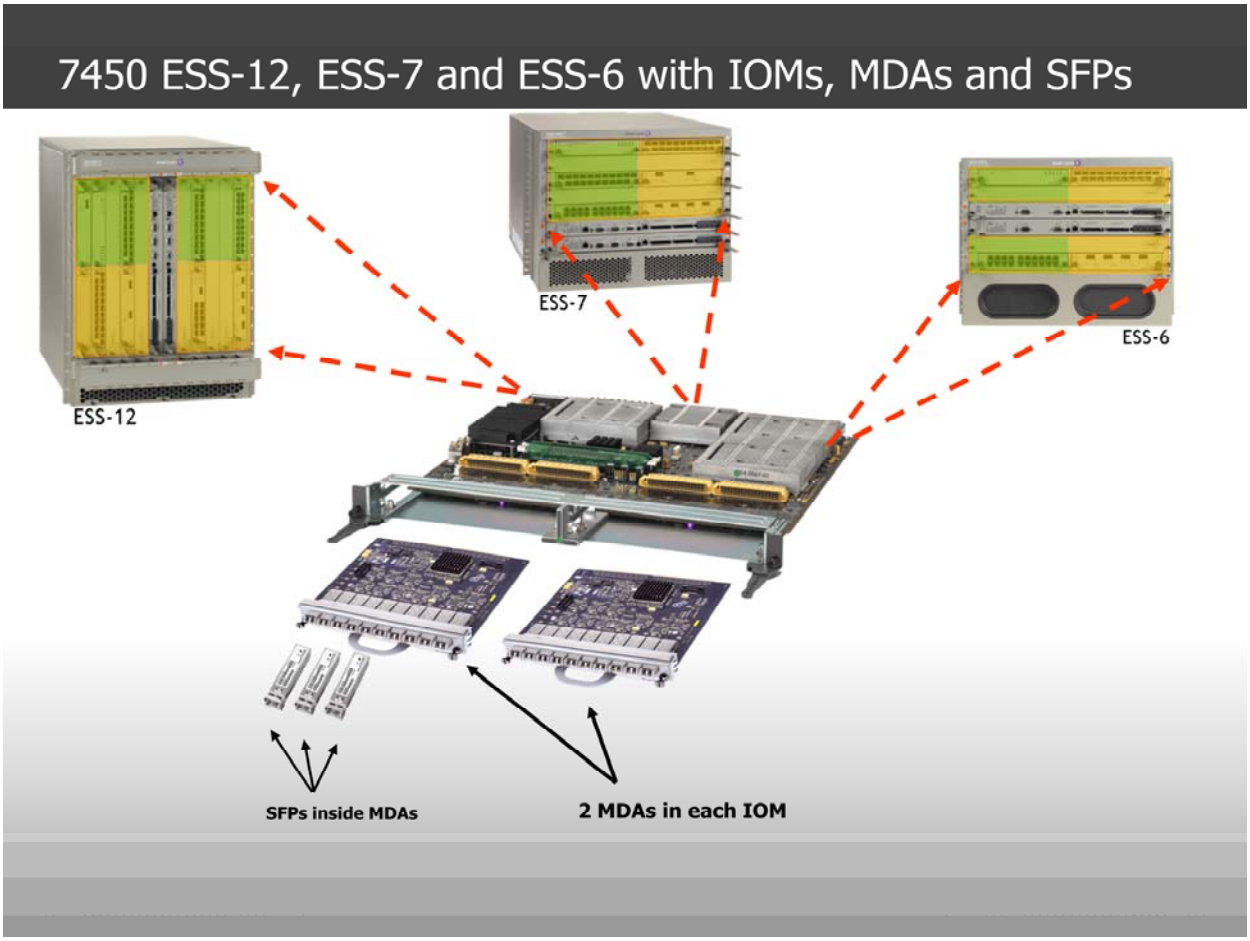
7450 ESS-12, ESS-7 and ESS-6 with IOM cards



All the components support in-service insertion and removal.

This is where the SF CPMs, IOMs and MDAs are located in the 7450 ESS-12, ESS-7 and ESS-6 models.

The slide does not show the 7450 ESS-6v, which has the same capacity and accommodates the same components as the ESS-6 model in a vertical form factor that looks like the ESS-12 chassis.



The IOMs are responsible for the queuing, processing and forwarding of data. The 7450 ESS-12 can accommodate 10 IOMs; the ESS-7, 5 IOMs; and the ESS-6 or ESS-6v, 4 IOMs.

Depending on the chassis type and the cards inserted in the IOMs, you can configure multiple operating modes with various feature sets and slot capacities.

The MDAs provide one or more physical interfaces, such as Ethernet or SONET/SDH. The wide variety of MDAs assures the flexibility to populate the 7450 ESS system according to your needs.

The SFPs are small optical or copper modules available in a variety of formats that can be installed in the MDAs.
The 7450 ESS can even have some 7750 SR MDAs installed, as long as the system is properly configured to function in mixed mode.

7450 ESS components at a glance

Switch Fabric and Control Processor Module (SF/CPM)
Multi-core control processor that runs routing, switching and OAM protocols

Input Output Module
Full slot module that can accommodate up to two MDAs or two ISAs, or one MDA and one ISA

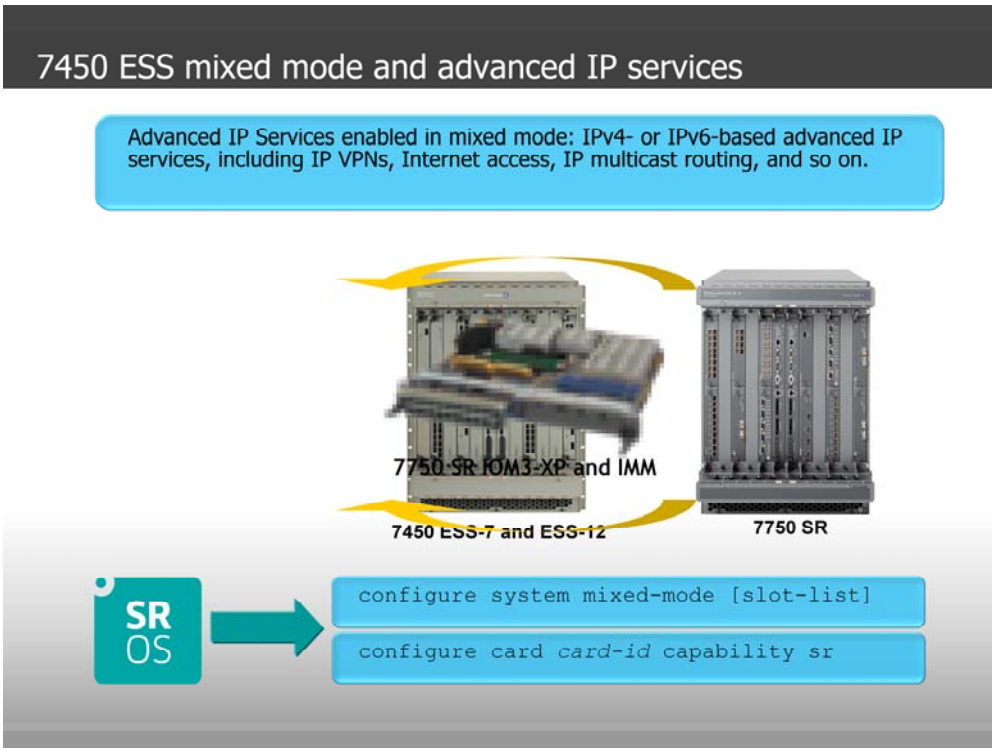
Media Dependent Adapters
Half slot module that provides various Ethernet adapter types, interface options and advanced services with SFPs and XFPs

Integrated Media Module
Full slot module with FP2 network processor complex equipped with integrated physical ports

Integrated Service Adapter
Half slot resource blade without physical ports that inserts into an input/output module and provides specialized services

The 7450 ESS uses the same types of components as the 7750 SR, but there are key differences in component support and functionality.

The 7450 ESS-7 and ESS-12 can be configured in a mixed mode of operation to support some of the 7750 SR components and services.



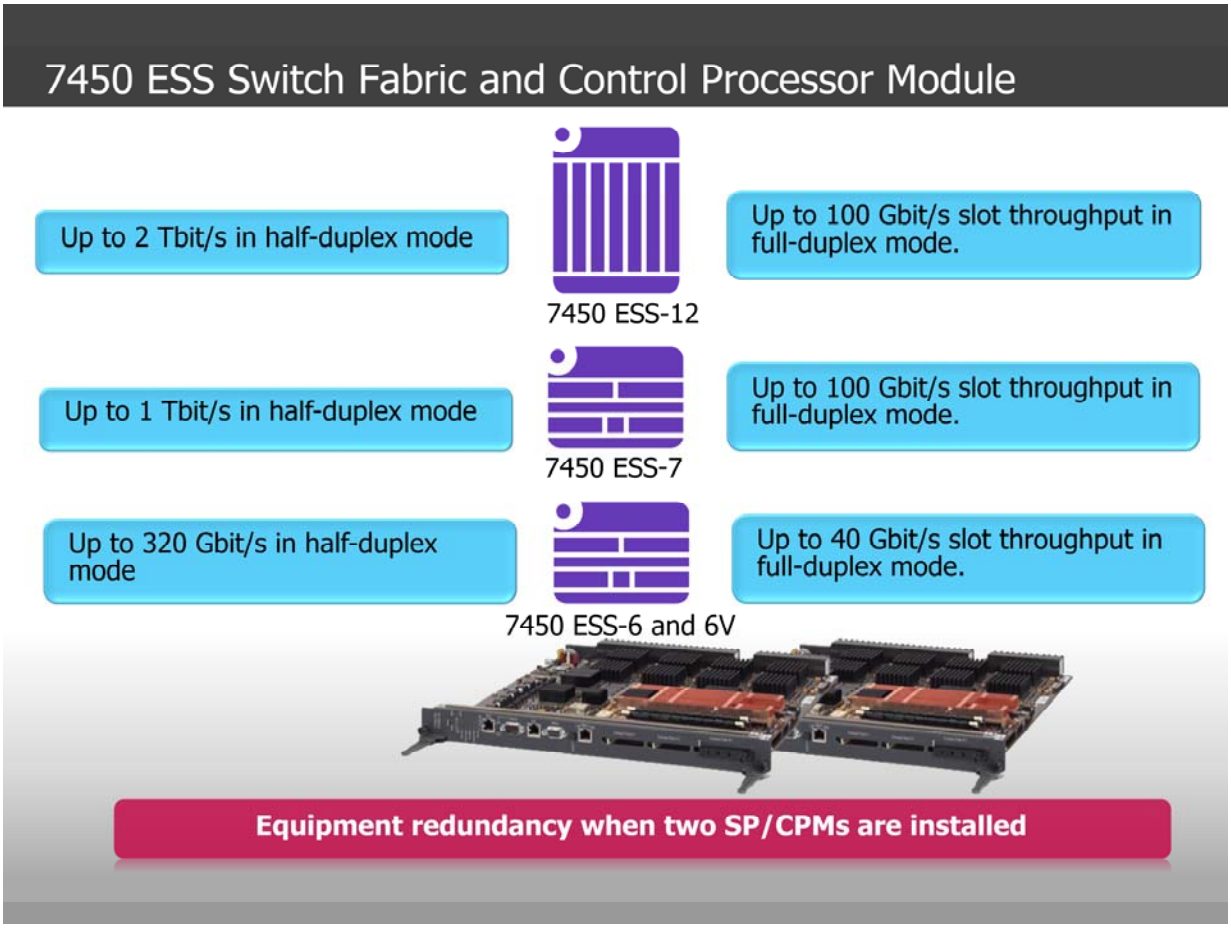
The 7450 ESS-7 and ESS-12 can operate in two modes: standard mode and mixed mode. The standard mode is the default mode of operation. In standard mode, the 7450 ESS uses the components designed to be used in a 7450 ESS and delivers the standard features and services. In mixed mode, the 7450 ESS can use some 7750 SR components that extend the 7450 ESS functionality to support 7750 SR features and services. The 7750 SR features are called Advanced IP Services, and can be enabled only on 7750 SR IOM3-XP cards.

The Switch Fabric and Control Processor Modules do not require an upgrade for the 7450 ESS to function in mixed mode.

To enable the mixed mode of operation, you need to configure the 7450 ESS and change card capabilities for the system to accept 7750 SR components and provide Advanced IP Services. You perform these tasks in the Service Router Operating System.

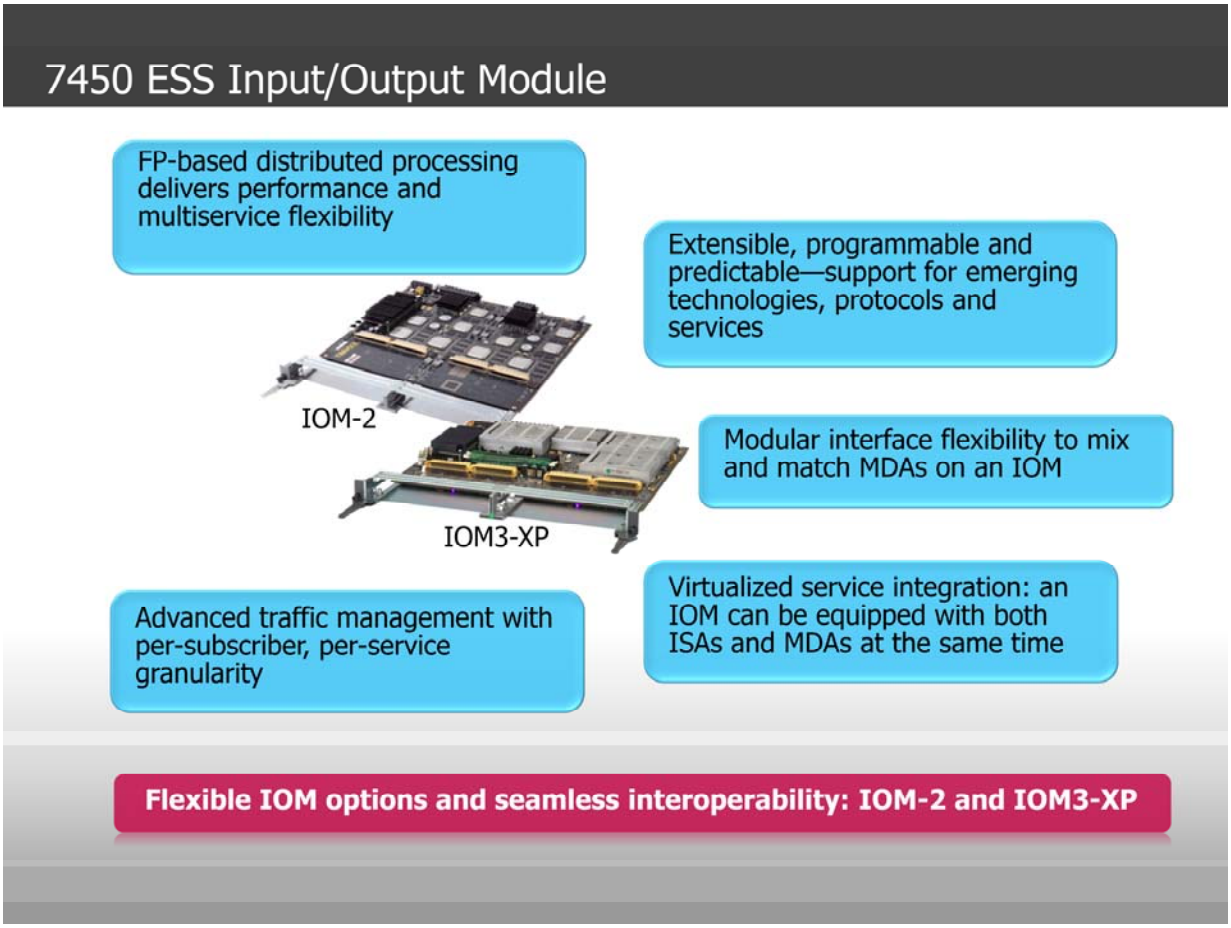
7450 ESS components

This section describes the 7450 ESS components.



The Switch Fabric and Control Processor Module integrates the switch fabric and control functions into one unit. All 7450 ESS models require one SF CPM to function. When the 7450 ESS is equipped with two SF CPMs, the switch fabrics are active on both cards and operate in a load-sharing arrangement, doubling the chassis switch fabric throughput.

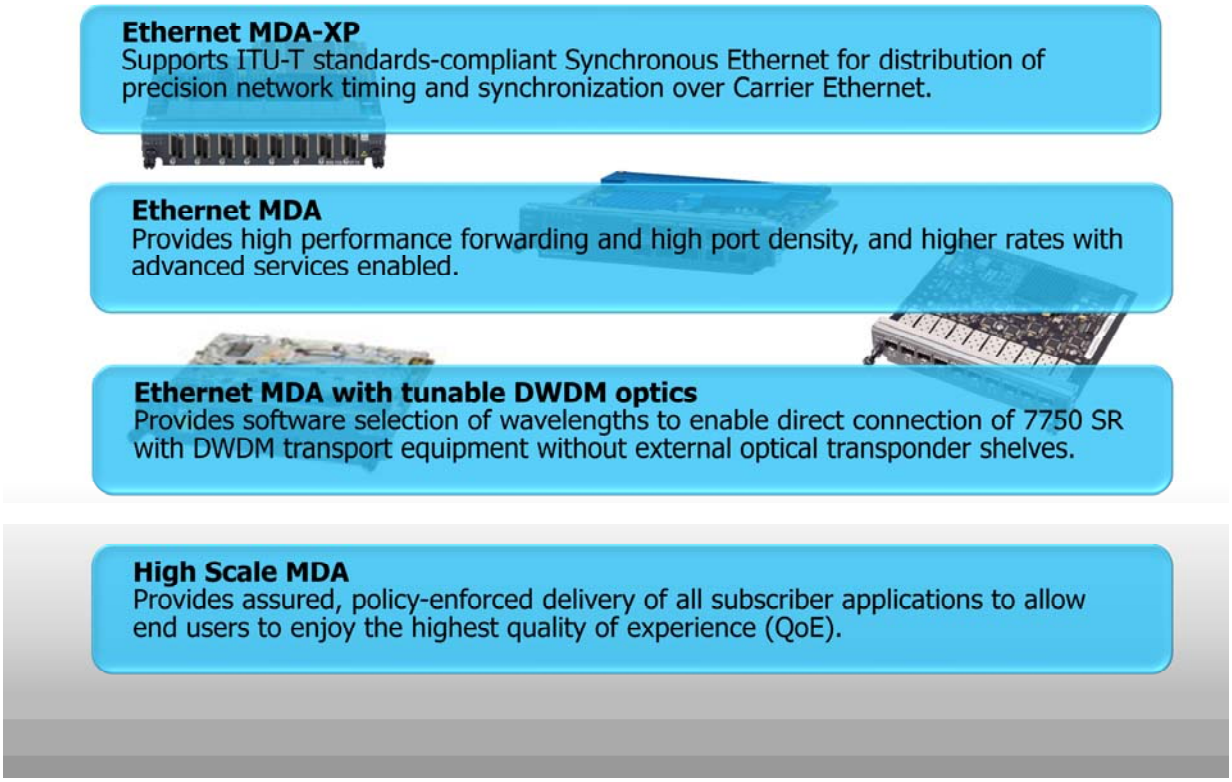
Each SF CPM contains ten 750-MHz core processors. The software that supports the SF CPM is enhanced for symmetric multi-processing. In a redundant configuration, the SF CPM allows for hitless switchover using non-stop routing and non-stop service capabilities.



The IOM is a full-slot board that inserts into a 7450 ESS chassis slot. Each IOM supports up to two MDAs or two ISAs, or a combination of both. The IOM is the distributed forwarding and packet services engine, while the MDA provides the physical network interfaces, and the two ISAs provide processing resources for advanced services.

The 7450 ESS supports some 7750 SR MDA types only when the 7750 SR IOM3-XP is installed, and the mixed mode is enabled in the SR OS.

7450 ESS Media Dependent Adapters—part 1 of 3

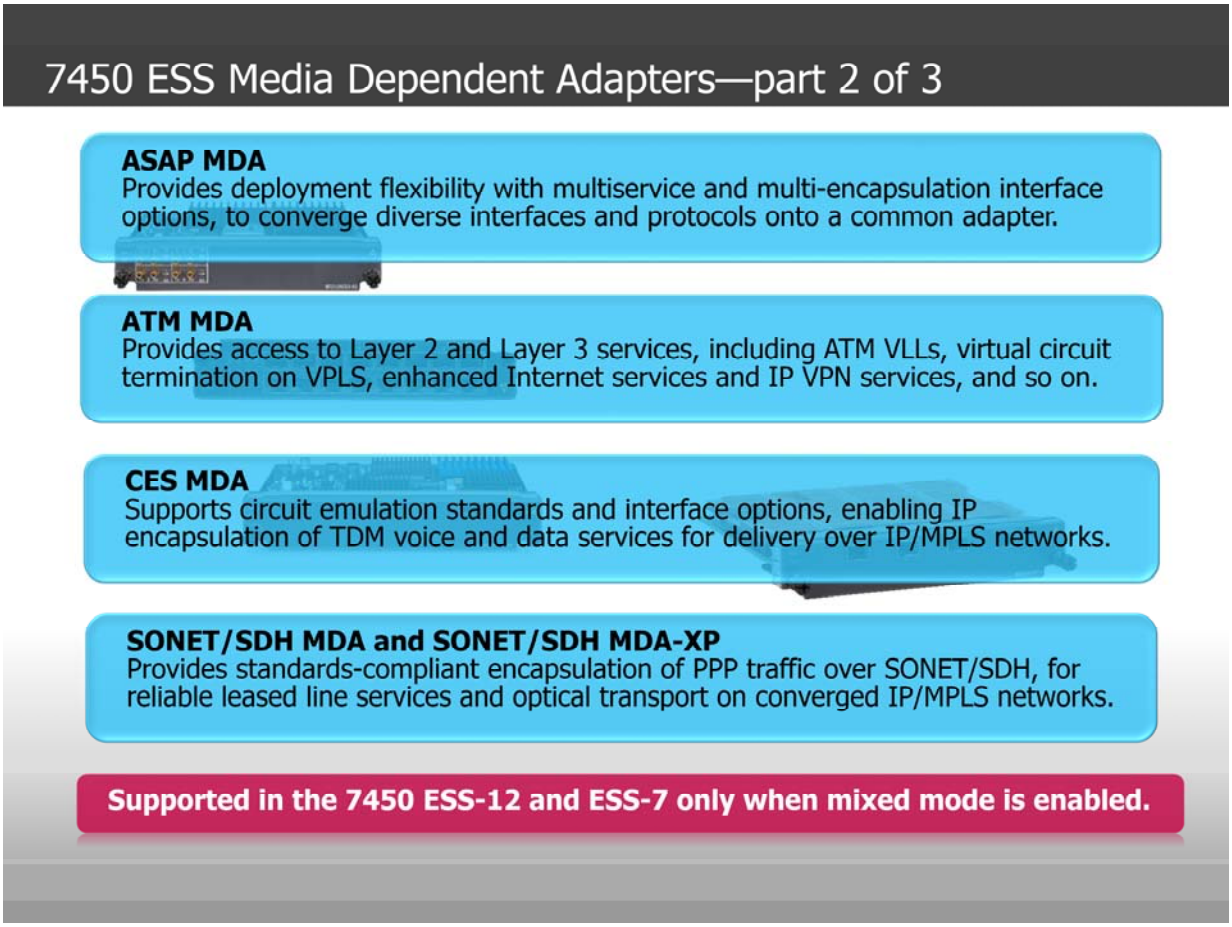


The 7450 ESS supports a wide range of MDA types, interface options and advanced service delivery capabilities.

The Ethernet MDAs support Carrier Ethernet applications and services, including mobile core and backhaul aggregation, Layer 2 and Layer 3 business virtual private networks and broadband residential services and content delivery networks, as well as traditional IPv4 and routing applications.

The Ethernet MDA-XP, Ethernet MDA, Ethernet MDA with tunable DWDM optics and High Scale MDA deliver high-density, high-performance Carrier Ethernet for business, mobile and residential applications.

All Ethernet MDA types support a wide range of SFP and XFP optical modules, and include variants with tunable optics to interface with DWDM transport equipment.



The 7450 ESS-12 and ESS-7 support the ASAP, ATM, CES, and SONET/SDH MDAs only when the 7750 SR IOM3-XP is installed and the mixed mode of operation is enabled. As a result, the 7450 ESS supports Advanced IP Services.

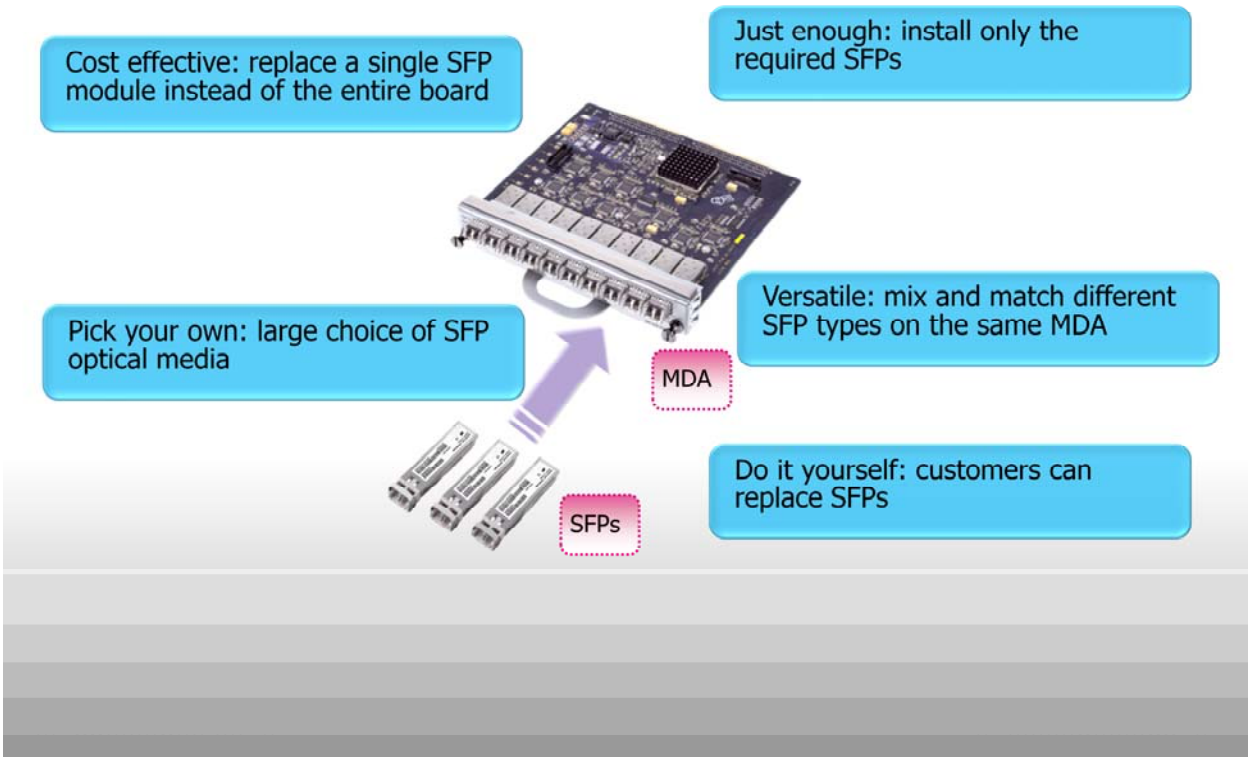
7450 ESS Media Dependent Adapters—part 3 of 3

VSM-XP
Enables the creation of new services by internally interconnecting services. Multiple VSM-XPs can be installed in a chassis and logically grouped, increasing the overall interconnect capacity through load sharing across VSM-XPs in the group.

Each VSM-XP occupies half a slot in an IOM and can interconnect up to 25 Gbit/s (half-duplex) of services when installed in an IOM3-XP/ IOM3-XP-B and 10 Gbit/s when installed in an IOM-20G.

The 7450 SR Versatile Services Module-XP provides high-performance, inter-service connectivity within a single chassis with advanced Quality of Service features. You can use the VSM XP to interconnect pairs of services internally and to create new service offerings. For example, you can interconnect a Layer 2 service and a Layer 3 service to create a "routed" Layer 2 service.

7450 ESS Small Form-Factor Pluggable transceivers

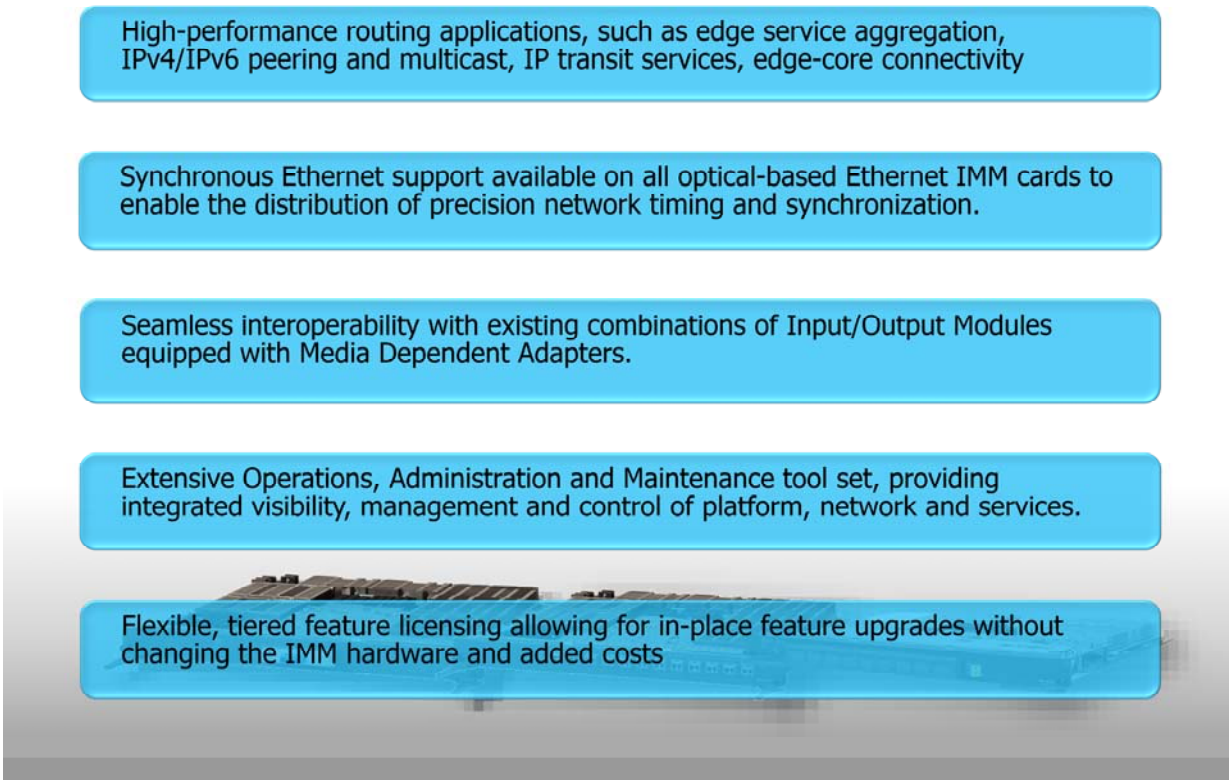


The SFP transceivers are available in a variety of formats. Alcatel-Lucent recommends you use the SFPs that have been tested, verified and validated for the 7750 SR. Alcatel-Lucent programs its SFPs with a type and a part number. As a result, the operator can easily determine the SFP type and replacement part number when troubleshooting.

The use of SFPs allows you to populate the MDAs with the required features on a per-port basis. The combination of MDAs and SFPs means that carriers can defer much of the capital expense of each point-of-presence customer interface until a firm order for service is received.

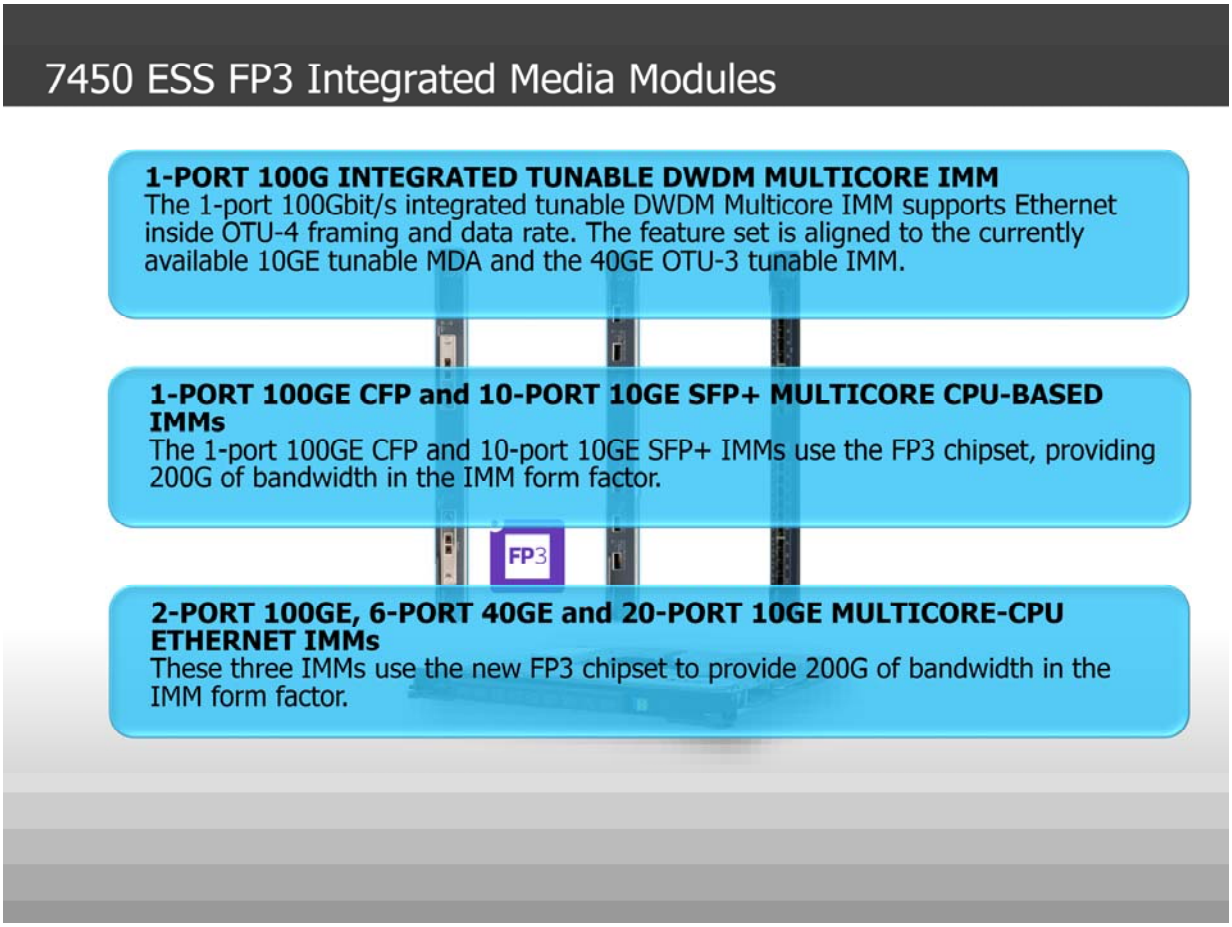
To differentiate them from the Gigabit Ethernet transceivers, the 10-Gigabit Ethernet transceivers are commonly referred to as XFPs.

7450 ESS Integrated Media Modules



The IMM is a full-slot module with an FP2 network processor complex, equipped with integrated physical ports. An IMM combines the functionality of an IOM board equipped with MDAs on a single interface module that you can install in the 7450 ESS. The IMMs provide high-density Gigabit Ethernet and 10 Gigabit Ethernet, high-speed 40-Gb/s and 100-Gb/s interfaces, and high-performance IP/MPLS routing and services. The IMMs support advanced traffic management with Hierarchical Quality of Service, a full range of Layer 2 and Layer 3 routing capabilities, Layer 2 and Layer 3 VPN services and residential services.

The IMM is also available in variants with tunable DWDM optics with Fixed Optic—LC connector interfaces. Refer to the product datasheet for details about the available IMMs for each chassis type.



The most recent SR-OS release introduces support for FP3-based IMMs on 7450 ESS-7 and ESS-12 chassis equipped with the latest SF/CPM.

• The one-port 100 Gigabit per second integrated tunable D-W-D-M multicore IMM extends optical reach in long-haul applications without requiring optical signal amplification or dispersion.

• The one-port 100 Gigabit Ethernet C-F-P and ten-port 10 Gigabit Ethernet SFP+ multicore IMMs provide flexible interface options for MDA-like capabilities.

• Each of the two-port 100 Gigabit Ethernet, six-port 40 Gigabit Ethernet and twenty-port 10 Gigabit Ethernet multicore IMMs support 200 Gbit per second throughput when two SF/CPM-4 cards are installed and operational.

7450 ESS Integrated Service Adapter—description

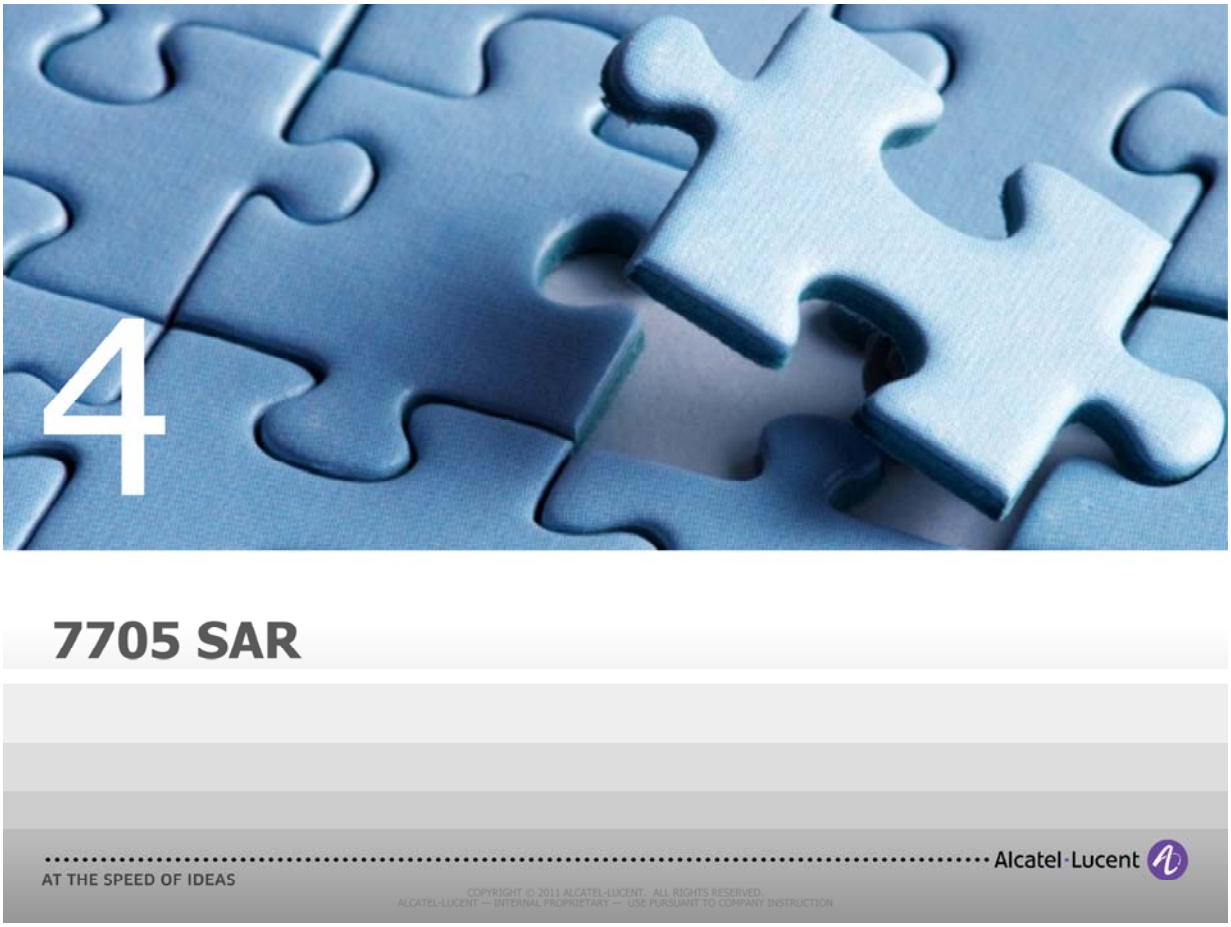
The Multiservice Integrated Service Adapter is a half-slot resource blade that inserts into a 7450 ESS Input/Output Module

A single Multiservice Integrated Service Adapter is equivalent to multiple external servers, offering scalability and resiliency advantages while reducing costs.

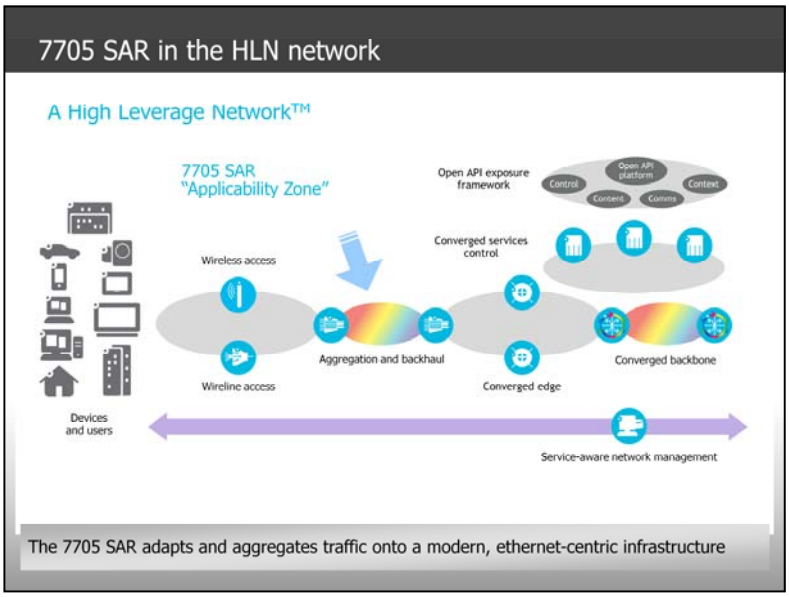
No support for the MS-ISAs when the advanced IP services are enabled.

The Alcatel-Lucent Multiservice Integrated Service Adapter delivers advanced business and residential services on the 7450 SR, eliminating the need for expensive external platforms to support these services. The MS-ISA virtualizes the services and makes them available to all ports across the 7450 ESS chassis. The MS-ISA reduces the need for standalone network elements, the space and cabling requirements, the power consumption, the topology churn and the network latency.

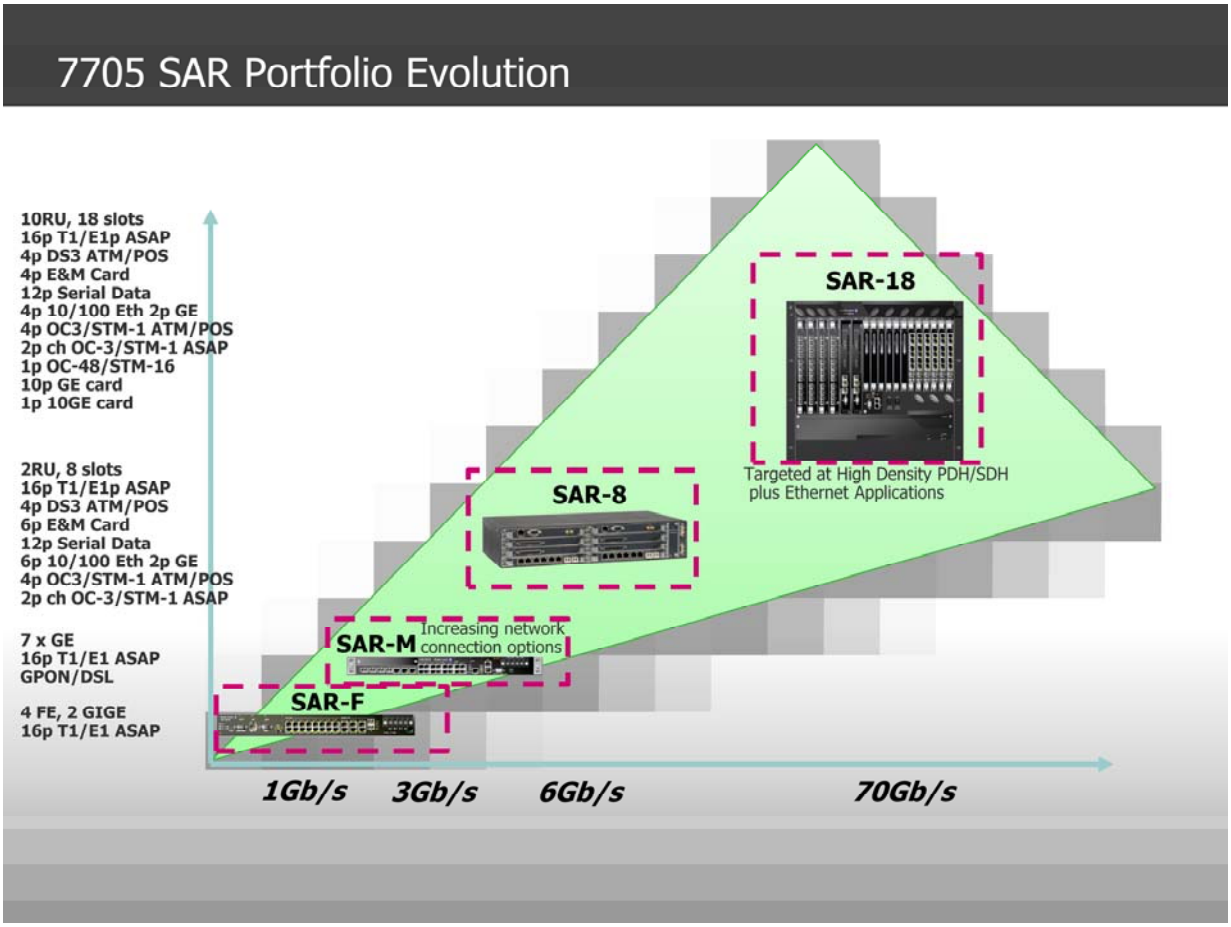
The 7450 ESS supports the MS-ISA Application Assurance and some Video Services when equipped with the IOM-20G or IOM3-XP.



Section 4 introduces the 7705 Service Aggregation Router or SAR.



The Alcatel-Lucent 7705 Service Aggregation Router (SAR) delivers industry-leading IP/MPLS and pseudowire capabilities in compact platforms with the ability to reliably groom and aggregate multiple media, service and transport protocols onto an economical packet transport infrastructure. The Alcatel-Lucent 7705 SAR is extremely well suited to the transport needs of the evolving mobile radio access network or RAN. The 7705 SAR delivers strong convergence capabilities in the mobile RAN with native service processing of 2G, 3G and 4G traffic.



The Alcatel-Lucent 7705 SAR portfolio is optimized for multiservice adaptation, aggregation and routing, especially on a modern Ethernet and IP/MPLS infrastructure. Leveraging the powerful Service Router Operating System (SR OS) and the 5620 Service Aware Manager (SAM), the 7705 SAR is available in compact, low power-consumption platforms that deliver highly available services over resilient and flexible network topologies.

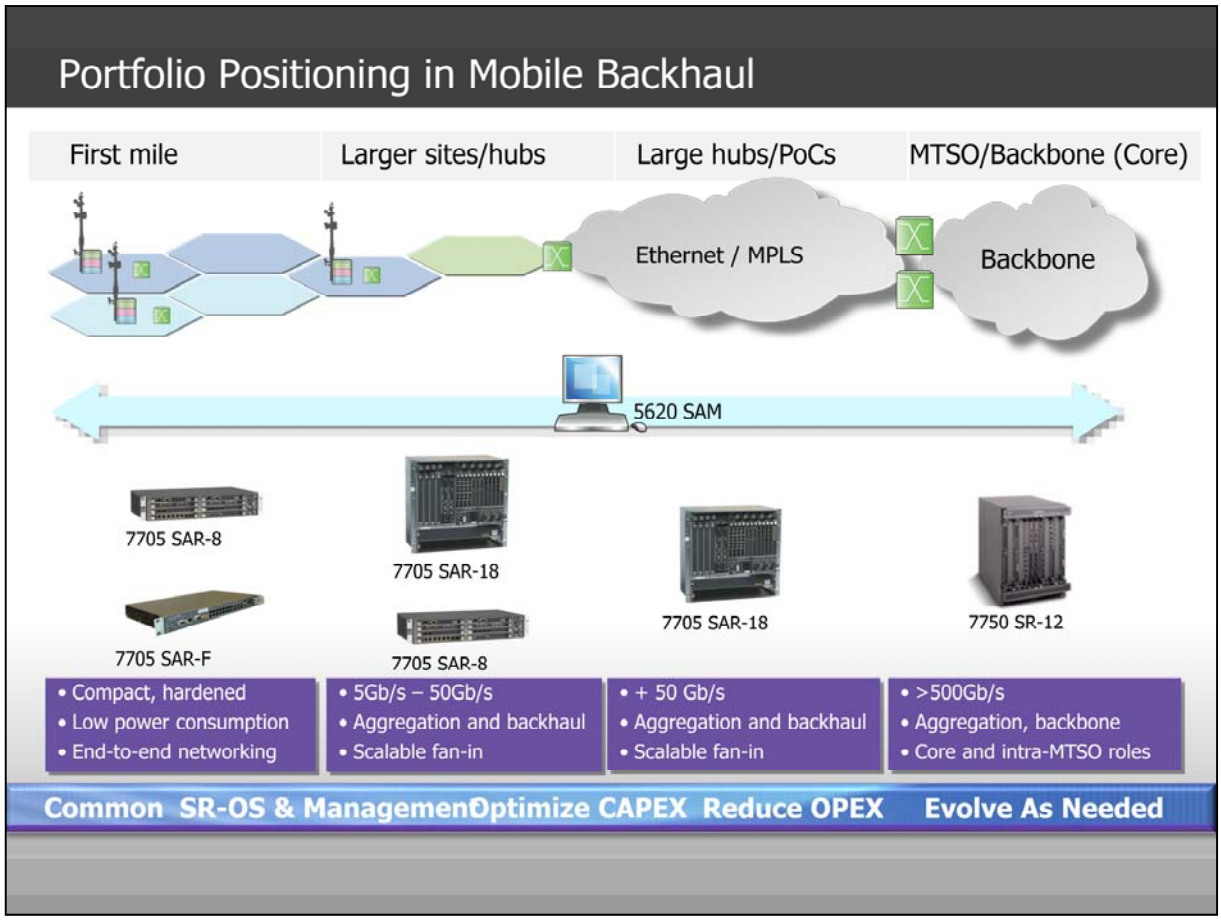
The Alcatel-Lucent 7705 SAR owes much of its development heritage to the Alcatel-Lucent Service Router product line.

The 7705 SAR-F is a fixed configuration version of the Service Aggregation Router packaged in a one-rack unit high form factor.

The 7705 SAR-M is a series of high-performance one-rack unit high form factor Service Aggregation Routers that are orderable in four different configurations.

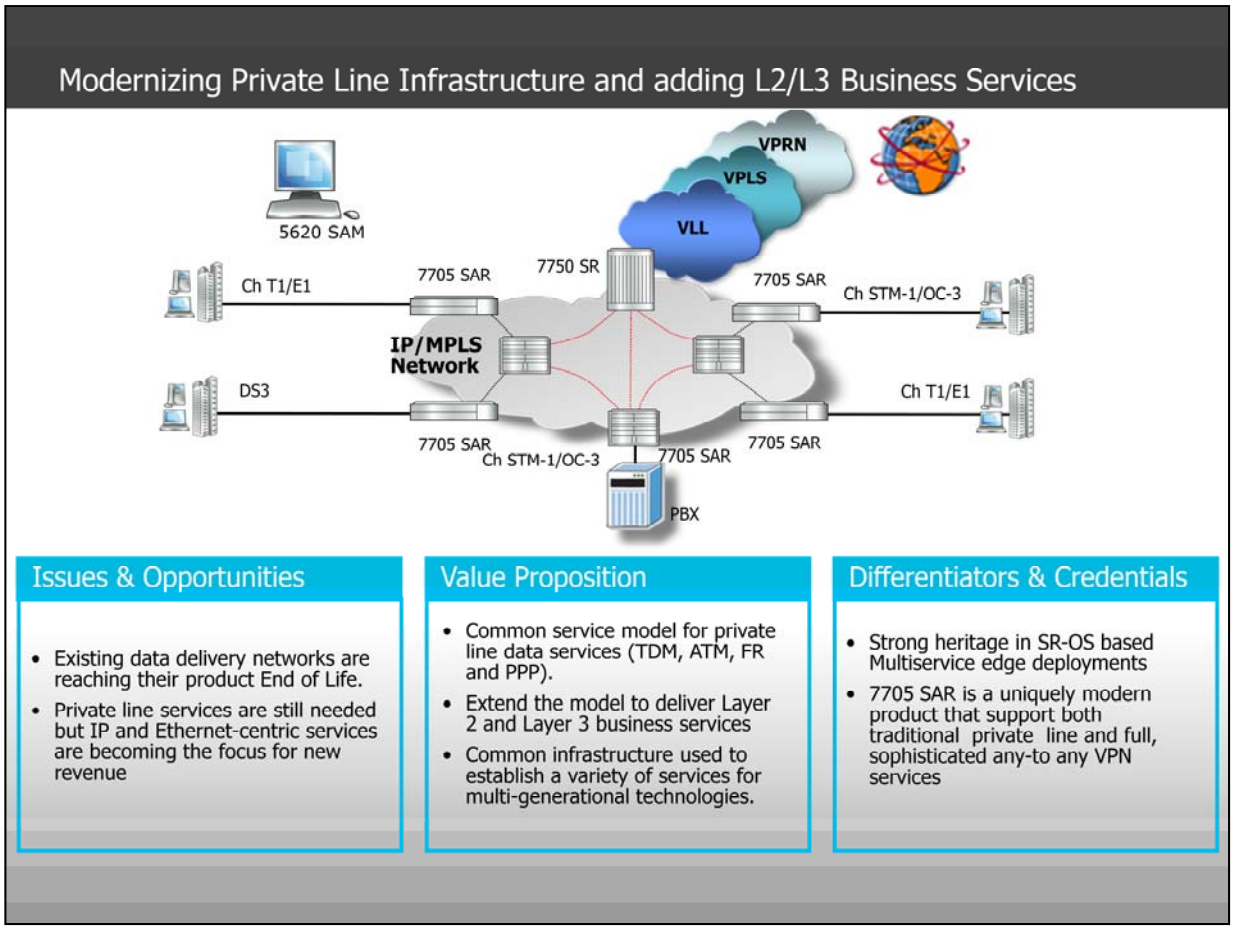
The 7705 SAR-8 is a two-rack unit version of the 7705 SAR with industry-leading access density.

The 7705 SAR-18 is a ten-rack unit version of the 7705 SAR with industry-leading scalability. The platform can be optionally configured with a redundant control and switch module and uplinks. The Alcatel-Lucent 7705 SAR-18 has 18 slots; two slots are allocated for control and switch modules (CSMs), with the remaining 16 slots being available for user traffic adapter cards.



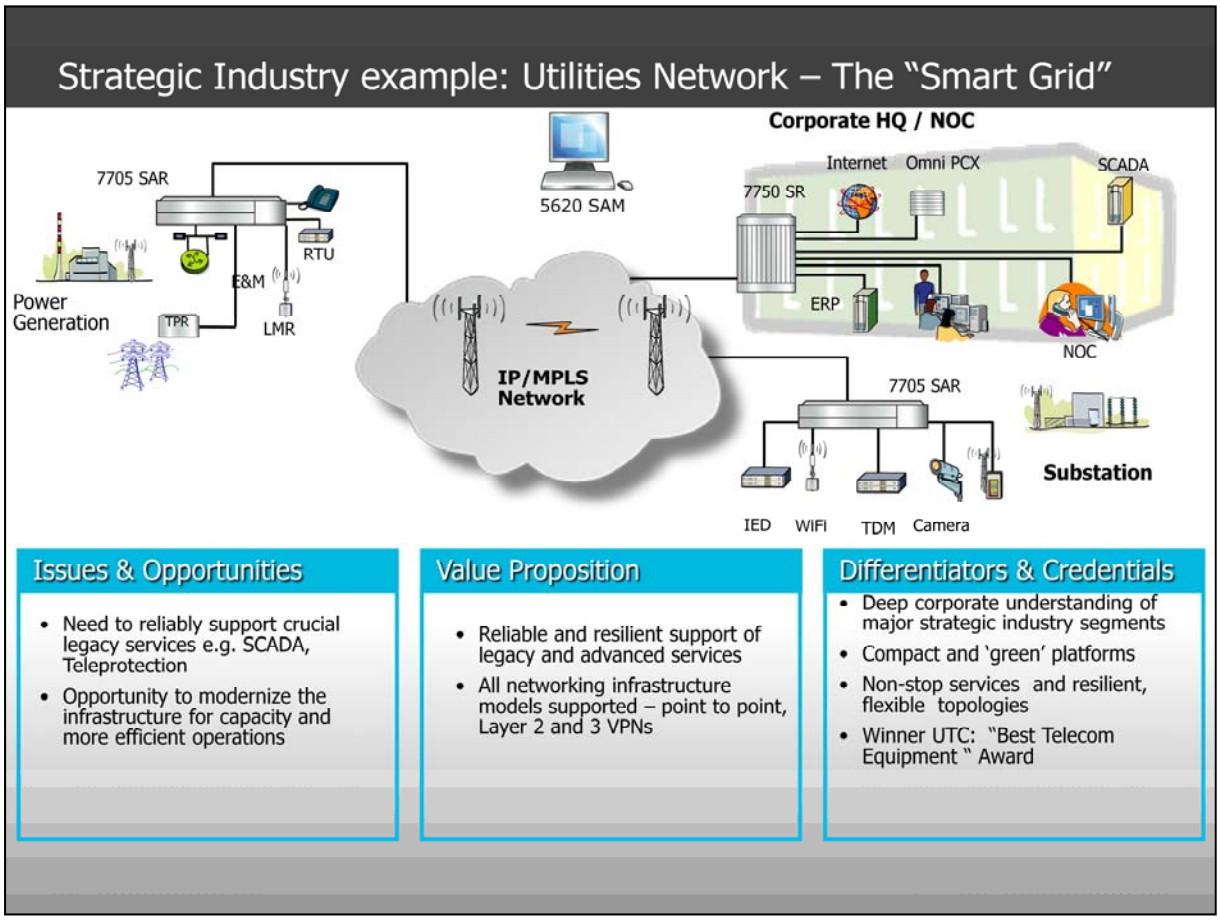
As mobile service providers prepare for the evolution to mobile broadband, they require a robust transport infrastructure that supports CDMA/EV-DO and GSM/UMTS/HSPA today and is well suited to support WiMAX, LTE and fixed-mobile convergence (FMC). Internet Protocol/Multi-Protocol Label Switching (IP/MPLS) has grown to become a foundation for many fixed, mobile and converged networks. In mobile networks, IP/MPLS consolidates disparate transport networks for different radio technologies, reduces operating expenditures (OPEX) and converges networks on a resilient and reliable infrastructure, while being ready to support further evolution to Fourth-Generation Mobile Network (4G) technologies.

The 7705 SAR product family enables an extensive set of broadband backhaul capabilities. Broadband backhaul is based on a series of supported interface variants like ADSL2, ADSL2+, SHDSL, VDSL2 or GPON uplinks. As such, the 7705 SAR product family specifically has the ability to migrate from copper access to fiber access uplinks within a single platform where that path may go from T1/E1 to Ethernet point-to-point fiber or from DSL to GPON, or any combination to fit an operator's needs.



Existing data networks are reaching their product end of life. While private line services like TDM, ATM, Frame Relay and PPP are still needed, new revenue are coming from IP and Ethernet services. This requires a common infrastructure and modern products that both support traditional private lines and the more sophisticated VPN services.

Business services modernization is supported with the 7705 SAR product family in the transition from legacy to consolidated, packet-based operation. Hugely reduced equipment footprints are achievable with reduced energy costs. Industries, enterprises and government organizations can deploy with confidence, achieving reliable and resilient support of legacy and advanced services.



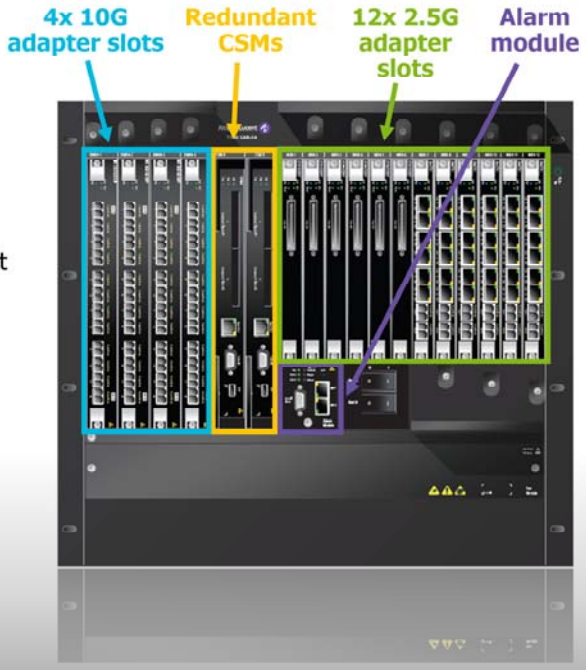
Communications network transformation to a highly available IP/MPLS infrastructure for mission-critical communications is a foundation to the Smart Grid architecture.

The implementation of Smart Grids is part of the new energy delivery strategy of many power utilities around the world. Smart Grid applications provide utilities with better automation and the benefits of reduced operating costs, increased power quality, and improved outage response.

The 7705 supports the critical legacy services like SCADA (Supervisory Control and Data Acquisition) and teleprotection, while at the same time transporting these legacy and advanced applications over a reliable and resilient IP/MPLS network.

7705 SAR-18

- 70G fully redundant solution
  - Redundant control and switching
  - Redundant synchronization input ports
  - Redundant power terminals
- Feature & functionality parity with 7705 SAR-8
  - Supports all SAR-8 line cards
- Self-regulating fan controller (8 per system)
  - Fan speed based on temperature sensing
- Support for PDH/SONET/SDH, ATM and Ethernet
  - +50 STM-1/OC-3 ports
  - 192 T1/E1 ports
  - 4 x 10G slots for 10G/GigE/FE and OC-48/STM-16 line cards
  - 12x 2.5G slots for existing and future SAR-8 adapter cards
- 10RU height, 4 to a standard rack
- 300-mm depth



The Alcatel-Lucent 7705 SAR-18 has 18 slots; two slots are allocated for control and switch modules (CSMs), with the remaining 16 slots being available for user traffic adapter cards. Twelve of the adapter card slots have full duplex 2.5 Gb/s connectivity to the switching fabric, while the remaining four slots have full duplex 10 Gb/s connectivity. The platform can optionally support 1+1 fully redundant CSMs for High Availability. The twelve 2.5 Gb/s adapter card slots support the same adapter cards as the 7705 SAR-8. Network connectivity options are: Ethernet, FE, GigE, n × T1/ E1 MLPPP or n × T1/E1 ATM IMA. Integrated DS3 point-to-point trunking is supported using the 4-port DS3 adapter card. OC-3/STM-1 trunking is supported using POS on the 4-port OC-3/STM-1 clear channel adapter card.

7705 SAR-8 : High Function, Compact Form Factor

Version 2 now available which can operate at 10Gb/s Full Duplex (top 2 shelves)

Compact, Modular Chassis for remote sites and hubs
- 2RU high, 19" width and 10" (300mm) depth
- 2 control module slots, 6 half-wide I/O slots
- Two feeds: -48/-60V DC or Two feeds: +24V DC•
- -40C to +65C normal operating temperature range

Flexible access options and services
- Up to 96 ports of T1/E1 , with Any Service, Any Port (ASAP)
- 10/100/1000 Copper and Fiber Ethernet Ports
- DS3 clear channel, ATM and PPP
- OC-3/STM-1 clear channel and channelized
- Serial Data Interface card (V.35 and RS-232)
- E&M Voice adapter card
- ATM, TDM, Ethernet and IP Pseudowires
- IP Routing and IP VPN support

High service availability
- Redundant control/switch modules
- Redundant power feeds & synch
- Uplink and tunnel redundancy
- Redundant pseudowires and FRR

Flexible network link options
- FE/GE
- NxT1/E1 MLPPP
- DS3 pt-to-pt trunking
- OC-3/STM-1

Flexible range of sync options
- Line, BITS(GPS), ACR, Synch Ethernet, 1588v2
- SSM for quality & resiliency

The Alcatel-Lucent 7705 SAR-8 has eight slots; two slots are allocated for control and switch modules (CSMs), with the remaining six slots being available for user traffic adapter cards. The Alcatel-Lucent 7705 SAR-8 has a compact, modular architecture, constructed to allow flexible use of line adapter cards so operators can optimize the configuration to meet the specific requirements of a site. With the modular architecture comes additional resilience and flexibility. The platform can optionally support 1+1 fully redundant CSMs. This industry-leading, independently validated High Availability feature has been inherited from the Service Router product line and is a strong contributor to overall network uptime. Network uplink connectivity options are: Ethernet, FE, GigE, n × T1/ E1 MLPPP or n × T1/E1 ATM IMA. Integrated DS3 point-to-point trunking is supported using the 4-port DS3 adapter card. OC-3/STM-1 trunking is supported using Packet over SONET/ SDH (POS) on the 4-port OC-3/STM-1 clear channel adapter card.

Note that the new SAR-8 Shelf version 2 is capable of supporting an increased bandwidth per adapter card slot compared to the pre Release 5.0 SAR-8 Shelf. The version 2 chassis has a new backplane with hardware capable of supporting up to 10 Gbps.

7705 SAR-F: Product Attributes in Brief

Dense, rugged 1 RU high form factor

- 16, T1/E1 ASAP ports
- 6, 10/100 Ethernet ports
- 2, 10/100/1000 Ethernet ports

Temp. range: -40°C to +65°C (-40°F to 149°F)

Dual Power Feeds: (-48V or +24 VDC)

Rich services set

- TDM (T1/E1) pseudowires)
- ATM and ATM IMA (n x T1/E1) pseudowires
- Ethernet (10/100/1000) pseudowires
- IP pseudowires
- IP Routing VPLS and IP VPN (VPRN)

Flexible, cost efficient network uplink options

- Ethernet; 10/100/1000
- NxT1/E1 MLPPP
- ATM/IMA
- DS3 pt-to-pt trunking via SFP plug-in
- GRE/DSL for HSPA offload in remote sites

Unequalled Synchronization Solution

- Broad range of techniques:
  - Line, BITS(GPS), ACR, Synch Ethernet, 1588v2
- Independent verification of performance
- Resiliency, on board ovenized Stratum 3
- SSM for quality & resiliency

The one-rack unit 7705 SAR-F supports up to 16 T1/E1 any-service-any-port (ASAP) ports. The ASAP ports can be configured to support ATM, ATM IMA, TDM and MLPPP. Six 10/100 Base-T autosensing Ethernet ports are provided, plus two extra ports supporting 10/100/1000 Base-TX with small form factor pluggable optics (SFPs). Network uplink connectivity options are: Ethernet, Fast Ethernet (FE), Gigabit Ethernet (GigE), n × T1/ E1 MLPPP or n × T1/E1 ATM IMA. Integrated DS3 point-to-point trunking is supported using a SFP device.

The SAR-M is similar as the SAR-F but with the support of an expansion module slot. This expansion slot supports a GPON ONT, xDSL Module or a CWDM Optical Add-Drop Mux (OADM) module.

ALCATEL-LUCENT 7705 SAR PORTFOLIO EXPANSION
THE 7705 SAR-M

- 2.5 Gb/s Full Duplex, 1 RU IP/MPLS aggregation router
- -48 V DC and +24 V DC support in a single variant
- Main features and functionalities
  - Routing: IPv6/4, OSPFv3, OSPF-TE , ISIS-TE , static routes, ….
  - IP/MPLS /transport: RSVP-TE , LDP, GRE/IP-based tunnels
  - Rich OAM feature set: Y.1731, 802.3ah, 802.1ag, VCCV, IP tools, SAA, etc.
  - Variety of synchronization options (synchronous Ethernet, 1588v2, ACR, ToD, Line, DCR, etc.)
  - Extensive traffic management capabilities

AT THE SPEED OF IDEAS

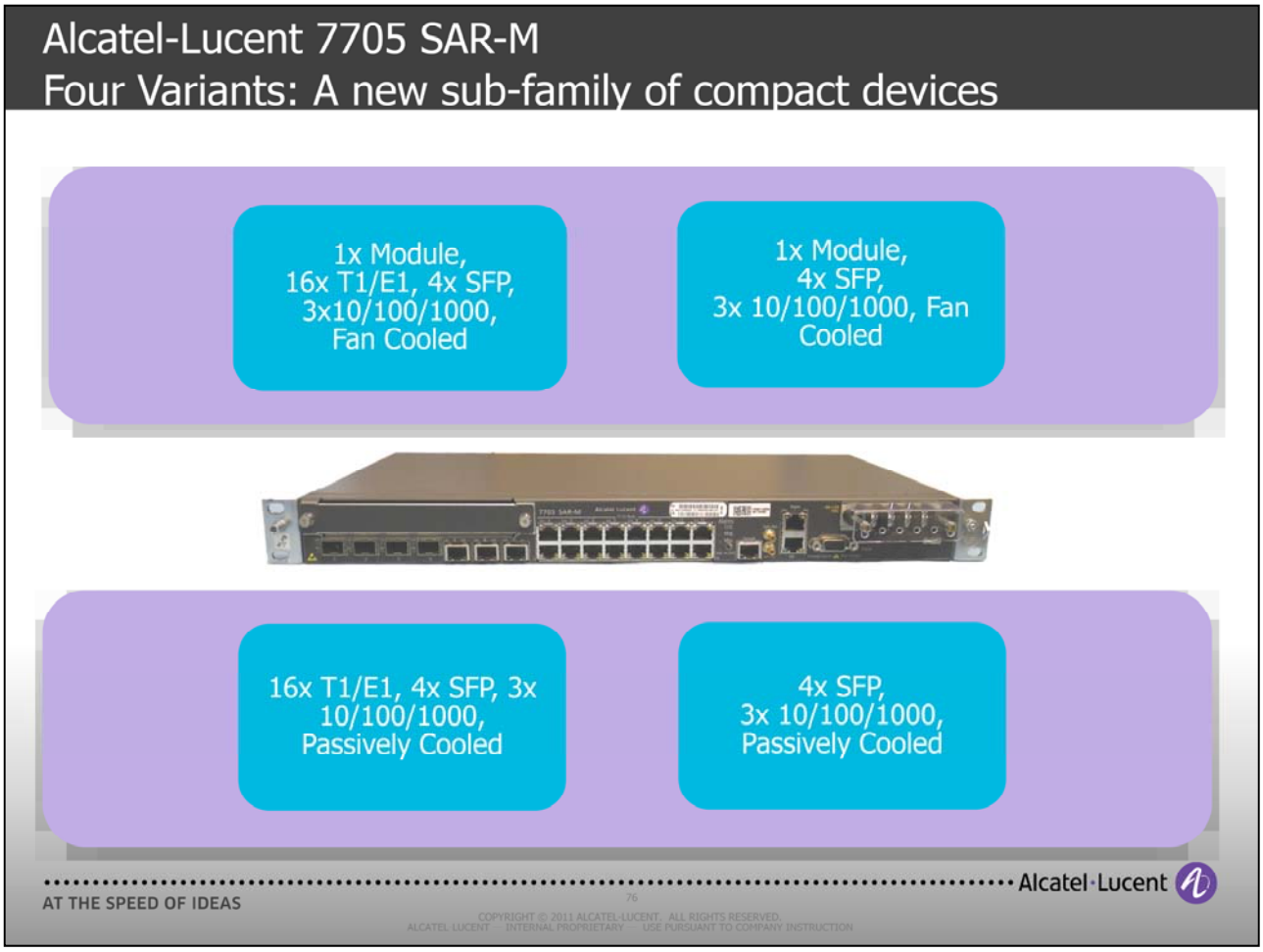
Alcatel·Lucent

75

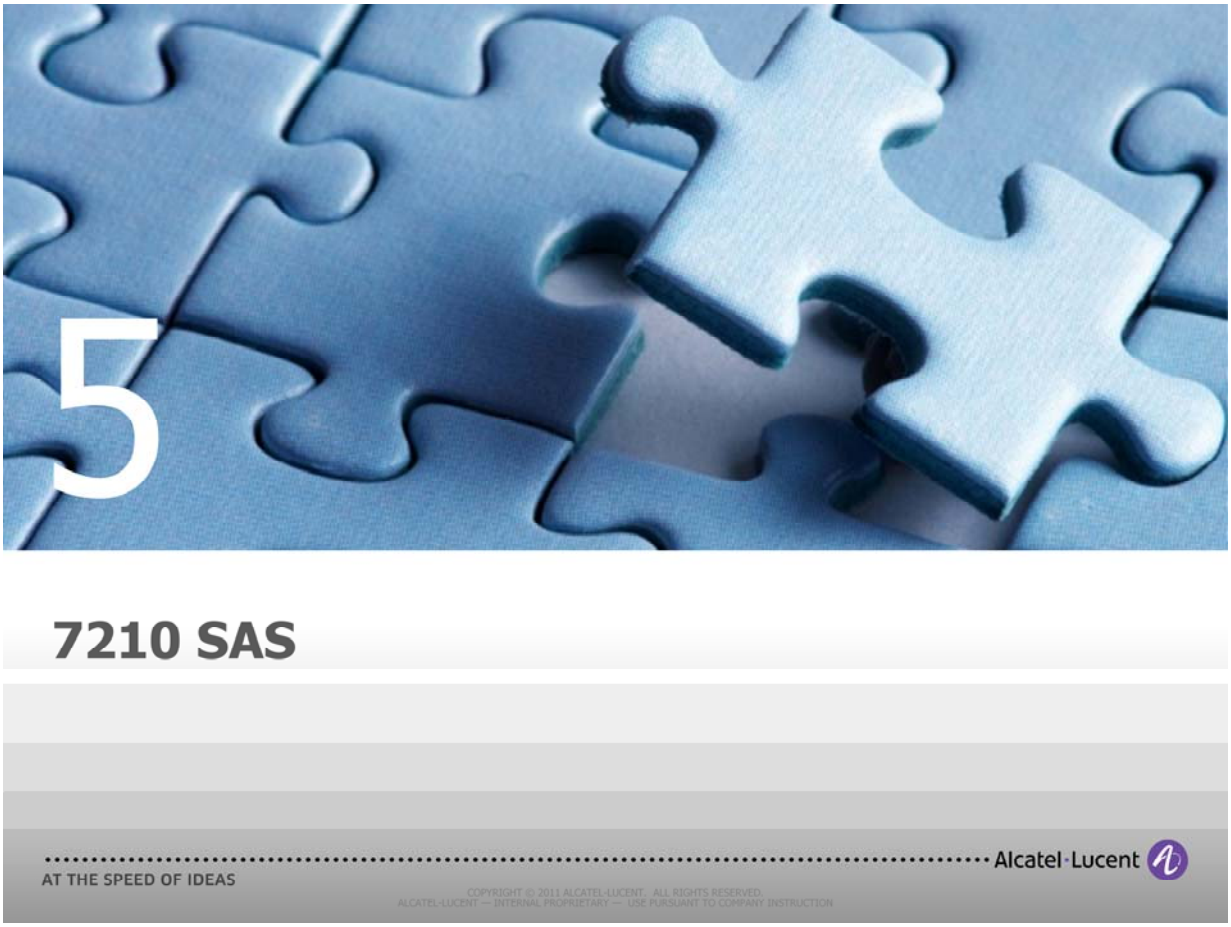
The "M" in the 7705 SAR-M stands for modular. It is very similar to the SAR-F but has an expansion module slot.

The F variant is a 2.5 Gbps full duplex IP/MPLS aggregation router working on -48VDC or +24VDC.

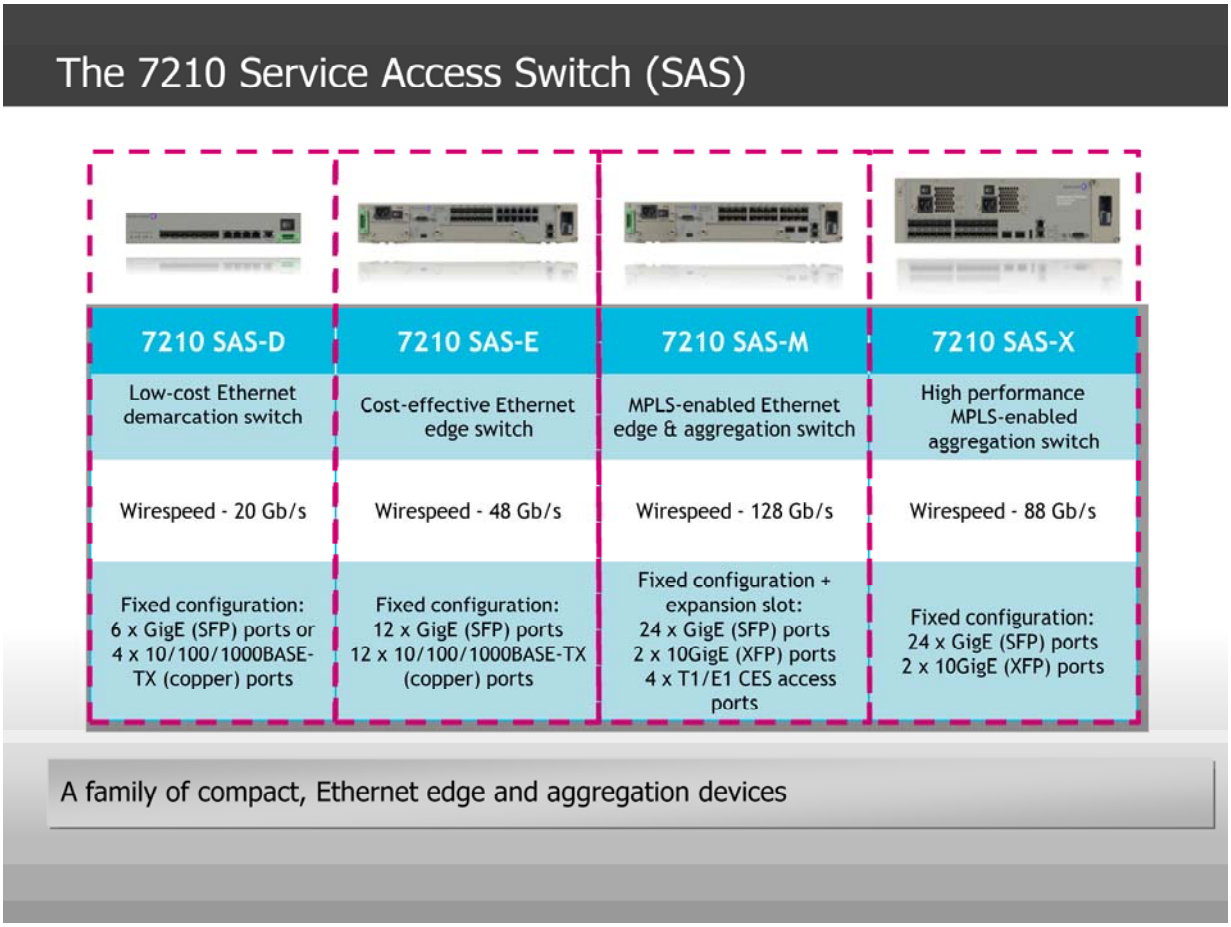
As indicated on the slide, the SAR-M supports all the major routing, MPLS, OAM, synchronization and extensive traffic management capabilities.



The expansion module slot allows for the installation of four different module variants, allowing for a wide variation of interfaces.



Section 5, the 7210 Service Access Switch



Designed for the customer edge, smaller central offices and distributed hub sites, the Alcatel-Lucent 7210 Service Access Switch portfolio of compact Ethernet-edge, demarcation and aggregation devices provides cost-effective Carrier Ethernet service delivery to the customer edge and extends the reach of MPLS-enabled Carrier Ethernet aggregation networks into smaller network locations.

The SAS is available in different compact form factor options.

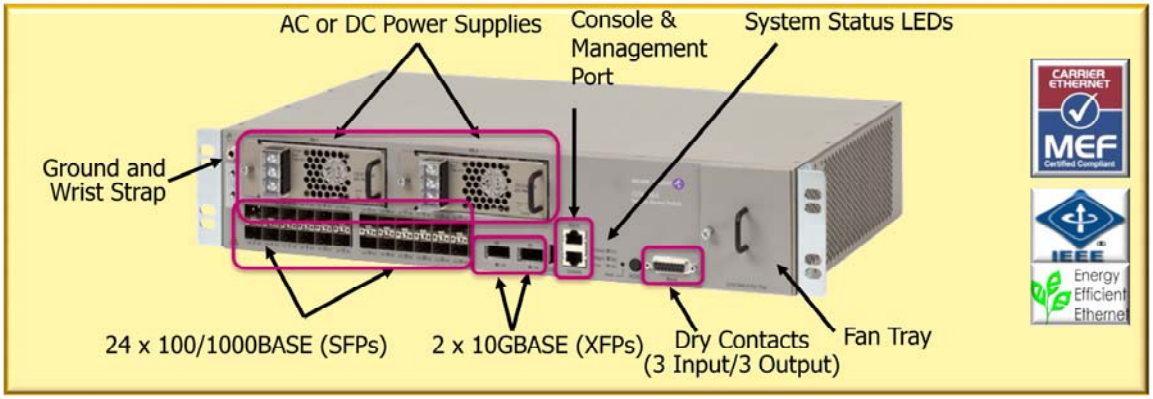
The 7210 SAS-D is an intelligent Ethernet-edge demarcation device that extends retail and wholesale SLA-based VPLS and VLL services to the customer edge. For service demarcation applications in 3G and LTE wholesale mobile backhaul service delivery networks, it is also available in a variant that supports ITU-T Synchronous Ethernet (Sync-E) and extended temperature ranges (ETR).

The 7210 SAS-E is an Ethernet edge access device with a fixed configuration twelve 10/100/1000BASE-T copper ports and twelve 100/1000BASE fiber ports, and is designed for VPLS and VLL services delivery in multi-tenant offices.

The 7210 SAS-M is an MPLS enabled Ethernet-edge and aggregation device. It supports IP, MPLS, Ethernet, PBB, 10GigE ports and is available in an ETR variant. As a customer edge device, it provides a common platform to enable VPLS, VLL and IP VPN services. It comes with a fixed amount of interfaces and one expansion slot for additional flexibility.

The 7210 SAS-X is a high performance MPLS-enabled aggregation device. Designed for more service intensive smaller network location applications, its powerful CPU is suitable where higher scale switching and routing capabilities are essential. In addition to supporting all the IP, MPLS and Ethernet capabilities of the Alcatel-Lucent 7210 SAS-M, the 7210 SAS-X supports a larger number of service queues, packet buffers, label switched paths (LSPs), counters, MAC addresses, routing table entries and ACLs for increased service scale.

7210 SAS-X: Hardware Overview



High performance MPLS–based device provides aggregation for more service-intensive smaller network locations

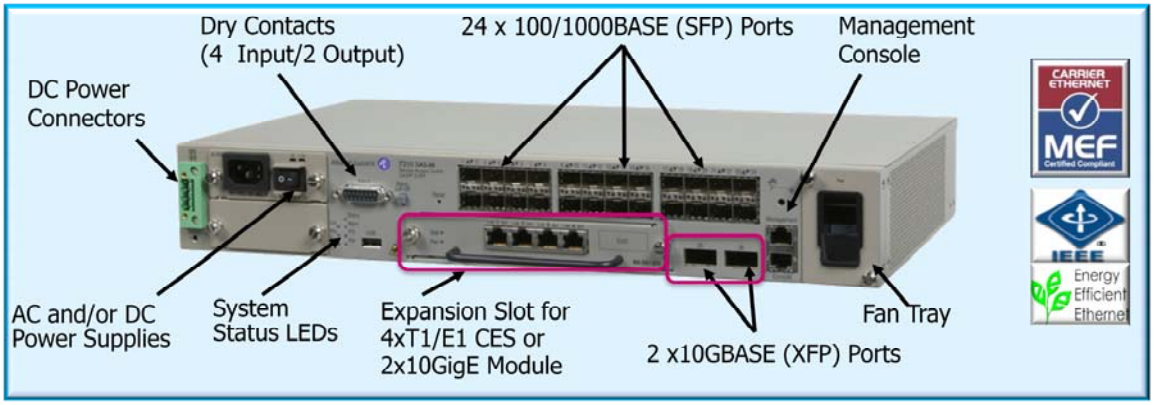
- Aggregates VPLS, VLL and IP VPN services over 10 GigE uplinks with burst capabilities
- Wirespeed (88 Gb/s half duplex), SR OS-based device, 2.0 RU, NEBS3 compliant
- Synchronization: ITU-T Sync-E, IEEE 1588v2 (future), stratum 3E (OXCO) clock

The SAS-X is the most feature rich product of the 7210 SAS product range and is MEF 9 and MEF 14 certified, the defining certification body for Ethernet. All interfaces are integrated. 24 SFPs can be installed to connect 100 or 1000 Base Ethernet. 2 available XFP based slots are available to connect 10 Gig Base Ethernet. AC or DC redundant power supplies are of modular design. Next to 3 input and 3 output connections for dry voltage contacts, the SAS-X is equipped with a console and management port for outbound management. The SAS aggregates VPLS, VLL and IP VPN services over 10 Gig uplinks with burst capabilities.

Burst capabilities are made possible with advanced H-QoS with ingress and egress queuing and shaping. The 7210 SAS-X also supports dynamic buffer allocation, allowing buffers to be allocated on a per-service basis for added SLA flexibility.

The SAS-X supports 888 Gigabits per second in half duplex and synchronization features that comply to the ITU-T Synchronous Ethernet and IEEE 1588v2 standards.

7210 SAS-M (10GigE w/ETR option): Hardware



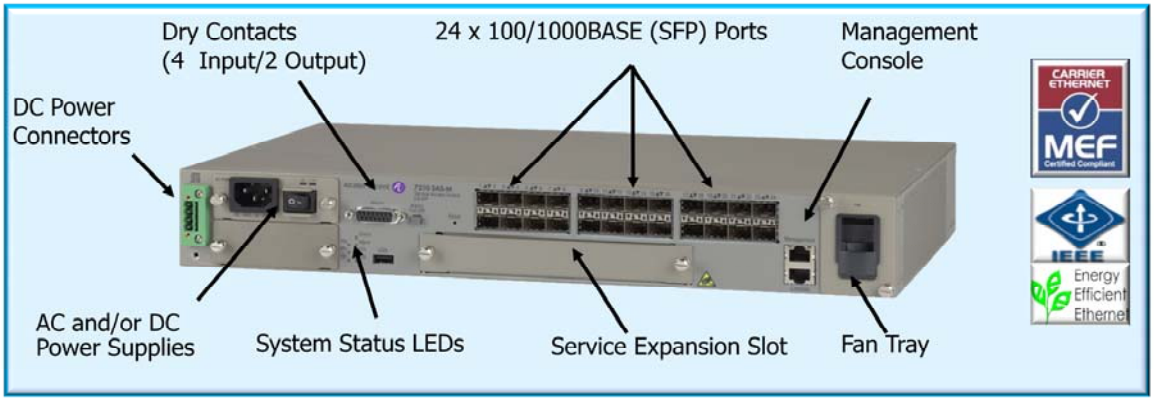
MPLS-based device for Ethernet-edge CPE and service aggregation

- CPE/aggregation device supports VPLS, VLL and IP VPN services over GigE/10GigE uplinks
- Wirespeed (128 Gb/s half duplex), SR OS-based MPLS device, 1.5 RU, NEBS3 compliant
- ETR variant supports Extended Temperature Ranges (ETR) - 40°C to +65°C
- Service expansion slot options: 4-port T1/E1 CES and 2 x 10GigE (XFP) modules
- Synchronization: ITU-T Sync-E, IEEE 1588v2 (future), stratum 3 (TXCO) clock

The SAS-M as a customer edge device provides a common platform to enable VPLS, VLL and IP VPN services. It also supports the configuration of VPLS and VLL services using PBB. As an aggregation device, it extends the reach of the MPLS-enabled Carrier Ethernet network into smaller offices and distributed hub sites and supports provider edge (PE) functionality to aggregate customer edge (CE) routers in IP VPN service delivery. The SAS-M is an SR-OS based MPLS device working at a wirespeed of 128 Gigabits per second. The extended temperature range variant supports temperatures of -40 to + 65 degrees Celsius.

For added flexibility, the 7210 SAS-M offers optional service expansion modules to support circuit emulation services (CES) over T1/E1 interfaces and two 10GigE ports for higher speed uplinks, larger 10GigE mesh configurations and support for 10GigE CPE. Both standards of synchronization, ITU-T Sync-E and IEEE 1588v2 are supported.

7210 SAS-M: Hardware Overview

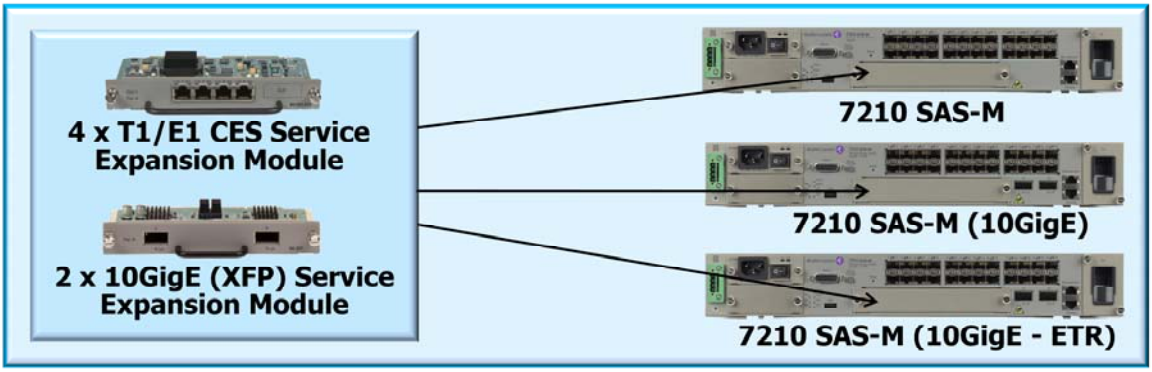


MPLS-based device for Ethernet-edge CPE and service aggregation

- CPE/aggregation device supports VPLS, VLL and IP VPN services over GigE/10GigE uplinks
- Wirespeed (88 Gb/s half duplex), SR OS-based MPLS device, 1.5 RU, NEBS3 compliant
- Service expansion slot options: 4-port T1/E1 CES and 2 x 10GigE (XFP) modules
- Synchronization: ITU-T Sync-E, IEEE 1588v2 (future), stratum 3 (TXCO) clock

This is the standard SAS-M, without any integrated 10 Gig interfaces supported and operating in temperatures from 0 to 50 degrees Celsius.

7210 SAS-M (All Variants): Service Expansion Modules



- **4 x T1/E1 CES module enables new business VPN revenue streams, transitions business services off SONET/SDH metro networks**
  - Allows an enterprise to run several services concurrently such Ethernet VPNs along with fractional nxT1/E1 services or clear T1/E1 services over a single physical connection
- **Scale service delivery to meet emerging 10GigE requirements**
  - Deploy 10GigE module to scale uplink toward 10GigE as required; add additional 10 GigE ports; use as access (customer) interface

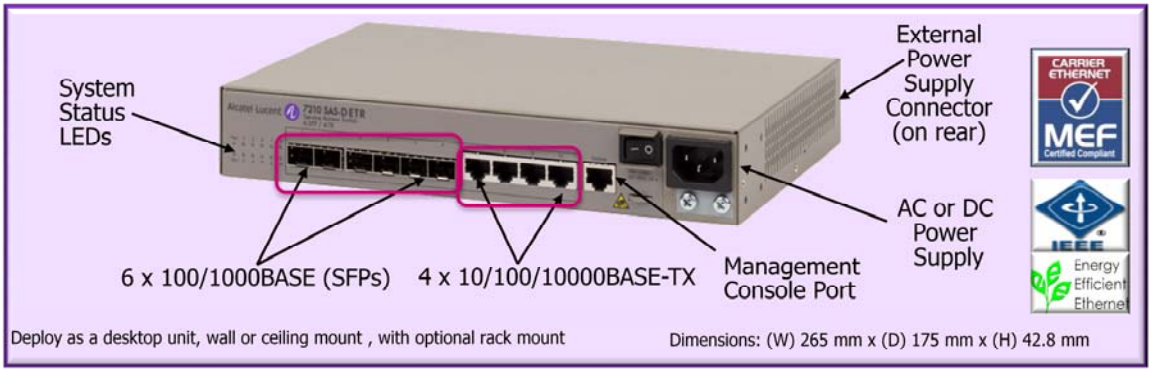
This slide shows all the options of the expansion modules with the available SAS-M variants.

The 4 times T1/E1 CES module allows an enterprise to run several services concurrently such as Ethernet along with fractional n times E1/T1 over a single physical connection.

The 2 times 10 Gig expansion module offers additional 10 Gig ports to be used as access or customer interface.

Any of the two expansion modules can be plugged into the expansion slot of any of the three SAS-M variants.

7210 SAS-D (with ETR option): Hardware Overview

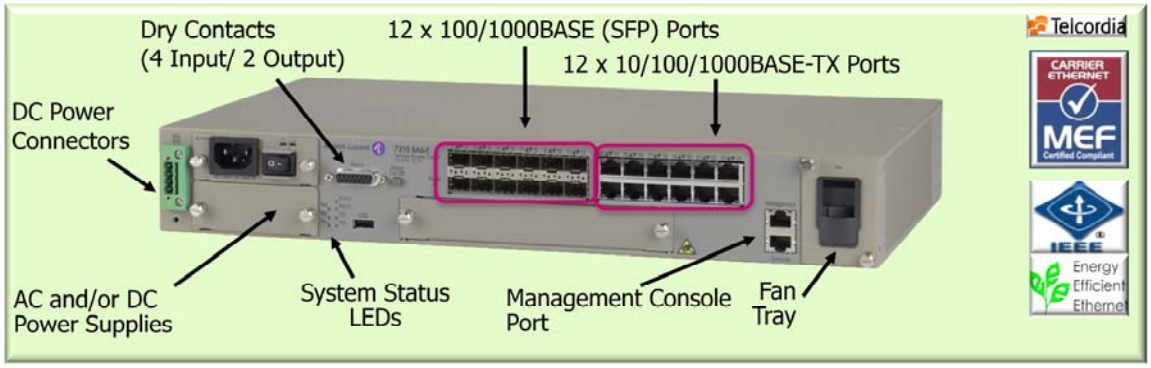


Intelligent Ethernet-edge demarcation device

- Wirespeed (20 Gb/s), non-blocking, SR OS-based device, 1 RU, NEBS level 3 compliant
- ETR variant allows for ITU-T Sync-E, IEEE1588v2 (future), 24 VDC or -48VDC power and redundancy option with external power supply (Note: any AC or DC combination between internal and external power supplies is supported)
- AC or DC power supply with fan-less operation

The SAS-D is a low cost 20 Gigabits per second ethernet switch with a fixed amount of interfaces. It supports 6 GigE ports and 4 10/100 or 1000 BaseT copper ports. The D variant of the SAS family works at speed of 20 Gigabits per second half duplex. This picture shows the variant with ETR. There is no MPLS implemented, so the uplink is solely ethernet.

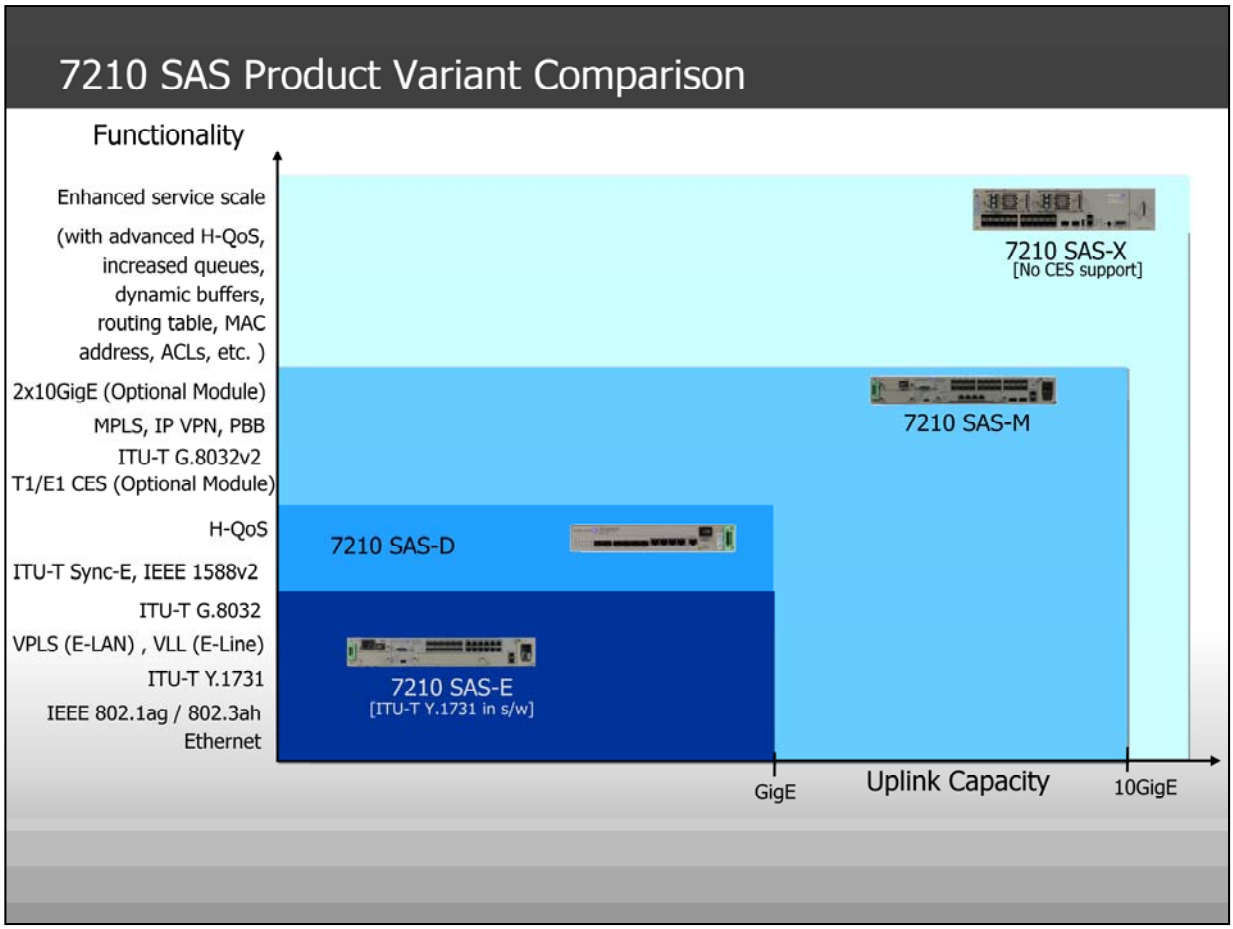
7210 SAS-E: Hardware Overview



Cost-effective Ethernet-edge device

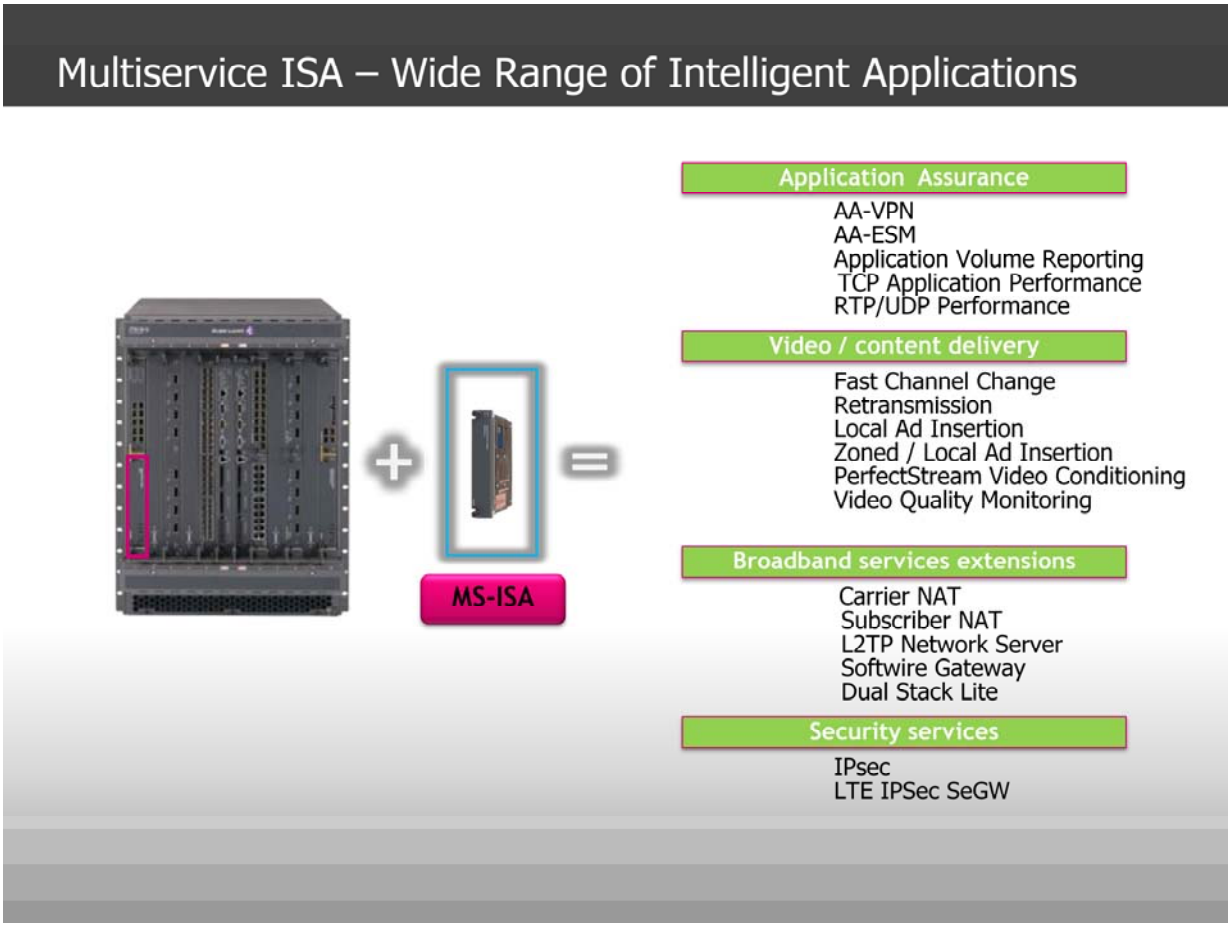
- Wirespeed (48 Gb/s HD), SR OS-based Ethernet switch (1.5 RU)
- Hot swappable AC and/or DC redundant power supplies and fan tray

The SAS-D is very similar to the SAS-E, except that it works at a wirespeed of 48 Gigabits per second in half-duplex and supports up to twelve 100/1000 Base SFP ports and twelve 10/100/1000 BaseT ports.



With platform capabilities that include IP, Multiprotocol Label Switching (MPLS), hierarchical quality of service (H-QoS), advanced operation, administration and maintenance (OAM) tools, leading resiliency and synchronization capabilities, and 10GigE interface support, the Alcatel Lucent 7210 SAS is an integral component, delivering value-added MPLS-enabled Carrier Ethernet services including Virtual Private Line Service (VPLS) and Virtual Leased Line (VLL) services. With Release 4.0, the required functionality and uplink capacity, the 7210 SAS product family offers the perfect solution for the customer edge, smaller central offices and distributed hub sites.

Section 6: MS-ISA, Velocix, OSS/BSS systems



The Alcatel-Lucent Multiservice Integrated Service Adapter (MS-ISA) extends the level of intelligence of the industry-leading Alcatel-Lucent 7750 Service Router (SR) and 7450 Ethernet Service Switch (ESS) platforms by virtualizing advanced service capabilities into a unified service edge. The Alcatel-Lucent MS-ISA provides purpose-built, extended functionality and enables deeper levels of integrated service capabilities with higher scale than typically available from costly dedicated appliances.

The Alcatel-Lucent MS-ISA is a half-slot, hot-swappable resource blade that inserts into an Input/Output Module (IOM). It features a flexible multi-core network processor designed for high-touch packet operations.

Services currently supported on the Alcatel-Lucent MS-ISA include Application Assurance, L2TP Network Server (LNS), Network Address Translation (NAT), Dual-Stack Lite AFTR services, WLAN Gateway, IPsec services, MS-ISA Threat Management System (TMS) and advanced video services.

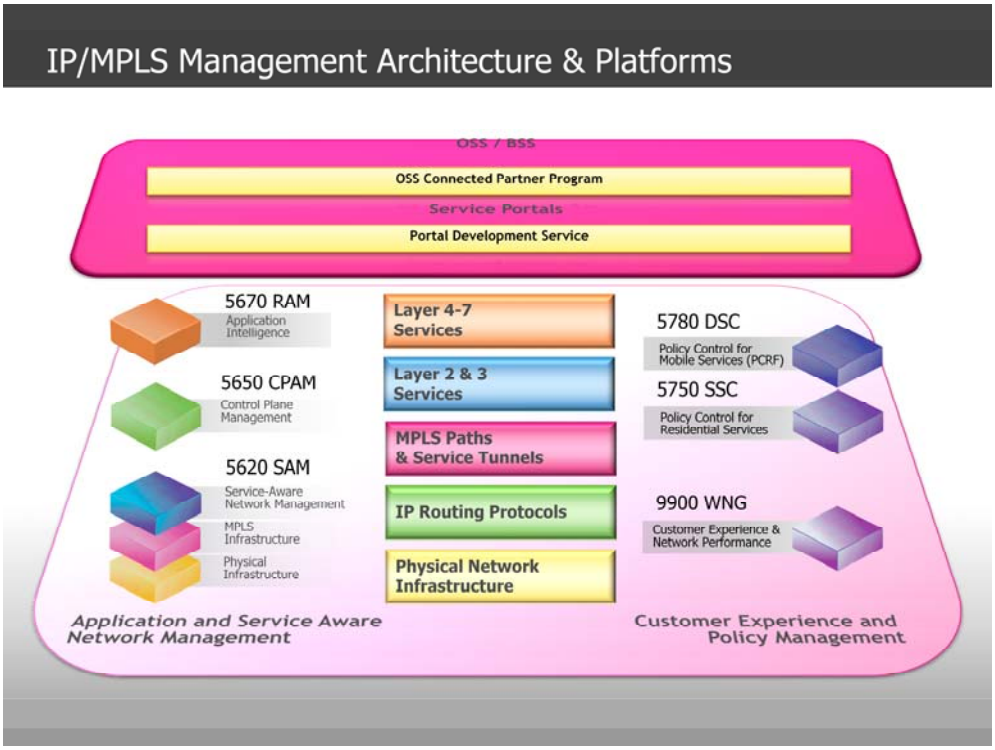
Application Assurance (AA) on the Alcatel-Lucent MS-ISA extends the service depth and functionality of the Alcatel-Lucent 7750 SR and 7450 ESS by enabling visibility and intelligent control for IP applications, with extensive per-application, per-subscriber, or per-VPN Layer 2 and Layer 3 service policies, providing application reporting and traffic management capabilities.

The Alcatel-Lucent MS-ISA can be used to deliver superior Internet Protocol Television (IPTV) QoE by incorporating advanced video services into the network elements, offering high scalability and flexible deployment scenarios. Advanced video services enabled by the MS-ISA include: Retransmission (RET) and Fast Channel Change (FCC) for IPTV running over RTP and linear TV Ad Insertion or ADI.

NAT, L2TP and Dual-Stack Lite are the three main features of the broadband service capabilities of the MS-ISA card.

NAT services on the Alcatel-Lucent MS-ISA add large-scale NAT capabilities to the Alcatel-Lucent 7750 SR, allowing service providers to conserve IPv4 addresses and maintain IPv4 Internet access while migrating to IPv6.

LNS services on the Alcatel-Lucent MS-ISA add L2TP Network Server functionality to the Alcatel-Lucent 7750 SR, allowing service providers to offer the same industry-leading Enhanced Subscriber Management (ESM) services to subscribers terminated on L2TP sessions that are already available



This slide provides an overview of all IP/MPLS management platforms.

The Alcatel-Lucent 5670 Reporting and Analysis Manager (RAM) delivers accurate, timely, application-level intelligence for service provider and enterprise business and operational planning, as well as insight into improving application performance and QoS optimization by collecting, warehousing, aggregating and analyzing application performance and volumetric data for application traffic flows.

The 5650 Control Plane Assurance Manager or CPAM is an IP/MPLS control plane management solution enabling service providers to assure network and service availability against control plane misconfigurations, malfunctions and undetected routing updates as well as accelerate service problems resolution over an IP/MPLS infrastructure. The Alcatel-Lucent 5650 Control Plane Assurance Manager offers real-time control plane visualization, proactive control plane surveillance, configuration validation and control plane diagnosis.

The Alcatel-Lucent 5620 Service Aware Manager or SAM takes service providers well beyond the traditional boundaries of element, network and service management. It enables unified, end-to-end management of IP/MPLS and Carrier Ethernet networks and the services they deliver to help service providers quickly gain the efficiencies they need to beat the competition. Rapid provisioning reduces time-to-market and increases flexibility when launching new services. Proactive troubleshooting helps resolve problems before they affect customers.

Faced with rapidly increasing mobile data traffic and growing subscriber expectations, mobile service providers are looking for ways to grow revenue and profitability while enhancing the subscriber experience. The Alcatel-Lucent 5780 Dynamic Services Controller or DSC enables service providers to monetize and optimize network resources while offering personalized choice for the subscriber in both a 3G and 4G environment.

The Alcatel-Lucent 5750 Subscriber Services Controller or SSC is a flexible, modular and pre-integrated subscriber and policy management solution for residential broadband services which is also positioned as part of Alcatel-Lucent's extensive IMS portfolio. The 5750 SSC enables innovative broadband services such as IPTV, enhanced high-speed Internet, video on demand and VoIP to be dynamically added, upgraded, initiated and controlled by subscribers to satisfy their personal needs.

The Alcatel-Lucent 9900 Wireless Network Guardian or WNG provides powerful capabilities for wireless data service providers to accurately design, engineer, optimize, manage, and price their networks. Currently, wireless service providers reuse IP network management tools designed for wireline networks, which typically view the traffic load on an IP network across all data applications to a single dimension: volume (or bandwidth).

WRAP-UP

Wrap-up: The SR-OS or Service Router Operating System is an industry leading product platform that plays a crucial role in the high leverage network. One of the key technology innovations is the introduction of the new Flex Path 3 chip. This chip is all about being faster, smarter and greener. The SR-OS products are also continiuously evolving and are found in different types of networks. In the residential, business and mobile networks, which are typically owned by the traditional service providers. But also in the strategic industry markets like energy, transportation, public sector, enterprises, defense and security markets as seen in the next modules.

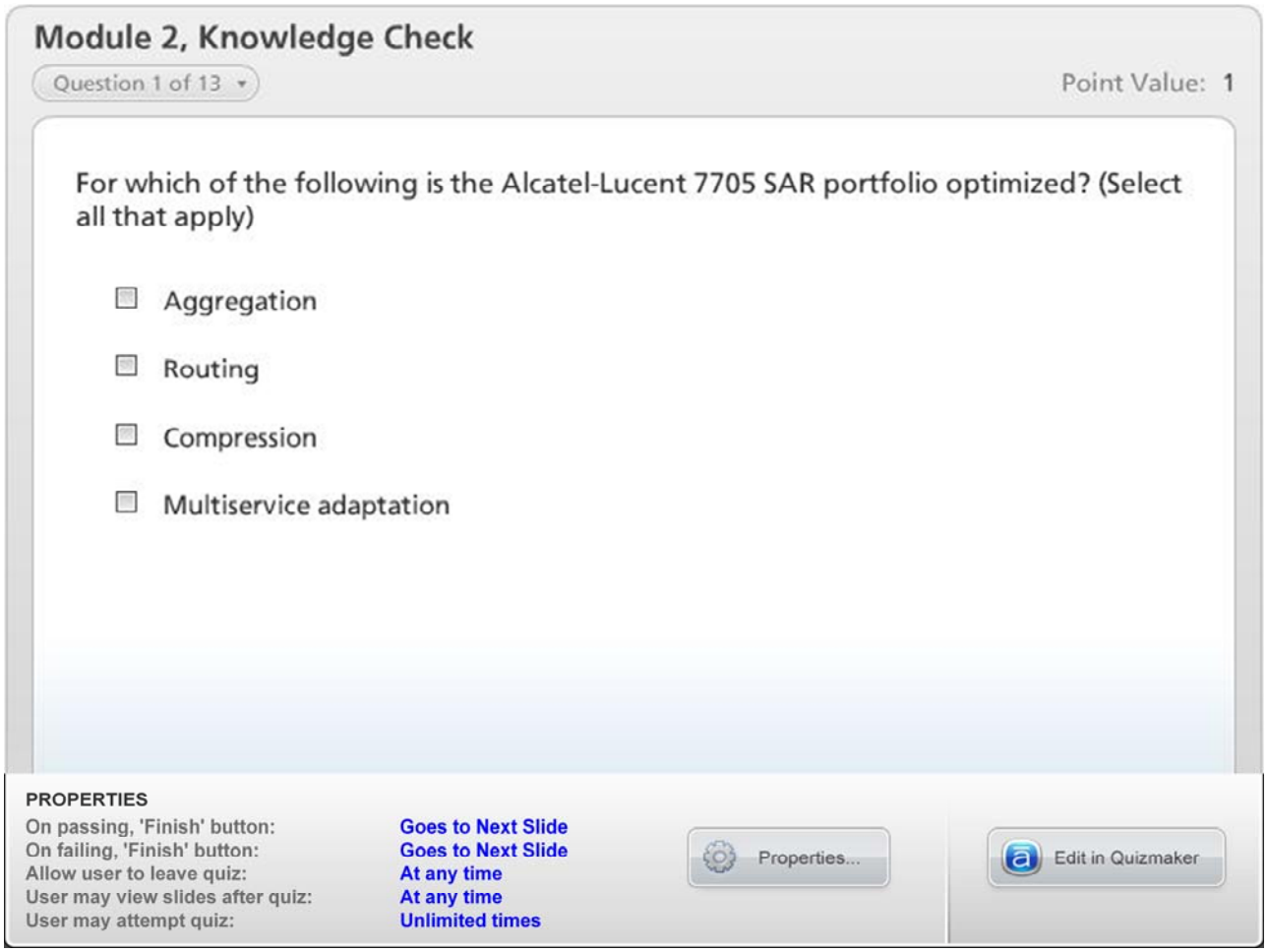
Module summary

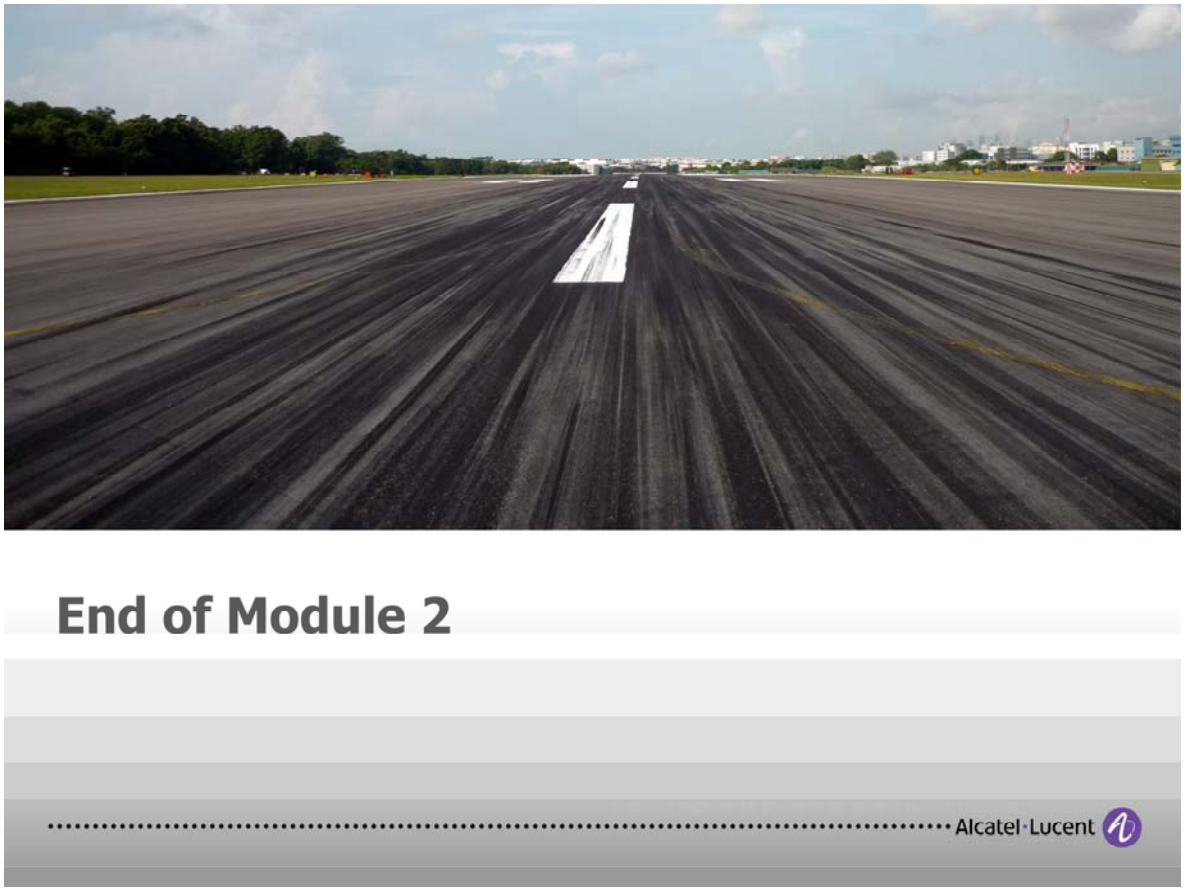
- The four major SR-OS product families are the 7750 service router, the 7450 Ethernet service switch, the 7705 service access router and the 7210 service access switch.
- The IP/MPLS products run at 9 Gigabits per second for the 7750 SR-c4 up to 400 Gigabits per second for the 7750 SR-12.
- The complete portfolio is managed by the 5620 Service Aware Manager (SAM).
- The key differentiators are the enormous service scale and performance, the high availability features like non-stop routing and non-stop service, the huge queue buffer space available and the deep packet inspection that allows a flexible way of forwarding.

In conclusion, this module has given an overview of the SR-OS product families. The 7750 SR with a focus on the new compact variants, the C4 and C12, the 7450 ESS, 7705 SAR and the 7210 SAS. This module highlighted also the key differentiators like high availability features, non-stop routing and non-stop service. At the end the network management platform was introduced.

Knowledge Checks

On the next slide are 13 questions that quiz you on the key points of this mod.





This completes module 2.

SR-OS Fundamentals
Module 3: CLI (Command Line Interface)

IPD Development

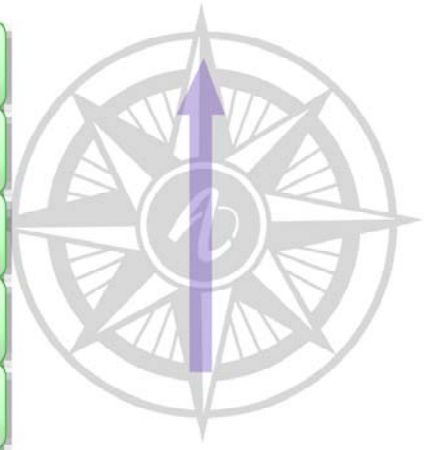
Available as PDF

Welcome to the third module of the SR-OS fundamentals course.

This module explains the basics of CLI and is mainly intended for network operators.

A separate course for administrators is available. This administrator course focuses more on the commissioning of a service router.

Table of Contents



Module 3 is divided into five sections.

Section 1 introduces the basics of CLI.

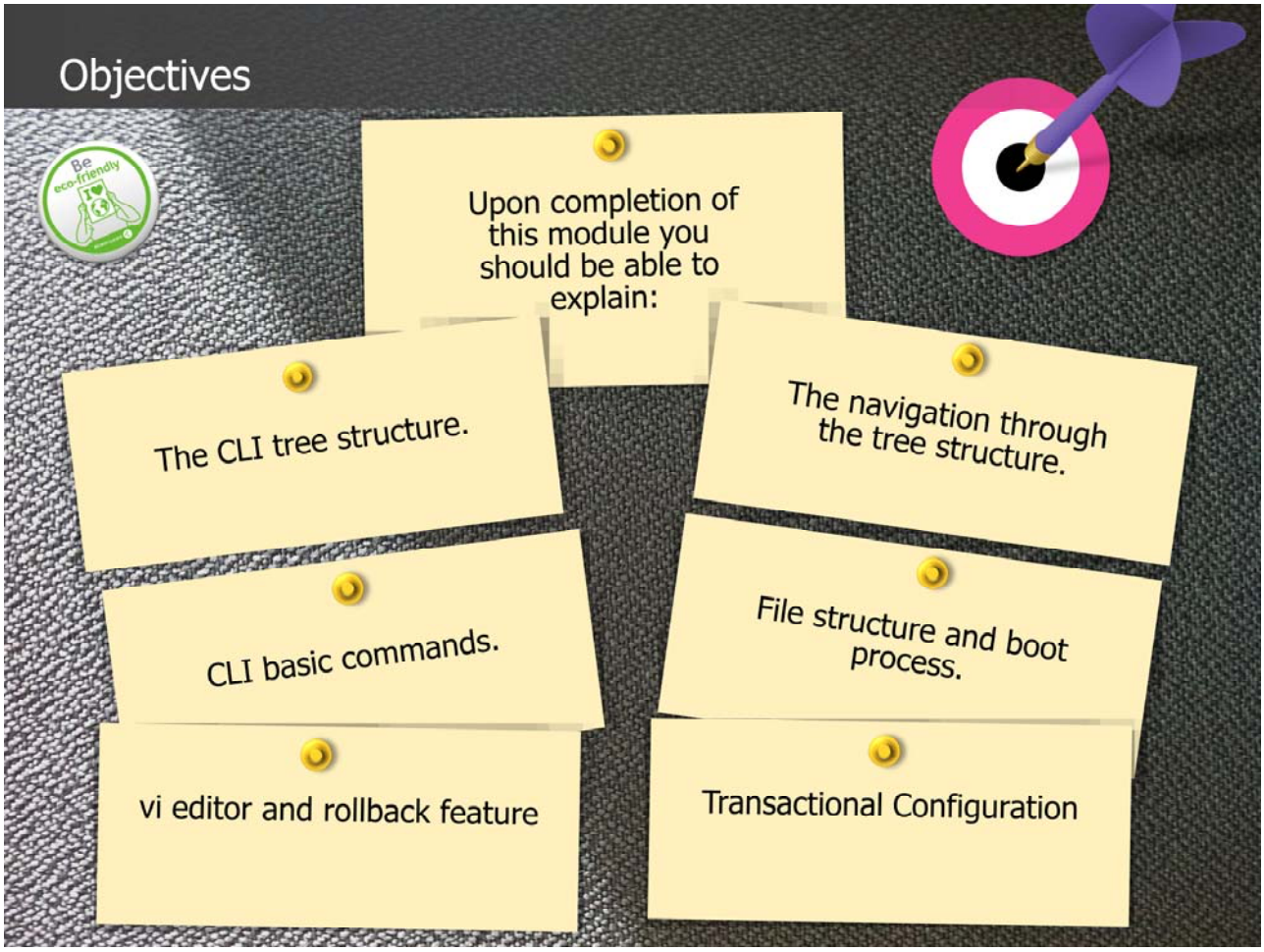
Section 2 explains how events are logged.

Section 3 shows the most important files in the file structure and explains the boot process.

Section 4 discusses some basic CLI configurations

Section 5 discusses the Vi editors and the rollback feature in CLI, and

Section 6 discusses the Transactional Configuration feature.



By the end of module 3 you will be able to explain:

The CLI tree structure

The navigation through the tree structure

CLI basic commands

File structure and boot process

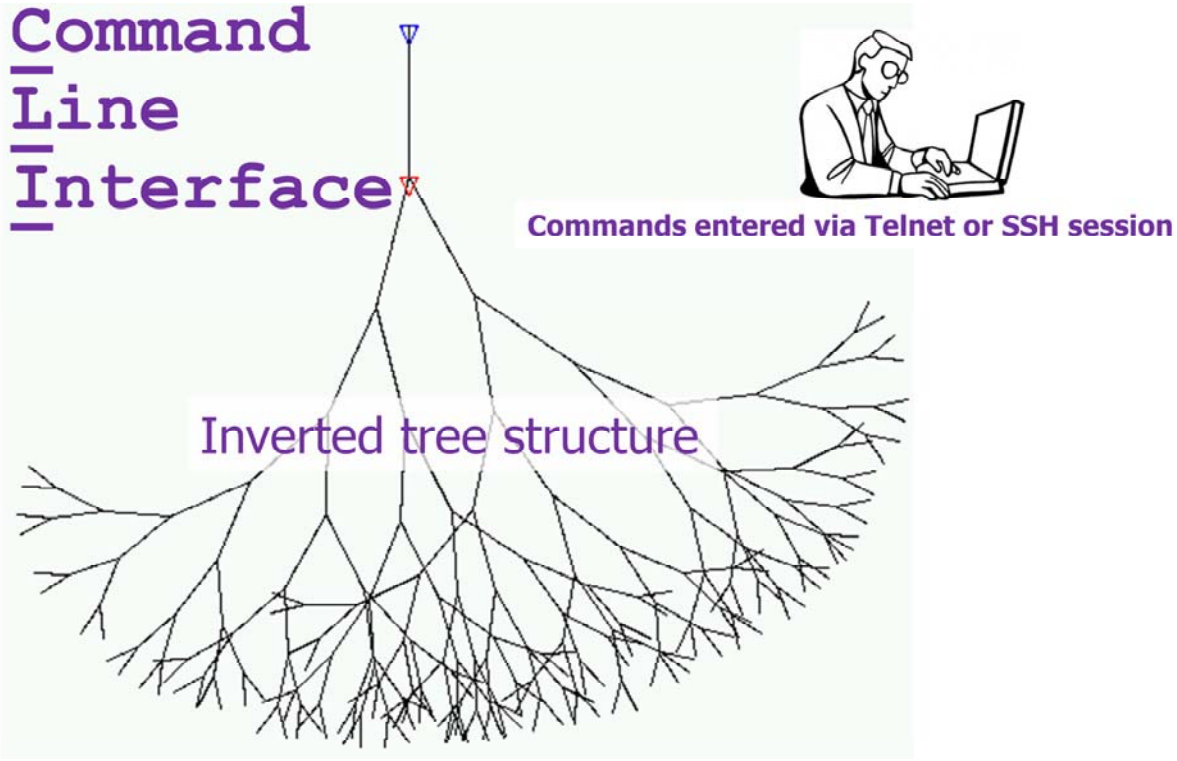
vi editor and rollback feature

Transactional Configuration

CLI basics

Section 1: The Basics of CLI

CLI Tree Structure



The CLI or the Command Line Interface is an interface available on a service router to send commands. The commands are typed in via a Telnet or SSH session. The parameters that can be changed are located in an inverted tree structure. On each branch of the tree parameters can be set, changed or analyzed.

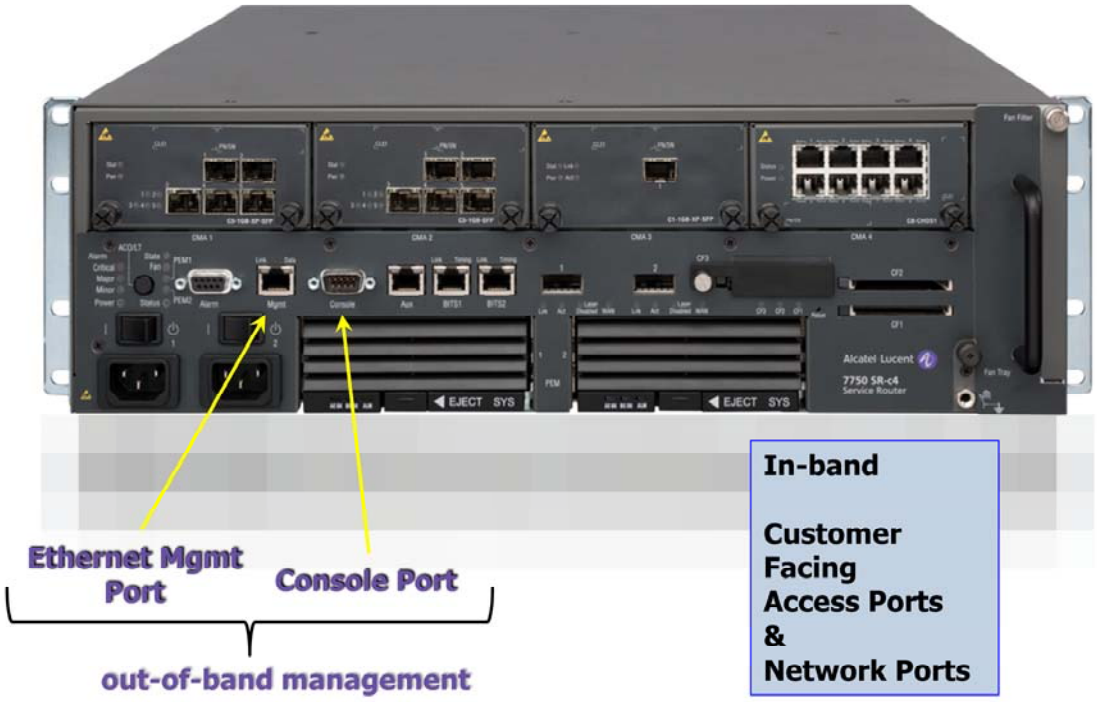
CLI tree structure

```
*A:GTC# tree

|
+---admin
|   |
|   +---application-assurance
|   |   |
|   |   +---upgrade
|   |
|   +---debug-save
|   |
|   +---disconnect
|   |
|   +---display-config
|   |
|   +---enable-tech
|   |
|   +---radius-discovery
|   |   |
|   |   +---force-discover
|   |
|   +---reboot
|   |
|   +---redundancy
|   |   |
|   |   +---force-switchover
|   |   |
```

```
*A:GTC# tree detail

|
+---admin
|   |
|   +---application-assurance
|   |   |
|   |   +---upgrade
|   |
|   +---debug-save [<file-url>]
|   |
|   +---disconnect [address <ip-address> |
username <user-name> | {console|telnet|ftp|ssh}]
|   |
|   +---display-config [detail|index]
|   |
|   +---enable-tech
|   |   no enable-tech
|   |
|   +---radius-discovery
|   |   |
|   |   +---force-discover [svc-id <service-id>]
|   |
|   +---reboot [active|standby|upgrade] [now]
|   |
|   +---redundance
```

The "tree" command shows the entire tree. This example shows the execution of the command from the root context of the tree, but this command can be given at any branch of the tree. The "tree detail" command provides more details on each part of the branch of the tree.

Physical access



The CLI commands are sent either in-band or out-of-band. When sent in-band, the CLI management traffic is sent together with user traffic. Out-of-band management traffic is sent either to the ethernet management port or via the console port.

Each of the service routers has two connectors available for out-of-band management. The 7750 C4 is shown as an example.

Ethernet management port (on SAR-8)

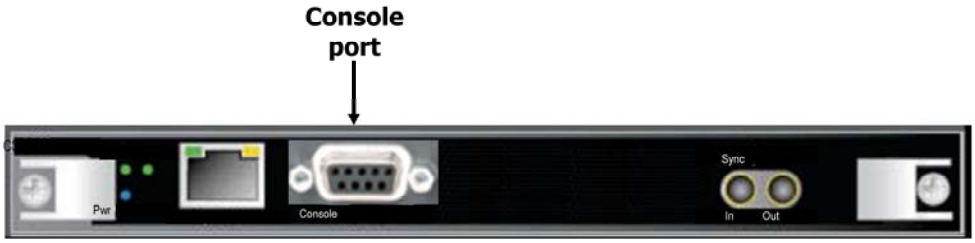
- Ethernet port is an RJ-45
- Ethernet port supports
  - full/half duplex mode
  - 10M/100M speed
  - autonegotiate

The Ethernet management port is an electrical RJ-45 port with full and half duplex mode supported. The speed of 10M or 100M and the port mode are auto-negotiable.

The example shows the management card of a 7705 SAR-8.

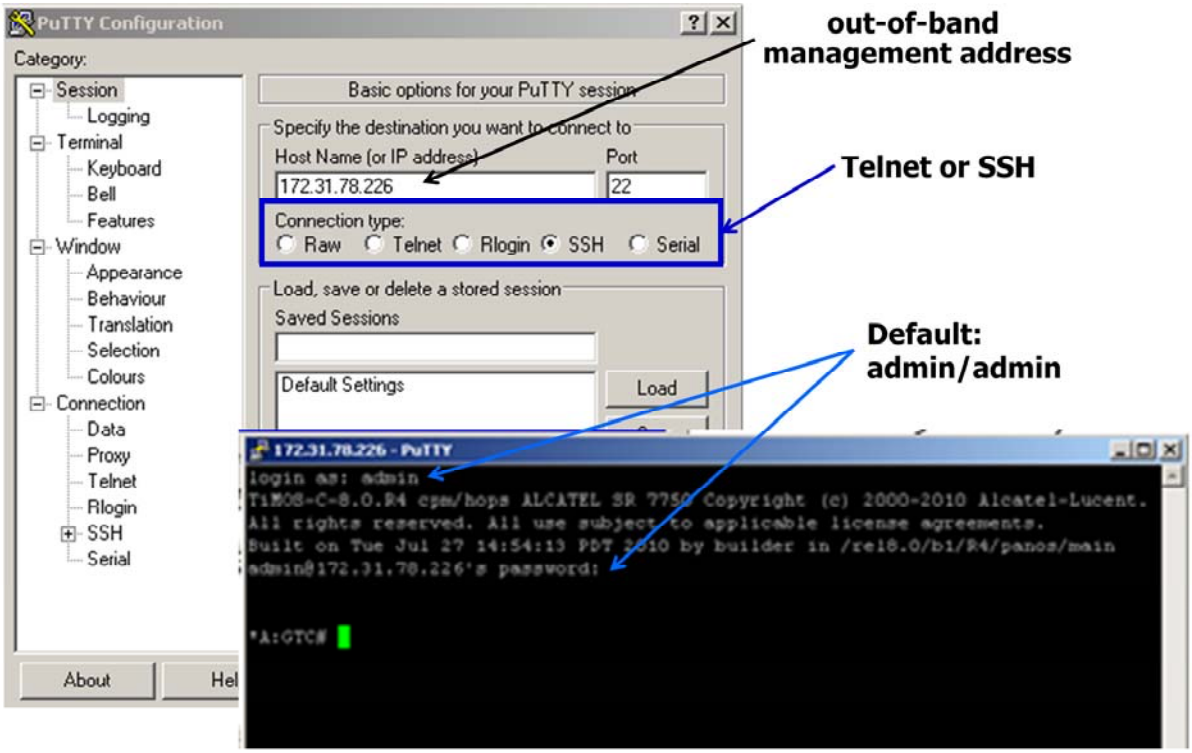
Console Port (on SAR-8)

- Console port is a DB-9 connector
- Console port supports
  - Baud rate of 9600, 19200, 38400, 57600 or 115200
  - Default is 115200, 8N1 and no flow control
  - 8 data bits, no parity and 1 stop bit (8N1 setup)



Next to the ethernet management port of the SAR-8, the console interface is found. The port needs a DB-9 connector and works at different configurable baud rates. The default baud rate is 115,200 Kbit/s. The console port uses 8 data bits, no parity bit and 1 stop bit.

Log-in



To log in into the service router any available Telnet or SSH client can be used. Putty is an easy to use freeware application found on the Internet. After typing in the out-of-band management address in Putty, Telnet or SSH has to be chosen as connection type. Note that SSH is by default supported on the service router. If needed, a Telnet server can be enabled.

After this, a prompt appears asking for the username and password. Admin/admin is the initial setting which can be changed by the administrator later on.

Console Port (on SAR-8)

Enables telnet

```
*A:GTC# configure system security telnet-server
```

Number of sessions

```
*A:GTC# configure system login-control telnet inbound-max-sessions
```

```
configure system security version 1 , 2 , 1-2
```

SSH version

```
show system security ssh

SSH is enabled
SSH preserve key: Enabled
SSH protocol version 1: Enabled
RSA host key finger print:e8:bd:fa:29:b6:39:81:a8:33:8d:ce:f2:25:d1:d9:2c

SSH protocol version 2: Enabled
DSA host key finger print:5e:65:d1:b8:ed:c4:db:b1:69:d3:43:e5:6e:95:c3:81
RSA host key finger print:fc:c7:44:60:df:1d:b5:e7:c8:65:20:93:76:70:f2:ea

=======================================================
        Connection            Username
=======================================================
        138.203.17.246        admin
```

If Telnet sessions need to be established between the user and the service router, the command "configure system security telnet-server" needs to be typed in.

If the number of simultaneous user sessions needs to be enhanced, the command "configure system login-control telnet inbound-max-sessions" increases the number of sessions. The SSH version is set by default to 2 but can be changed if necessary. All the SSH settings can be verified by "show system security ssh".

Console Port (on SAR-8)

Example: Configuring OSPF

Node>config>router>ospf#

Host name Node — Context separator

Example: Creating a new interface

Node>config# router interface Toronto
Node>config>router>if$ address 131.131.131.1/30

At the end of the prompt, there is either a pound sign # or a dollar sign $:

- A # indicates the context is an existing context
- A $ indicates the context is newly created

To view the present working context and the user configured entities:

Node>config>router>if# pwc

The prompt in a Telnet or SSH shows, unless otherwise configured, the current location or context in the hierarchical tree. The prompt starts on the left with the name of the router which represents the Root context and ends on the right with the current context. At the end of the prompt, you will see either a pound "#" sign or a dollar "$" sign. The # sign indicates you are working in an existing context. An existing context is one previously created, such as an interface you created earlier in the day. The $ sign indicates you have created a new context. A newly created loopback interface would create a new context. You work within this new context to configure the new interface.

The prompt doesn't show the user configured entities, only context's. To view the present working context and the user configured entities such as an interface name type in "pwc".

The CLI Navigation

```
Node>config# router

Node>config>router>ospf


back       - Brings you back one context
exit       - Returns to previous worked in context
exit all   - Brings you back to the root level
<Ctrl-z>   - Acts like hitting <ENTER> and exit all sequentially
<Ctrl-c>   - Clears the command prompt
up/down arrow      - Lists previous command(s) to be repeated
tree       - Shows available commands from context
```

Simply moving down the hierarchy can be accomplished by entering the next context at the prompt to move down one step in the tree. For example by typing the command router at the root "config#" prompt. Or by entering a string of context at the prompt to move down several levels at once. For example by typing the command "config router ospf at the root# prompt.

Moving up the hierarchy can be accomplished by several commands.

Entering the **back** command to go up one context in the tree.

Entering the **exit** command to return to the previous worked in context.

Entering the **exit all** command returns you directly to the root context.

Entering <Ctrl-z> which represents a special keyboard sequence that acts like pressing the **<Enter>** key and entering **exit all** to return the user to the ROOT context.

Entering <Ctrl-c> clears a typed-in string at the prompt instead of backspacing.

The <up> and <down> arrow keys list the previously typed in commands.

The tree command can also be of help to show the available commands in the context.

The Help Command

Help	Displays a brief description of the help system
?	Lists all commands in the current context
string **?**	Lists all commands available in the current context that start with *string*
command ?	Displays command's syntax and associated keywords
command keyword ?	Lists the associated arguments for *keyword* in *command*
Help Edit	Displays help on editing (editing keystrokes) Lists available editing key strokes

Pressing "help" displays a brief description of the help system.
The question mark is one of the most useful commands available. It lists all the commands available in the current context but also gives the variable fields that can be filled-in after the command.
The question mark can be given after a string, command or key work to see what to type next.
"Help edit" displays help on editing and lists the available editing key strokes.

Help	Displays a brief description on the Help functionality. This command can be used in any context.
?	This command displays all the possible commands in the current context.
"string" ?	Lists all the possible commands in the current context that start with *"string"*.
Command ?	Displays a command's syntax and associated keywords.
Command keyword ?	Lists all the associated arguments for a keyword in a command.
Help Edit	Displays help on editing (editing keystrokes) and lists all available editing key strokes.
Help Globals	Displays help on the global commands and lists all available global commands.

The Global Commands

Help Globals	Displays help on global commands Lists available global commands

logout	Terminates the CLI session
oam	OAM test suite. See the Service OAM section of the SR OS Services Guide.
ping	Verify the reach ability of a remote host
pwc	Displays the present or previous working context of the CLI session.
sleep	Causes the console session to pause operation (sleep) for one second or for the specified number of seconds. Primary use is to introduce a pause within the execution of an exec file.
ssh	Open a secure shell connection to a host
telnet	Telnet to a host
traceroute	Determine the route to a destination address
tree	Displays a list of all commands at the current level and all

The Global Commands are commands that can be entered at any context. The entire list can be viewed with the "help globals" command and, since help is one of the global commands, this can be done at also any context. The most frequently used global commands are listed here.

Back	- Go back a level in the command tree
Echo	- Echo the text that is typed in
enable-admin	- Enable the user to become a system administrator
Exec	- Execute a file - use -echo to show the commands and prompts on the screen
Exit	- Exit to intermediate mode - use option all to exit to root prompt
Help	- Display help
History	- Show command history
Info	- Display configuration for the present node
Logout	- Log off this system
Mrinfo	- Request multicast router information
Mstat	- Trace multicast path from a source to a receiver and display multicast packet rate and loss information
Mtrace	- Trace multicast path from a source to a receiver
Oam	- OAM Test Suite
Ping	- Verify the reachability of a remote host
Pwc	- Show the present working context
Sleep	- Sleep for specified number of seconds
Ssh	- SSH to a host
telnet	- Telnet to a host
Traceroute	- Determine the route to a destination address
Tree	- Display command tree structure from the context of execution
write	- Write text to another user

The CLI Environment Commands

alias	Enables the substitution of a command line by an alias
create	Enable create parameter check
more	Configures whether CLI output should be displayed one screen at a time awaiting user input to continue
reduced-prompt	**Configures the number of higher level CLI contexts to display in the CLI prompt**
terminal	Configures the terminal screen length for the current CLI session
time-display	Specify whether time should be displayed in local or UTC

The CLI environment commands, issued from the **root>environment** context are used to customize session preferences for a single CLI session and are lost when logged off. Let's take the reduce-prompt as an example. To configure the number of higher level CLI context's to display in the CLI prompt, type in the environment context "reduced-prompt 2" if the maximum number of context's that need to be displayed is 2. This simplifies the look of the prompt.

Node>environment# alias	Substitutes a long command line with an alias. The example shows a way to let an alias to replace a larger command.
Node>environment# create	Enables the create parameter check.
Node>environment# more	Allows the CLI output to be displayed one screen at a time.
Node>environment# reduced-prompt	Configures the number of higher level CLI contexts to display in the CLI prompt.
Node>environment# terminal	Configures the terminal screen length for the current CLI session.
Node>environment# time-display	Specifies whether the time should be displayed in local or UTC mode.

The Auto-completion Commands

Command completion can be achieved by:

Abbreviation, if keystrokes entered are unique enough.

```
Node>config# ro    <ENTER>
Node>config>router#
```

Tab Key or Space Key to auto-complete the command.

```
Node>config# ro    <TAB>
Node>config# router

Node>config# ro    <SPACE>
Node>config# router
```

If match is not unique CLI will display possible matches:

```
Node>config# r [TAB]
redundancy router
```

The CLI is able to auto-complete partially entered commands. This auto-complete function is invoked by hitting "ENTER", "TAB" or "SPACE".

As long as the partial command can uniquely identify a command or context that is allowed in the current context "TAB" and "SPACE" will auto-complete the entire command.

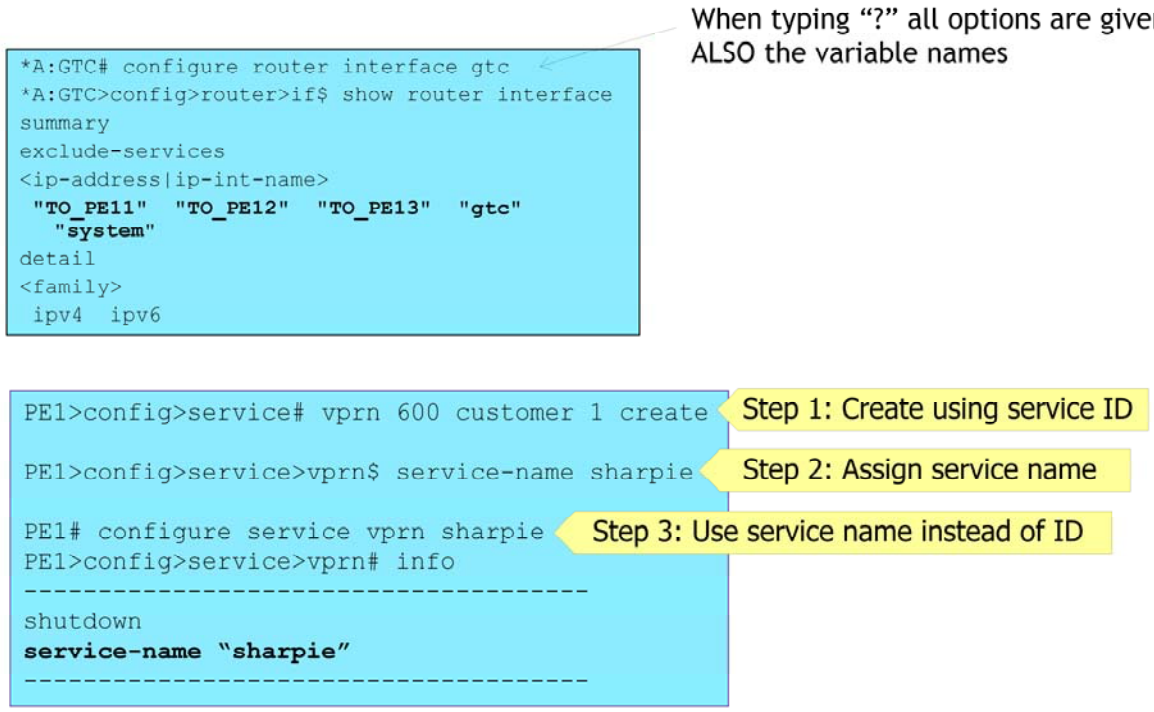
"ENTER" will auto-complete the command AND also execute it.

If the partial command is ambiguous and indicates more than one option, "TAB" and "SPACE" will list all the remaining possible matches and waits for the operator to add more letters.

"ENTER" will result in an error message in this case.

The system maintains a history of 30 previously entered commands that can be displayed with the "history" command from any context.

Auto-Complete of Variables



Not only commands but also variables are subject to be auto completed. When typing "?", not only the commands but also all previously used variable names are displayed. This feature gains time when configuring a service router.

Another variable is the name of a service. All services have ID's but can also have an additional service name. When addressing the service, the name can be used instead of the ID. A name is always easier to remember than a number.

The Info Command

- The info command displays information during configuration without the need to use the show context.

```
Node>config>router# interface Toronto
Node>config>router>if# info
----------------------------------------------
            address 131.131.131.1/30
            port 1/1/1
----------------------------------------------
```

- More detail can be viewed by using the info detail command.
- To view the entire configuration file:

```
Node# admin display-config
```

There are two options to see what has already been configured.

The "admin display-config" command displays the entire present configuration.

Or the "info" command can be used in any sub-context of the **config** context.

The "info" command only shows the configuration that is different from the default configuration. To view the entire configuration of the current context including the default settings, the "info detail" command can be used.

Show system information

```
A:GTC# show system information
```

An overview of the system can be seen by typing “show system information”.

Show Chassis

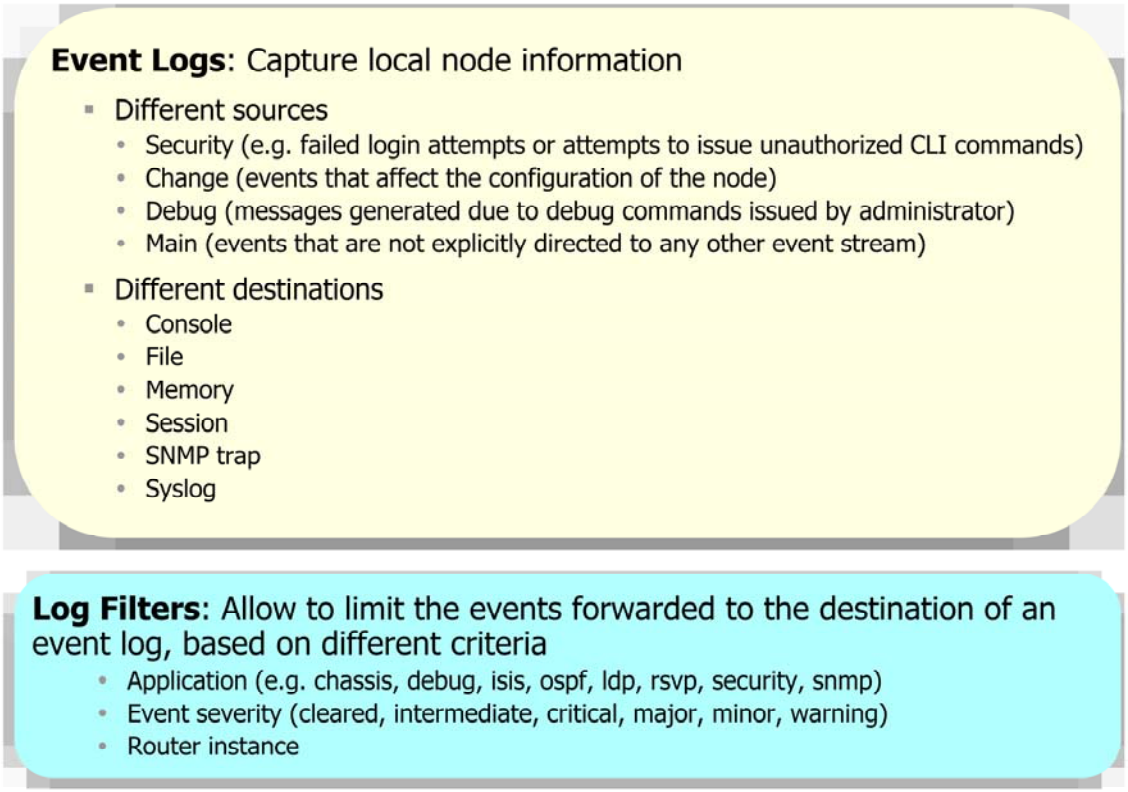
```
A:GTC# show chassis
```

All relevant chassis and hardware data can be seen by typing “show chassis”.

Event Logging

Section2: Event Logging.

Logs



Event logging controls the generation, dissemination and recording of system events for monitoring status and troubleshooting faults within the system. The 7750 SR groups events into four major categories or event sources.

- Security - Events that pertain to attempts to breach system security.
- Change - Events that pertain to the configuration and operation of the node.
- Debug - Events that pertain to trace or other debugging information.
- Main - Events that pertain to applications that are not assigned to other event categories/sources.

An event log within the 7750 SR OS associates the event sources with logging destinations. Examples of logging destinations include, the console session, a specific telnet or SSH session, memory logs, file destinations, SNMP trap groups and syslog destinations.
A log filter policy can be associated with the event log to control which events will be logged in the event log based on combinations of application, severity, event ID range, VRF ID, and the subject of the event.

Severity Number	Severity Name
1	Cleared
2	Indeterminate
3	Critical
4	Major
5	Minor
6	Warning

Logs - continued

Event Logs

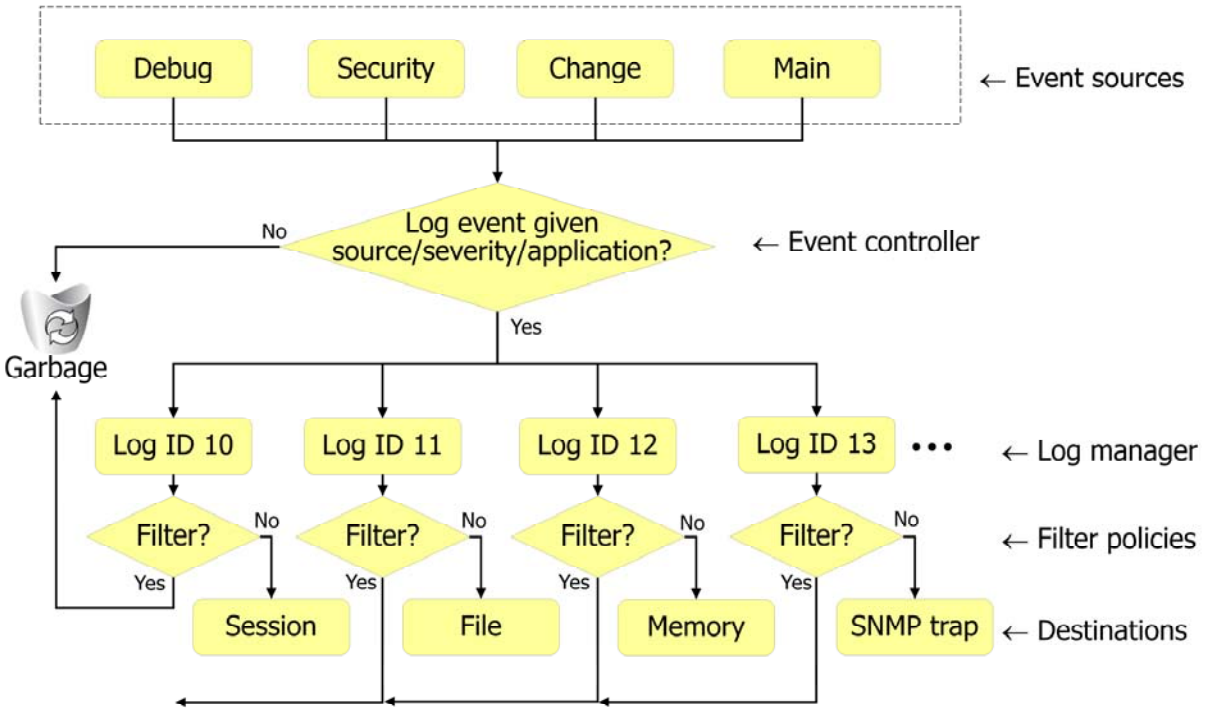
- There are automatic event logs
  - Log 99 - All severity levels of events from main
  - Log 100 - Severity greater than major events from main
- There are user-defined event logs that can be created as needed
  - Log 98 - Typically reserved for SAM (best practice)
  - Log IDs from 1 to 97

```
*A:GTC# show log log-id

===============================================================================
Event Logs
===============================================================================
Log Source Filter Admin   Oper      Logged    Dropped    Dest      Dest   Size
Id         Id     State   State                         Type      Id
-------------------------------------------------------------------------------
 99 M      none   up      up          203         0     memory          500
100 M      1001   up      up           10       193     memory          500
===============================================================================
```

Each event log needs to be created by the operator, except log 99 and log 100. These logs are always available on a router and can't be changed. They keep up to 500 major and minor main sourced events. Log 100 keeps the log events with a severity greater than major from main. Log 99 keeps all the other main events.

Log 98 is by default used for SAM snmp events, although any other number between 1 and 97 can be used. This is specified in the SAM network management platform.

Logs - continued

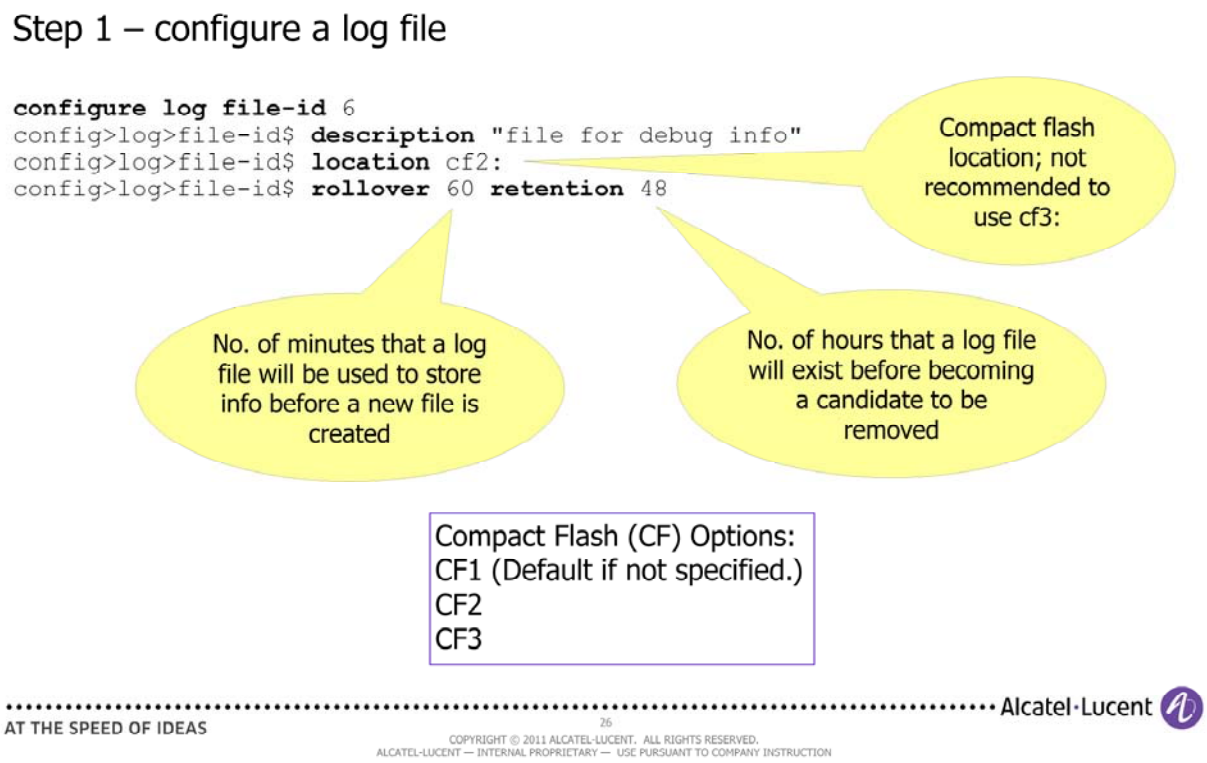


The event control pre-processes the events generated by applications before the events are passed into the log manager. Event control assigns a severity level for each application event and determines if the event should be generated or suppressed depending on whether there is a configured log expecting such an event type. Events that are suppressed by event control will not generate any event log entries since they never reach the log manager.

Events that are forwarded by event control are sent to the log manager. The log manager manages the event logs in the system as well as the relationships between the log sources, event logs, log destinations, and log filter policies.

The log manager uses event filter policies to allow fine control over which events are forwarded to the log destination or dropped based on various criteria. Filter policies include a number of entries that are identified with an entry ID and define specific criteria plus a forward or drop action for events that match the criteria. Matching criteria include the event's application, event number, severity, and subject conditions. Filter policies also have a default action, which applies to events that do not match more specific criteria. The default action can also be to forward or to drop the event.

Example 1 - Sending debug information to a local file



Let us look at an example. If the destination of the events is a log file on a router compact flash of the router, the file has to be created first.

After assigning an ID to the file, the location has to be specified. Three compact flashes options are possible, cf1, cf2 or cf3. The location command is optional. If not explicitly configured, log files will be created on cf1:

Logs can consume lots of flash space thus it is advised to never log to compact flash 3, which is reserved for system.

The Rollover Interval, defined in minutes, determines how long a file will be used to store log information before it is closed and a new log file is created. The default rollover interval is 1440 minutes or 1 day.

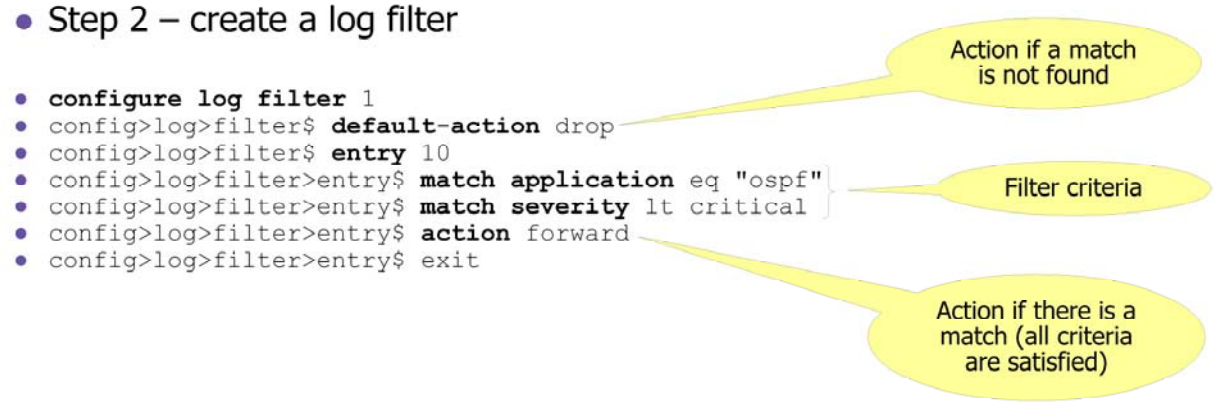
The Retention Interval, defined in hours, determines how long a log file will exist before becoming a candidate to be removed if space is needed to store new files. The file becomes a candidate for removal once the creation timestamp + rollover time + retention time is less than the current timestamp. The default retention interval is 12 hours.

Log files are created automatically in the specified location. The file name, also generated automatically, reflects the file ID and the log ID that they are associated with, as well as the date and time the file is created. Each file will be used to store log information during a period of time equal to the rollover interval, and then a new file will be created to continue storing information. Old log files may be removed after their retention interval if space is needed for newer files.

Example 1 - Sending debug information to a local file

- Step 2 – create a log filter

```
configure log filter 1
config>log>filter$ default-action drop
config>log>filter$ entry 10
config>log>filter>entry$ match application eq "ospf"
config>log>filter>entry$ match severity lt critical
config>log>filter>entry$ action forward
config>log>filter>entry$ exit
```

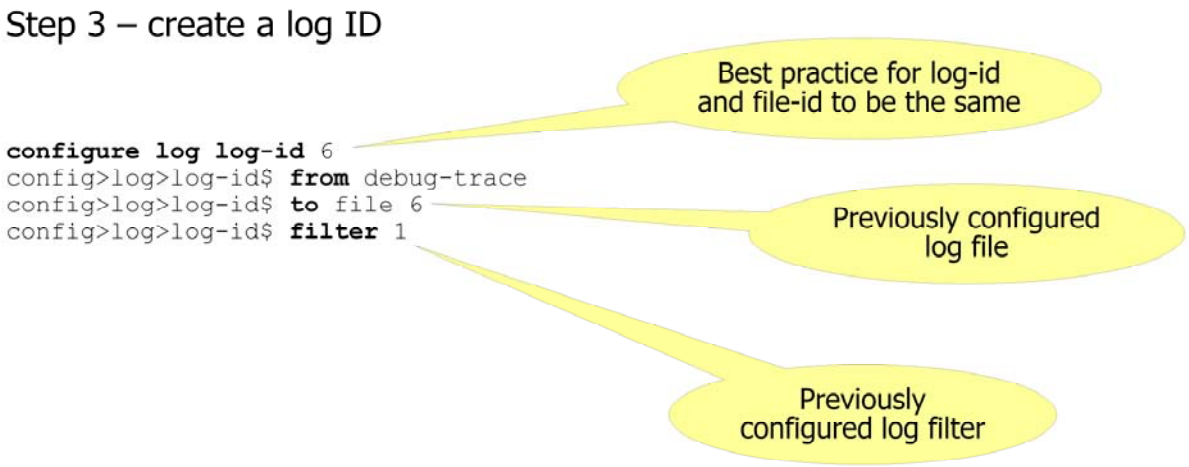


Optionally a filter log can be used to allow fine control over which events are forwarded.

Multiple entries may be created for a given filter using unique *entry-id* numbers. Entries are evaluated in an ascending entry-ID order. When a match is found, the SR OS implementation executes the specified action, drop or forward, and exits the filter without evaluating the subsequent entries. For that reason, when multiple entries are created, they should be arranged sequentially from the most explicit entry to the least explicit one.

An entry's action applies if all of the criteria that correspond to such an entry are satisfied by the event being evaluated.

If a match has not been found after going through all of the filter entries, the default action is executed. If the default action is not explicitly configured, then the default action will be to forward the relevant event.

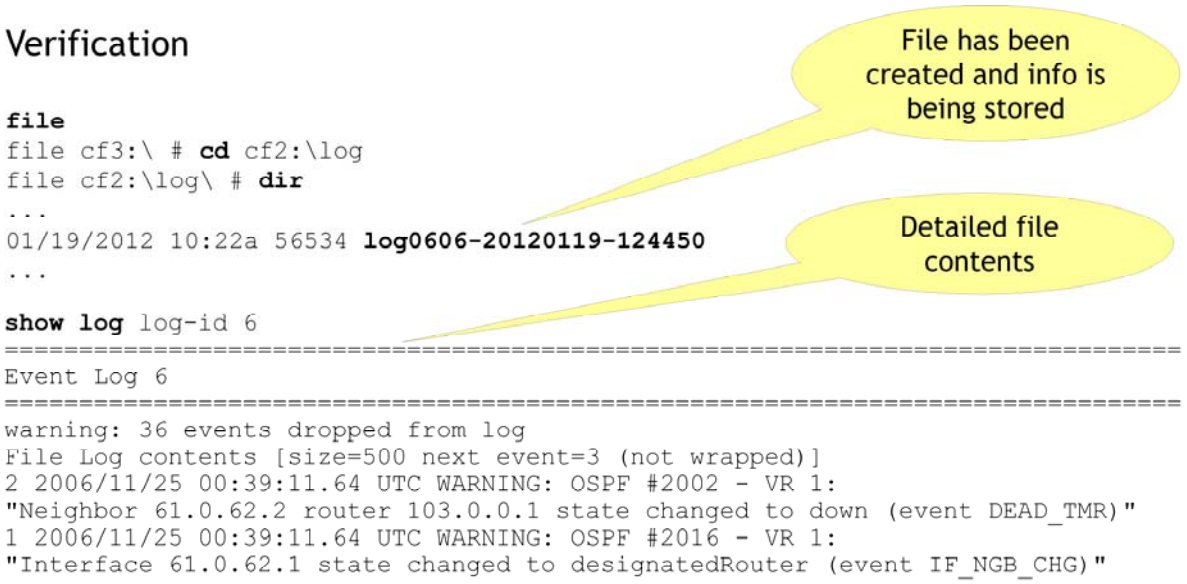


It is considered best practice to configure the log-id and file-id to be the same to make troubleshooting easier when needed.

Example 1 - Sending debug information to a local file

- Step 4 – activate debug
- **debug router ospf packet detail**

Verification



Log files are created automatically in the specified location. The file name, also generated automatically, reflects the file ID and the log ID that they are associated with, as well as the date and time the file is created.

In this example, for the file name generated:

- **log0606** represents log-id 06, file-id 06.
- **20120119** represents the calendar date of file creation on the node.
- **124450** represents a timestamp

To turn off the debug, just type "no debug"

This command should be used with caution, however, since it will disable all debugging globally. Debug output can also be turned off by shutting down the log that was created for the debug.

configure log log-id 6 shutdown

The debug statement will remain valid and if the log is re-enabled using the "no shutdown" command, the debug output will resume.

Example 1 - Sending debug information to a local file

- Verification (continued)

- `show log log-id 99 application chassis` ← Chassis related alarms
- `show log log-id 99 subject 1/1/1` ← Port specific
- `show log log-id 99 subject "Card 1"` ← IOM specific
- `show log log-id 99 subject "Card A"` ← SF/CPM
- `show log log-id 99 subject "Mda 1/1"` ← MDA specific

When showing information stored in an event log, we can select specific information of interest, instead of the entire log content.

Example 2 - Sending debug information to terminal

- Step 1 – create a log ID

```
configure log log-id 22
config>log>log-id$ from debug-trace
config>log>log-id$ to session
```

Step 2 - activate debug

```
debug router ospf packet detail
```

Detailed debug info displayed on the same terminal

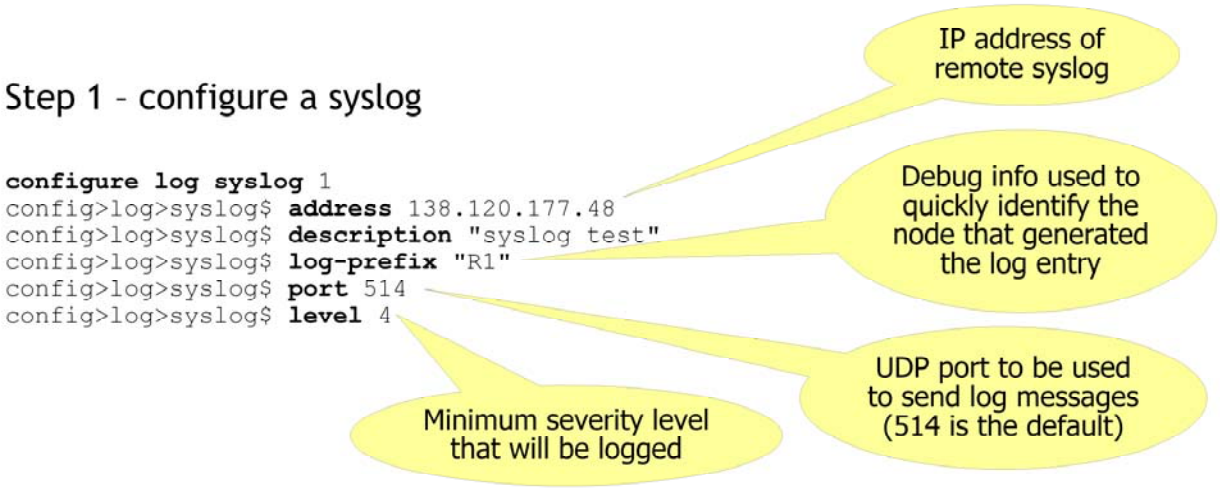
```
378 2010/02/10 21:33:39.14 UTC MINOR: DEBUG #2001 Base OSPFv2
"OSPFv2: PKT

>> Outgoing OSPF packet (on I/F toR5)
OSPF Version      : 2
Router Id         : 172.16.1.1
Area Id           : 0.0.0.1
Checksum          : 4e94
Auth Type         : Null
Auth Key          : 00 00 00 00 00 00 00 00
Packet Type       : HELLO
Packet Length     : 44
Network Mask      : 255.255.255.248
Hello Interval    : 10
Options           : 02
Rtr Priority      : 1
Dead Interval     : 40
Designated Router : 0.0.0.0
Backup Router     : 0.0.0.0
"
```

The output of the debug is posted to the session terminal. This is an easy option, good for short debug exercises.

When sending debug traffic to the session, the prompt may not be visible. Typing an "enter" brings back the prompt. If information is being sent to the session at the same time that you type a command, you may not be able to see what you are typing.

Example 3 - Sending debug information to a syslog

- A **syslog** is a central repository of management data, potentially from different sources
- It allows a separation of functions between the device that generates messages and the system that stores them, which can be remotely located for added redundancy
- It also allows keeping track of a longer history of events since locally-stored logs are circular (old info is overwritten)



Example 3 sends the debug information to a syslog server.

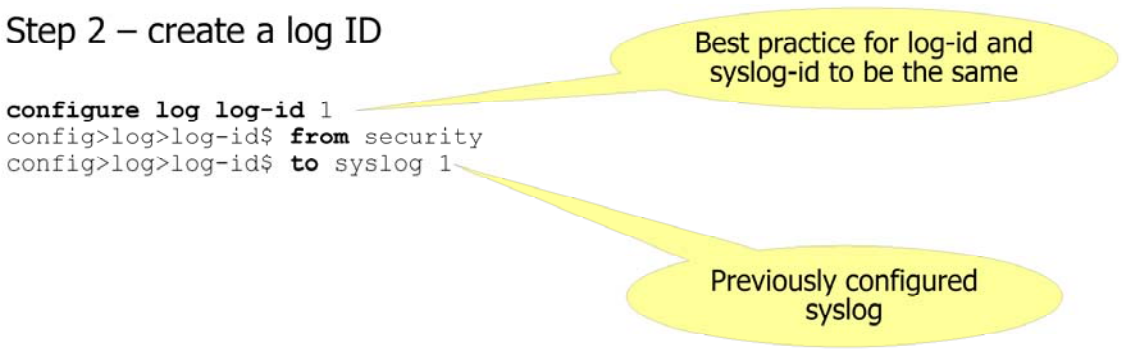
The remote syslog server, located at the IP address given, must also be properly configured. There are multiple vendors of syslog server software. To configure a syslog server please refer to its own user's guide.

The log-prefix is used to quickly identify the node that generated the log entry.

Port 514 identifies the UDP used.

All events are assigned a specific severity level. The definition of a severity level for syslog purposes is different from the definition in other log-related contexts. Level 4 in this example will be the minimum severity level that will be logged.

Example 3 - Sending debug information to a syslog

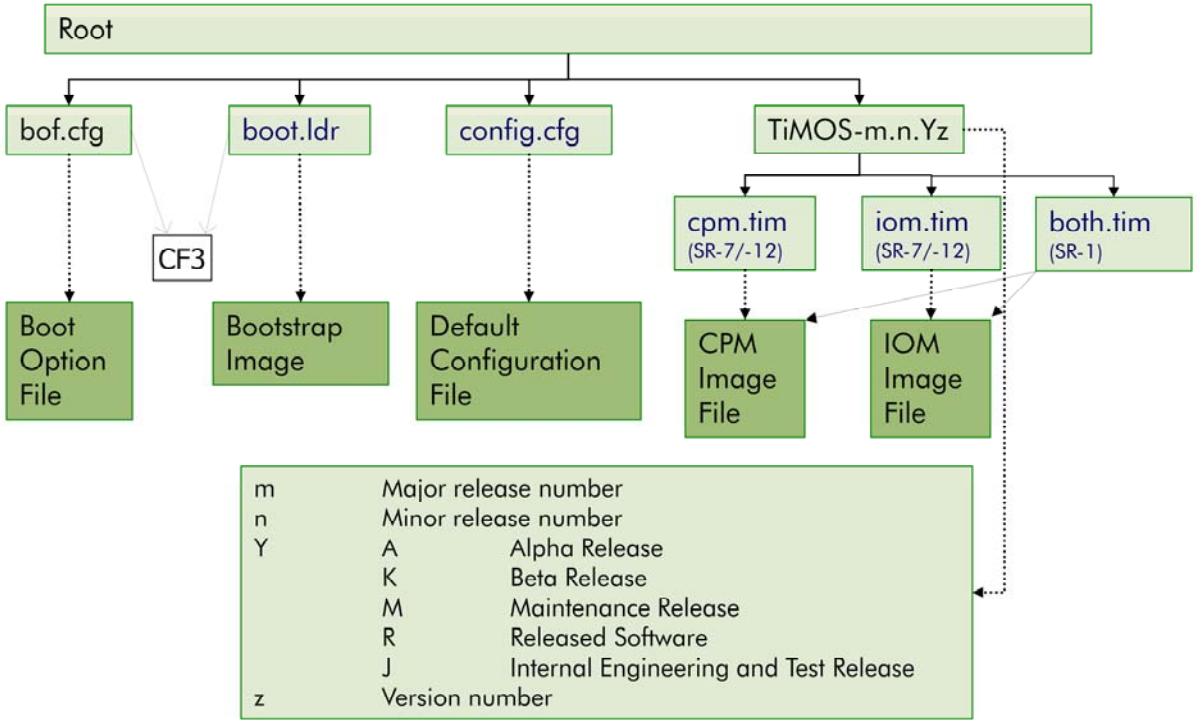


It is considered best practice to configure the log-id and syslog-id to be the same to make troubleshooting easier when needed.

File structure

Section 3: File structure

Software Release Media

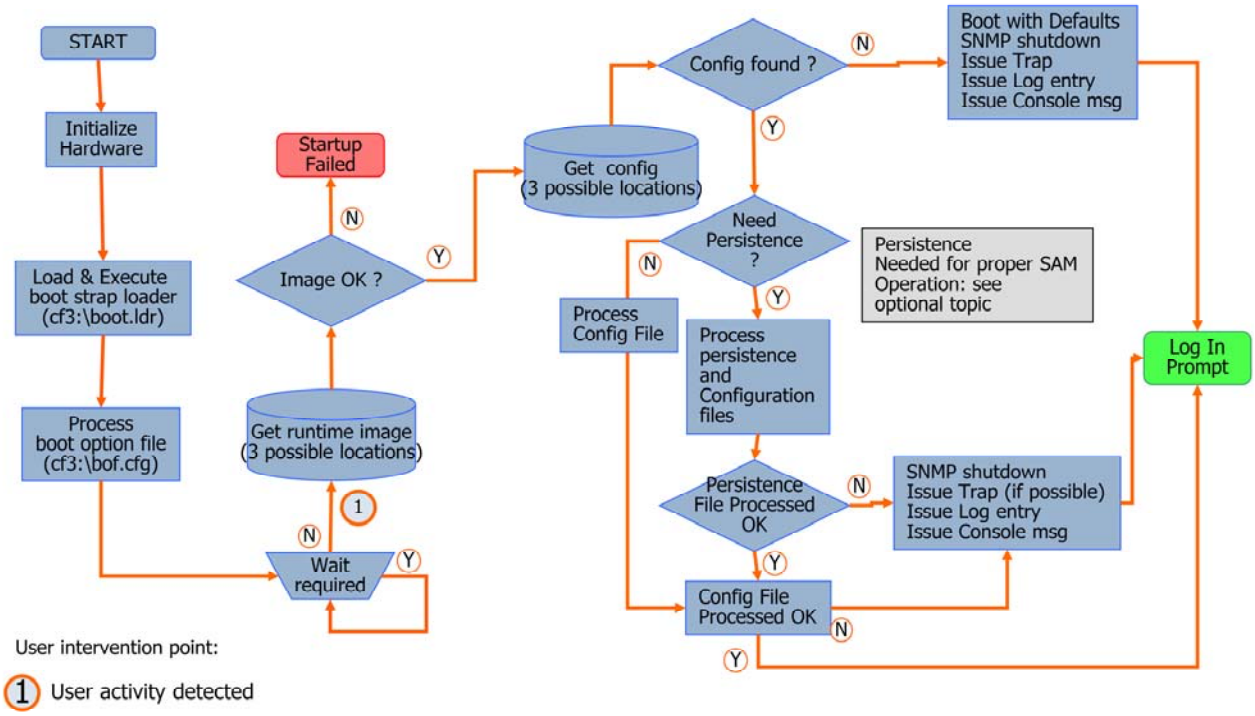


The following files are required:

- **bof.cfg:** This user-configurable file contains a set of pointers that indicate locations like the image files, the configuration files and some initial parameters like the management port IP address and the serial interface characteristics.
- **boot.ldr:** Contains the system bootstrap image file.
- **config.cfg**: This file contains the default configuration file. The default configuration file is very basic and provides just enough information to make the system operational. Other configuration files can be created later by the user.
- **TiMOS-m.n.Yz:** This is a directory named according to the major and minor software release, type of release and version. For example, if the software release is Version 6.0 of a released software version the name would be "TiMOS-6.0.R.0". On an SR-7 and SR-12, this directory contains two files, "cpm.tim" and "iom.tim", for the SF/CPM and IOM cards respectively. An SR-1 has an integrated SF/CPM and IOM, there is only one file, named "both.tim".

The initialization process requires the Boot.ldr and the Bof.cfg to be present on the Compact Flash Card 3.

System Initialization



When turning up the box, the hardware is initialized and the boot strap is loaded and executed. After this, the Boot Option File or BOF is analyzed. First the router waits before to continue. This wait time is specified in the bof file and allows the operator to intervene and change some critical parameters like the used configuration file. After the wait timer has expired, the image or software is loaded. This software could be located on three possible locations. If the image is OK, the configuration file is loaded. No configuration file found results in the log in prompt with the default settings, referred to as base config. If a configuration file is found, the need of persistence is checked. Persistence is only needed when working with the SAM. After the persistency process is finished, the configuration file is processed and the log in prompt appears.

Persistence specifies whether the system will preserve system indexes when a save command is executed. During a subsequent boot, the index file is read along with the configuration file. Indexes uniquely identify objects in the router and are used by SNMP tools to identify these objects. When configuration changes on a device are saved, a corresponding index file (.ndx extension) is created.

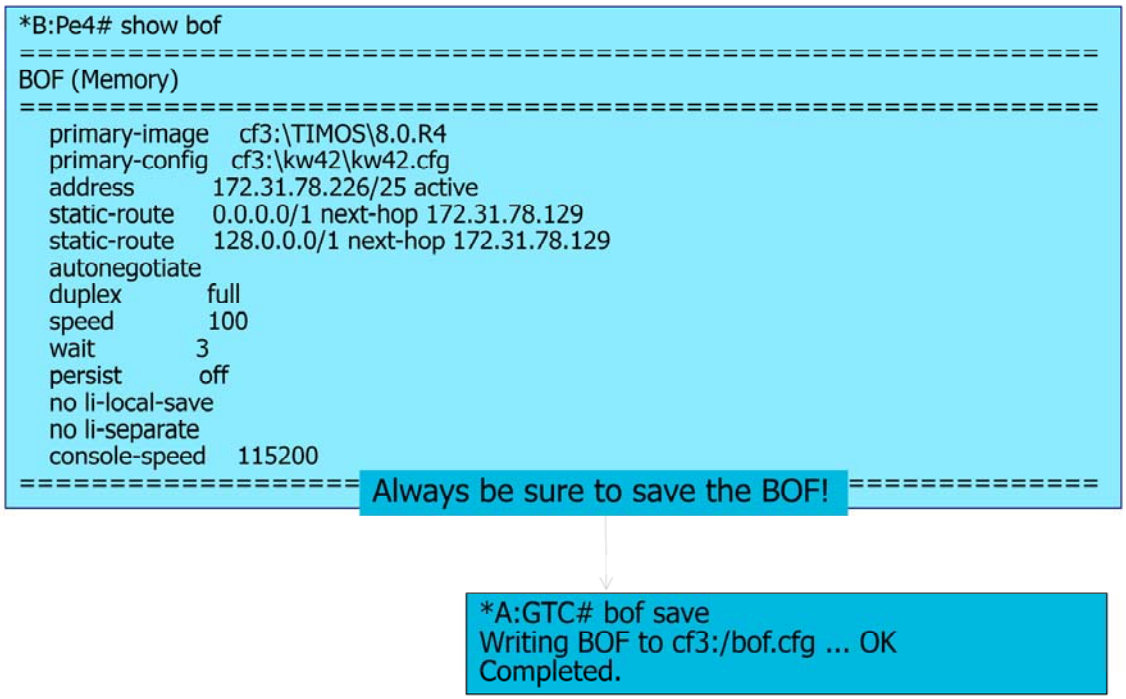
During a reboot of the network element, the configuration file is compared to the index file to ensure that indexes remain the same, thus establishing a persistent index.

System initialization

```
*A:GTC# show boot-messages
===============================================================================
cf3:/bootlog.txt
===============================================================================
Boot log started on CPU#0
  Build: X-8.0.R4 on Tue Jul 27 14:43:51 PDT 2010 by builder
  CPUCTL FPGA version: 1A
Boot rom version is v31
Booted from Control PROM 1
>>>Testing mainboard FPGA chain...
>>>Validating SDRAM from 0x7fe00000 to 0x80000000
>>>Testing SDRAM from 0x02200000 to 0x7fe00000
>>>Testing Compact Flash 1... Slot Empty
>>>Testing Compact Flash 2... Slot Empty
>>>Testing Compact Flash 3... OK (SILICONSYSTEMS INC 256MB)
Peripheral FPGA version is 0x14
Chassis Serial Number is 'NS065063158'
Board Serial Number is 'NS064940141'
Searching for boot.ldr on local drives:
Searching cf3 for boot.ldr...
*********************************************

Total Memory: 2GB  Chassis Type: sr7  Card Type: belarus_r1_200G
TiMOS-L-8.0.R4 boot/hops ALCATEL ? 7xxx Copyright (c) 2000-2010 Alcatel-Lucent.
All rights reserved. All use subject to applicable license agreements.
Built on Tue Jul 27 14:49:32 PDT 2010 by builder in /rel8.0/b1/R4/panos/main

TiMOS BOOT LOADER
```

The complete boot process can be followed in real time when connected to the console interface. However this process is recorded and can be watched later on with command “show boot-messages”. The start-up is the job of the administrator, but the operator gets an overview of all the critical set parameters.

Managing The BOF

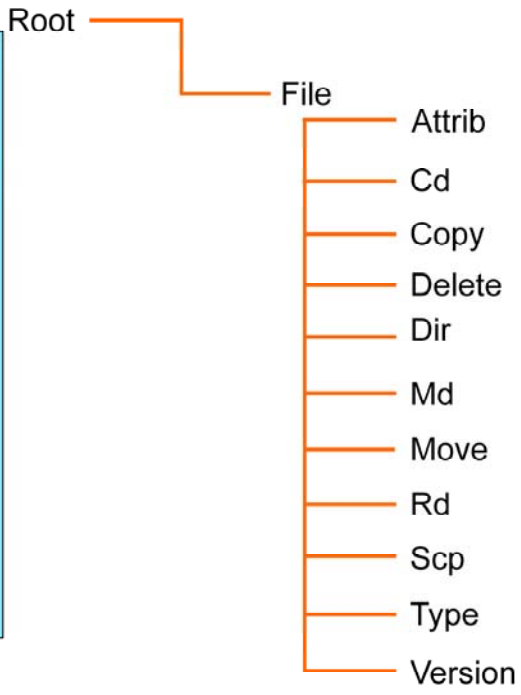


The bof file is an important file. It is a text file which contains the references to the images files and configuration files. It contains the static-routes, the management IP address, DNS domain names, CPM Ethernet port settings, persistent requirements, lawful intercept options and console ports settings. The bof is also loaded into the memory, so any change needs to be saved so that after a reboot the changes are taken into account.

Node# bof	Enter the CLI BOF context to change or create a bof file.
Node>bof# address 10.10.10.2/24 primary	Change or create the CPM Ethernet Port IP address.
Node>bof# speed 100	Set the CPM Ethernet Port speed to 100 Mbps.
Node>bof# primary-image cf3:/TIMOS.5.0.R0	Set the primary image directory.
Node>bof# primary-config cf3:/test.cfg	Set the primary configuration file (eg. test.cfg).
Node>bof# save	Save the bof.
Node# show bof	Display the in-memory bof file (last used).

File System CLI Context

```
*A:GTC# file dir

 Volume in drive cf3 on slot A has no label.

 Volume in drive cf3 on slot A is formatted as FAT32.

Directory of cf3:\

05/01/2012  02:51p                969 bof.cfg
06/21/2011  01:29p                0 SR7_4
07/24/2009  02:37p    <DIR>       Erik
03/09/2011  03:34p    <DIR>       Geert
12/03/2008  08:17a    <DIR>       Steven
12/12/2010  05:57p                boot.ldr
12/13/2010  01:42p    <DIR>       aat
07/15/2008  01:04p    <DIR>       SRC
04/15/2009  10:45p    <DIR>       Erdinc
04/18/2012  02:04p                4312 bootlog.txt
04/18/2012  02:02p                969 bof.cfg.1
12/28/2010  01:01a                base.cfg
04/18/2012  01:49p                bof.cfg.2
04/18/2012  12:42p                bootlog_prev.txt
08/12/2010  05:11p    <DIR>       TiMOS
04/18/2012  12:40p                bof.cfg.3
04/03/2012  01:42p                bof.cfg.4
04/03/2012  01:14p                bof.cfg.5
*A:GTC#
```



The files of CF3 can be seen by executing the "file dir" command. All known DOS commands can be used to copy, delete, move,… the files.

Node>file cf3:\ # delete	Delete the specified file. The optional wildcard (*) can be used to delete multiple files that share a common (partial) prefix and/or (partial) suffix.
Node>file cf3:\ # move	Move a local file, system file, or a directory. If the target already exists, the command fails and an error message displays.
Node>file cf3:\ # scp	Copy a file from the local files system to a remote host on the network. scp uses ssh for the data transfer, and uses the same authentication and security as ssh.
Node>file cf3:\ # type	Display the content of a text file.
Node>file cf3:\ # version	Display the version of an OS "cpm.tim" or "iom.tim" file.

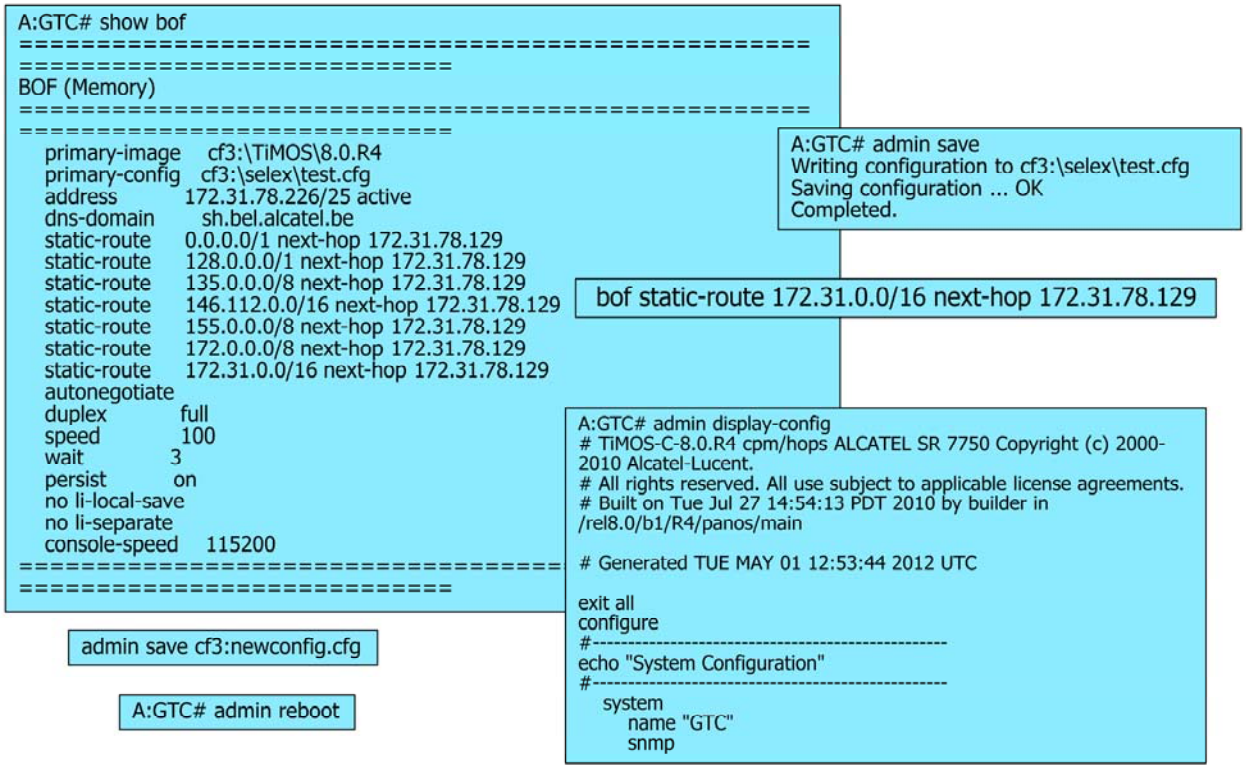
Basic Configuration

AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 4: Basic CLI Configuration

The config



The "admin save" commands saves the running configuration to the configuration file specified in the bof file.

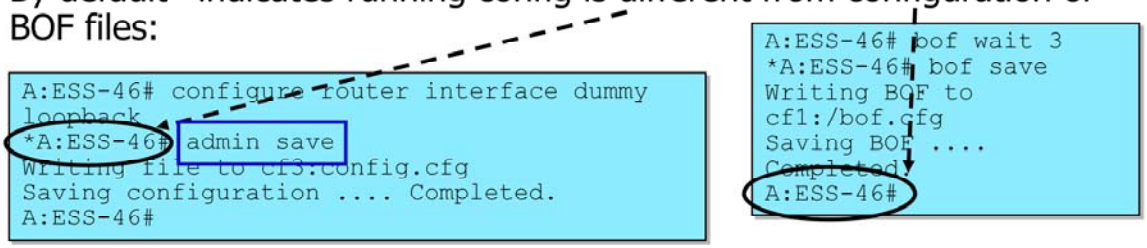
Users that connect from the service router, to the ethernet management port, will use a certain source IP address. A Telnet or SSH session is a bi-directional communication. So the service router needs a route back to this source address. The route back is specified in the static-route configuration of the bof. A dedicated routing table is maintained for management IP traffic.

To show the whole configuration, type "admin display-config".

Saving the configuration is performed with "admin save".

Running or saved config

By default* indicates running config is different from configuration or BOF files:



An environment variable allows you to toggle between the default behavior and no unsaved changes indication:

```
A:ESS-46# environment saved-ind-prompt ?
  - no saved-ind-prompt
  - saved-ind-prompt    (default)
```

The "*" before the node name indicates that the running configuration is different from the configuration file. This means that changes have been made to the running configuration without saving them to the actual configuration file. The command "Admin save" will remove the "*" indicator. If this notification is not wanted, "no saved-ind-prompt" command will remove this.

Last operating change

A new “last operating state change” timestamp is added in the show router interface detail display:

```
A:SR-221# show router interface system detail

===============================================================================
Interface Table (Router: Base)
===============================================================================

-------------------------------------------------------------------------------
Interface
-------------------------------------------------------------------------------
If Name      : system
Admin State  : Up                       Oper (v4/v6)     : Up/Down
Protocols    : OSPFv2 ISIS MPLS RSVP PIM
IP Addr/mask : 10.0.0.221/32            Address Type     : Primary
IGP Inhibit  : Disabled                 Broadcast Address: Host-ones
-------------------------------------------------------------------------------
Details
-------------------------------------------------------------------------------
If Index     : 1                        Virt. If Index   : 1
Last Oper Chg: 01/03/2007 13:52:05      Global If Index  : 128
Port Id      : system
TOS Marking  : Trusted                  If Type          : Network
...
===============================================================================
```

When operators are changing parameters, a last operator change state can now be tracked with the added timestamp.

Config system login-control

```
A:PE1 config system
            login-control
                pre-login-message "This is the pre-login-message"
                motd text "This is the message of the day - motd "
                no login-banner
             exit
```

```
This is the pre-login-message

Login: admin
Password:

This is the message of the day - motd

A:PE1#
```

Routers can change their pre-login message whenever an operator logs in.

Absolute path

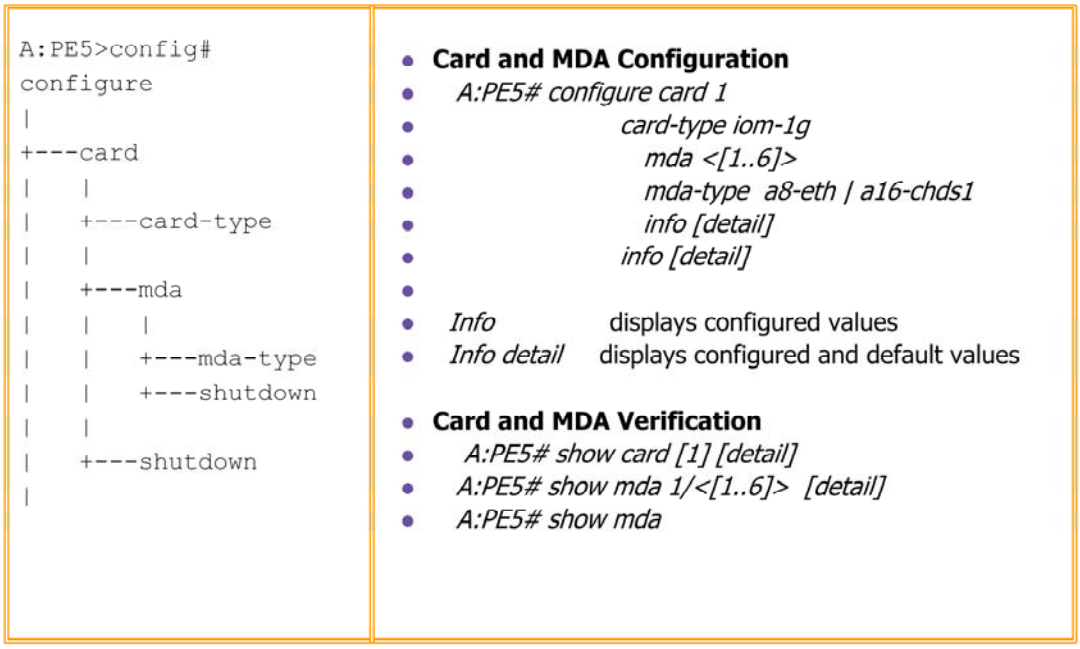
Execute an out-of-context command enter:

- forward slash `/` or
- backward slash `\`

```
A:SR-221>config>service# /admin save
Writing file to cf3:/config.cfg
Saving configuration .... Completed.
A:SR-221>config>service# /clear port 1/1/1 statistics
A:SR-221>config>service#
```

Absolute path CLI commands can now be executed out of context by specifying the full path from the CLI root. To execute an out-of-context command enter a forward slash `/` or backward slash `\` at the beginning of the command line. The commands are then interpreted as absolute path. Spaces between the slash and the first command will return an error.

MDA Configuration Verification



To configure a card and MDA, the configuration and shows commands can be used as mentioned on this slide.

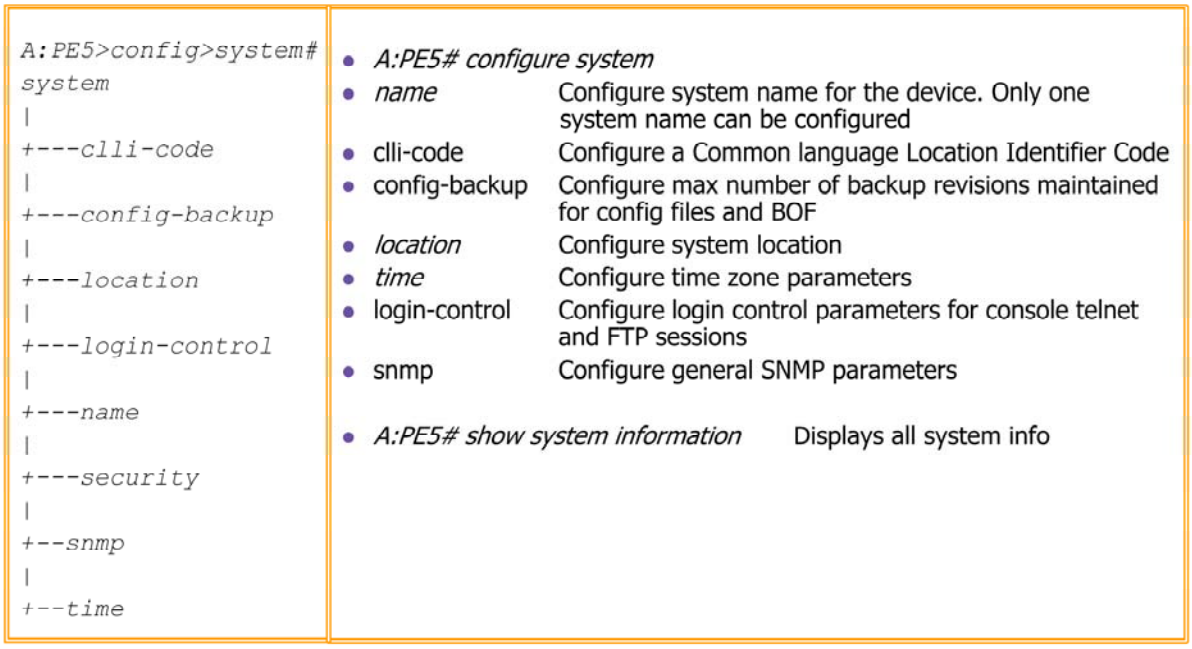
Card and MDA Configuration Example

```
Configure Card and MDAs
A:PE5# configure card 1
A:PE5>config>card# card-type iom-sar
A:PE5>config>card# mda 1
A:PE5>config>card>mda# mda-type a8-eth
A:PE5>config>card>mda# exit
A:PE5>config>card# mda 2
A:PE5>config>card>mda# mda-type a16-chds1

A:PE5# show card
===============================================================================
Card State
===============================================================================
Slot/  Provisioned   Equipped     Admin   Operational     Num          Num      Comments
Id     Type          Type         State   State           Ports        MDA
-------------------------------------------------------------------------------
1      iom-sar       iom-sar      up      up                           6
1/1    a8-eth        a8-eth       up      up              8
1/2    a16-chds1     a16-chds1    up      up              16
A      csm-1g        csm-1g       up      down                                   Standby
B      csm-1g        csm-1g       up      up                                     Active
===============================================================================
```

This is an example of a configuration of an IOM of type iom-1g with two mda's.

The "show card" commands show the card state and the state of the MDA's.

System parameters Configuration and Verification



This the CLI tree structure of the main system configurations of a service router. It configures the name, CLLI code, the maximum number of backup revisions maintained for the config files and bof, the system location, time zone, log-in control parameters and general snmp parameters.

"show system information" displays all system info.

System Configuration Example

Configure the system name
- A:PE5# configure system
- A:PE5>config>system# name "PE5"

Configure idle timeout for console telnet and FTP sessions
- A:PE5>config>system# login-control idle-timeout 1440

Configure time zone
- A:PE5>config>system# time zone UTC

Enable the SNMP daemon
- A:PE5>config>system# snmp no shutdown

Four basic configuration examples.

How to configure the system name.

How to increase the login idle timeout timer to 1440 seconds.

How to set the time zone.

How to issue a snmp no shutdown.

System Configuration output

A:PE5# show system information

```
===============================================
============
System Information
===============================================
============
System Name              : PE5
System Type              : 7705 SAR-8
System Version           : B-0.0.I321
System Contact           :
System Location          :
System Coordinates       :
System Active Slot       : B
System Up Time           : 2 days, 00:29:13.06 (hr:min:sec)
SNMP Port                : 161
SNMP Engine ID           : 0000197f00007cedff000000
SNMP Max Message Size    : 9216
SNMP Admin State         : Enabled
SNMP Oper State          : enabled
SNMP Index Boot Status   : Persistent
SNMP  Sync State         : OK
Telnet/SSH/FTP Admin     : Enabled/Enabled/Disabled
Telnet/SSH/FTP Oper      : Up/Up/Down
BOF Source               : cf3:
Image Source             : primary
Config Source            : primary
Last Booted Config File  : cf3:\config.cfg
Last Boot Cfg Version    : TUE MAR 18 23:24:55 2008 UTC
Last Boot Config Header  : # TiMOS-B-0.0.I321 both/hops ALCATEL SAR 7705
#
          Copyright (c) 2000-2008 Alcatel-Lucent. All rights reserved.
           # Built on Mon Feb 11 01:15:45 EST 2008 by
          csabuild in /rel0.0/I321/panos/main # Generated TUE
          MAR 18 23:24:55 2008 UTC
Last Boot Index Version  : TUE MAR 18 23:24:55 2008 UTC
Last Boot Index Header   : # TiMOS-B-0.0.I321 both/hops
      ALCATEL SAR 7705  #
      Copyright (c) 2000-2008 Alcatel-Lucent. All rights
      reserved. # All use subject to applicable license
       agreements. # Built on Mon Feb 11 01:15:45 EST 2008
by
       csabuild in /rel0.0/I321/panos/main # Generated TUE
       MAR 18 23:24:55 2008 UTC
Last Saved Config        : cf3:\config.cfg
Time Last Saved          : 2008/03/19 09:49:52
Changes Since Last Save  : No
Max Cfg/BOF Backup Rev : 5
Cfg-OK Script            : N/A
Cfg-OK Script Status     : not used
Cfg-Fail Script          : N/A
Cfg-Fail Script Status   : not used

Management IP Addr       : 138.120.224.31/24
Primary DNS Server       : 138.120.252.56
Secondary DNS Server     : N/A
Tertiary DNS Server      : N/A
DNS Domain               : ca.alcatel.com
BOF Static Routes        :
  To                   Next Hop
  138.120.0.0/16       138.120.224.1
ATM Location ID          :
01:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00
===================================
=================
```

All the major system information settings are shown by the "show system information" command.

vi Editor+Rollback

Section 5: vi editor and rollback

SROS vi editor

SROS has a full screen text editor used to:

- Create and modify any ASCII files on the CF cards of your 7x50

The integrated text editor is a simplified version of the comprehensive vi editor (well-known in most UNIX systems)

Example application:

Edit your config.cfg file without any need to transfer the file via ftp/scp to a desktop PC or server

The SROS has a full screen editor to create and modify any ASCII files on the compact flash of the 7x50 service routers.

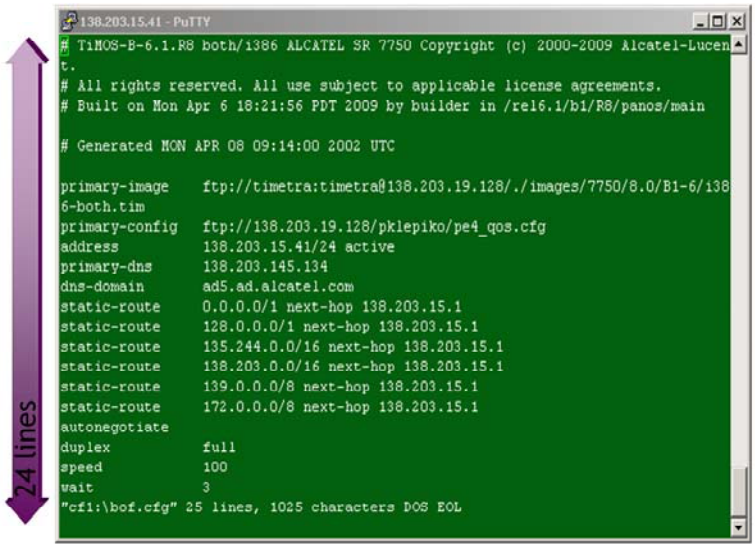
The integrated text editor is a simplified version of the comprehensive vi editor.

The configuration file can now be edited on the service router without sending it to a desktop PC or server.

SROS run vi editor from CLI

Run vi editor:

- VT100 terminal mode needs to be supported
- Editor's resolution is fixed to 24 lines with 80 characters
- Only files on the local CF card are applicable for editing (no remote URLs)




```
# file vi ?
  - vi <local-url>

 <local-url>  : [<cflash-id>/]<file-path>
                cflash-id      - cf1:|cf2:|cf3:
```

To run the vi editor vt100 terminal mode needs to be supported.

The editor's resolution is fixed to 24 lines with 80 characters.

Only files on the local CF card are applicable for editing.

Why rollback?

Purpose of rollback functionality is to undo changes without:

- reverting everything manually (shutdown xxx, no xxx, back, ...)
- rebooting the whole node (maybe after editing the config)

"Checkpoints" can be set whenever you feel that one might be needed later on to:

- Return to these checkpoints with a single command
- Useful in case of accidental misconfigurations or for lab environments

Not within scope of rollback: CLI transactions

- CLI commands become active immediately
- When <Enter> key is depressed, it happens...

One of the latest CLI features is the rollback functionality.

This rollback allows you to undo changes to the configuration without reverting everything manually or rebooting the whole node.

The point where you return to is called the checkpoint. Checkpoints can be set whenever you feel you might possibly need one later on.

A single command returns to a particular checkpoint.

This is extremely useful in case of accidental misconfigurations or for lab environments.

Note that CLI commands that are typed also take effect immediately.

How does it work?

You need to enable rollback first by configuring it

- configure a location for the rollback files

At suitable times, you create a rollback checkpoint

- the system creates sort of a running config file in the background
- this includes all settings made via CLI and SNMP
- does **not** include BOF, LI and SNMP index settings

When necessary, you can return to any of your existing checkpoints

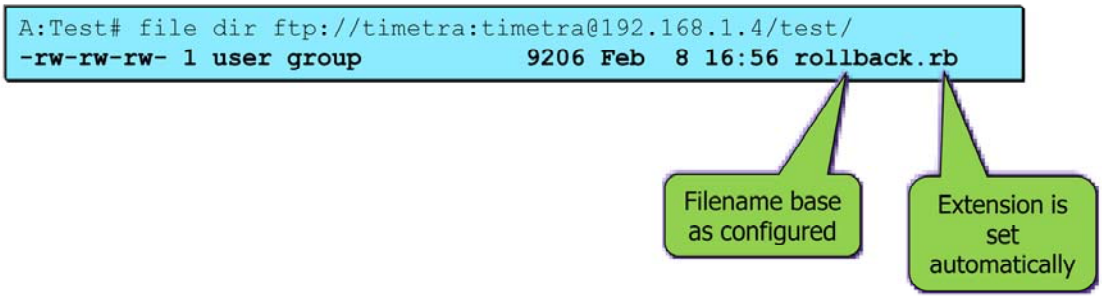
Before rollback is configured, a location to store the rollback files has to be set.

Whenever a configuration is found stable, a checkpoint can be set. The system creates a sort of a running configuration file in the background. This file includes all settings made via CLI and SNMP. But it doesn't include BOF, LI and SNMP index settings.

When necessary, you can return to any of the set checkpoints.

Step-by-step: What happened inside?

A new rollback file is created:

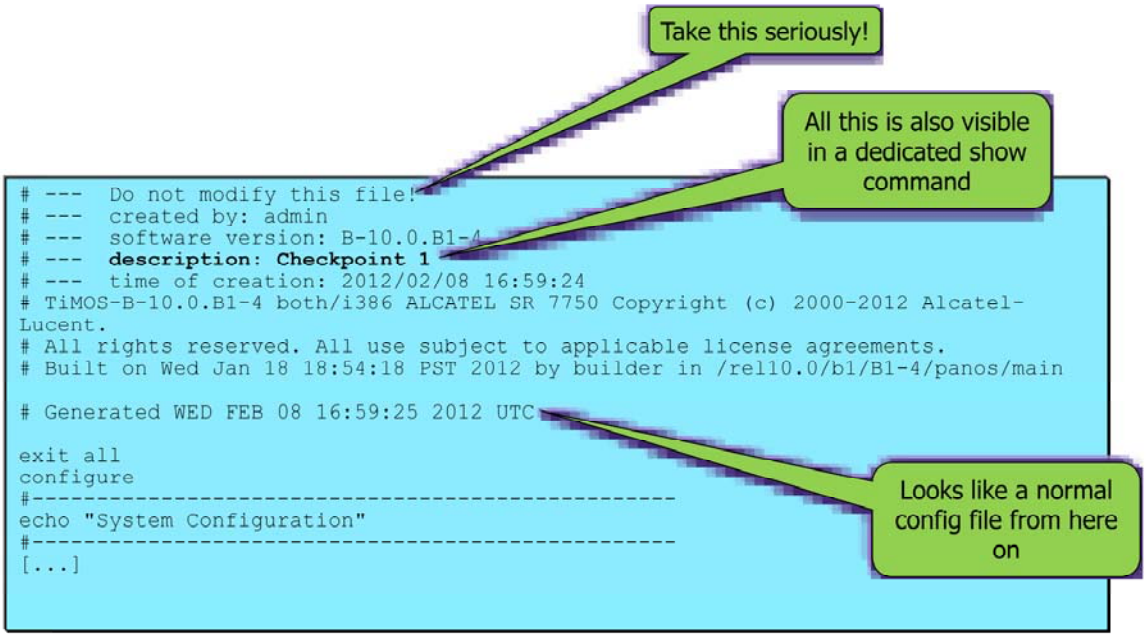


The system rotates the existing rollback files automatically to make space

- Amount of checkpoints to be maintained is configurable (*shown later in course*)

A new rollback file is created on the configured location. The system rotates the existing rollback files automatically. Whenever there is a new checkpoint created, it will be labeled "rollback.rb" in this case and the older one gets "rollback.rb.1" as name. The file without a number as an extension will be the latest checkpoint.

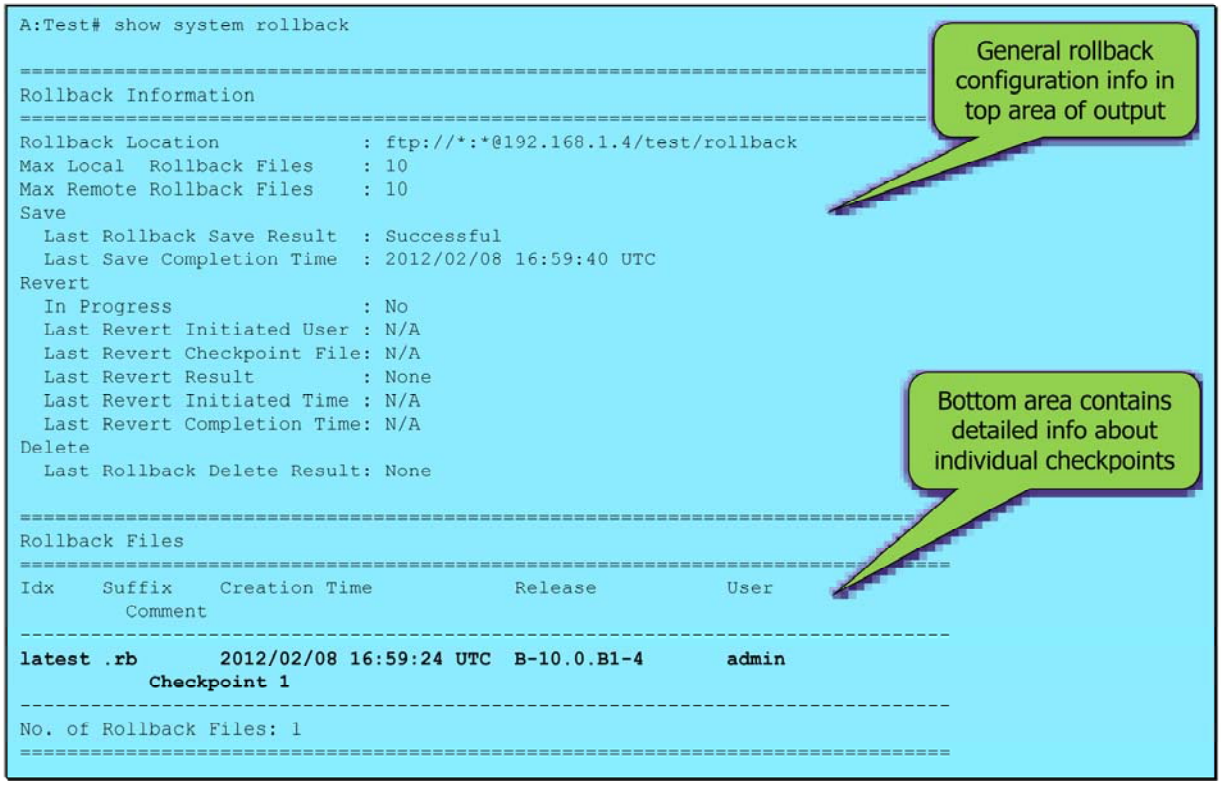
Step-by-step: Rollback of contents



A rollback file is created by the service router and should never be changed manually. The file can be viewed by entering a dedicated show command and except for the header, looks like a normal configuration file.

Step-by-step: CLI view

Checkpoint is now visible in dedicated CLI command:



The check point is visible with the CLI command "show system rollback. After the general information, all info can be found about the individual checkpoints.

Step-by-step: Change and create another checkpoint

Now configure something, and create another checkpoint:

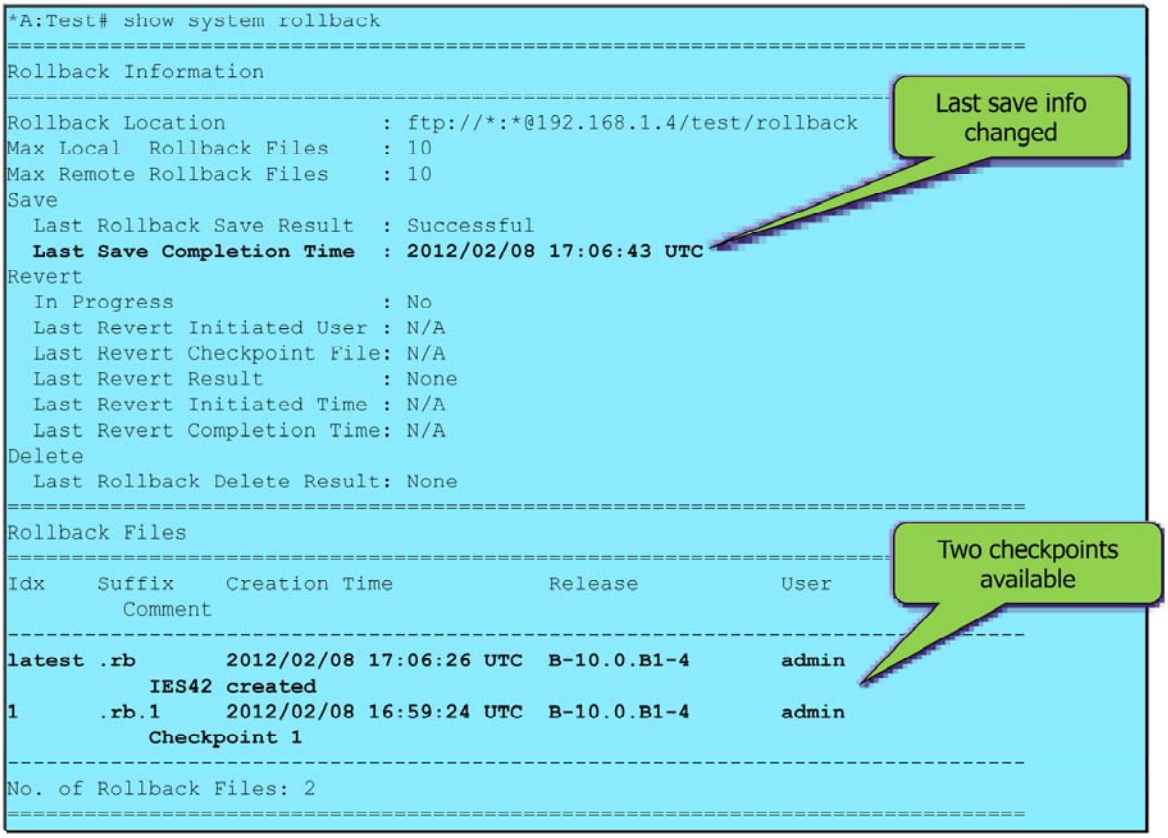
```
A:Test# configure service ies 42 customer 1 create
[...]
*A:Test>config>service>ies# info
----------------------------------------------
            shutdown
            interface "T42" create
                address 42.42.42.42/30
                sap 1/1/2 create
                exit
            exit
----------------------------------------------
[...]
*A:Test# admin rollback save comment "IES42 created"
Saving rollback configuration to ftp://*:*@192.168.1.4/test/rollback.rb ...
OK

*A:Test# file dir ftp://timetra:timetra@192.168.1.4/test/
-rw-rw-rw- 1 user group           9550 Feb  8 17:03 rollback.rb
-rw-rw-rw- 1 user group           9206 Feb  8 16:56 rollback.rb.1
```

Rollback files are rotated automatically

If changes are made to the configuration , a checkpoint with new content can be made

This example shows a creation of an IES service 42.

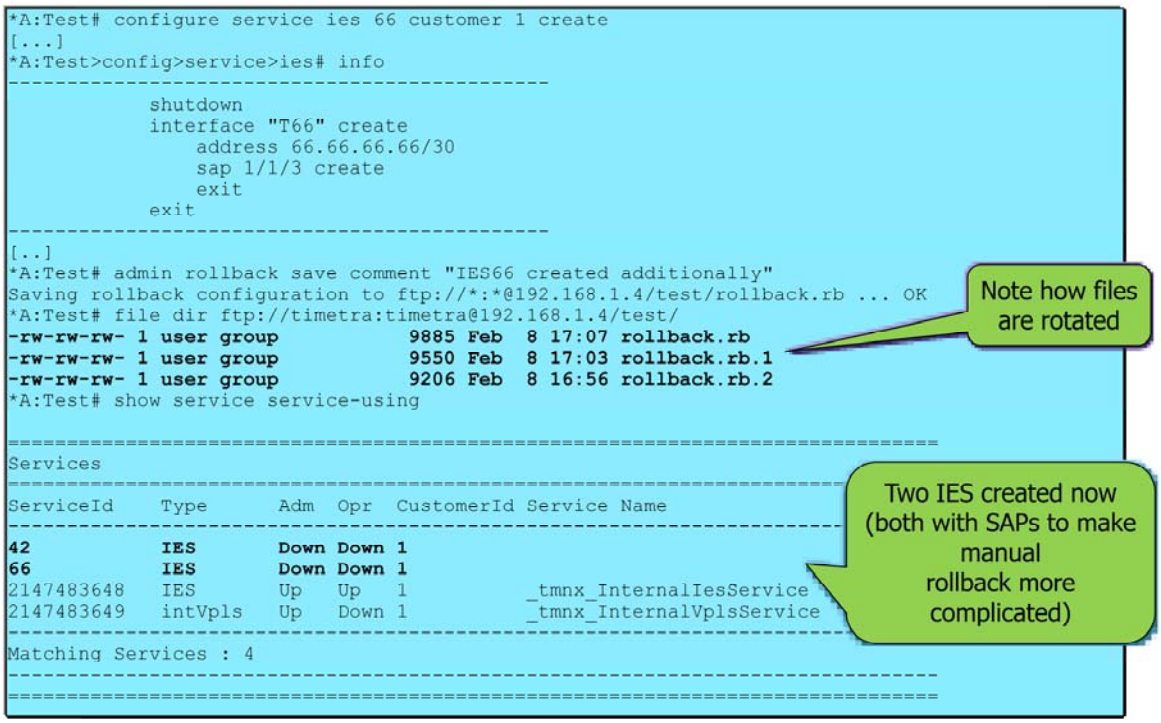
Step-by-step: CLI view with new Checkpoint



Two checkpoints are now available. The latest one created has been added to the list.

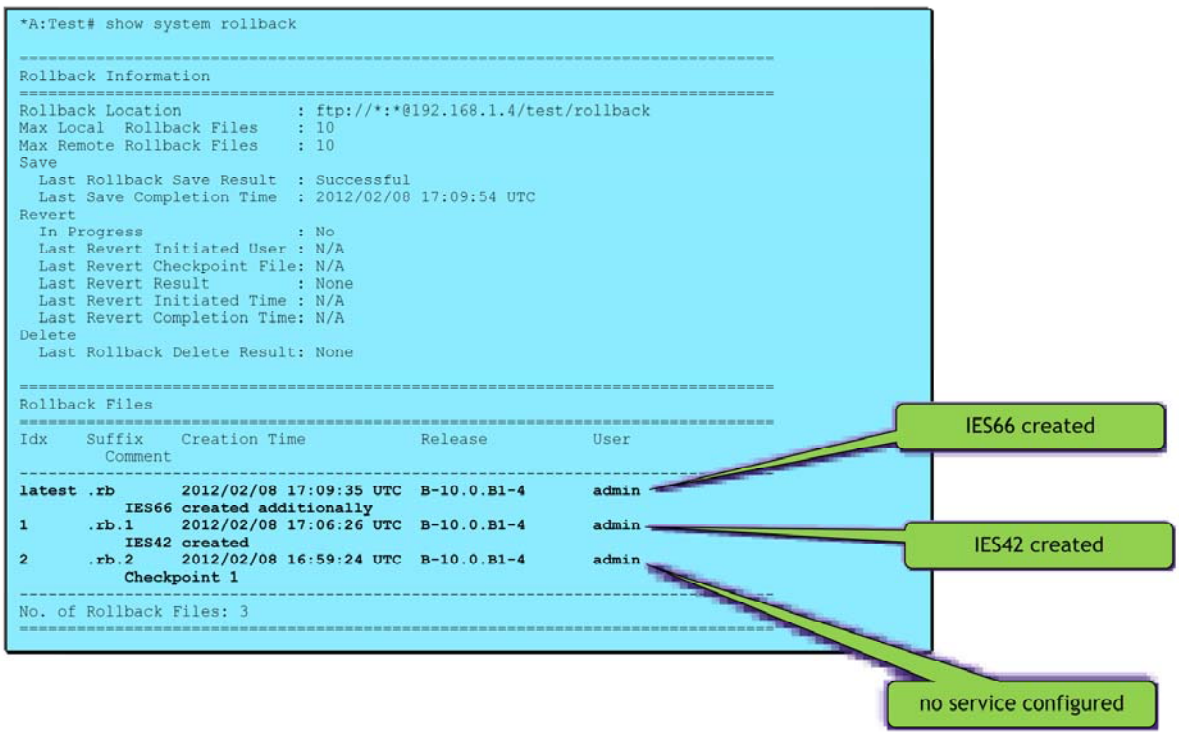
Step-by-step: Create third checkpoint

Again configure something, and create the third checkpoint:



After a third change a new checkpoint is created. A new IES has been added to the configuration.

Step-by-step: CLI view of our Playground



So in total three (3) checkpoints have been created.

Step-by-step: Comparing checkpoints

The CLI has a special command to check which changes happened between two checkpoints.

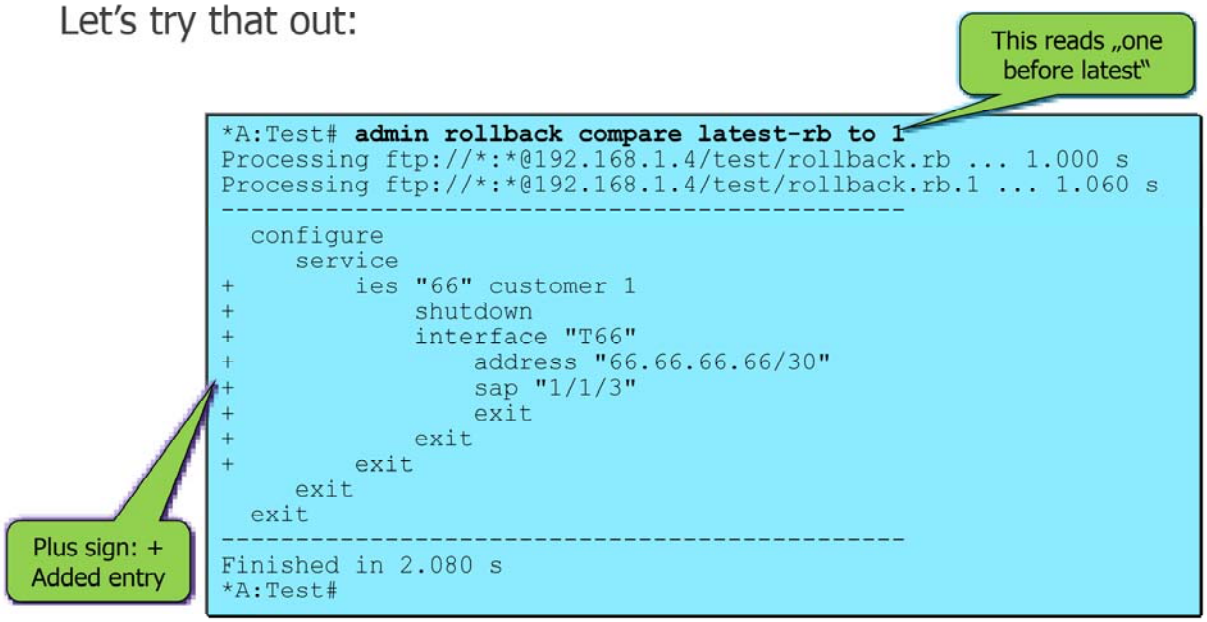
```
*A:Test# admin rollback compare
  - compare [to <checkpoint2>]
  - compare <checkpoint1> to <checkpoint2>

 <checkpoint1>         : active-cfg|rescue|latest-rb|<checkpoint-id>
                           active-cfg      - Current running config (Default)
                           rescue          - Rescue config file
                           latest-rb       - Most recent checkpoint file
                           checkpoint-id   - [1-10] - Checkpoint files
 <checkpoint2>         : active-cfg|rescue|latest-rb|<checkpoint-id>
                           active-cfg      - Current running config
                           rescue          - Rescue config file
                           latest-rb       - Most recent checkpoint file
(Default)
                           checkpoint-id   - [1-10] - Checkpoint files
```

Note: You cannot use the comment fields to refer to a specific checkpoint – you must always give its numeric index. Do a "*show system rollback*" beforehand when you belong to those individuals to whom names mean more than a number.

But what is the difference between two checkpoints? A comparison between checkpoints can be seen by typing "admin rollback compare". Note that the ID of the checkpoint is used and not the description. The relation between the checkpoint ID and the description can be seen by "show system rollback".

Step-by-step: Comparing checkpoints

Let's try that out:



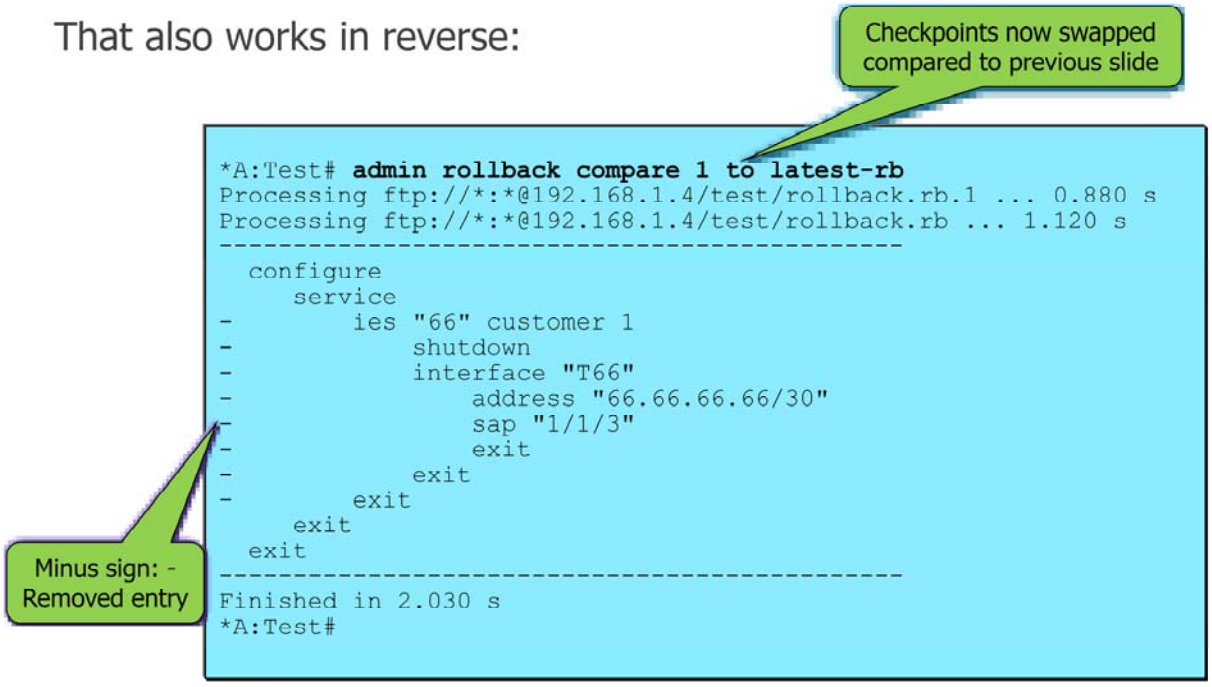
Between the last checkpoint ("1") and the one we just created ("latest"), only IES 66 was added.

Let us compare the two checkpoints. A keyword latest-rb can be given. This addresses the latest created checkpoint. "1" is the one before the latest.

The difference is the ies 66 service as expected. An added plus sign identifies an added line of configuration.

Step by Step: Comparing checkpoints

That also works in reverse:



We could also make a comparison between an older and the latest one. This results in minus signs, removed entries.

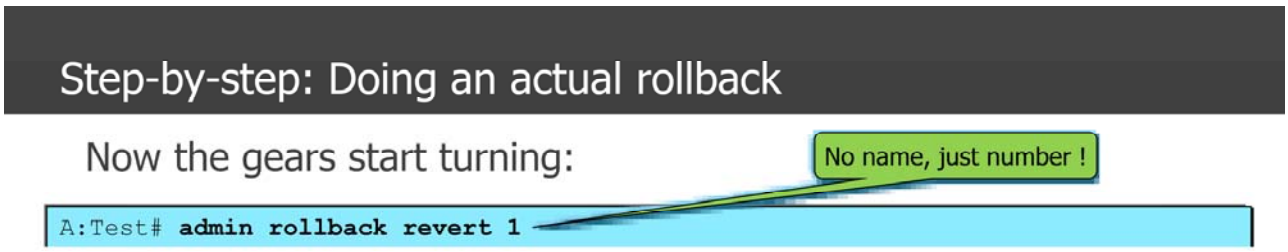
Step-by-step: Doing an actual rollback

So this was it before we try a rollback:

```
A:Test# show system rollback

===============================================================================
Rollback Information
===============================================================================
[...]
===============================================================================
Rollback Files
===============================================================================
Idx    Suffix    Creation Time            Release          User
         Comment
-------------------------------------------------------------------------------
latest .rb       2012/02/08 17:09:35 UTC  B-10.0.B1-4      admin
           IES66 created additionally
1      .rb.1     2012/02/08 17:06:26 UTC  B-10.0.B1-4      admin
           IES42 created
2      .rb.2     2012/02/08 16:59:24 UTC  B-10.0.B1-4      admin
           Checkpoint 1
-------------------------------------------------------------------------------
No. of Rollback Files: 3
===============================================================================
A:Test# show service service-using

===============================================================================
Services
===============================================================================
ServiceId    Type      Adm  Opr  CustomerId Service Name
-------------------------------------------------------------------------------
42           IES       Down Down 1
66           IES       Down Down 1
2147483648   IES       Up   Up   1          _tmnx_InternalIesService
2147483649   intVpls   Up   Down 1          _tmnx_InternalVplsService
-------------------------------------------------------------------------------
Matching Services : 4
-------------------------------------------------------------------------------
===============================================================================
A:Test#
```

Before any rollback, the latest configuration contains now two IES services, IES 42 and IES 66.

"Show service service-using" is the command to check the services running on a node.



Let's revert to the one older checkpoint. Perform an "admin rollback revert 1".

After the processing, it shows the lack of IES 66 as expected. This was a successful rollback.

Rolling back a rollback

Rollback is just called "**rollback**", should not be taken literally.
- Use the term "revert" in any discussions, as this has less of a reference to the actual timing of checkpoints

It actually rolls forward as well between checkpoints when required.

Rollback checkpoints are not deleted when reverting.
- Not the one you are reverting to, nor
- Any checkpoints that have been taken/occurred in between

"Revert from a revert" is also possible:
- To create a checkpoint just before reverting
- Leaves option to come back to where you already were, in case revert is worse in its actual outcome

The command rollback might create the perception that there can only be reverted back in time. However, also a roll forward can be possible in time.

Rollback checkpoints don't get deleted when reverting, neither the one you are reverting to, nor any checkpoints that have been taken in between.

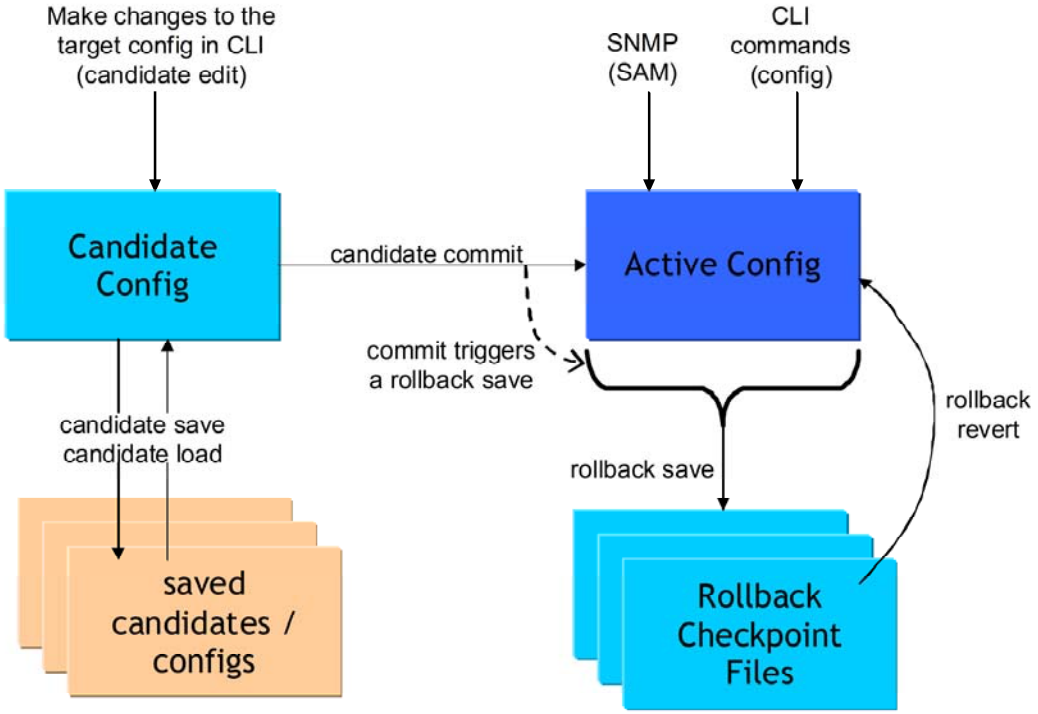
So a "revert from a revert" is also possible. You might want to create a checkpoint just before reverting.

That leaves you the option to backup to where you already were, in case the revert that was performed turns out to be worse in its actual outcome.

Transactional Configuration

Section 6: Transactional Configuration

Transactional Configuration Model



Transactional Configuration and Configuration Rollback, together, provide the operational model as shown here.

Transactional configuration allows a user to generate a candidate configuration.

The user enters commands as normal in CLI. This candidate configuration could then be saved to be executed at another time or committed to execute the candidate configuration immediately.

During the creation of a candidate configuration, syntax will be verified and confirmed; however, resource allocation will not. If a candidate configuration will cause the exhaustion of a system resource, memory for example, this will not become apparent until the candidate configuration is executed.

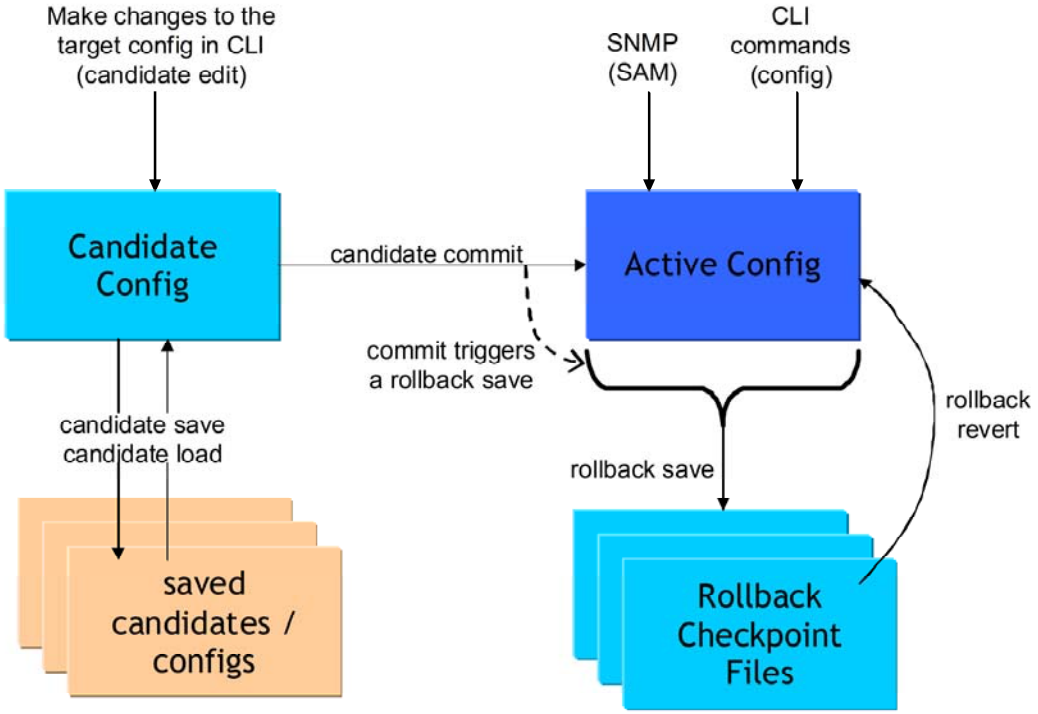
A new set of commands is provided for this functionality under the global "candidate" command. Many candidate commands are only visible once the operator is in edit-cfg [edit-c-f-g] mode by typing "candidate edit".

The operator can enter and quit the configuration mode as many times as desired, before finally committing the candidate.

Caution should be exercised if multiple users are concurrently modifying a candidate configuration. If a second user were to commit before the first user completed configuration changes, all configuration from both users, up to that point, would be loaded into the router. This situation can be avoided by including exclusive in the candidate command string, i.e., "candidate edit exclusive".

Once the candidate configuration is complete, the operator can explicitly commit the changes and cause the entire new configuration to become active.

Transactional Configuration Model [cont'd]



In edit-cfg mode, the operator builds a set of candidate configuration changes using the same CLI tree as standard (line-by-line non-transactional) configuration. Tab completion and keyword syntax checking is available.

If a commit operation is successful, then all of the candidate changes will take operational effect and the candidate is cleared. If there is an error in the processing of the commit, or a 'commit confirmed' is not confirmed and an auto-revert occurs, then the router will return to a configuration state with none of the candidate changes applied. The operator can then continue editing the candidate and try a commit later.

All commands in the candidate configuration must be in the correct order for a commit to be successful. A set of candidate editing commands (for example, copy, insert) are available to correct and reorder the candidate configuration.

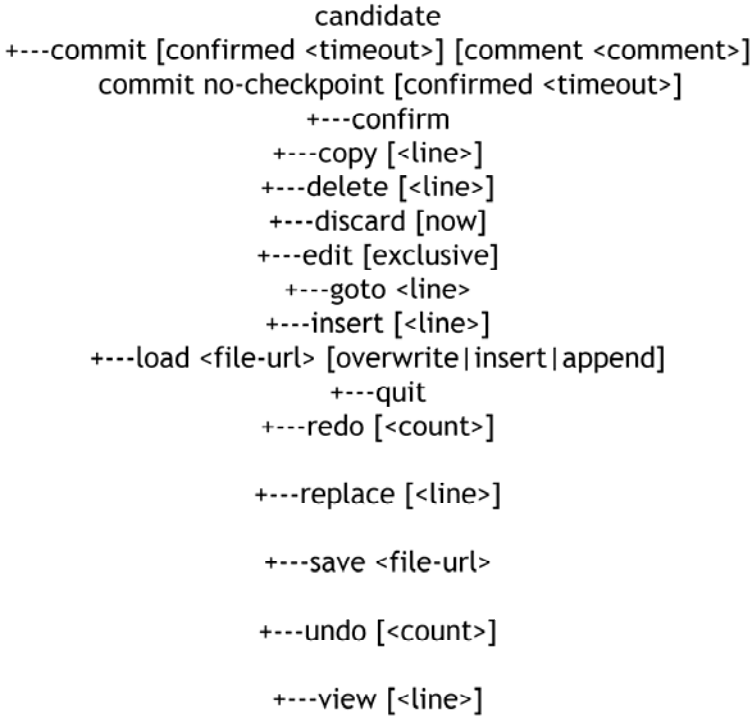
There is no SNMP [S-N-M-P] access to the candidate configuration and no SNMP management of candidates, although any configuration changes done via a transaction are reported via the standard SR-OS [S-R-O-S] SNMP change traps, and basic candidate status information is available via SNMP.

Standard line-by-line (immediate operational effect upon pushing the enter/return key) non-transactional CLI and SNMP commands are not blocked during the creation/editing of a candidate or the processing of a commit. These commands take immediate effect as normal.

If a command causes an error or resource exhaustion, the router will be returned (using rollback) to the state it was before the execution of the candidate configuration.

By default, SR-OS will automatically create a new rollback checkpoint after a commit operation. The rollback checkpoint will include the new configuration changes made by the commit. An optional "no-checkpoint" keyword can be used to avoid the auto-creation of a rollback checkpoint after a commit. If the commit fails, then no new rollback checkpoint is created.

Transactional Configuration CLI Tree



The CLI tree for Transactional Configuration is shown here.

Transactional Configuration Example

```
A:196# candidate edit exclusive

INFO: CLI Entering exclusive mode. Leaving the candidate config without
committing will discard the content of your candidate config.

A:196>edit-cfg#
A:196>edit-cfg# configure system name PE1
A:196>edit-cfg# configure card 1 card-type "iom3-xp"
A:196>edit-cfg# configure card 1 mda 1 mda-type m20-1gb-xp-tx
A:196>edit-cfg# configure router interface "system" address 10.10.10.1/32
A:196>edit-cfg# configure router interface toPE2
A:196>edit-cfg>config>router>if# address  10.1.2.1/30
A:196>edit-cfg>config>router>if# port 1/1/1
A:196>edit-cfg>config>router>if# no shutdown
A:196>edit-cfg>config>router>if# exit all
A:196>edit-cfg# configure port 1/1/1 no shutdown
A:196>edit-cfg# configure port 1/1/2 shutdown
A:196>edit-cfg# configure port 1/1/2 ethernet mode access
A:196>edit-cfg# configure port 1/1/2 ethernet encap-type dot1q
A:196>edit-cfg# configure port 1/1/2 no shutdown
A:196>edit-cfg# candidate save cf1:\ExampleCandidate
Writing candidate-cfg to cf1:\ExampleCandidate.. OK
A:196>edit-cfg# candidate commit no-checkpoint
Processing current config... 0.010 s
INFO: CLI Successfully executed 32 lines in 0.090 s.
```

A candidate configuration example is shown here.

WRAP-UP
Module Summary

Wrap-up.

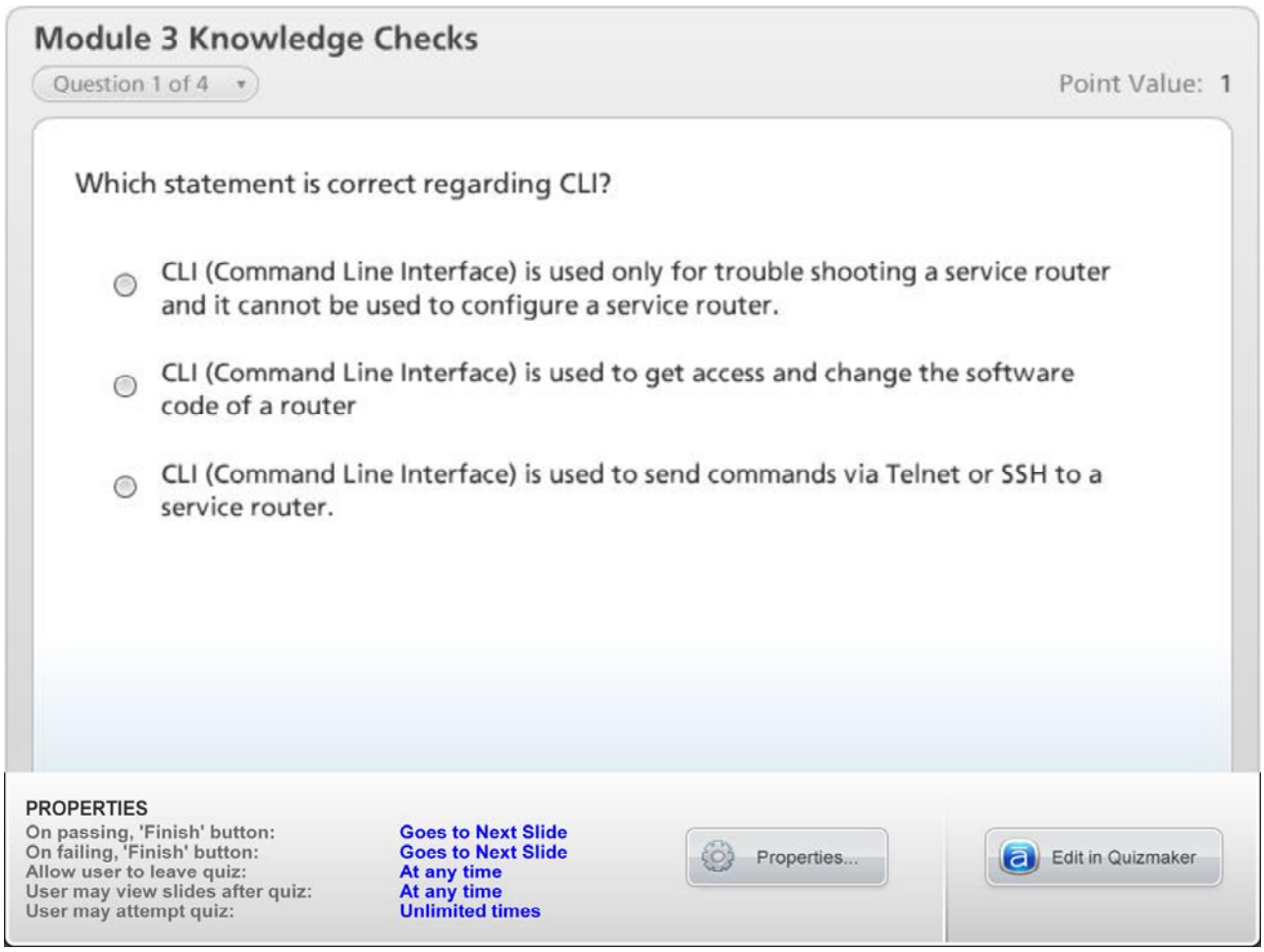
Module Summary

- A CLI is an inverted tree structure
- A number of commands are available for navigation through the tree
- "?": is an important key
  -: provides suggestions
  +: options for variables
- "admin save" and "bof save" are key and fundamental commands
- Events can come from four sources:
  - main
  - security
  - change and
  - debug
- Log destinations are:
  - Console
  - File, Memory
  - Session
  - SNMP trap and
  - Syslog
- Vi text editor is available for file text editing
- Rollback allows to go back to any configuration checkpoint
- Transactional Configuration allows a user to generate a candidate configuration

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempt at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 3

Learning experience powered by Alcatel-Lucent University

Alcatel·Lucent

This completes module 1.

SR-OS Fundamentals

Module 4: Policies

IPD Development

Available as PDF

Welcome to the fourth module of the SR-OS fundamentals course.

Agenda

- Module 4:
  - The concept of a policy on a service router
  - Filter policy example
  - Router policy example
  - SAP-ingress policy example
  - Redistribution policy

The agenda of Module 4.

By the end of this module you will be able to explain:

The usage of a policy on a service router

How to create a policy

How to verify a policy

Policies

- Generally used concept to enforce rules on a service router
- Defined globally on a service router
  - Locally applied in a context
    - On SAP's, interfaces, Ports, protocols,…
- Policy can be applied multiple times
  - Become active when applied
- Use the concept on "match" -> "action"
  - A matching criteria results in an action
  - Wording may be differed per policy
- Router policies have "edit" mode
  - Edit mode allows changes without taken effect on local instances
    - Changes take effect with "commit"

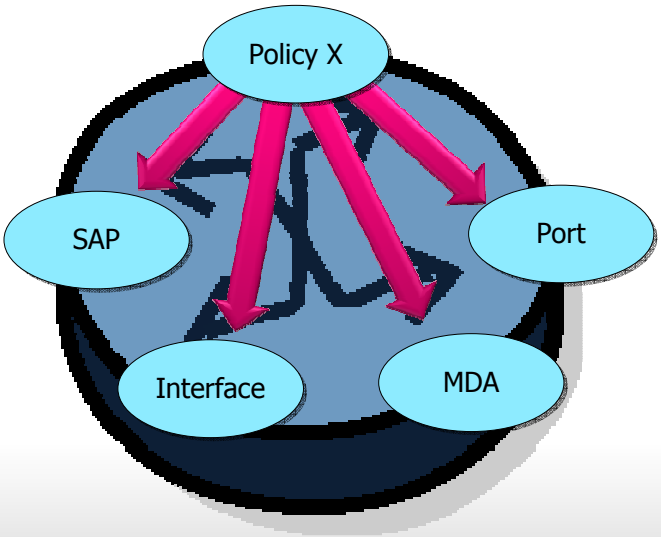
A policy is a very important concept on a service router. It could be seen as a set of rules or statements to influence the behavior of a data or control packet. Policies are used in different area's. They are used for example in quality of service, routing distribution or to filter data packets. For each of these application area's, a specific policy is available and therefore the look and feel is different per application area, but the concept stays the same.

A policy is first globally defined and then applied to certain contexts like SAP's, interfaces, ports or protocols. The creation of the policy globally does not have any effect. From the moment it is applied to a certain context, it is active to that specific instance or context. The policy can be applied multiple times. A change to any of the policy statements inside a policy will have an effect to all the instances the policy is applied to.

Although the naming is different between different kind of policies, a policy uses the concept of "match" and "action". Certain packets must match a criteria before a certain action is taken.

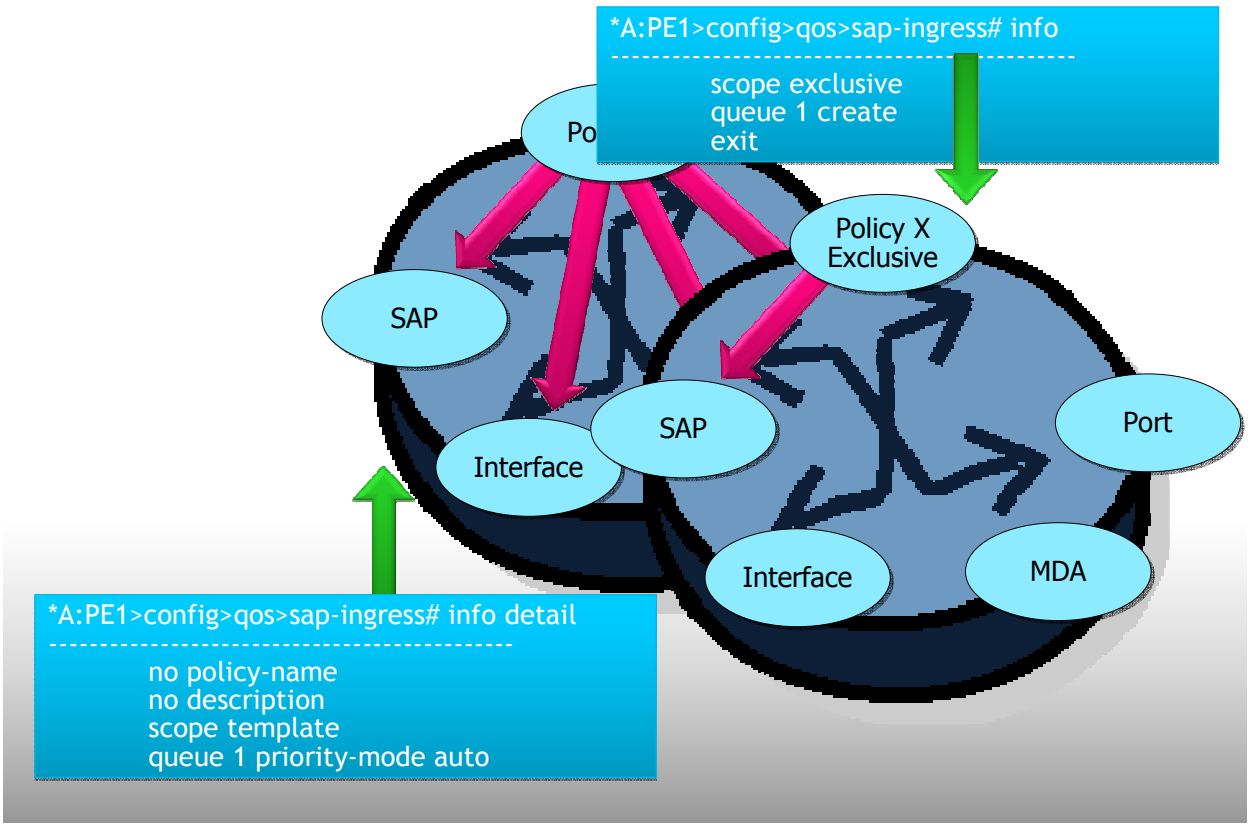
One of the available policies, a router policy, does have in addition an edit mode. The edit mode allows changes to be made to a policy without any affect to the locally applied context. From the moment all changes have been made, the policy can be "committed" and the new policy becomes active.

Policy



A policy is identified by a name or number. After globally defined, it can be applied multiple times. For example, it can applied under a SAP, port, interface or MDA.

Scope: template <-> exclusive

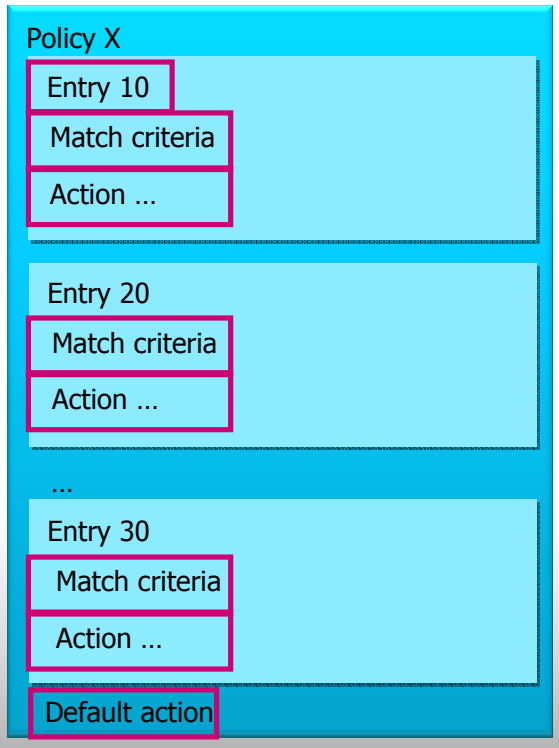


If the administrator that created the policy doesn't want that this policy can be applied multiple times, it can change its scope to "exclusive".

The scope exclusive allows the policy to be applied only once. This is mainly for security or engineering purposes.

Policy processing

- Top down processing
  Entry per entry
- Lowest integer value first
  - Suggest to take intervals of ten
  - Allows interleaving of entries later on
  - The process stops when the first complete match is found and executes the action defined in the entry.
  - Default action when no matching is found in the entries



Each policy has one or multiple entries and uses a top down processing mechanism. Starting with the entry that has the lowest numerical value. Each of the entries has a "match" criteria and an "action". The process stops when the first complete match is found and the action defined in the entry is executed, thus the order of the entries is important. It is advisable to take integer values of ten to allow interleaving of entries later on. If no match is found in any of the entries, the default action is executed. There is always a default action, even when there is no default action explicitly configured. There is a default setting of a default action.

Example Filter policy

- Packets with destination IP address 10.23.12.1 or 10.33.33.1 will be dropped!
- All other packets will be forwarded as per default action.
- When no match -> default action

```
*A:PE1>config>filter# info
----------------------------------------------
        ip-filter 2 create
            default-action forward
            entry 10 create
                match
                    dst-ip 10.23.12.1/32
                exit
                action drop
            exit
            entry 20 create
                match
                    dst-ip 10.33.33.1/32
                exit
                action drop
            exit
        exit
----------------------------------------------
```

```
A:PE1# configure router interface "toPE2-1"
ingress filter ip 2
```

Let us take an IP filter policy as a first example of a policy.

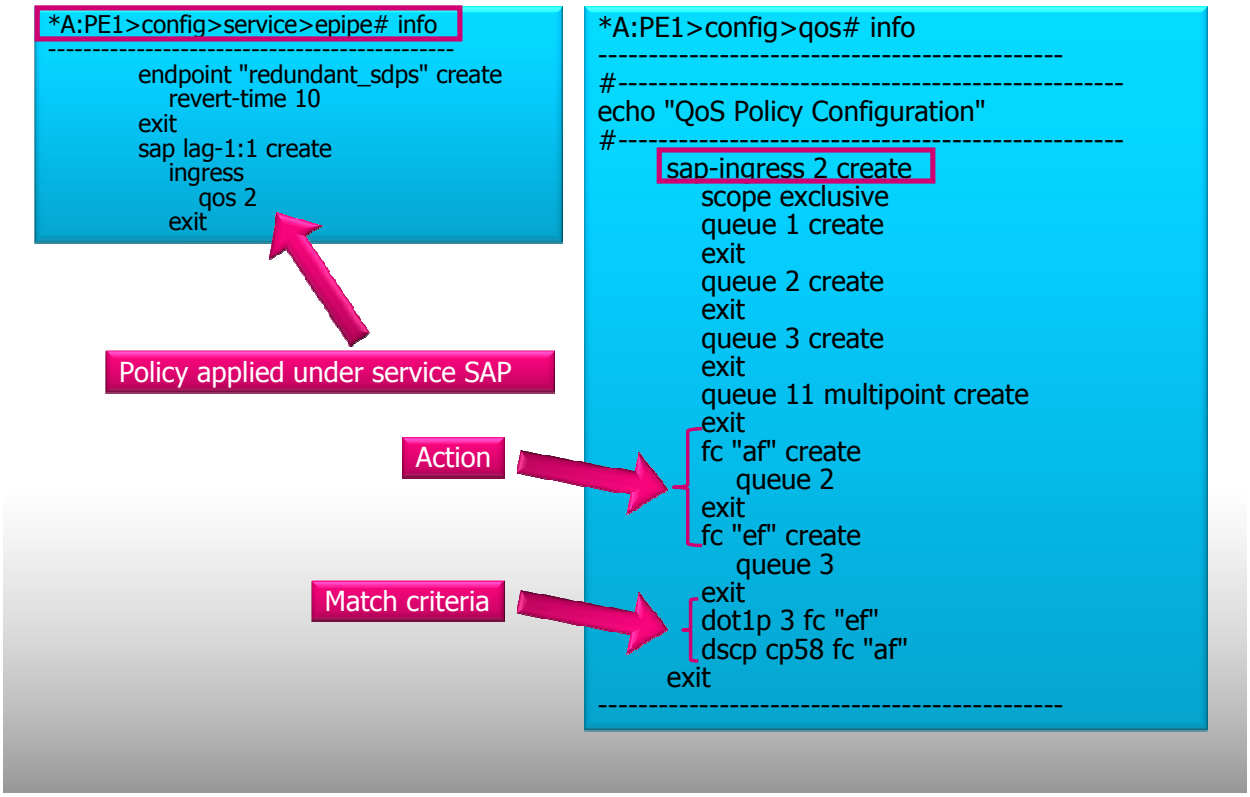
The filter is globally defined and gets an identification number of 2.

The default action is forward and although it is stated in CLI before all the entries, it is executed only when no match has been found in the different entries.

A filter policy work on the data plane and in this case, it filters IP packets with a destination IP address of 10.23.12.1 or 10.33.33.1. These packets will be dropped. All other packets will be forwarded based on the default action.

This IP filter 2 is applied on interface "toPE2-1" for packets entering the interface. The same or a different IP filter can be applied to the egress direction. A maximum of 1 filter can be applied per direction.

Example SAP-ingress QoS policy



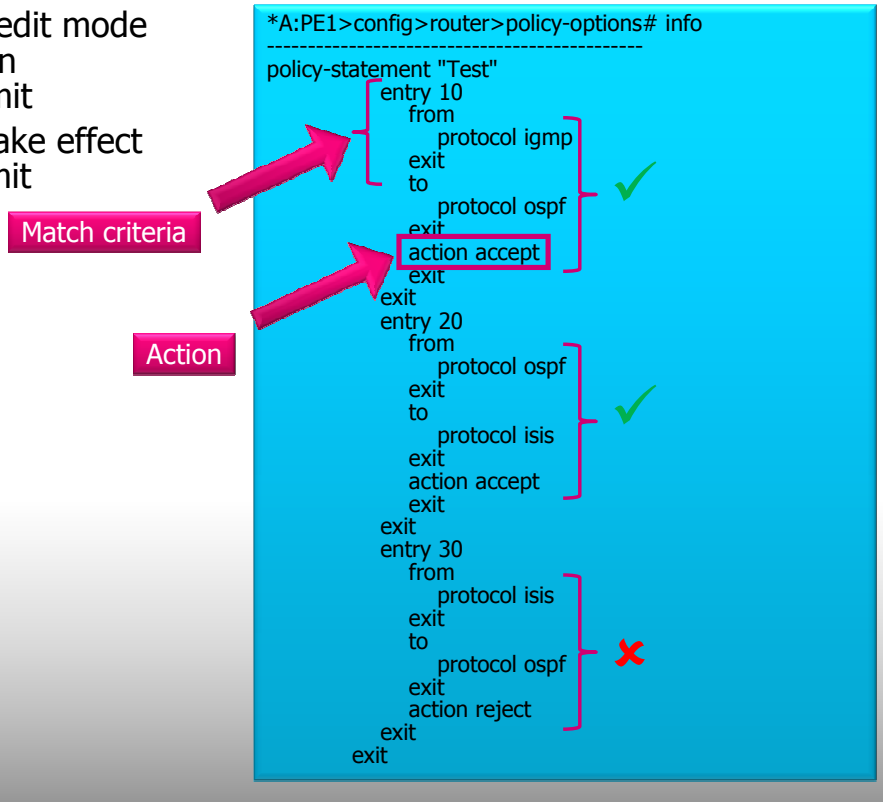
The principle of match and action is also applicable to a SAP ingress policy, although it does not use this naming explicitly in the policy. The packets entering a service on a SAP will hit a SAP ingress policy. If none is configured, the default SAP ingress policy 1 is active. When another SAP ingress policy is configured globally and applied to a SAP, it overrides this default policy of 1.

In this example, SAP ingress policy 2 is configured and applied to an epipe service SAP in the ingress direction. The same rule applies here as well that there is a maximum of 1 policy per direction per SAP. For traffic leaving the service via the SAP, an egress policy can be configured and applied.

A couple a different match criteria are available for a SAP ingress policy and use a fixed order for evaluation. Because of the fixed order on a service router, there is no need to put them in an entry type of form. In this example, two match criteria are configured, a dot1p and dscp criteria. DSCP takes precedence over dot1p settings. If a packet with dscp value of "cp58" enters the SAP, it will be placed in forwarding class "af" and queue 2. If no match on dscp value, the dot1p value is analyzed and if it equals 3, it will get "ef" and queue 3 as action.

Example router policy

- Router policy has edit mode
  - start with begin
  - end with commit
- Change will only take effect when typing commit



The router policy is a different example of a policy. The policy in this example is used to influence the routing control advertisements to other routers. The "to" and "from" command are seen as the match criteria. The action is here either accept or deny.

This policy is also referred to as a re-distribution policy. It re-distributes routes between routing protocols. Here between igmp and ospf and between ospf and is-is, but not between is-is and ospf.

Note that a policy might be applied without taking effect because of no match. It is only there for security reasons in case unwanted traffic is arriving.

A router policy has an "edit" mode. This mode allows to make changes without having direct impact on the instances the policy is applied to. Typing "commit" exits this edit mode and applies the policies to all its instances.

Different router policy action

Action	Description
[no] as-path	Assign a BGP AS Path list to routes matching the entry
[no] as-path-prepend	Prepending a BGP AS number to the AS Path attribute of routes that match the entry
[no] community	Apply a BGP community list to routes matching the entry
[no] damping	Configure a damping profile to be used for routes matching the entry
[no] fc	Configure forwarding-class to routes matching the entry
[no] local-preferen*	Assign a BGP Local Preference to routes matching the entry
[no] metric	Assign a metric to routes matching the entry
[no] next-hop	Assign a next hop IP address to routes matching the entry
[no] next-hop-self	Advertising a next hop IP address belonging to this router to routes matching the entry
[no] origin	Set the BGP origin assigned to routes exported into BGP that match the entry
[no] preference	Assign a route preference to routes matching the entry
[no] tag	Assign an OSPF RIP or ISIS tag to routes matching the entry
[no] type	Assign an OSPF type metric to routes matching the entry

Not only protocols can be set as action criteria. A wide range of options are available as actions.

Route re-distribution

- Routing protocols are essential in the discovery of paths to various destinations in the network.
- Routing information taken from the routing protocols are used to build the routing table and forward packets.
- An administrator can control the routing information by applying route filtering or route redistribution policies
- Re-distribution is the process of passing the routing information from one routing domain to another.
- Goal: To provide full IP connectivity between different routing domains.

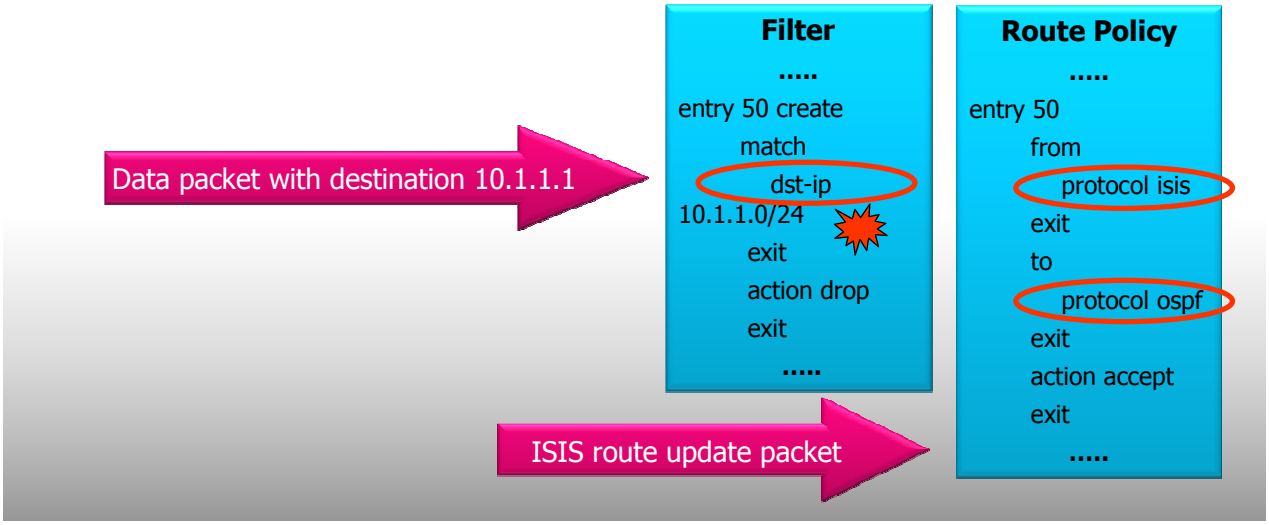
The dynamic routing protocols were developed and implemented to aid the network administrator by creating a network topology automatically instead of manually with static routes. This simplifies configuration and maintenance of the network and the convergence times are improved in networks that have frequent topology changes. The disadvantage, or trade-off, is a loss of control as to what traffic is allowed in the network.

An administrator can control the routing information by applying route filtering or route re-distribution policies.

Redistribution is the process of passing the routing information from one routing domain to another to provide full IP connectivity between different routing domains.

Filters and Route Policies

- Dynamic Routing Protocols:
  - Advantage: builds the Routing Tables automatically
  - Disadvantage: no control for the administrator
- Solution:
  - Filters on the data plane
  - Route Policies on the control plane



Filters and route policies can be used to control traffic traveling through the network. Filters and route policies contain simple forward/accept and drop/deny actions that can be applied to certain traffic/routes as determined by the administrator.

While filters, or Access Control Lists, are active on the data plane and monitor incoming or outgoing traffic, route policies are active on the control plane and monitor the route updating process.

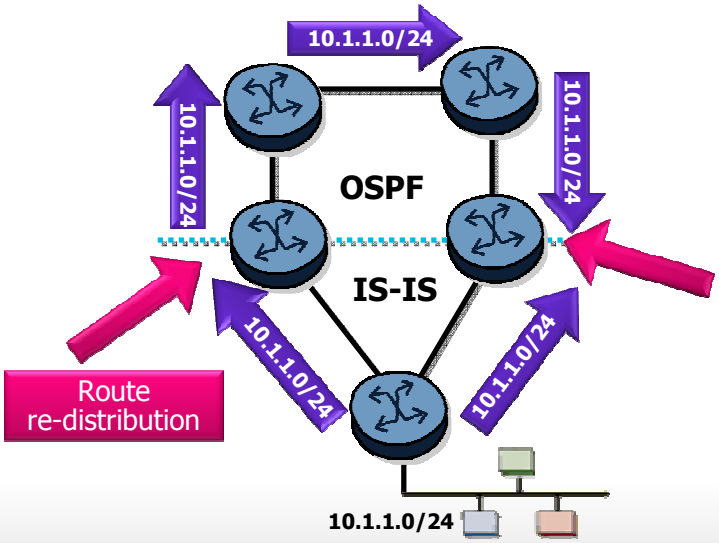
This filter policy drops all the packets with destination address in the range of 10.1.1.0/24.
The route policy works upfront, before any data packets are sent. It controls the way traffic is re-distributed between is-is and ospf.

Route Re-distribution

Protocol	Import (incoming)	Export (outgoing)
OSPF	Not applicable. All OSPF routes are accepted from OSPF neighbors and cannot be controlled via route policies.	• Internal routes (Native): All OSPF routes are automatically advertised to all neighbors. • External routes (Foreign): By default all non-OSPF learned routes are not advertised to OSPF neighbors
IS-IS	Not applicable. All IS-IS routes are accepted from IS-IS neighbors and cannot be controlled via route policies	• Internal routes: All IS-IS routes are automatically advertised to all neighbors. • External routes: By default all non-IS-IS learned routes are not advertised to IS-IS peers.
RIP	By default, all RIP-learned routes are accepted.	• External routes: By default all non-RIP learned routes are not advertised to RIP peers.
BGP	By default, all routes from BGP peers are accepted and passed to the BGP route selection process.	• Internal routes: By default all active BGP routes are advertised to BGP peers • External routes: By default all non-BGP learned routes are not advertised to BGP peers

The chart shows the default behavior of the different dynamic routing protocols.
A good example to demonstrate the difference between the Link State Protocols and the Distance Vector Protocols is the distribution of the system IP-address. In the link state protocols the system interface is added to the protocol instance and is therefore advertised to the adjacent routers. The Distance Vector Routing protocols don't allow this and need an export policy that will advertise the system IP-address over the Distance Vector Routing Protocol.
When a router runs OSPF and IS-IS simultaneously for example, there is no way the routes from one routing protocol can merge with the routes of another since they have separate link state databases. Therefore a route redistribution is necessary. Some way must be used to determine what metrics are used and which protocol will be preferred.

Preferences and External Route Tagging



- Which route is used?
  → based on preference
- Preference decides between same route from different protocols
- OSPF has lower value than IS-IS
  → so preferred
- But because of re-distribution
  → OSPF gets pref. of *150*

A router should never have two routes to the same destination in its routing table. Only the best route, learned through a routing protocol, should be installed in the routing table. When a router runs a single routing protocol, the metric is used to identify this best route, but when multiple routing protocols run simultaneously on a router, the metric alone is no longer an indicator of the best route since different routing protocols use different metric systems. A Preference value is assigned to each protocol for this purpose.

The preference is a value that creates a hierarchy among the routing protocols. If different protocols offer the same destination to the Routing Table Manager, the route with the lowest preference is chosen. After the lowest preference route is chosen, the lowest metric will decide the best route if there are more than one. Now, a single route can be installed for a single destination in the routing table.

But, what if there was a re-distribution from one routing protocol to another routing protocols via a re-distribution policy? In this case from is-is to ospf. What can the router that gets the same route via different interface decide? Basically the route that got re-distributed by a policy was tagged and will get a higher preference.
The route from ospf will have a higher and therefore worst preference that the is-is learned one and is less preferred.

The preferences can be altered if desired by the administrator, for example allowing IS-IS to dominate over OSPF. But if there is a conflict, the router will go to the default preference table.

Default preference settings

- Default Preference:
  - One destination
  - More then one routing protocol
  - Select best path based on lowest preference

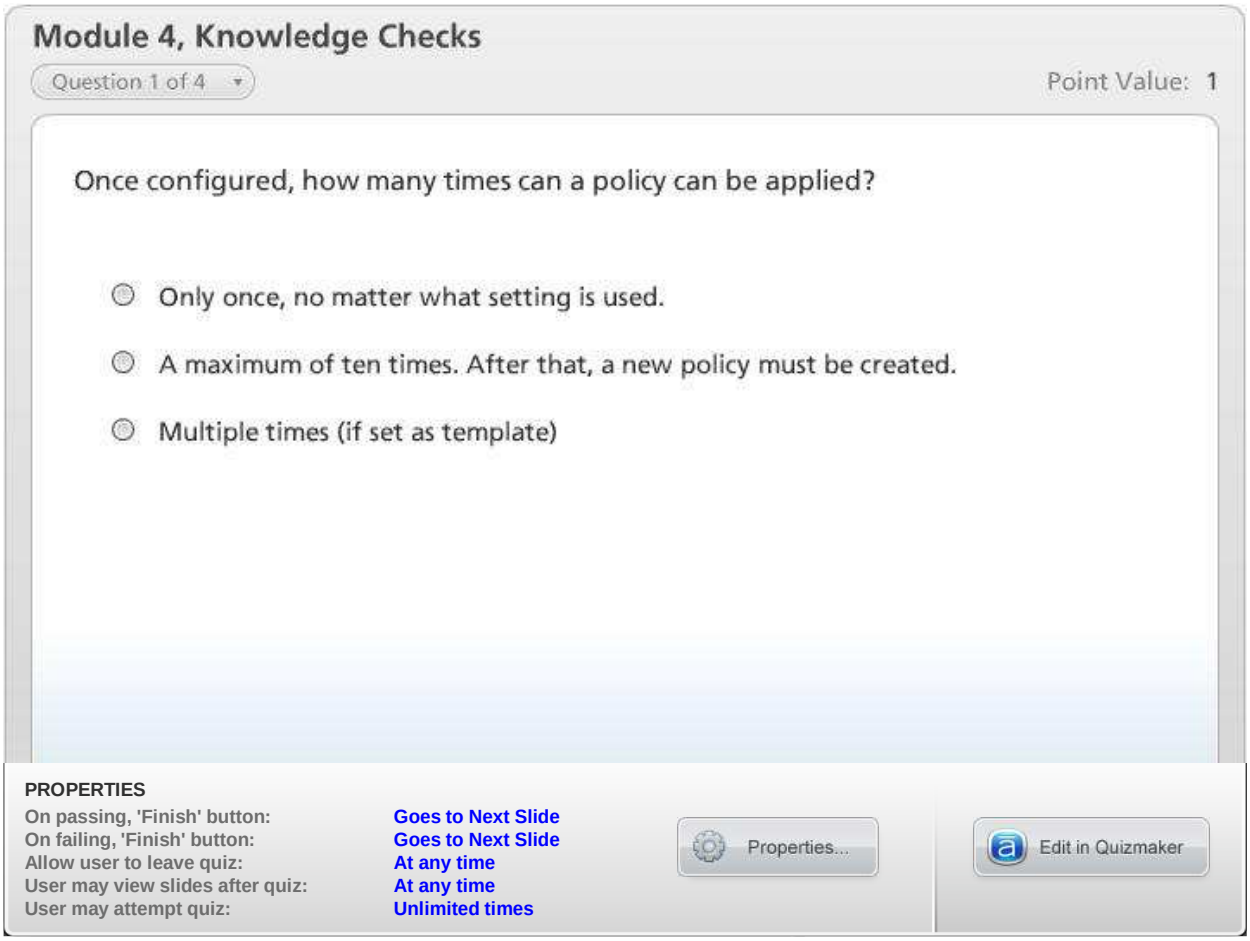
Protocols	Default preference
Directly connected	0
Static routes	5
OSPF internal	10
IS-IS level 1 internal	15
IS-IS level 2 internal	18
RIP	100
OSPF external	150
IS-IS level 1 external	160
IS-IS level 2 external	165
BGP	170

This table indicates the default preference values.

KNOWLEDGE CHECKS

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempt at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 4

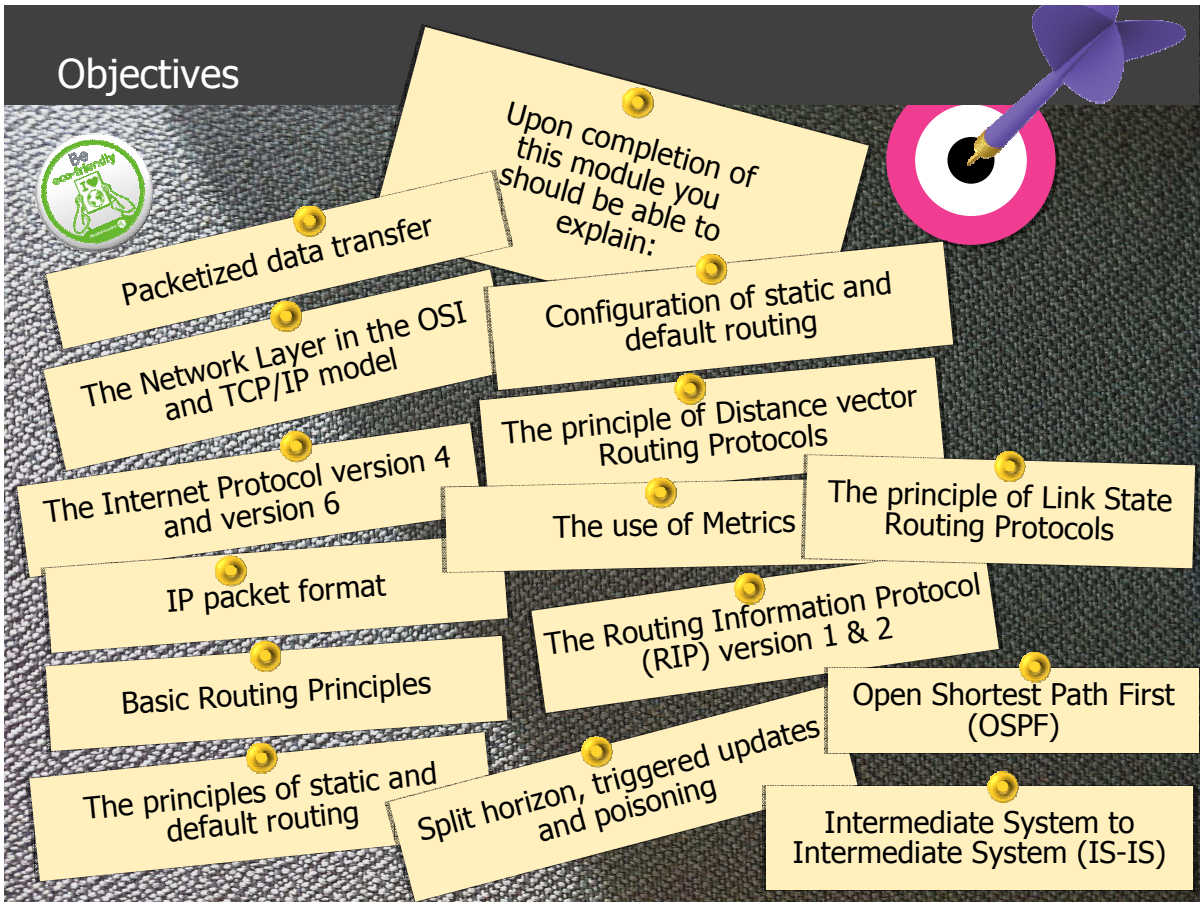
This completes module 4.

SR-OS Fundamentals

Module 5: IP Routing

IPD Development

Welcome to the fifth module of the SR-OS fundamentals course.



By the end of this module you will be able to explain:

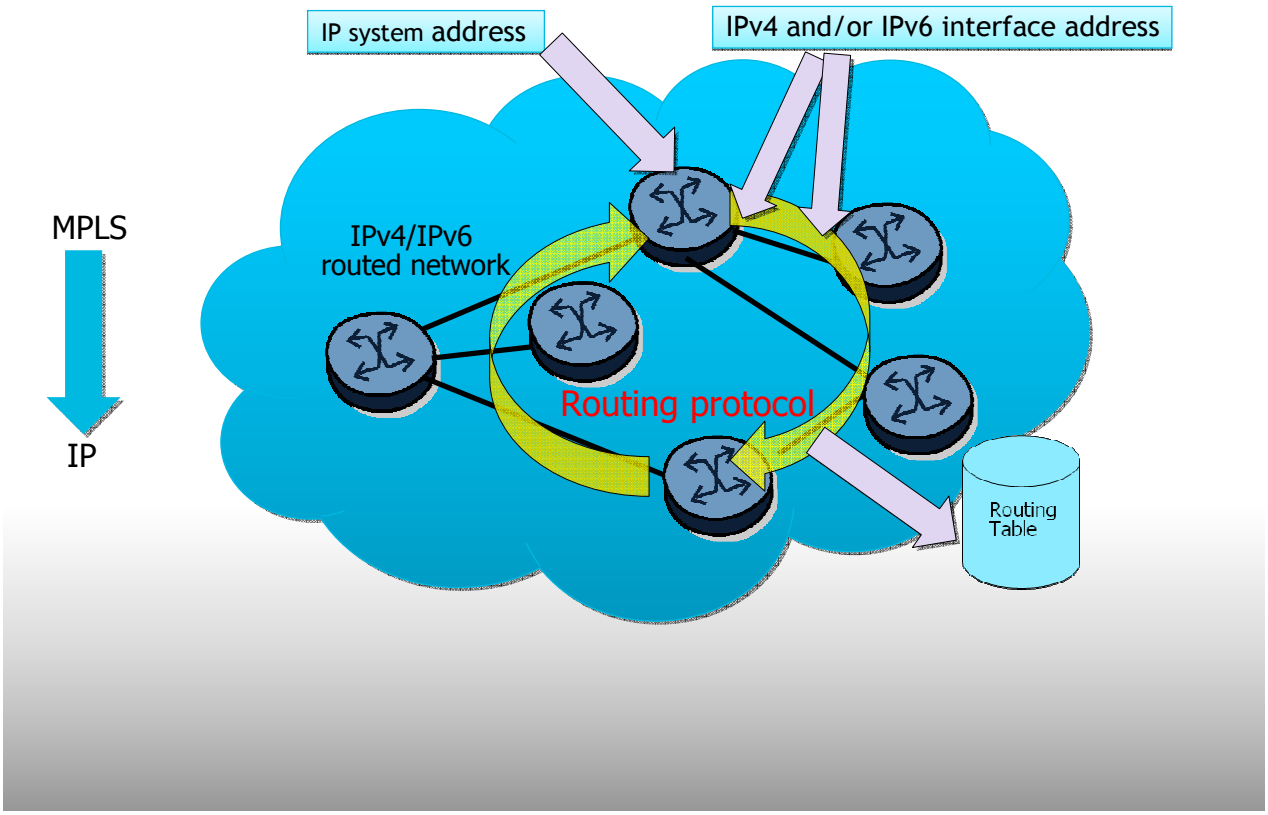
- Packetized data transfer
- The Network Layer in the OSI and TCP/IP model
- The Internet Protocol
- IP packet format
- Basic Routing Principles
- The principles of static and default routing
- Configuration of static and default routing
- The principle of Distance vector Routing Protocols
- The use of Metrics
- The Routing Information Protocol (RIP) version 1 & 2
- Split horizon, triggered updates and poisoning
- The principle of Link State Routing Protocols
- Open Shortest Path First (OSPF)
- Intermediate System to Intermediate System (IS-IS)

1

IP Protocol

Chapter 1 – XXX.

Introduction



While MPLS is the dominant technology used for services on a service router, IP routing still plays a very important role. The MPLS control plane signaling makes use of the IP routing to set up the tunnels. Once set-up, the MPLS data forwarding is independent from IP. But a change in IP routing can lead to a re-signaling of MPLS labels.

A routing table should be available for regular IP forwarding, but also for MPLS label signaling. How this routing table is constructed depends on the used routing protocol.

This module introduces the fundamental differences between the different routing protocols and discusses its pro's and con's.

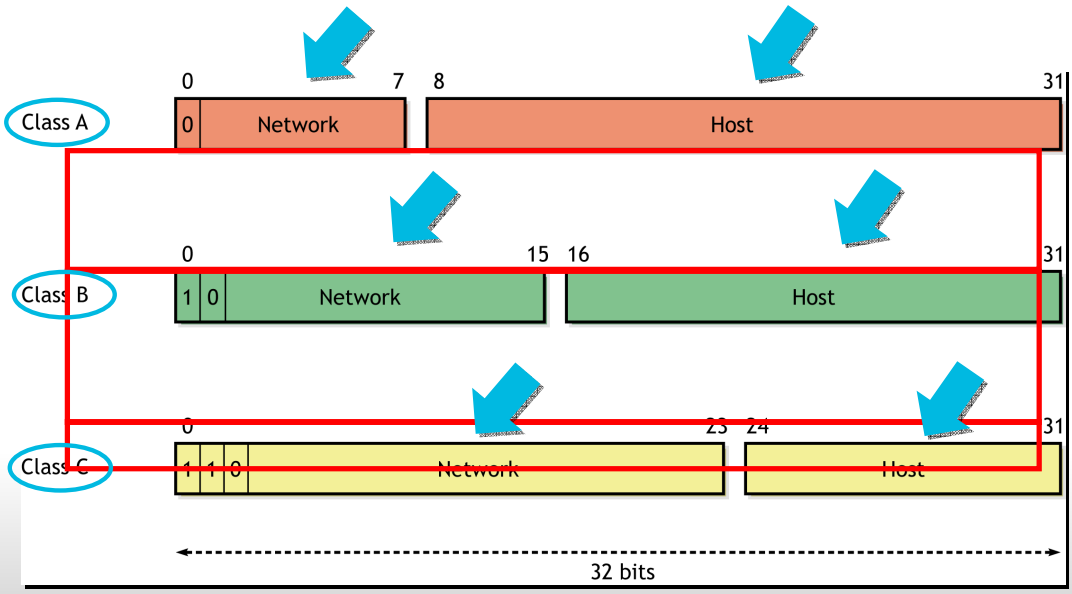
Routing is the forwarding of packets from a source address to a destination address.

The goal of a routing protocol is to build a routing table.

This routing table instructs the packet on a hop by hop basis how to go from the source to the destination. But before any routing protocol can run, all interfaces and routers need to have an address. An IPv4 and/or IPv6 address.

These addresses are used by the routing protocols as input to build the complete routing table. This is a dynamic process and does not require manual intervention from the administrator. However, there are options to statically configure routes through the network.

The Internet Protocol – IP Addressing



*Minus the Private address ranges

IP version 4 is the first widely used version for address assignment.

32 bits for an IPv4 address results in 2 to the power of 32 possible addresses. These addresses are divided into ranges.

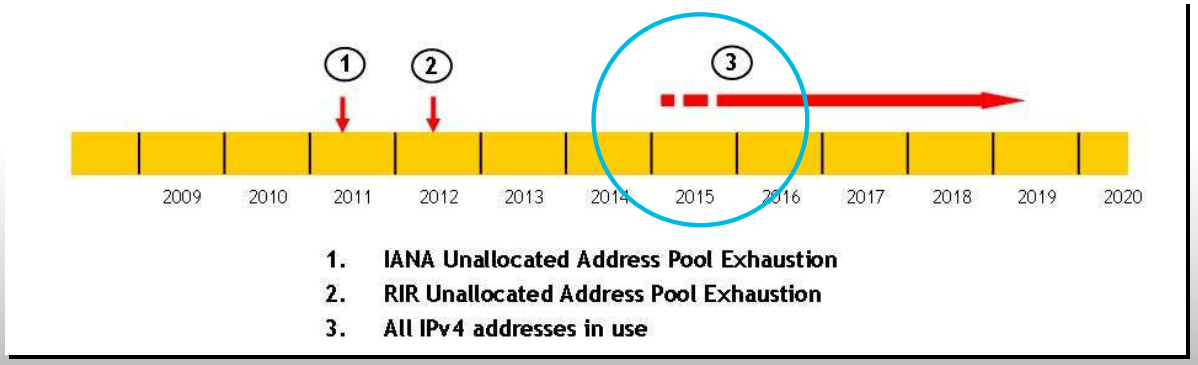
The public range are addresses assigned to an organization and routable over the Internet. The private addresses may be used freely within an organization but are blocked in the Internet. The multicast addresses are mainly used to stream audio and video traffic.

Except for the addresses in the multicast address range, each IP address is split into two portions. A portion that identifies the network and a portion that identifies the host within this network. The size of the network and host portion can be different. A smaller network portion results in a higher number of hosts. The size of the network portion is decided by its class. Each class has a different number of bits assigned for the network and host portion as explained. The first four bits indicate the class range. Although, these days, people do not talk about classes anymore as there is a flexibility method of indicating the network and host portion.

IPv4 Pool Address – The END

- IANA may release additional reserved IPv4 blocks (/8s)
- Service Providers may deploy NAT

4.3 billion IPv4 addresses vs. 6.4 billion population



IPv4 addresses are widely used, but because of the popularity of IP, there are almost no IP addresses available anymore.

To solve this, IANA or RIR, the authorities allocating addresses, might release some reserved blocks or service providers might deploy Network Address Translation, but this would only be a temporary solution.

The best future proof solution is the use of IP version 6 addresses. These addresses are using more than 32 bits per address.

Alcatel-Lucent service routers do fully support IPv6.

What Does IPv6 Offer?

- Provides huge address space
  - More than 3.4x10e38 addresses
- Hierarchical address allocation provides efficient routing
  - Small routing table
- Support Anycast addresses and eliminate broadcast addresses
- Efficient IP header, 40 bytes header with 8 fields
  - Less fields
- Stateless (no DHCP6) and stateful (DHCP6) address configuration
- Built-in Security, IPsec implemented in IPv6
  - Authentication Header (AH) and Encapsulation Security Payload (ESP)
- Better QoS support provided by IPv6 (Flow Label)
- Improved neighbouring node interaction (e.g. ICMPv6 replaces ARP)
- Flexible by using Extension Headers
  - Daisy chain of next headers
- IGMP is replaced with ICMPv6: Multicast Listener Discovery (MLD)

So what does IPv6 offer?

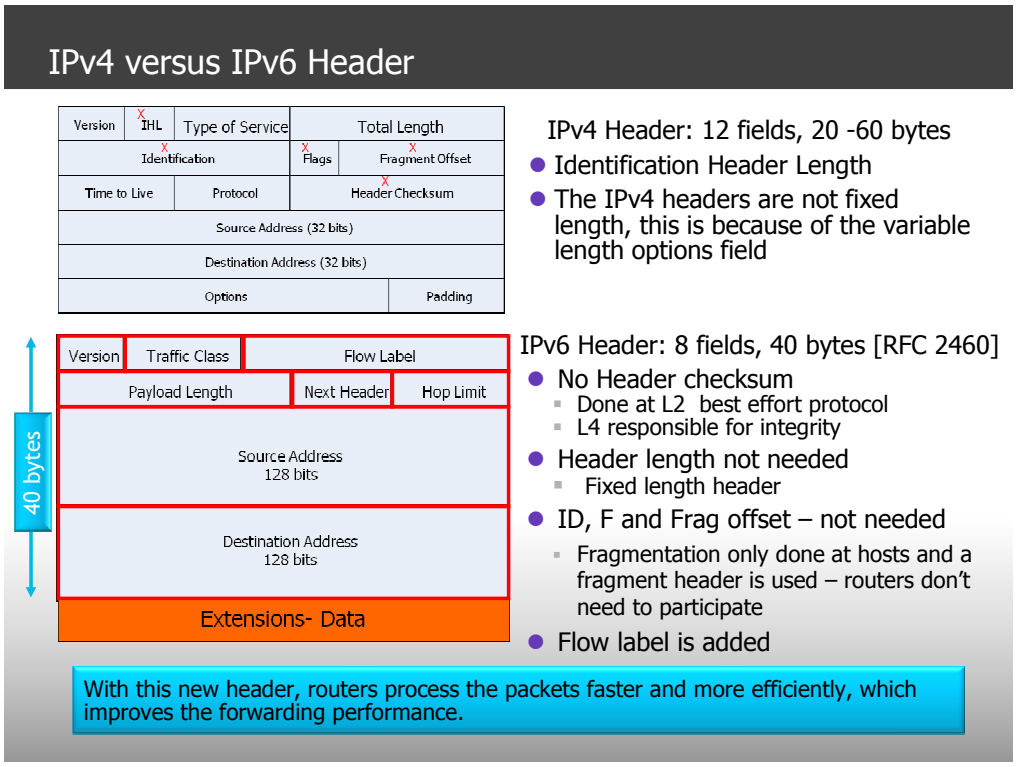
IPv6 offers a huge address space because of the 128 bits used in an IPv6 address. A more intelligent and hierarchical allocation of IP addresses will result in smaller routing tables. There is standard support for anycast addresses and broadcast addresses are eliminated. Less fields are in the header. 8 fields compared to 12 fields in IPv4.

IPv6 also implements additional features not present in IPv4 like the IP address configuration; stateless and stateful. Stateless IPv6 hosts can configure themselves automatically when connected to a routed IPv6 network using Internet Control Message Protocol version 6 router discovery messages. Stateless provide plug and play networking for host while stateful uses a DHCP server to get its IPv6 address. Both IP address assignments methods are complementary to each other.

IPv6 has a build in security. Authentication Headers provide connectionless integrity and data origin authentication for IP datagrams and provides protection against replay attacks. Encapsulating Security Payloads provide confidentiality, data-origin authentication, connectionless integrity, an anti-replay service and limited traffic-flow confidentiality. Better QoS is provide by the flow label in the header. The usage of the Flow Label field enables efficient IPv6 flow classification based only on IPv6 main header fields in fixed positions.

ICMP version 6 replaces the need for ARP and IGMP is replaced with ICMPv6 multicast listener discovery protocol.

There is also a flexible way of pointing to the next header. Extension headers are additional headers that can be present in the IPv6 packet. Packets requiring additional information at the network layer can encode this information into additional headers



Less fields and a fixed amount of bytes are one of the main differences between the two versions.

The 4-bit Version field contains the number 6. This field is the same size as the IPv4 version field that contains the number 4.

The traffic class field in the IPv6 header can assume 16 different values. It enables the source node to differentiate packets it generates by associating different delivery priorities to them.

The 24-bit Flow Label field in the IPv6 header can be used by a source to label a set of packets belonging to the same flow. A flow is uniquely identified by the combination of the source address and of a nonzero Flow Label.

The Payload Length field contains the payload length—that is, the length of the data field following the IPv6 header, in octets.

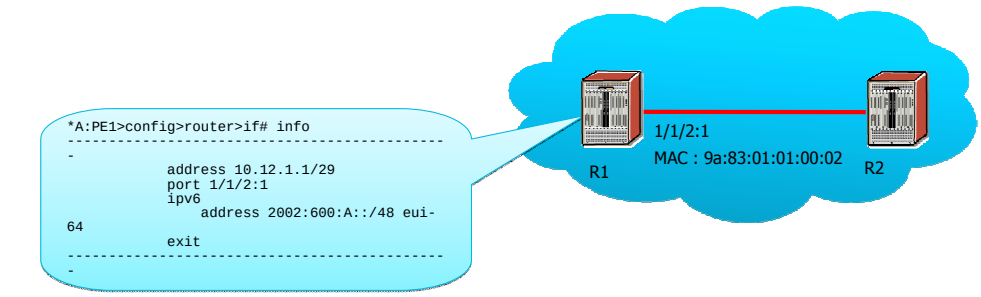
The Next Header field identifies the type of header immediately following the IPv6 header and located at the beginning of the data field of the IPv6 packet.

The 8-bit Hop Limit field is decremented by one by each router that forwards a packet. If the Hop Limit field is decremented to zero, the packet is discarded. The main function of this field is to identify and to discard packets that are looping because of erroneous routing information. Clearly, between two IPv6 nodes, we cannot have more than 255 hops, which means no more than 254 routers.

The 128-bit Source Address field contains the IPv6 address of the node originating the packet. The 128-bit Destination Address field contains the IPv6 address of the node recipient of the packet.

With the IPv6 new header, routers process the packets faster and more efficiently, which improves the forwarding performance.

IPv4 and IPv6 Interface Configuration – dual stack




```
*A:PE1# show router interface

===============================================================================
Interface Table (Router: Base)
===============================================================================
Interface-Name                   Adm       Opr(v4/v6)  Mode      Port/SapId
   IP-Address                                          PfxState
-------------------------------------------------------------------------------
toPE2-1                          Up        Up/Up       Network   1/1/2:1
   10.12.1.1/29                                         n/a
   2002:600:A::9883:1FF:FE01:2/48                                 PREFERRED
   FE80::9883:1FF:FE01:2/64                                       PREFERRED
```

IPv4 and IPv6 can run simultaneously on an Alcatel-Lucent service router. An interface can have the two addresses supported in dual stack. An additional context is available under the interface context to configure an IPv6 address.

Note that this is only possible within the supported chassis mode.

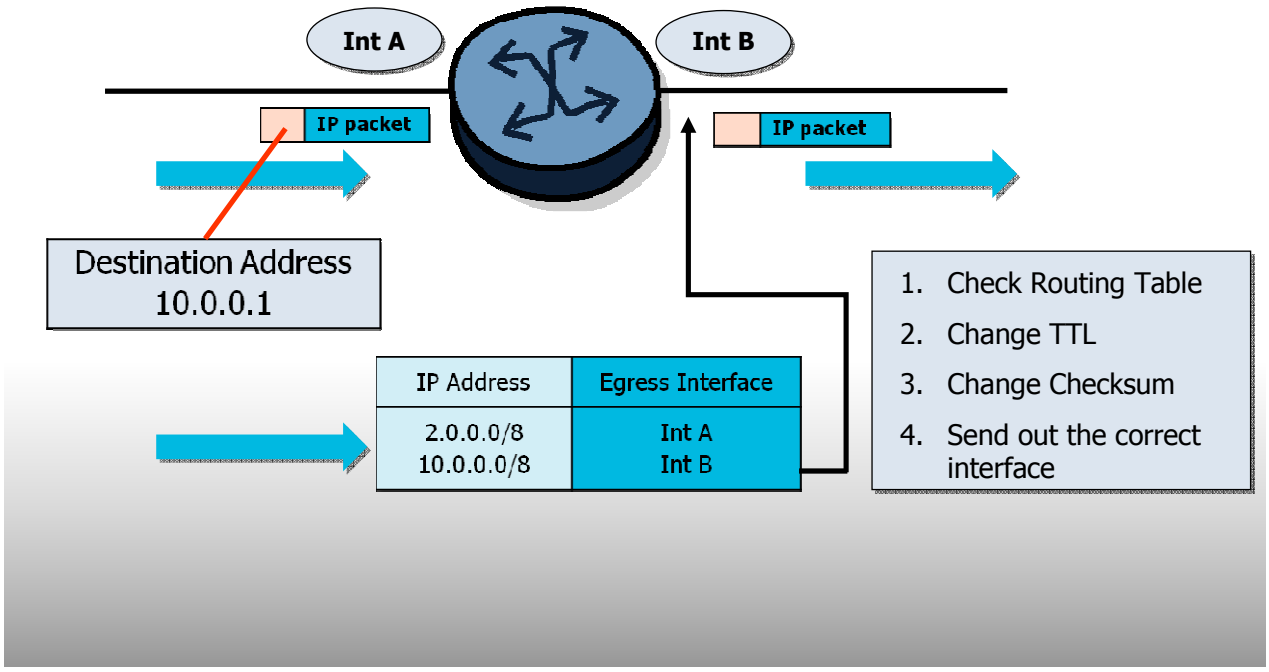
To verify if the IPv4 and v6 address are operational up, type "show router interface".

IPv6 Address - Using Manual Configuration

```
*A:SARF-72>config>service>ies>if$ ipv6 address 4004:bac:3344:6788:fecf:bd34:b45c:4312/64

*A:SARF-72>config>service>ies>if$ show router interface ipv6 test

===============================================================================
Interface Table (Router: Base)
===============================================================================
Interface-Name                   Adm         Opr(v4/v6)  Mode    Port/SapId
   IP-Address                                                    PfxState
-------------------------------------------------------------------------------
test                             Up          Down/Up     IES     1/2/1
   4004:BAC:3344:6788:FECF:BD34:B45C:4312/64                     PREFERRED
   FE80::223:3EFF:FE43:392C/64                                   PREFERRED
-------------------------------------------------------------------------------
Interfaces : 1
===============================================================================
```

Another example of manually creating an IPv6 address, but not an IPv4 address. This time within an IES service.

Routing Principles

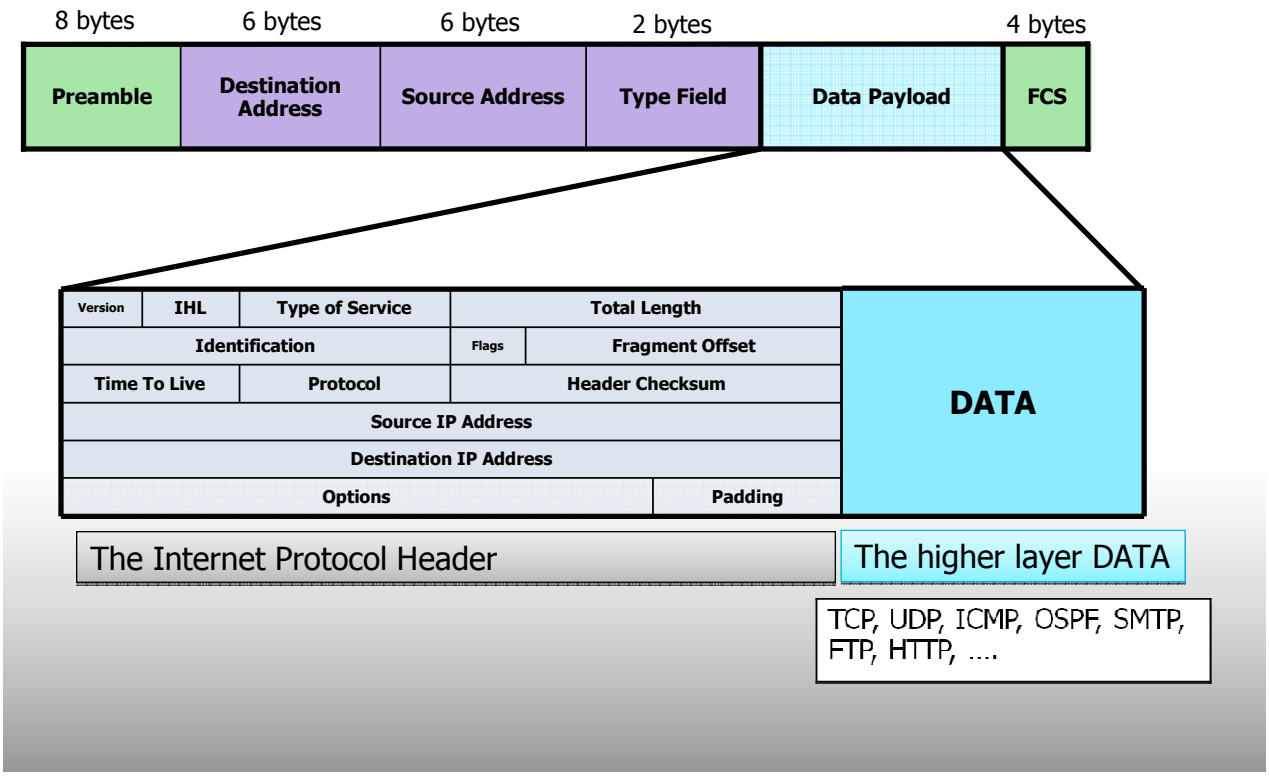


Routing an IP packet from a sender to a receiver is the process of transporting the packet from node to node. Every node receives packets through one of its interfaces and must decide where to send the packet next by choosing an outgoing interface. All the routers perform the same process.

Deciding what interface to send a packet to depends on the destination IP address contained in the IP header of each IP packet. The router checks a list, the routing table, which stores destination addresses mapped to egress interfaces. If a packet arrives on the ingress with a destination address not stored in the routing table, the packet is thrown away or dropped. Listing every individual destination addresses can use up a lot of router resources so usually the router groups addresses according to networks or sub-networks.

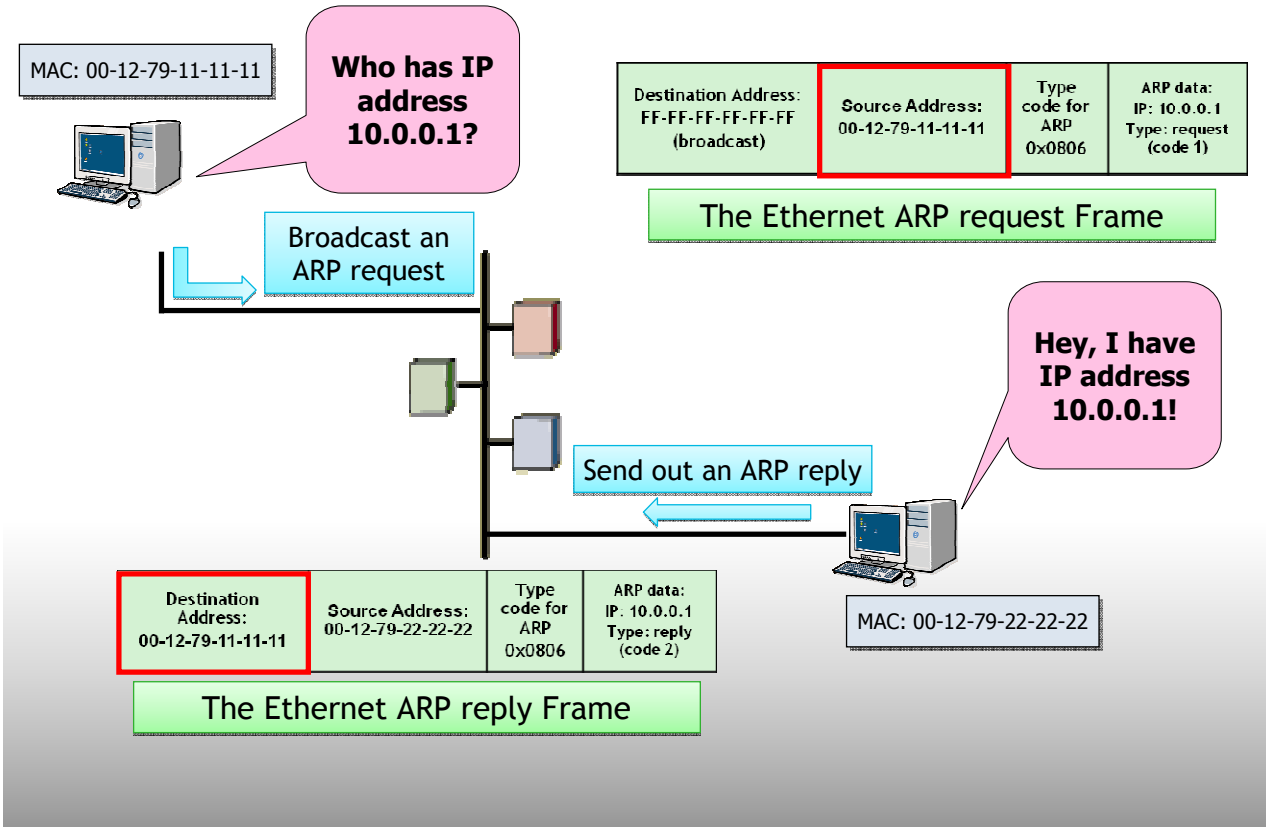
In summary:
1.A packet arrives on an ingress interface of a router.
2.The IP destination address is read and compared to the routing table entries.
3.If no match is found, the packet is dropped.
4.If a match is found, the **TTL** is decremented, the header checksum is recalculated and the packet is sent out the egress interface listed in the routing table.
5.Each router in the path from source to destination performs the same operation.

Ethernet Header



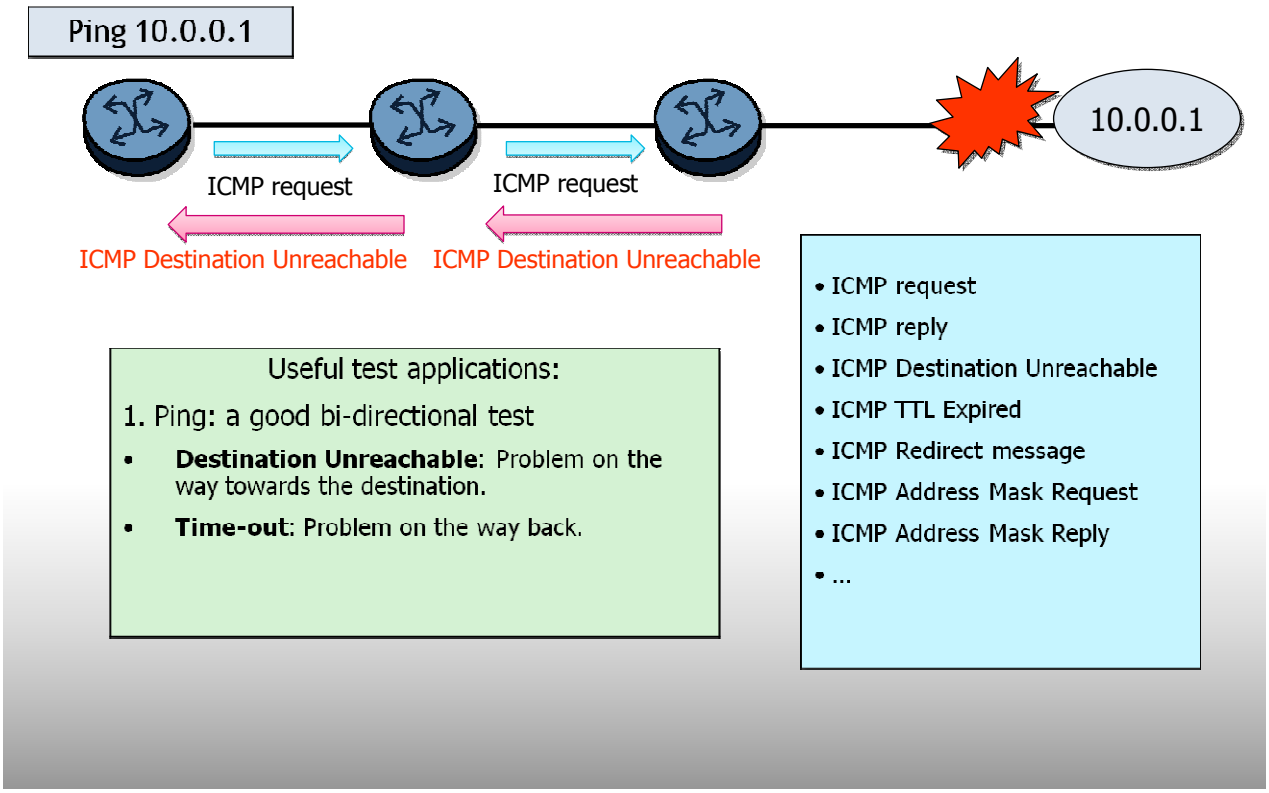
When encapsulated in Ethernet, the type field will be 0x0800 for IPv4 as in this example.
For IPv6 this will be 0x86DD.

Other Protocols – ARP



The TCP/IP protocol suite is made up of many protocols, one of these is the Address Resolution Protocol (ARP).
When an Ethernet Card needs to send out an Ethernet Frame onto the link it needs to know the destination MAC address of the device it is trying to reach. It uses the ARP to determine the MAC address associated with an IP address.
When an ARP request is broadcast on the LAN all devices connected to the LAN receive the broadcast. The broadcast attempts to find a device with a particular IP address and determine its MAC address. The device that receives the broadcast and recognizes its own IP-address will send out an ARP reply towards the original sender containing its own MAC address. Now, the sender will be able to construct the Ethernet Frame with the destination Address being the Source MAC address of the ARP reply.
It is time consuming to send out an ARP request for the same address over and over, so the sending device stores a table of IP addresses with their corresponding MAC addresses in an ARP cache. Individual ARP entries will be flushed out of the cache after a certain period of inactivity.

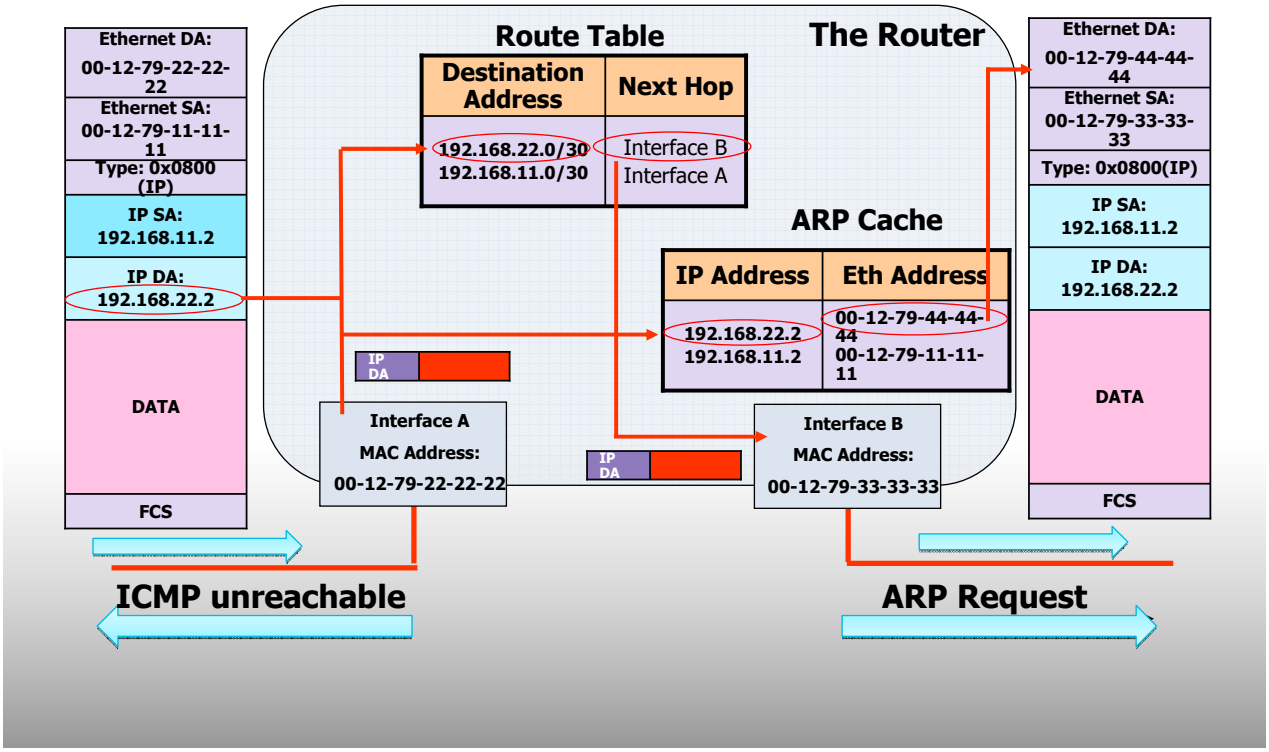
Other Protocols – ICMP



Another protocol in the TCP/IP suite is the Internet Control Message Protocol (ICMP). The ICMP provides several functions:

1. ICMP request and reply, commonly used in the PING application. This function allows a network node or operator to determine if a certain address is reachable. It works by sending an ICMP request to a certain destination (IP-address) and waiting for a reply. A device is reachable if it replies to the ICMP request.
2. ICMP Destination Unreachable. When a packet arrives at a node that has no further connection to the destination (for example the link is down or the router has no entry for this destination in its routing table) the packet is dropped and an ICMP Destination Unreachable message is sent back to the sender.
3. ICMP TTL expired. When a router is forced to drop a packet, because it's TTL has reached zero, it will send an ICMP TTL Expired message back to the sender. This functionality is used in the common application TRACEROUTE where ICMP requests are sent with incrementing TTL's. The TRACEROUTE application uses the saved ICMP TTL Expired messages to obtain information on the path to a certain destination.

The Full Routing Cycle



Combining the L3 and L2 functions, the Routing Cycle is complete:

1. An Ethernet Frame arrives at an ingress interface with the Destination Address set to the MAC-address of the NIC of that interface.
2. Since the NIC recognizes the Ethernet-DA as its own MAC address, it strips the Ethernet Frame from the link and since the Ethernet Type Field indicates that the content is an IP packet the router analyses the IP-DA in the IP header.
3. The IP-DA is compared with all the entries in the routing table to find the longest match. If no match is found, the router discards the packet and sends an ICMP Destination Unreachable message to the IP-SA.
4. The Routing Table entry tells the router where this packet should be sent next. The routing table lists an egress interface and the IP-address of the interface of the next router.
5. This IP-address is compared to the ARP cache of the router. If there is no entry, an ARP request is sent out on the same link. When the entry is known, the MAC address can be used.
6. A new Ethernet Frame is created, the SA is now the MAC address of the egress interface, the DA is the MAC address that was found in the table or determined by the ARP request, the Type Field is 0x0800 for IP and the FCS is calculated on the new Frame. The Data Payload of the Frame is only altered a little at the IP header where the TTL is decremented and the Checksum recalculated.

These 6 steps are the basic steps for the routing process. It is important to note that the routing table is a crucial item in this process.

The Routing Table

Dest Address	Next Hop	Type	Metric
192.168.11.0/30	Interface A	Local	0
10.12.1.0/29	Interface B	Local	0
192.168.22.0/30	10.12.1.2	Remote	10

```
================================================================================
========
Route Table (Router: Base)
================================================================================
========
Destination Prefix    Next Hop              Type          Protocol   Age      Preference Metric
                      [Interface Name]
--------------------------------------------------------------------------------------------------------------------------------
--
1.1.1.1/32            system                Local         Local      02d18h18m  0                0
2.2.2.2/32            10.12.1.2             Remote        OSPF       02d04h10m  10   1000
4.4.4.4/32            10.12.1.2             Remote        Static     02d18h18m  5                1
10.12.1.0/29          B                     Local         Local      02d18h18m  0                0
192.168.22.0/30        10.12.1.2            Remote        OSPF       02d04h10m  10   1000
--------------------------------------------------------------------------------------------------------------------------------
--
No. of Routes: 5
================================================================================
========
```

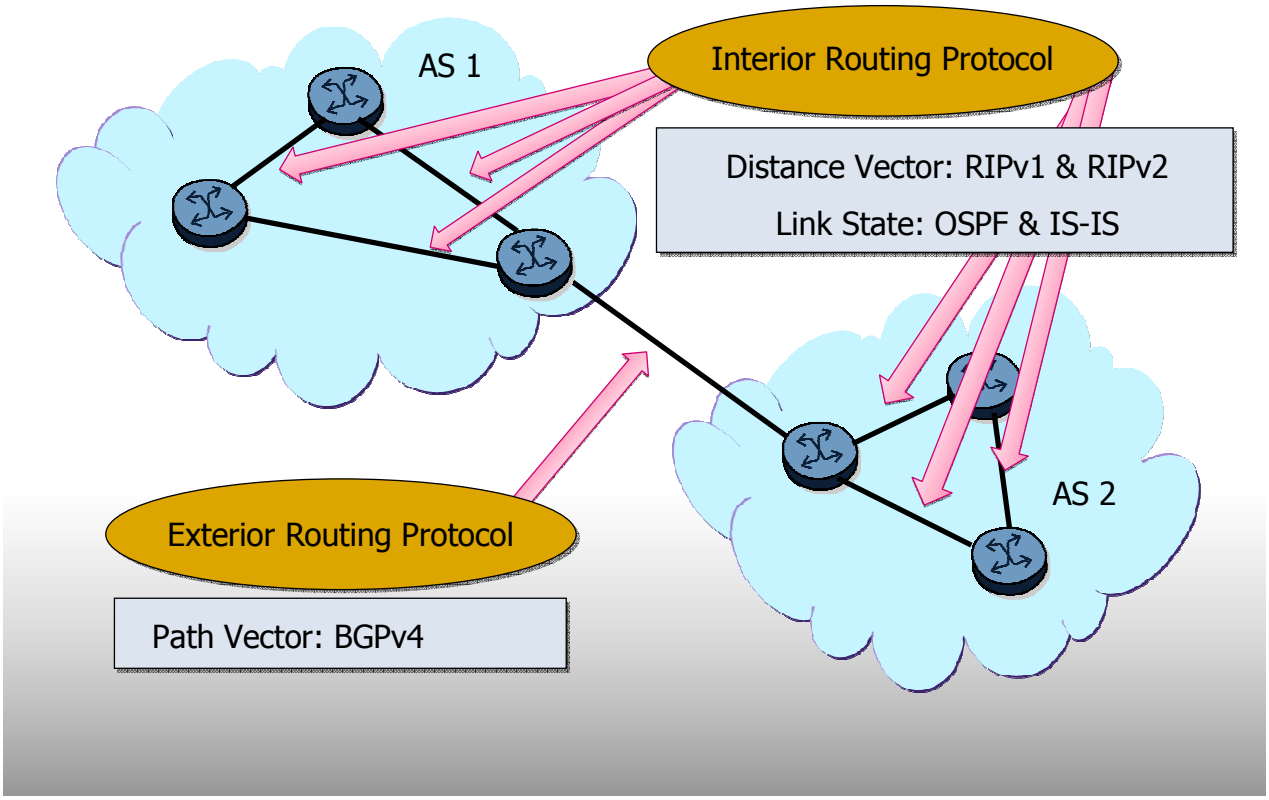
There are two techniques used to build the Routing Table:

1. Someone can create the table manually, by logging in and typing all the entries into the Routing Table. This process is referred to as: **Static.**
2. A program can create the Routing Table automatically. Such a program is called a Routing Protocol and this process is referred to as: **Dynamic.**

Static routing is very time consuming and when a failure occurs, the administrator has to be aware of it and reconfigure all the nodes to avoid the failed link or router. Dynamic routing using routing protocols simplifies the process. The routing protocols build all the routing tables in the routers and when there is a topology change (a Router or network has been added, or a failure occurs) these routing protocols update all the Routing Tables dynamically without requiring any human intervention.

Although dynamic routing protocols make things easier, static routing is still used for some purposes, for example when an administrator wants to keep certain routes under strict control.

The Routing Protocols



The Routing protocols fall into two categories, Interior Gateway Protocols or IGP's and Exterior Gateway Protocols or EGP's. A gateway is the legacy name of a router. The distinction is clear, while the one is inside "something" the other is outside "something". To answer the question on what this "something" is, an Autonomous System (AS) needs to be introduced. An Autonomous System (AS) is a collection of networks and routers under a single administration. This basically means that all the routers under the administration of a single company represent an AS.
The IGPs are used within (intra) an AS and are of two types: Distance Vector and Link State.

Distance Vector IGPs:

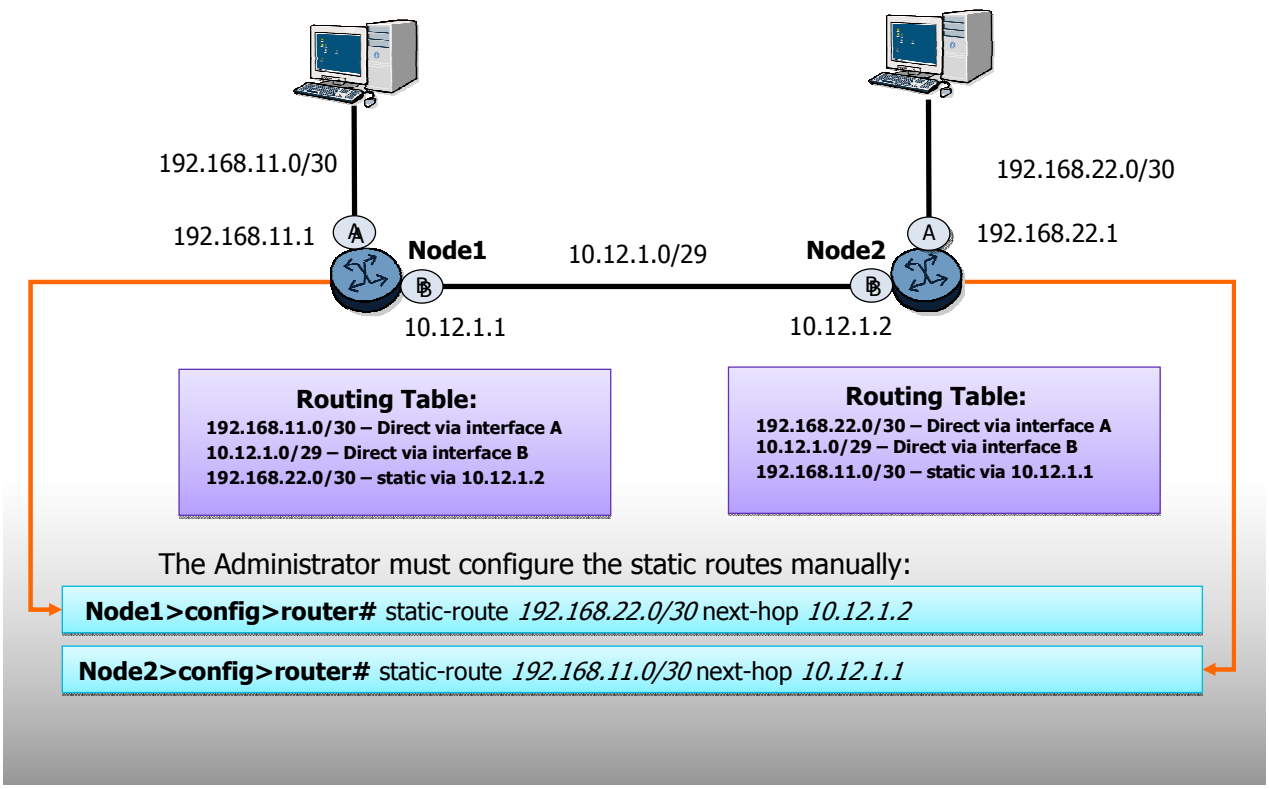
1. Routing Information Protocol version 1: RIPv1
2. Routing Information Protocol version 2: RIPv2

Link State IGPs:

1. Open Shortest Path First: OSPF
2. Intermediate System – Intermediate System: IS-IS

The EGPs are normally used between (inter) AS's. The Border Gateway Protocol (BGPv4) is the best known EGP.

Static Routing



Static routes are entries in the routing table that were manually entered by someone, typically the network administrator. The administrator must have a good understanding of the existing network topology and be confident that any changes in the network topology will not affect these static routes. Static routes are often used to connect two AS's or to provide a default route for Stub Networks.

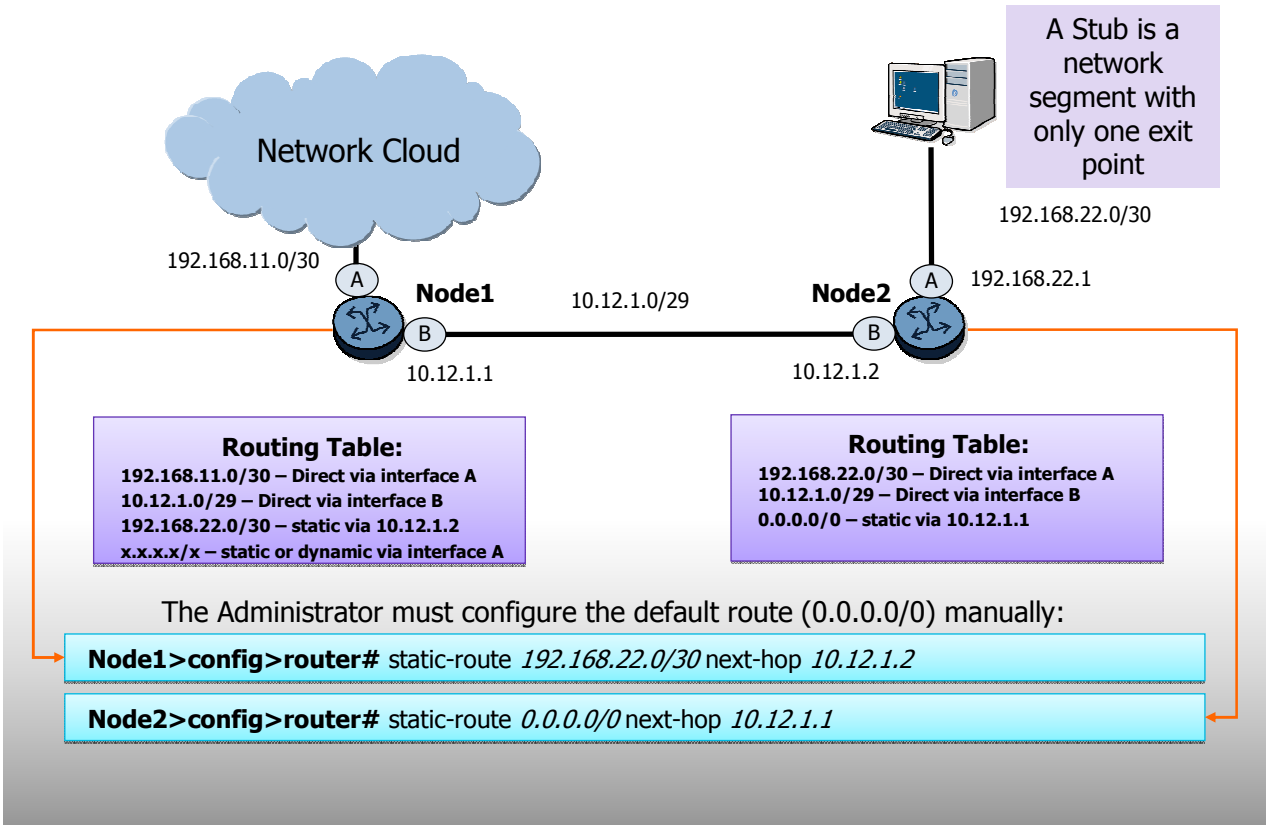
There are three types of static routes:

1. Next-hop: specifies the IP-address of the interface of the next hop router on a directly connected link.
2. Indirect: specifies the IP address of the interface of the next hop router, not directly connected, but at least 1 hop away.
3. Black Hole: used to silently discard an IP packet with the specified IP-DA.

Node 1 has a static route pointing to a route of the other PC's subnet. Note that this is a subnet and not a host address. The next hop is the far end interface address of the other router. Node 2 has a comparable static route configured so that bidirectional communication is possible between the two PC's.

Static routes can be viewed using the following command: `show router static-route`

Default Routing



A static route can be used as a default route. The default route 0.0.0.0/0 provides a route for addresses not listed in the routing table. Instead of dropping routes that don't match a routing table entry, they will follow the path indicated by the default route.

A default route is very useful in stub networks. A stub network is a network with only one exit point. The default route can be used to point all the traffic towards this exit point.

IPv6 Static Routing

```
*A:SARF-41# configure router static-route 3004:BAC:3344:6788::/64 next-hop
   2002:600:A::223:3EFF:FE43:3D2D

*A:SARF-41# show router static-route ipv6
===============================================================================
Static Route Table (Router: Base)  Family: IPv6
===============================================================================
Prefix                                    Tag         Met    Pref Type Act
   Next Hop                                Interface
-------------------------------------------------------------------------------
3004:BAC:3344:6788::/64                   0           1      5    NH   Y
   2002:600:A::223:3EFF:FE43:3D2D          to-sar72
-------------------------------------------------------------------------------
No. of Static Routes: 1
===============================================================================
```

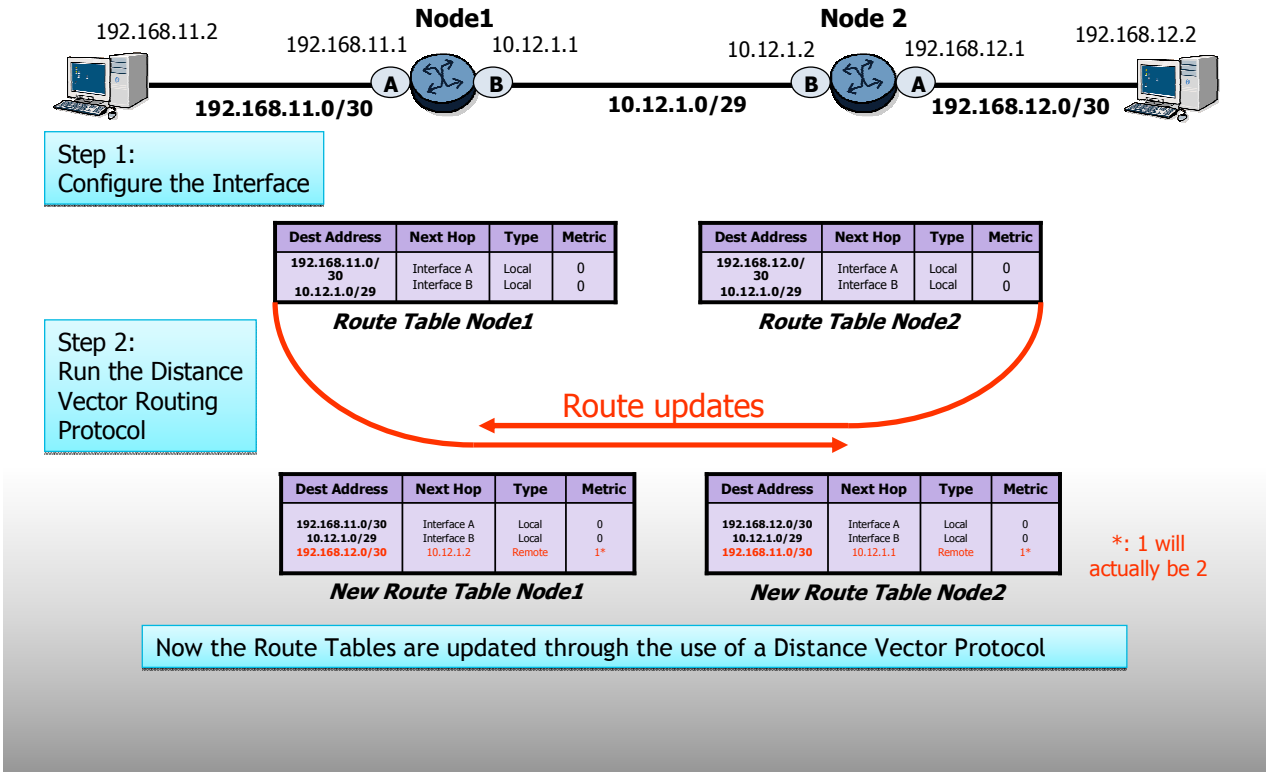
An IPv6 static route is configured in the same way as an IPv4 address. To verify the static route, don't forget to put the keyword "IPv6" after the verification command.

IPv6 Routing Protocols

```
*A:SARF-41# show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix                                 Type    Proto    Age          Pref
       Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
2002:600:A::/48                             Local   Local    04d01h41m   0
       to-sar72                                                      0
3004:BAC:3344:6788::/64                     Remote  Static   00h00m42s   5
       2002:600:A::223:3EFF:FE43:3D2D                                1
4004:BAC:3344:6788::/64                     Local   Local    03d22h06m   0
       to-sar72-ies                                                  0
-------------------------------------------------------------------------------
No. of Routes: 3
===============================================================================
```

Once a static IPv6 route is configured and valid, meaning that the next hop is reachable. It will appear in the routing table.

Distance Vector Routing Protocols



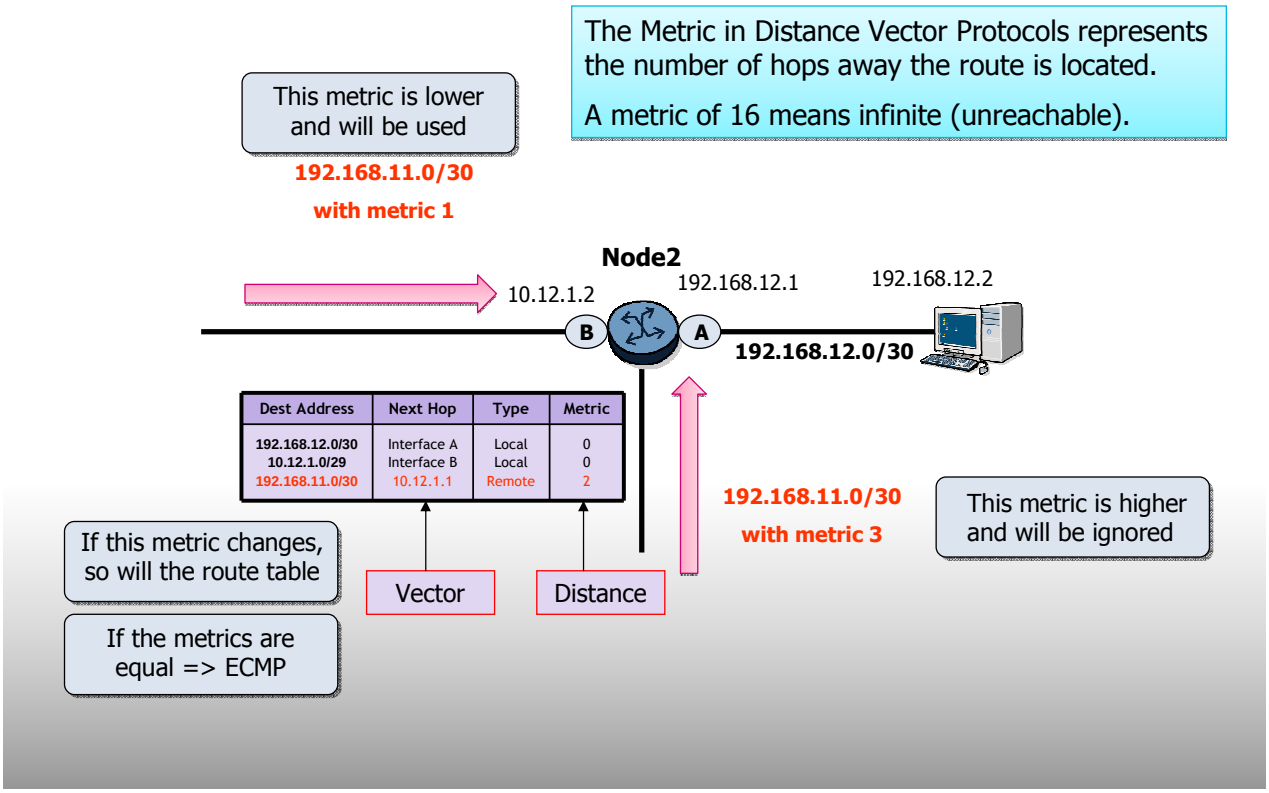
Manually populating routing tables with static routes is not a scalable solution for medium and large networks. Routing protocols are used to find routes dynamically. The Routing Information Protocol was the first IGP. It is based on a Unix program called "Routed". RIP is a Distance Vector Protocol that applies the Bellman-Ford algorithm and periodically sends copies of the Routing Table to neighboring routers.

RIP works as follows:

1. The Administrator creates the interfaces on the router with their respective IP address and subnet mask. Once these interfaces are created, the router knows its directly connected networks and adds these into its routing table with a distance metric of zero.
2. The router periodically sends its entire routing table over these interfaces to its neighbours and receives the routing tables of its neighbours.
3. This way the router will not only know its directly connected networks, but also those of its neighbors. Every time a new route is learned, it is added to the routing table with the metric incremented by 1. The neighbors will learn these new routes with the next update.

This way the entire network will learn all the routes in the network automatically.

Metric



The metric in Distance Vector Protocols represents the distance, how far, of a route as measured by the number of hops away it is. The Vector, in what direction, is indicated by the next-hop value.

When a route is learned that was already in a router's routing table, it will check the metric of the route.

1.If the metric of the new route is better then the one in the routing table, the new route will overwrite the old route.
2.If the metric is worse, the route-update is ignored.
3.If a different metric is received from a neighbor that previously mentioned another metric value, the route in the route table will be updated with the new value.
4.If a metric is received from more then one neighbor with the exact same metric and Equal Cost Multi Path (ECMP) is enabled, duplicate destinations from different neighbors are allowed to co-exist in the route table. This allows load-balancing to occur.

RIPv1 versus RIPv2

The RIP version 1 update packet

Command	Version	Must be zero
Address family identifier		Must be zero
IP address		
Must be zero		
Must be zero		
Metric		

- Uses broadcast
- Up to 25 network entries per update
- Classfull addressing
- No authentication

The RIP version 2 update packet

Command	Version	Unused
Address family identifier		Route Tag
IP address		
Subnet Mask		
Next Hop		
Metric		

- Uses multicast
- Up to 255 network entries per update
- Classless addressing
- Allows Authentication

Alcatel-Lucent implemented RIP version 2, but RIP version 1 characteristics can be simulated to allow interoperability with other RIPv1 routers.

RIP is available in Version 1 and Version 2.

RIP uses UDP to send or receive route updates. The update packet has two major functions:

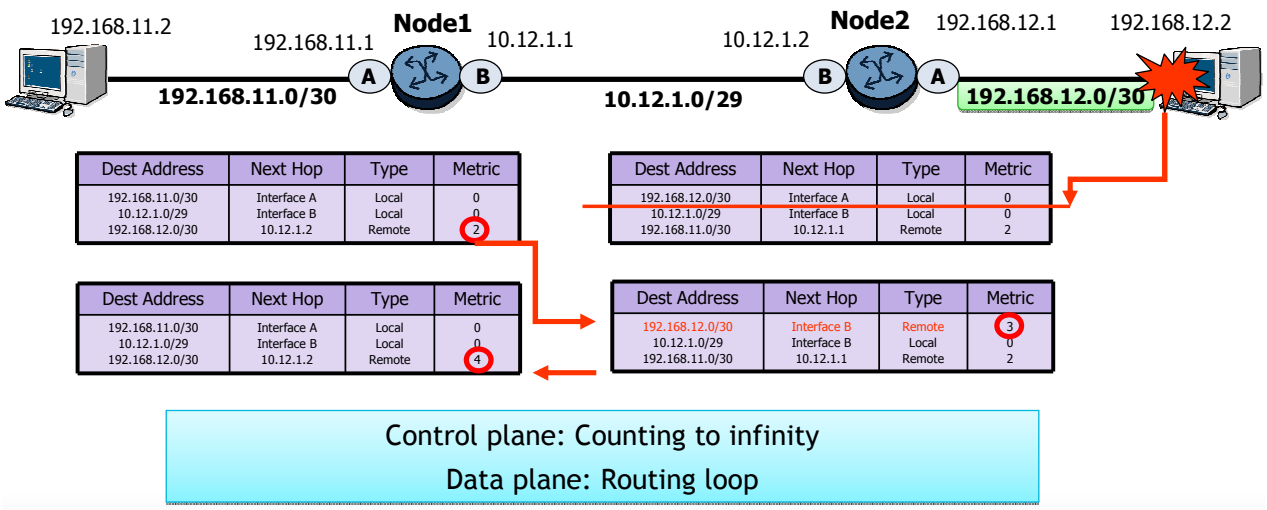
1. To update the routes to its neighbors.
2. To verify if the RIP neighbour is still up and running.

Every 30 seconds a RIP enabled router sends the content of its route table towards its neighbours through the use of update packets. If, after a certain amount of time an update hasn't been received from a neighbour, the routes from that neighbour are removed from the route table.

RIP version 2 was developed with additional functionality, such as classless addressing (allowing VLSM and CIDR), authentication (security), route tagging, and an increase in the maximum number of route entries per packet up to 255.

Alcatel-Lucent routers support RIPv2, but RIPv1 characteristics can be simulated to allow interoperability with other RIPv1 routers.

Split Horizon, Triggered Update, Poisoning



- Split Horizon: Don't send a route update over the same interface on which it was received
- Triggered update: Send an extra update when something changes
- Poisoning: Send a route update with metric 16 (infinity)

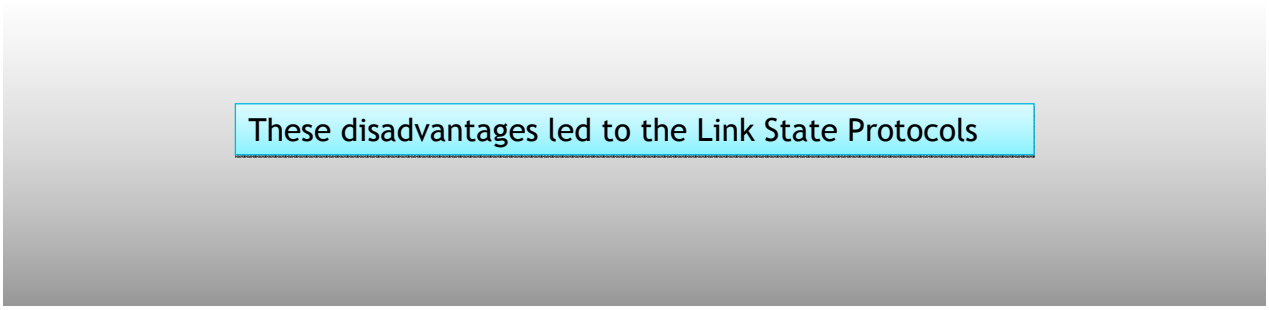
Assume that the link to network 192.168.12.0/30 fails and Node 2 removes its local route directly after the interface detects the failure (layer 1). If Node 2 then receives a route update from Node 1 before sending its altered route table, Node 2 will add the just removed destination but with next hop Node 1 and metric 3. Later, when Node 2 sends its update to Node 1, Node 1 will learn the new metric and increment it by one, the result is 4. This process is called counting to infinity for the control plane and results in a routing loop between Node 1 and Node 2 on the data plane for this route.

This is a major weakness in Distance Vector Protocols. A few remedies were developed to try and solve the problem:

1. **Split horizon**: states that a router can't send a route update out over the same interface on which it has previously learned. In our example this means that Node 1 will never update Node 2 about the route in question, avoiding the routing loop.
2. **Triggered updates**: this allows a router to send an extra route update when something has changed. This greatly reduces the chance of routing updates passing each other.
3. **Poisoning**: instead of not sending a route update at all when a route is no longer valid, the router sends the route with metric 16, so other routers know that something is wrong with this route.

Distance Vector Protocols - Disadvantages

1. A "gossip" principle
2. No bandwidth considerations in the hop-count metric
3. No solution for larger networks (max. 15 hops)
4. Inefficient use of bandwidth by the route update packets
5. Slow convergence (timers)



Distance Vector Protocol have a number of disadvantages:

1. Distance Vector Protocols update the directly connected neighbors with information that a router knows of itself or learned from other routers. This can be compared to gossiping neighbors in a street spreading the news over the fences in the garden. While one router updates his neighbor with information about a route, the original information may already have changed resulting in different network overviews by different routers. This "gossip" behavior is a direct result of the Distance Vector Protocol principle.

2. The hop-count metric does not take link bandwidth into account when calculating the shortest route.

3. The hop count is limited to 15 hops which doesn't scale well for larger networks.

4. There is considerable overhead in the route update packets. Every 30 seconds the entire route tables of all the routers travel through the network, even though nothing may have changed in that period of time.

5. Route convergence time, the time required for all the routers to have the same current network topology in their route table, is quite slow because of the periodic nature of the updates.

These disadvantages of Distance Vector IGPs lead to the development of a new type of IGP, the Link State Protocols.

Link State Protocols Overview

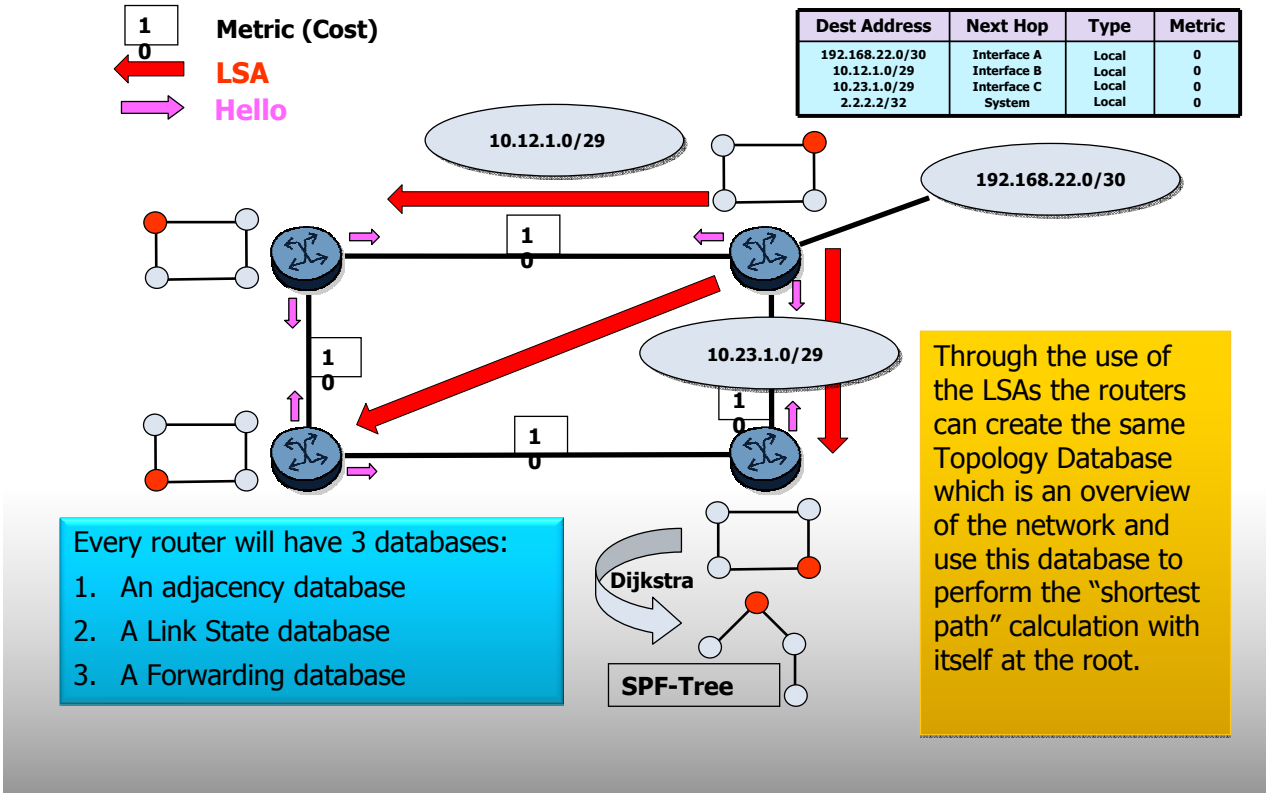
1. A new principle: Link State Update (LSU), directly sent to all the other routers (flooding) in the AS, contain information on the local "Links" (metric) and their "State" (Active or not-Active)
2. A new metric system, cost- or bandwidth based, no longer limited to a hop-count of 15 resulting
3. No more periodical updating, triggered updating only => Fast Convergence
4. A dual level hierarchy
5. Scalability: allows up to thousands of nodes in an AS
6. Two protocols simultaneously developed:
    1. TCP/IP model: OSPF
    2. OSI model: IS-IS

Routing information distributed by Distance Vector protocols is not very reliable since it was not received directly from the source, but from a third party, a direct neighbor. It would be better if every router could pass local information directly to all other routers.

This idea resulted in a new generation of IGPs, the Link State Protocols. Routers running these protocols distribute local information to every other router in the network using Link State Advertisements (LSAs). This "flooding" of LSAs assures that every router has information on routes coming from the most reliable source, the updating router itself. This new approach radically changed Intra-AS Routing and resulted in two IGPs, Open Shortest Path First and Intermediate System to Intermediate System. In addition to changing the way that routing information is updated, these Link State Protocols offer other advantages:

1. A new metric system, cost or bandwidth based, that is no longer limited to 15 hops.
2. Instead of sending complete updates periodically, only topology changes are sent out, at the time they occur. This reduces the overhead of route updates compared to the Distance Vector Routing Protocols.
3. A dual-level hierarchy, together with the unlimited metric allows the network to contain many more nodes so the new IGPs are scalable for large Autonomous Systems.

Link State Protocols – Link State Advertisements



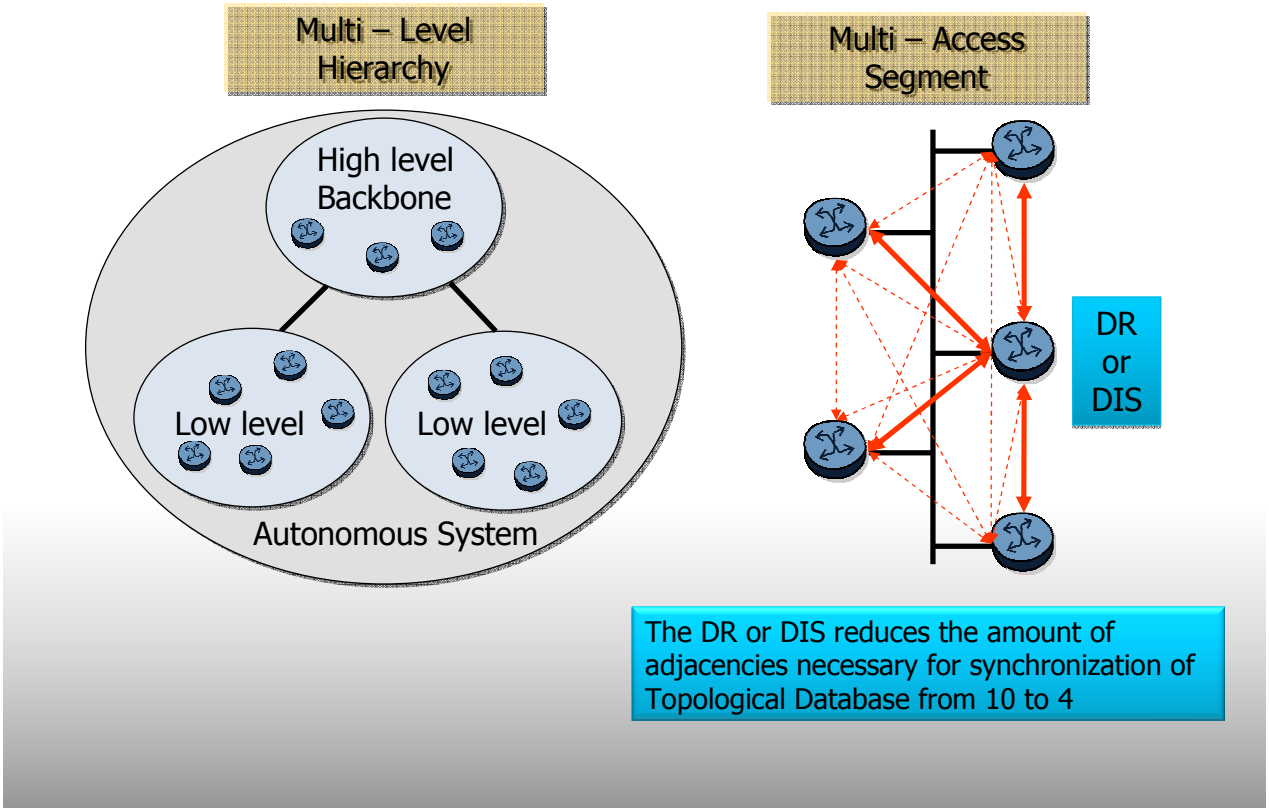
LSAs are sent using reserved multicast addresses. The LSAs tell all the other routers what Links are directly connected and in what State they are. This way all the routers, using the LSAs from the other nodes, can create a network topology, represented by the Link State or Topological Database. This is a crucial difference from the Distance Vector Routing Protocols because now every node has the exact same Link State Database. Each node can now perform a calculation, known as the Dijkstra Algorithm, to create an SPF-tree with itself at the root. Using this SPF-tree the shortest paths to all destination are known resulting in a single entry in the Forwarding Database: The Route Table.

LSAs:

- Are sent out and acknowledged by all the other routers.
- Have a sequence number
- Are sent out when:
  - a Router or its IGP instance comes up
  - a new Link (interface) is added (Active) to the IGP instance
  - a Link fails or is removed from the IGP instance (Inactive)
  - the Hello Protocol has found or lost a neighbor
  - a keep-alive timer has expired.

A third database, called the Adjacency or Neighbor Database, contains all the directly attached neighbors or adjacencies, that were discovered by the Hello-protocol.

Link State Protocols – Hierarchical Structure



Link State Protocols introduced the idea of hierarchies. A router will only learn routing information at its own level, information between levels is handled by specially designated routers. This concept was introduced to reduce the size of link state databases and the number of LSAs propagated in the network.

Designated Routers for OSPF or Designated Intermediate System Routers for IS-IS are found in multi-access segments in the network. There are also backup DRs in case of a DR failure.

Say, "n" nodes are connected on such a segment. It is necessary that all these nodes are synchronized between each other to obtain the same topological Database which implies a full mesh of adjacencies, that will exponentially increase with the number of nodes on this multi-access segment: n*(n-1)/2. To reduce this number of adjacencies, the DR or DIS is introduced. This Router acts as a single source for all routing updates, which means that all the routers do not have to constantly update each other, and can now get all their updates from a single source.

2

OSPF

Chapter 1 – XXX.

OSPF Overview

- OPSF workgroup started in 1987, lead by John Moy
- Present standard: RFC 2328: OSPFv2
- Cost = reference bandwidth (100Gbps) / link bandwidth
- IP and MAC multicast addressing
  - 224.0.0.5 equals MAC 01-00-5E-00-00-05 for any OSPF speaker
  - 224.0.0.6 equals MAC 01-00-5E-00-00-06 for any DR/BDR speaker
- No Transport Layer Protocols (TCP/UDP): OSPF works on top of IP (Protocol field: 89)
- Router ID address is necessary, and may be
  - Last 32 bits of chassis MAC address
  - The system IP address
  - The router-id in the config>router# context
- Authentication
  - No Authentication
  - Simple Authentication
  - MD5 Authentication

Open Shortest Path First (OSPF) is a link state protocol. The present standard is the OSPFv2 based on RFC 2328 for IPv4..

OSPF Link Cost = a reference bandwidth (100Gbps) divided by the link bandwidth. For example a 10 Gbps link will have a cost of 10, but the costs are configurable if the administrator prefers other values.
OSPF uses IP and MAC multicast addresses, instead of broadcast, to distribute LSAs. This reduces the overhead on other devices on the same segment that are not running OSPF.
224.0.0.5 equals MAC 01-00-5E-00-00-05 for any OSPF speaker
224.0.0.6 equals MAC 01-00-5E-00-00-06 for any DR/BDR speaker
OSPF works on top of Layer 3 in the OSI model. It uses no transport layer protocol such as TCP or UDP but duplicates their functions through the use of acknowledgements and sequence numbers.
OSPF requires each OSPF speaker to have a Router ID (RID) to identify the node in the Link State Databases. This RID can be:
The last 32 bits of the chassis MAC address (default),
The system IP address (overrides the default),
The Router ID defined in the **config>router#**router-id context (overrides the system IP address).
OSPF supports three types of authentication
None (default)
Simple authentication and
MD5 authentication

OSPF Packet Types

- OSPF Hello
- OSPF Database Descriptor
- OSPF Link State Request
- OSPF Link State Update
  - Type 1: Router
  - Type 2: Network
  - Type 3: Summary
  - Type 4: ASBR
  - Type 5: External
  - …
- OSPF Link State Acknowledgement

OSPF uses 5 different types of packets to establish and maintain router connectivity and network convergence:

1. **Hello Packets:** generated by all OSPF speaking routers. They are used to discover neighbors, and to form and maintain adjacencies. They are propagated periodically (10 seconds).
2. **Database Descriptors:** used to distribute a summary of all the networks in the router's database. Typically this will be the classless network, routers cost to access, and the sequence number associated with that network entry.
3. **Link State Requests:** used to request additional information on a particular network entry on which the present information is non-existant or outdated after comparing the Database Descriptors.
4. **Link State Updates:** respond to Link State Requests by using the requested LSAs. There are many forms of LSAs available.
5. **Link State Acknowledgments:** acknowledge every newly received LSA. Many acknowledgments may be grouped together into a single Link State Acknowledgment packet.

OSPFv3

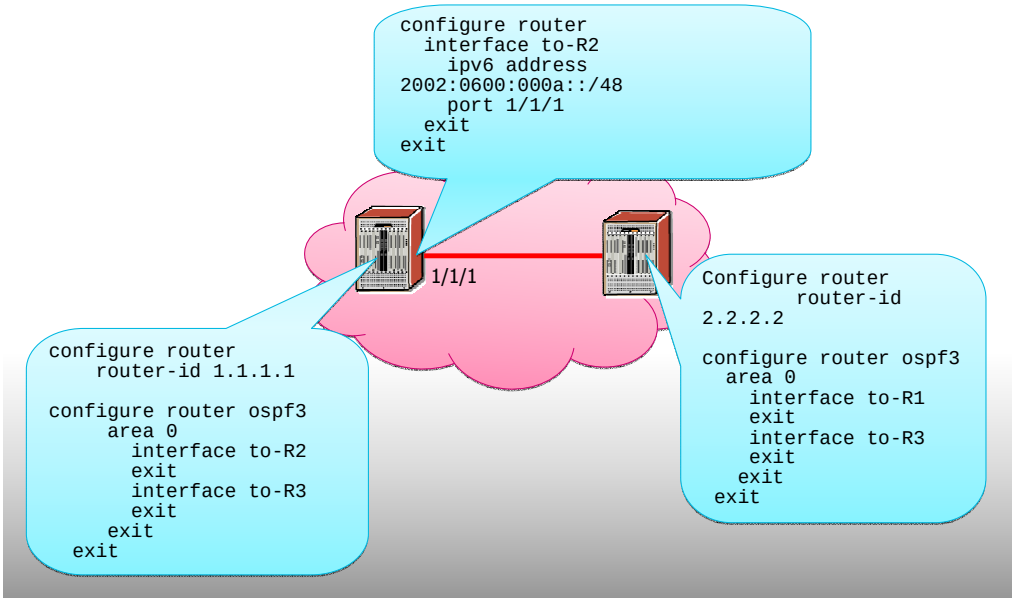
IPv6 -> new version of OSPF is introduced!

- OSPFv3 is defined in RFC 2740
- Fundamental mechanism of OSPF remain unchanged
  - Area support
  - SPF calculations
  - LSA flooding
  - DR/BDR election
- OSPFv2 and OSPFv3 can run independently on the router.

When IPv6 was introduced, a new version of OSPF was needed. However the fundamental mechanism of OSPF remained unchanged.

The IPv4 version of OSPF, version 2, can run simultaneously on a router and work independently.

OSPFv3 Configuration Example



A basic configuration of OSPF version 2 or version 3 is straightforward. Once the router interface is configured, meaning it contains at least an address and port assignment, it can be put under the OSPF context and under a certain OSPF area. All other signal messaging, LSA and route distribution goes automatically.

OSPFv3

```
 show router ospf3 neighbor

===============================================================================
OSPF Neighbors
===============================================================================
Interface-Name                   Rtr Id          State      Pri  RetxQ   TTL
-------------------------------------------------------------------------------
to-r2                            2.2.2.2         Full       1    0       39
to-r3                            3.3.3.3         Full       1    0       39
-------------------------------------------------------------------------------
No. of Neighbors: 2
```

The state of "full" is a final state where neighbors have successfully interchanged their databases and they are fully aligned. States other than "full" or "two way" may exist because of mis-configuration or mismatches.

OSPFv3

```
show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================

Dest Prefix                                 Type    Proto    Age         Pref
       Next Hop[Interface Name]                                  Metric
-------------------------------------------------------------------------------
2002:C0A8:6301::/48                         Local   Local    02d01h55m   0
       [to-r2]                                                   0
2002:C0A8:6302::/48                         Local   Local    01d23h50m   0
       [to-r3]                                                   0
2002:C0A8:6303::/48                         Remote  OSPF3    00h09m00s   10
       FE80::C29D:FFFF:FE00:0-"to-r3"                            2000
-------------------------------------------------------------------------------
No. of Routes: 3
===============================================================================
```

After the interfaces are in “full” state, route will appear automatically in the routing table.

3

ISIS

Chapter 1 – XXX.

IS-IS Overview

- Developed by DEC, later handed over to International Standards Organization (ISO)
- Cost = 10 by default, configurable (reference bandwidth setting can be applied)
- MAC multicast addressing
  - Level 1 updates use 01-80-C2-00-00-14
  - Level 2 updates use 01-80-C2-00-00-15
- No Network or Transport Layer Protocols: IS-IS works on top of Ethernet 802.3 LLC Encapsulation (DSAP/SSAP: 0xFEFE)
- IS-IS uses an NSAP (Network Service Access Point) addressing
- System ID address is necessary (4 bytes), and may be
  - The chassis MAC address
  - Derived from the system IP address
  - Derived from the router-id in the config>router# context
- Authentication
  - No Authentication
  - Simple Authentication
  - MD5 Authentication

Another Link State Protocol is Intermediate System to Intermediate System or IS-IS. An Intermediate system, in comparison to an end-system, is a router. This protocol was developed by Digital Equipment Corporation and later handed over to the International Standards Organization who implemented it into the OSI protocol stack. Later this OSI protocol was extended to support IP and called Integrated IS-IS. OSPF and IS-IS are extremely similar Link State Protocols and are evenly used amongst networks worldwide.

Characteristics of IS-IS.

•The cost of the links is by default 10 but is configurable.

•IS-IS uses L2 multicasting instead of L3 multicast in OSPF:

- Level 1 updates use 01-80-C2-00-00-14
- Level 2 updates use 01-80-C2-00-00-15

•IS-IS works on top of layer 2 in the OSI model, it uses no IP or Transport Layer Protocol but takes care of their responsibilities itself through the use of acknowledgements and sequence numbers.

•IS-IS uses an ISO standardized network addressing method, called NSAP or Network Service Access Point addressing.

•IS-IS requires a System ID of 4 bytes which can be derived from the chassis MAC address, the system address, or the router-id, configured in the **config>router#** context.

•IS-IS supports three types of authentication

- None (default).
- Simple authentication.
- MD5 authentication.

IS-IS Protocol Data units (PDU)

- IS-IS Hello PDU
- IS-IS LSP (Link State PDU)
- IS-IS PSNP (Partial Sequence Number PDU)
- IS-IS CSNP (Complete Sequence Number PDU)

IS-IS uses four different types of packets, called Protocol Data Units or PDU's to establish and maintain router connectivity and network convergence:

1. Hello PDUs: generated by all devices running IS-IS mainly to discover neighbors, to form and to maintain adjacencies.

2. Link State PDUs or LSP's: generated to distribute information about neighbors and links used to create the topological database.

3. Partial Sequence Number PDUs or PSNP's: used to request specific information about a network. A PSNP can:

1.contain a subset of LSPs in the database

2.acknowledge one or more LSPs

3.request transmission of specific LSPs

4. Complete Sequence Number PDUs or CSNP's: are multicast periodically by the ISIS routers to maintain database consistency. The CSNP lists every LSP in the database. There can be several CSNPs sent at one time when the database is quite large.

ISIS IPv6

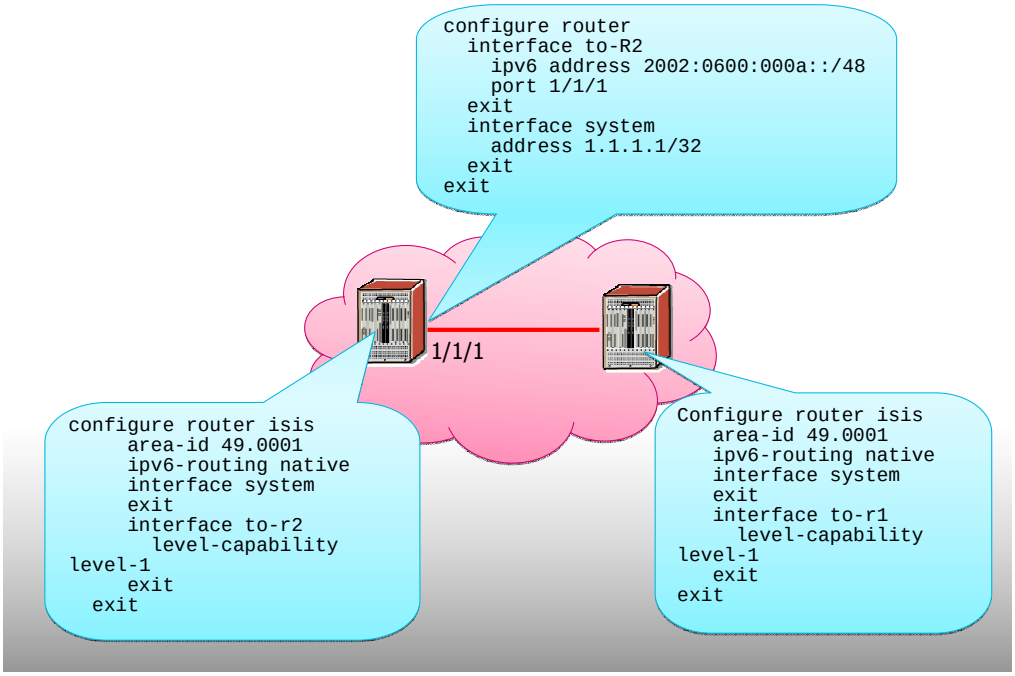
ISIS for IPv6 is defined in routing IPv6 with IS-IS "draft-ietf-ipv6-05.txt"

New TLV's are created for IPv6

- IPv6 Reachability (0xEC)
  - Equivalent to IPv4 "IP Internal Reachability Information" and "IP External Reachability Information" with External bit set
  - IPv6 routing prefix, metric information
- IPv6 Interface Address (0xE8)
  - Equivalent to "IP Interface Address"
  - Contain 16 Octet IPv6 interface address
  - Hello contain "link-local address", LSP contain "non-link-local address".
- New Network layer protocol Identifier (0x8E) is used in IPv6 ISIS.

ISIS can also be used for IPv6. Because ISIS works with TLV's or type/length/value fields, no new protocol stack was needed as with OSPF. Only new TLV's were defined. The IPv6 reachability, IPv6 interface address and the new network layer protocol identifier TLV are the new TLV's.

ISIS IPv6 Configuration Example



Once the interfaces are configured under the router context, they can run ISIS by bringing them under the CLI context of isis.

This procedure is the same for IPv6 and IPv4.

In addition, IPv6 routing is enabled by specifying "ipv6-routing native" or "ipv6-routing mt" or multi topology.

ISIS IPv6

```
show router isis adjacency

===============================================================================
ISIS Adjacency
===============================================================================
System ID                       Usage State Hold Interface
-------------------------------------------------------------------------------
simpc-2                         L1    Up    7    to-r2
simpc-3                         L1    Up    22   to-r3
-------------------------------------------------------------------------------
Adjacencies : 2
===============================================================================
```

Before checking the routing table, verify that the adjacency is up.

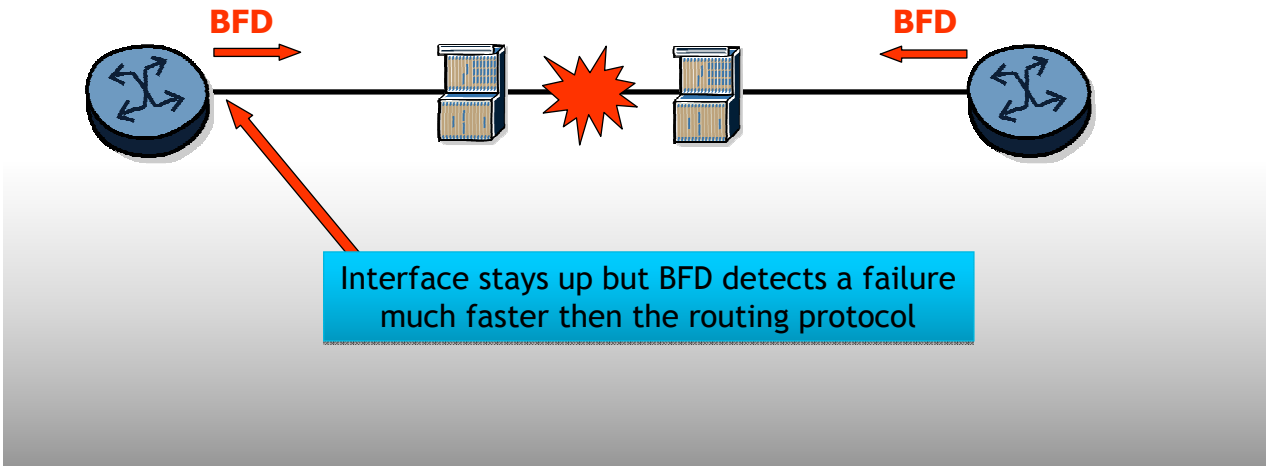
ISIS IPv6

```
show router route-table ipv6

===============================================================================
IPv6 Route Table (Router: Base)
===============================================================================
Dest Prefix                                     Type    Proto    Age         Pref
       Next Hop[Interface Name]                                       Metric
-------------------------------------------------------------------------------
2002:C0A8:6301::/48                             Local   Local    00h33m29s   0
       [to-r2]                                                          0
2002:C0A8:6302::/48                             Local   Local    00h28m27s   0
       [to-r3]                                                          0
2002:C0A8:6303::/48                             Remote  ISIS     00h19m46s   15
       FE80::C2B7:FFFF:FE00:0-"to-r2"                                   20
-------------------------------------------------------------------------------
No. of Routes: 3
===============================================================================
```

IPv6 learned routes will appear in the routing table.

Bidirectional Forwarding Detection (BFD)

- BFD sends control messages periodically to the far end
  - Sub-second detection of path failures
  - Supplements IGP hello and discovery mechanisms
  - Failure to receive the expected number of messages in the configured time interval causes the far end node to be declared down
- Increases convergence rate of IGP protocols; detects static route failures



Bidirectional Forwarding Detection is a high availability feature that speeds up convergence time when a link, a node, or a protocol fails by rapidly detecting the failure. If, for example, a fiber link fails between two Add Drop Multiplexers and the interface isn't aware of the failure, it is up to the routing protocol to discover the failure. This can be longer than is desirable.

BFD is an IETF feature designed to improve this detection time. It is a lightweight, low-overhead protocol that creates sessions between the routing protocol instances of two adjacent routers and sends very small packets back and forth to evaluate the condition of the link. If something should go wrong, BFD will detect this in a sub-second time interval and update the routing protocol it works for. BFD is implemented in the control plane of routers and other systems. A network failure detected by BFD can be corrected by the forwarding plane, or by the control plane.

BFD can detect failures at many different layers, and can therefore be employed to monitor the validity of Ethernet networks, MPLS Label Switched Paths (LSPs), Generic Routing Encapsulation (GRE), or virtually any other type of transport.

A separate BFD session is created for each routing protocol. BFD on the 7750 currently supports OSPF, IS-IS, PIM and static routes.

BFD CLI Commands

Configure BFD intervals on the interface

PE# configure router interface <interface-id>
bfd <transmit-interval> [receive <receive-interval>] [multiplier <multiplier>]
no bfd

Enable BFD for the static route

PE# configure router static-route <ip-prefix> next-hop <interface> bfd-enable

Display BFD interfaces and sessions

PE# show router bfd
interface [<interface-id>]
session [src <ip-addr> [dest <ip-addr>] | [detail]]

Clear BFD sessions and counters

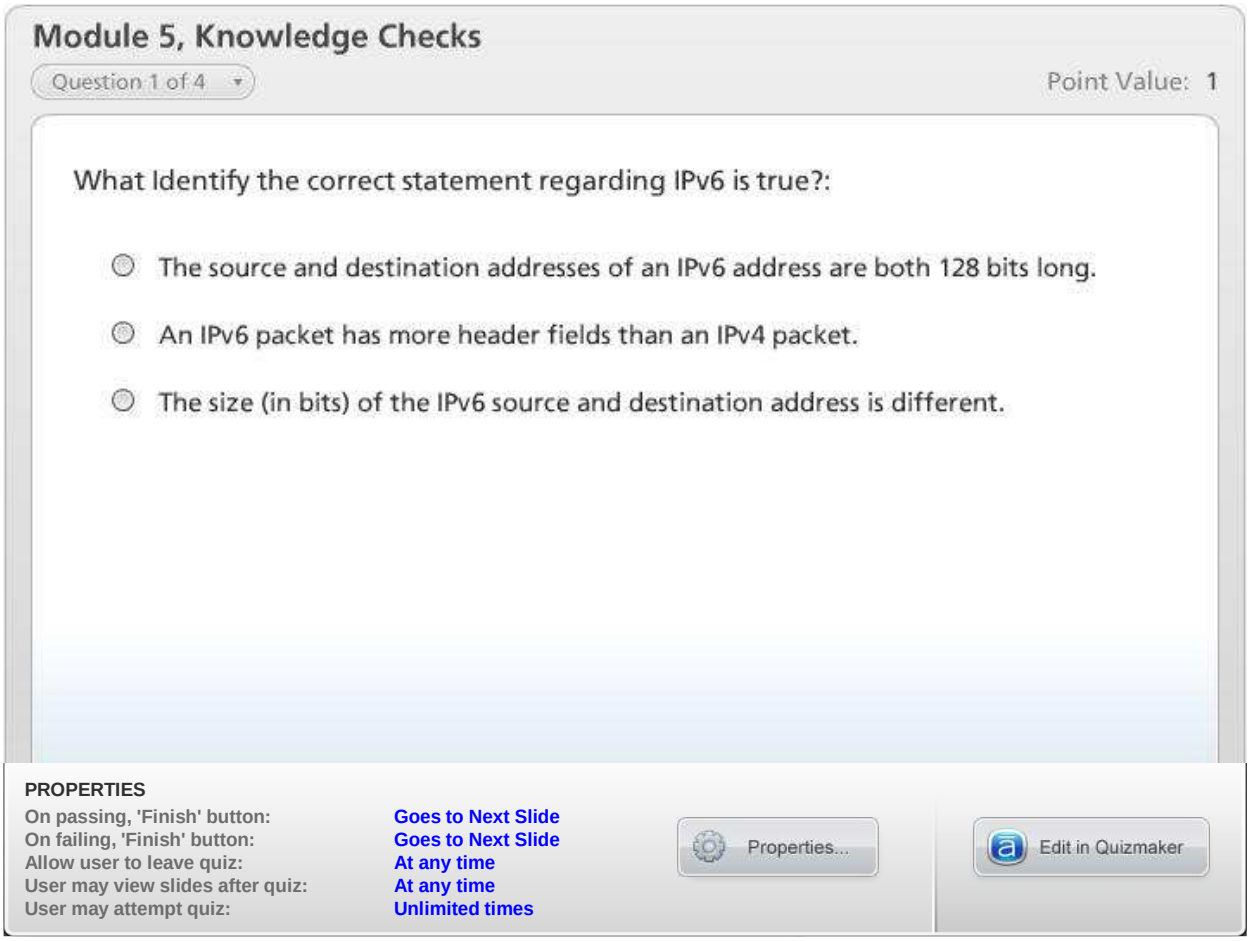
PE# clear router bfd
session src-ip <ip-addr> dst-ip <ip-addr>
session all
statistics src-ip <ip-addr> dst-ip <ip-addr>
statistics all

The commands to configure, enable, display and clear a BFD session.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 5

Alcatel·Lucent

This completes module 5. This module explained the basics of IP routing and highlighted the difference between a link state protocol and a distance vector protocol. While MPLS is the dominant technology for building L2 or L3 services, there is always a good IP infrastructure needed. The SR-OS product family also fully supports IPv6 and the all of the major routing protocols like OSPF, IS:IS, RIP and BGP.

SR-OS Fundamentals
Module 6: MPLS

IPD Development

Available as PDF

Welcome to the sixth module of the SR-OS fundamentals course.

This module explains the fundamentals of MPLS and the two Label Distribution mechanisms.

Table of Contents

Module 6 is divided into five sections.

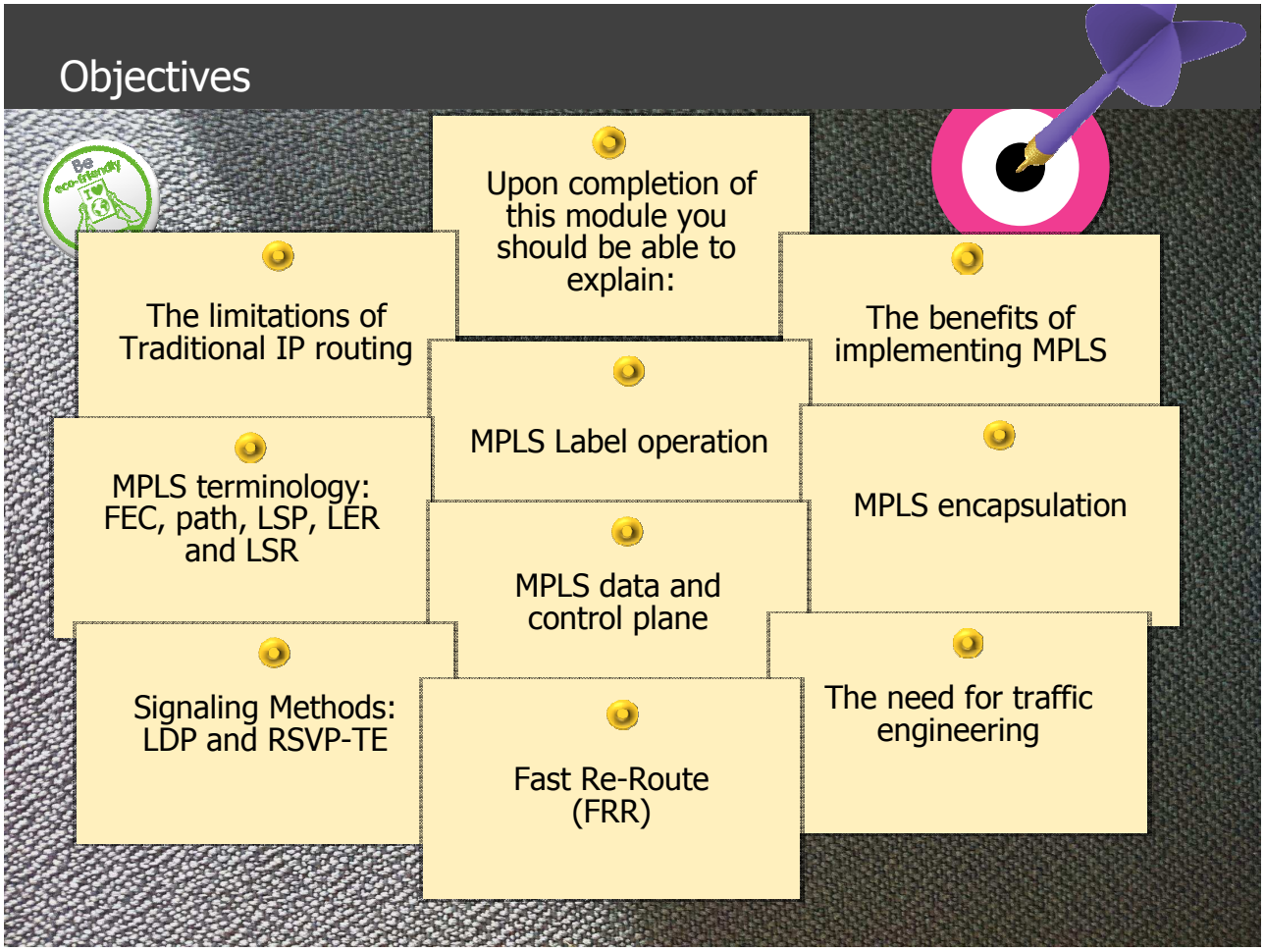
Section 1 introduces the fundamental principles of the MPLS (MultiProtocol Label Switching)

Section 2 talks about static label distribution

Section 3 introduces the first dynamic label distribution protocol, LDP (Label Distribution Protocol)

Section 4 explains the more advanced label distribution protocol RSVP-TE (Resource Reservation Protocol – Traffic Engineering) with and without CSPF (Constraint Shortest Path First)

And the last section talks about the fast re-route capabilities of MPLS



By the end of this module you will be able to explain:

The limitations of traditional IP routing.

The benefits of implementing MPLS (MultiProtocol Label Switching).

MPLS Label operation.

MPLS terminology: FEC (Forward Error Correction), path, LSP (Label Switched Path), LER (Label Edge Router) and LSR (Label Switching Router).

MPLS encapsulation.

MPLS data and control plane.

The MPLS Sgnaling Methods: LDP (Label Distribution Protocol) and RSVP-TE (Resource Reservation Protocol – Traffic Engineering).

The need for traffic engineering.

And fast reroute.

1

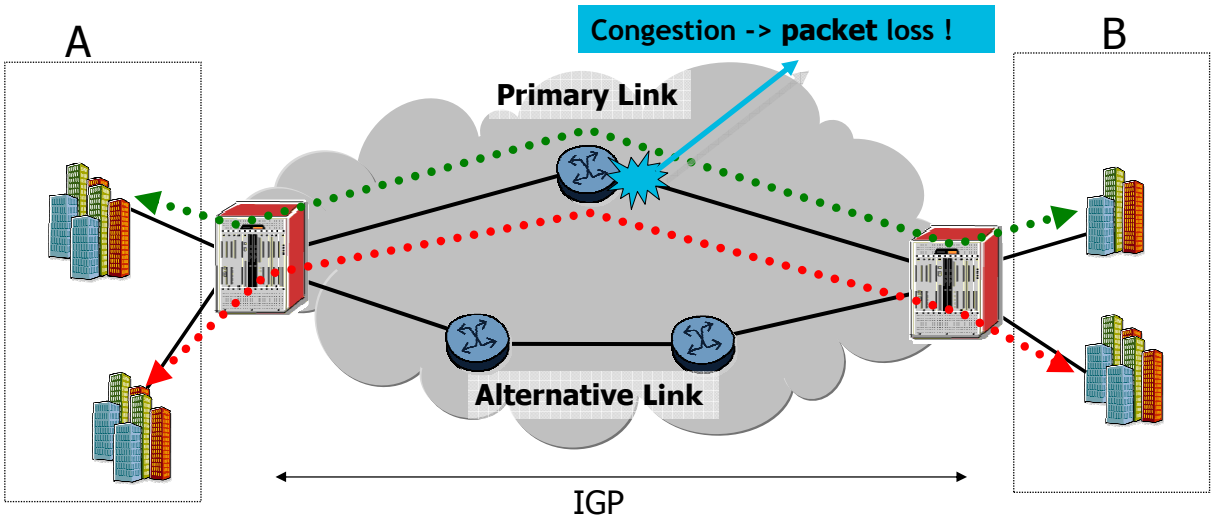
MPLS fundamentals

Section1: MPLS fundamentals

Limitations of Traditional IP Routing: Hyper-Aggregation

IP has a limited ability to alleviate hyper aggregation which leads to network and link congestion.

- All network traffic will flow via the primary path
- No traffic will use the Alternate Link → Inefficient use of resources



One of the major limitations in a classical IP routing network is that packets entering the network will follow the IGP (Interior Gateway Protocol) shortest path based on the metric calculations.

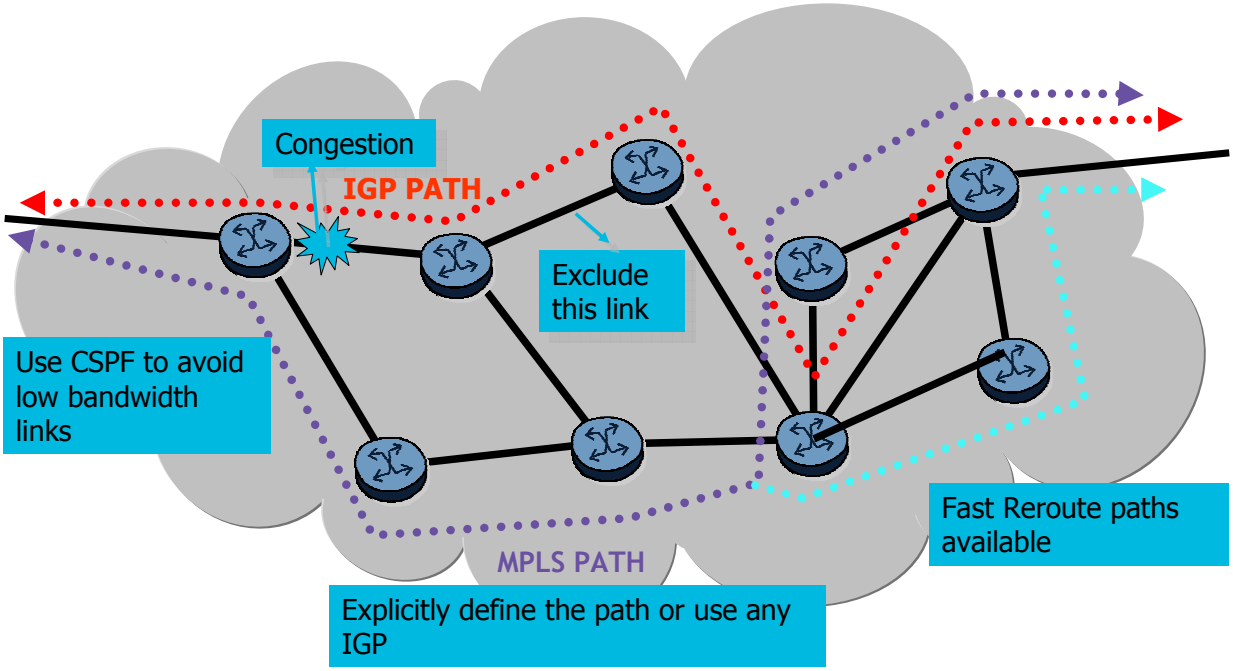
This implies that packets from different routers at side A that need to send traffic to side B over the IP network, will follow the same path defined by the IGP topology. This can lead to congestion and under-usage of other available resources. MPLS provides a whole array of tools to allow a network architect to direct traffic over other paths depending on criteria such as bandwidth, link colors, hop limits and manual configuration of the path preference.

The IP protocol alone is also unable to provide guaranteed service levels across a network end to end. Certain applications may require a guaranteed amount of bandwidth and service providers want to be able to offer service level agreements to their customers guaranteeing bandwidth and resource availability.

Traditional IP routing protocols does not consider the available bandwidth in the network across the primary or alternate links. This means that routers in the network are not aware of what bandwidth resources are available over primary or alternate paths.

As shown in the example, if all packets take the primary path the link may soon become congested, resulting in packet loss.

Benefits of MPLS



MPLS will not send data packets faster through the network compared to IP packets, but the main advantage is the way traffic is sent through the network. MPLS has more capabilities and flexibility to influence the direction in which packets are travelling through the network. MPLS can avoid certain links, use links with a certain min available bandwidth and use explicitly defined paths.

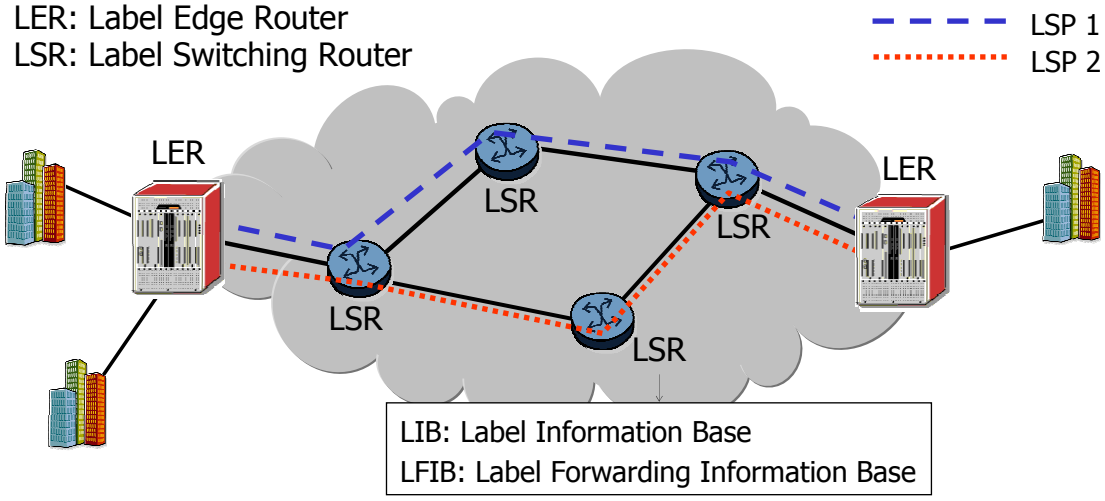
Another main advantage of MPLS is the fast re-route back-up mechanism that allows traffic switchover in less than 50 milliseconds.

A major difference compared to IP is that after path setup, data forwarding is done independently of the destination address.

MPLS Terminology

Multi Protocol Label Switching, MPLS:

- Allows the network administrator to create end-to-end LSPs for data forwarding instead of using hop-by-hop IP routing.
- Alleviates the limitations of traditional IP routing



MPLS allows the network administrator to create "tunnels" from the ingress to the egress of the MPLS domain, usually a provider's network. These end-to-end tunnels alleviate the limitations of traditional IP hop-by-hop routing, since the tunnel does not have to follow the path chosen by the IGP and it can be created with certain characteristics such as bandwidth requirements, allowing SLAs to be met.

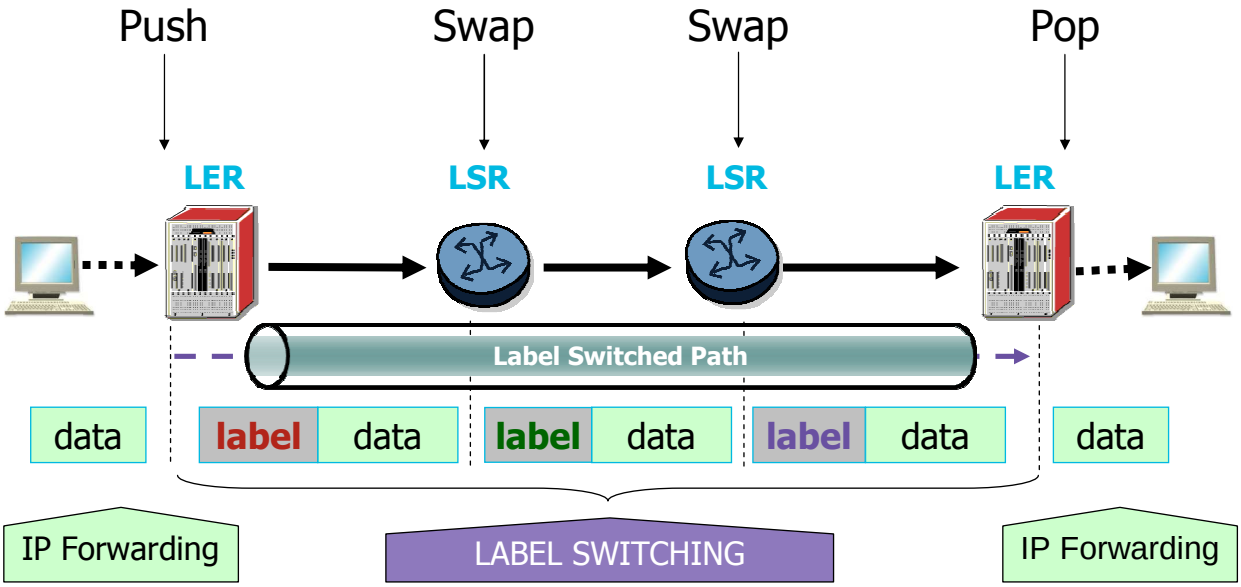
An LER sits at the edge of an MPLS domain, either at the beginning or at the end.

An LSR is a device that typically resides somewhere in the middle of an MPLS network and is capable of forwarding datagram's based upon a label. An LSR will develop a Label Information Base or LIB and a Label Forwarding Information Base or LFIB to switch packets according to certain label values assigned to each packet.

The unidirectional tunnel set up across the MPLS network between two LERs is called an LSP or label switched path. Label Switch Paths are established by network operators for a variety of purposes, such as to guarantee a certain level of performance, to route around network congestion, or to create IP tunnels for network-based virtual private networks or VPN's.

MPLS is the preferred tunnel mechanism for VPN's.

MPLS Basic operation



The MPLS LSPs are virtual "tunnels" created through the use of labels signaled between MPLS speaking routers. Once a label is chosen at the ingress, the label header is put at the front of the packet header so that the label value can be carried across the network with the packet. At each subsequent hop, the Label Switch Router simply looks up the label value in a LFIB table to make the forwarding decision. There is no need here to parse the IP header. Since the label is a fixed length, label look-up is fast and simple.

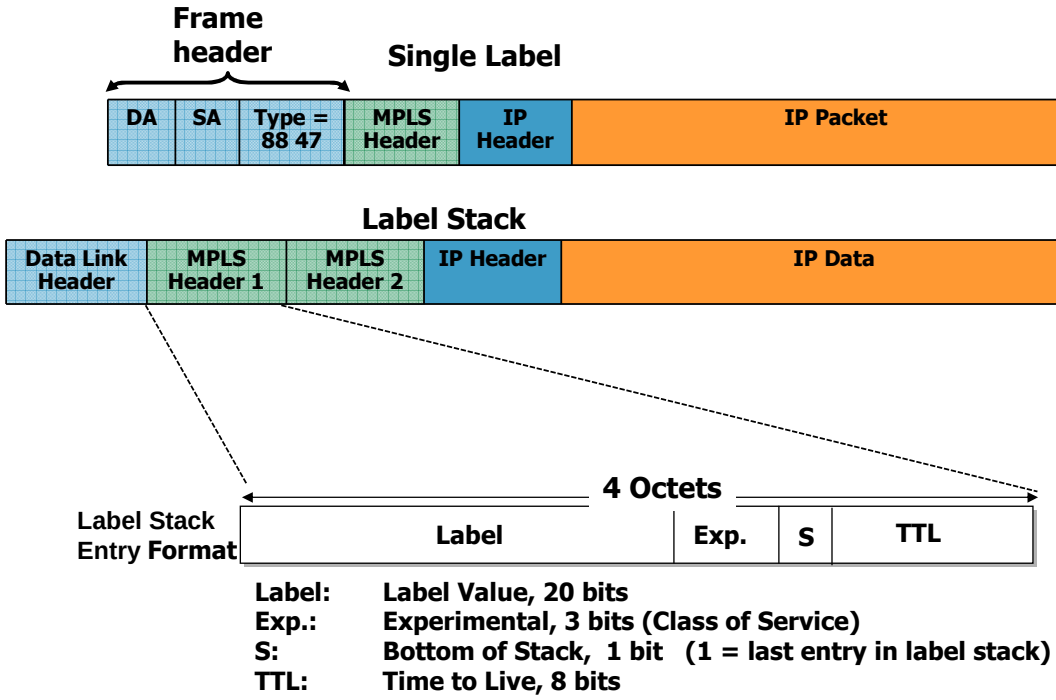
MPLS performs three basic operations:

Push - Puts a label onto the packet.

Swap - Swaps or switches a label received at an ingress interface with a label at an egress interface of the node. The label swapping is done in accordance to the Label Forwarding Information Base.

Pop - Removes the label from the packet.

MPLS Label



An MPLS label consists of 4 octets where 20 bits are reserved for the label value itself.

The next three bits after the label, the EXP or experimental bits are dedicated for quality of service.

After the EXP bits, there is one bit that indicates if there are other labels stacked.

The last 8 bits are used to encode a TTL or time to live value.

In MPLS, packets can carry not just one label, but a set of labels in a stack. An LSR can swap the label at the top of the stack, pop the stack, or swap the label and push one or more labels into the stack. The processing is always based on the top label, without regard for the possibility that some number of other labels may have been above it in the past, or that some number of other labels may be below it at present.

MPLS Label

```
A:S141# show router mpls label-range

===============================================================
Label Ranges
===============================================================
Label Type      Start Label     End Label       Aging         Total Available
---------------------------------------------------------------------------
Static-lsp           32            1023          -            992
Static-svc         2048           18431          -          16384
Dynamic           18432          131071          0          112640
===============================================================
```

To show the label range use the command; "show router mpls label-range"

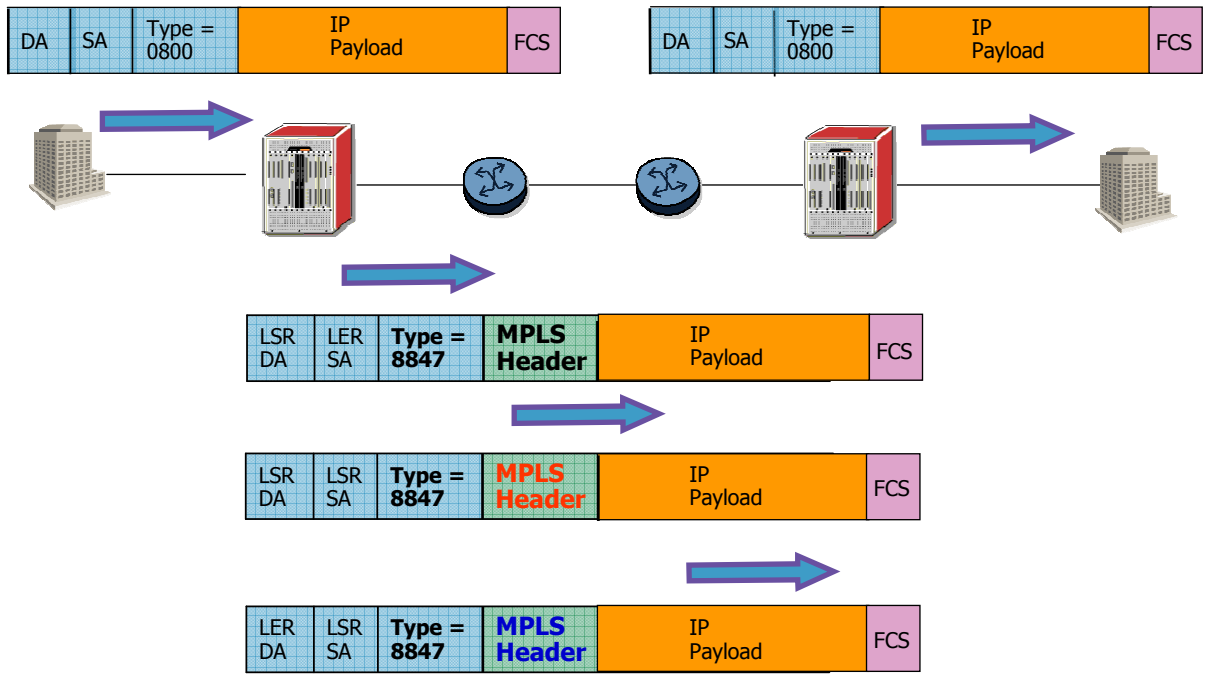
Labels are divided in to three main ranges.

The smallest range is used to create static lsp's.

More than 16 thousand labels are used as static-service labels.

And the last range of labels is the most important one as these are used by the two label distribution protocols, LDP and RSVP-TE, but also used as dynamic service labels.

MPLS Encapsulation: The Shim Header



The MPLS label values are carried in an MPLS Shim header. This means that the label is inserted in between the Ethernet and the payload.

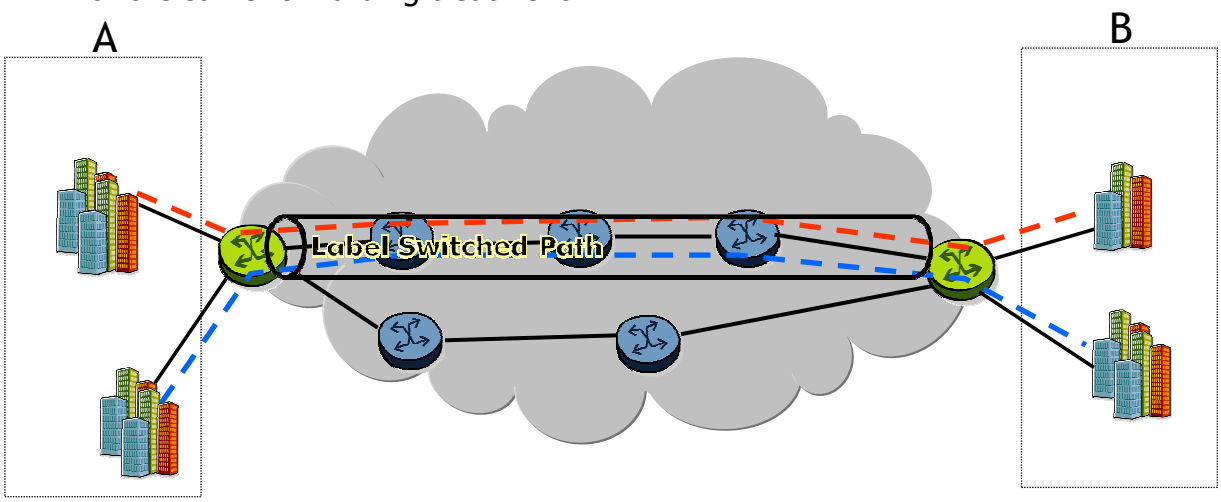
As defined by the Internet Assigned Numbers Authority IANA the Ether type field of an Ethernet frame must be 0x8847 for MPLS Unicast traffic.

This example illustrates that the payload of the customer is encapsulated into an MPLS frame and sent to the egress LER. Each time the frame is passing an LSR router, the MPLS label is changed or swapped.

Forwarding Equivalence Classes

Forwarding Equivalence Class (FEC)

A group of IP packets forwarded in the same manner, over the same path, with the same forwarding treatment



MPLS packets are assigned to a FEC at the network ingress

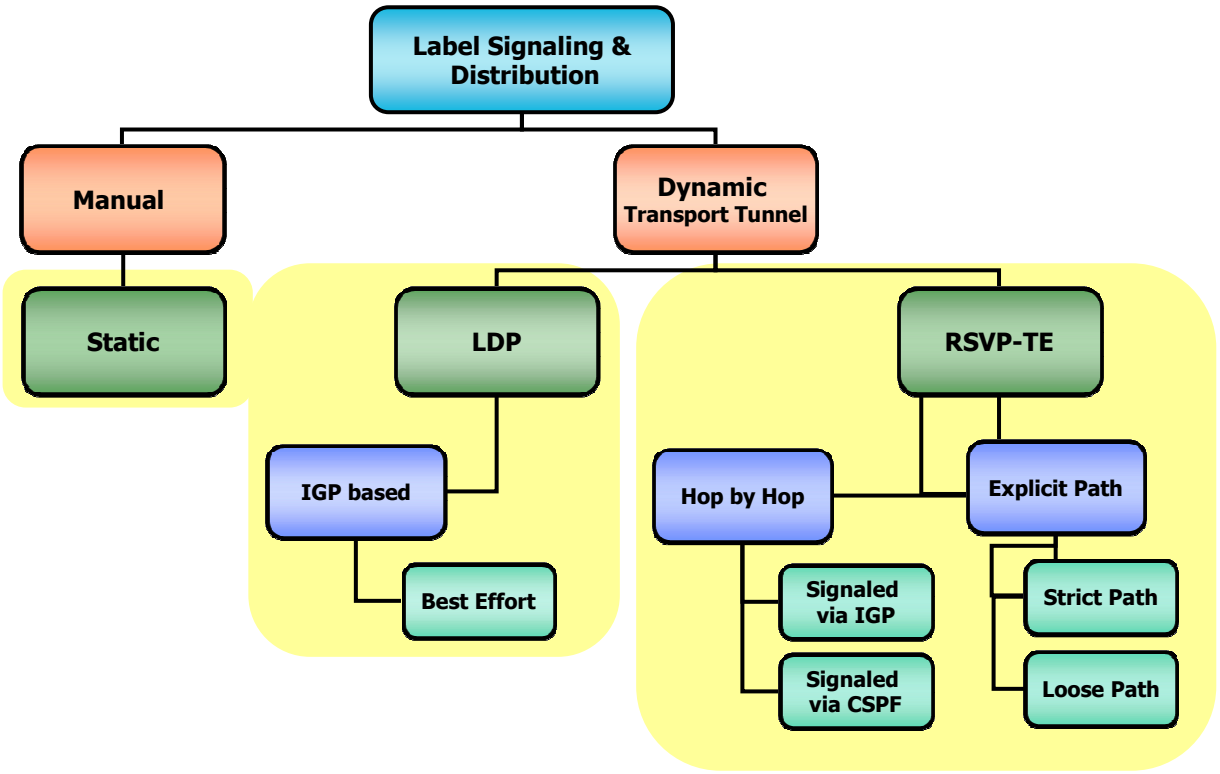
- FEC packets are transported by the LSP which has a label to identify it
- Label switching is used to route the labelled packet

The FEC is an important concept in MPLS. A FEC is any subset of packets that are forwarded in the same way by a router. When a packet enters the MPLS network at the ingress node, the packet is mapped to a particular FEC and forwarded in the same way across the network. Packets belonging to the same FEC will get the same label.

When a packet enters the MPLS network at an iLER, the packet is mapped into a FEC. FEC assignment is performed only once at the ingress and the packet follows the LSP assigned to the FEC until network egress.

Using MPLS, the aggregation of packet flows into FECs provides scalability that meets the demands of the public Internet as well as enterprise applications.

Label Signaling (Distribution) Protocols Overview



The flow diagram outlines the different MPLS LSP configuration options available. As in IGP, there are static and dynamic protocols available. These protocols make sure the labels get distributed along the network representing the required LSPs.

Alcatel-Lucent products support static, Label Distribution Protocol or LDP and Resource Reservation Protocol or RSVP as label distributing protocols.

Static label distribution requires an operator to manually configure each router with appropriate label values as well as push, swap, and pop operations.

LDP and RSVP-TE assign and distribute labels automatically. LDP was standardized after MPLS was developed, whereas RSVP existed long before MPLS. RSVP was initially designed to do end-to-end QoS IntServ, in an IP network. It did not gain popularity until the introduction of MPLS when it was extended with traffic engineering capabilities (RSVP-TE).

Static Label distribution

Section 2: Static label distribution.

Static Label distribution

Static LSPs

- Manually configured on each LER/LSR
- Labels are assigned on each router individually
- Allows the administrator to have full control of the path

Disadvantages:

- Labor intensive operation
- Changes need a reconfiguration on every node
- No back-up or Fast Re-Route

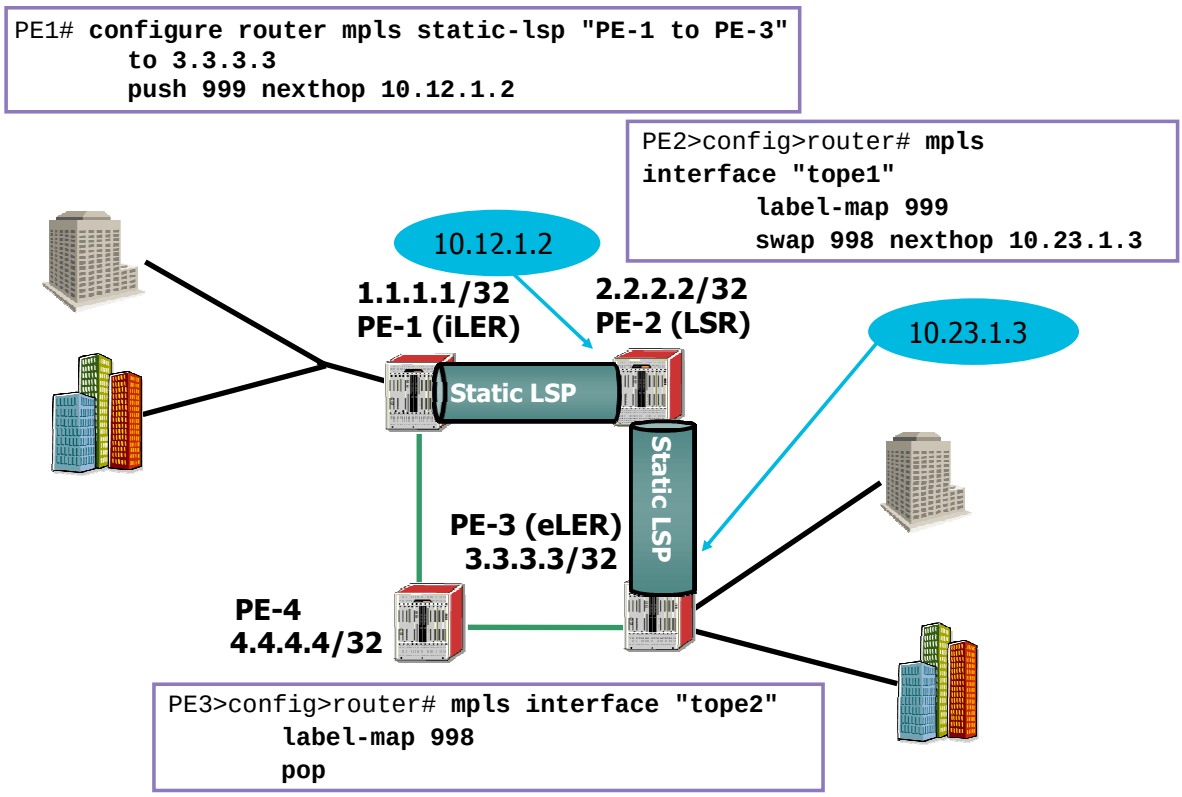
- Static label distribution is mainly intended for engineering purposes

Static LSPs are manually configured on each LER and LSR. Labels are assigned on each router individually. No signaling such as RSVP or LDP is required. This reduces processing power on the service routers and allows the administrator to have full control of the path.

However, creating LSP's manually is labor intensive and absolutely not scalable if a large number of LSP's needed to be created. A change in the network does require a reconfiguration on every node along the path of the static LSP. There is also no back-up or fast re-route capability.

So, static label distribution is mainly intended for engineering purposes.

Static Label distribution



The configuration above shows the static LSP configuration of PE-1. The static LSP transport tunnel is configured to forward traffic across the provider core from PE1 to PE3.

The Static LSP is configured between these devices, from system address 1.1.1.1 to the device with system address 3.3.3.3.

PE1 performs a PUSH operation and forwards the incoming packets to the next-hop address of 10.12.1.2, LSR2, with a label of 999.

The transit LSRs perform SWAP operations and forward the packet to the manually defined next-hop.

PE3 performs a POP operation and forwards unlabeled packets external to the MPLS domain.

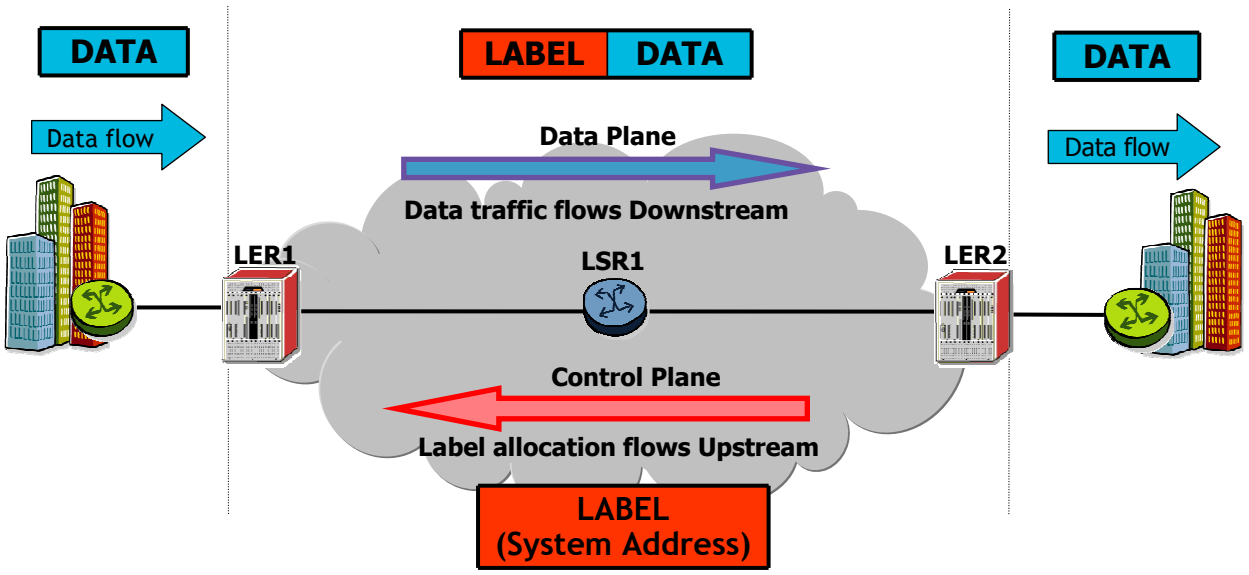
An LSP is always unidirectional, therefore after the configurations shown above; only traffic from PE1 has an LSP available to reach PE3, not the other way around.

These configuration steps demand the operator to log in and manually enter all the necessary commands in every node, an extremely heavy task that can be accomplished by dynamic signaling protocols instead.

LDP - Label Distribution Protocol

Section 3: LDP or Label Distribution Protocol.

Label distribution: Traffic/Control flow



Data packets flow in the downstream direction

Control driven model

- Label bindings are formed before the arrival of user generated data packets
- Alcatel-Lucent products use the control driven model

Label bindings are distributed from the downstream to the upstream direction

Before going into the details of the LDP protocol, the terminology of downstream and upstream need to be explained.

Data packets flow in the downstream direction.

Control packets follow the upstream direction.

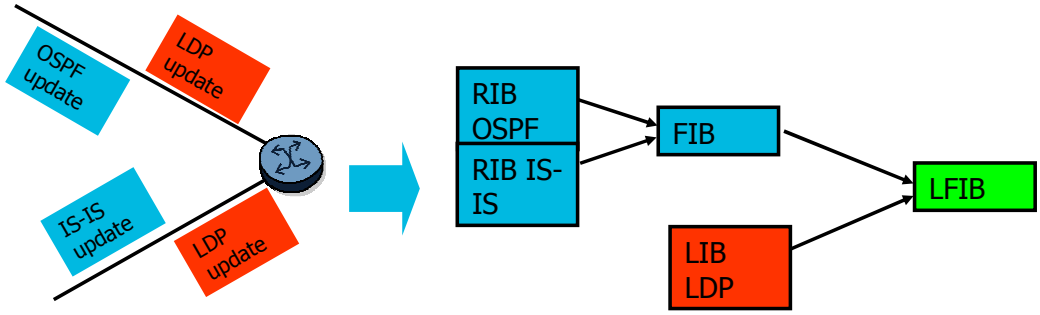
In the initial phase, LER2 is initiating the label distribution by sending a label advertisement message to LER1 of its system. The system address is actually the FEC here. Label distribution is initiated at LER2 and follows the upstream direction.

After label distribution, data flows in the opposite direction of the label distribution path.

Packets flow from the ingress Label Edge Router, LER1 to the egress Label Edge Router, in this case LER2.

LIB and LFIB

Table Name	Meaning	Contents	Populated By
RIB	Routing Information Base	Routing updates received	Routing Protocol Exchange - Each routing protocol has a separate RIB
FIB	Forwarding Information Base	Active routes	RTM selects the active routes from all protocol "Best" routes
LIB	Label Information Base	Locally generated and received MPLS labels	MPLS Label Exchange
LFIB	Label Forwarding Information Base	Labels used by the LSR	The labels assigned to the active routes (for each next-hop)



The Label Information Base or LIB is populated based on the label exchange process. The Label Forwarding Information Base or LFIB is built from the LIB and the Forwarding Information Base of the IP routing. The LFIB contains only the labels that will be used for label switching packets. If a label is received from a neighbor for a FEC, and the FIB contains a route for the same FEC for which the same neighbor is the next-hop, then it is populated into the LFIB.

The difference between the LIB and the LFIB is comparable to the difference between the Routing Information Base or RIB and the Forwarding Information Base or FIB of a routing protocol.

In the RIB, all the routes learned by any routing protocol running on the router are stored in that specific routing protocol's RIB. The RIB is a collection of all the possibilities learned through a routing protocol. Out of all these possibilities only one route, assuming no ECMP (Equal Cost Multipath), can be used, this route is selected by the Route Table Manager or RTM through the preference and metric procedures and stored in the FIB. In other words, the FIB represents the best routes of all the RIBs. The FIB has the actual outgoing port defined where the traffic needs to be send.

LIB/LFIB

The LIB relates to the LFIB as the RIB relates to the FIB. The LIB contains all the labels learned while the LFIB only stores the best label that will actually be used to forward traffic.

Because the LFIB is based on the FIB of IP routing, LDP is based on IGP to find its label path in the network.

LIB: Show Router LDF Bindings

```
PE1# show router ldp bindings
===============================================================
LDP LSR ID: 1.1.1.1
===============================================================================
Legend:  U - Label In Use,  N - Label Not In Use
         E - Epipe Service, V - VPLS Service, M - Mirror Service
         A - Apipe Service, F - Fpipe Service
===============================================================================
LDP Prefix Bindings
===============================================================================
Prefix             Peer            IngLbl  EgrLbl EgrIntf       EgrNextHop
-------------------------------------------------------------------------------
1.1.1.1/32         2.2.2.2         131071U   --     --            --
1.1.1.1/32         3.3.3.3         131071U   --     --            --
1.1.1.1/32         4.4.4.4         131071U   --     --            --
2.2.2.2/32         2.2.2.2           --    131071 1/1/2         10.12.1.2
2.2.2.2/32         3.3.3.3         131069U 131069   --            --
2.2.2.2/32         4.4.4.4         131069U 131064   --            --
3.3.3.3/32         2.2.2.2         131070U 131069   --            --
3.3.3.3/32         3.3.3.3           --    131071 1/1/4         10.13.1.3
3.3.3.3/32         4.4.4.4         131070U 131065   --            --
4.4.4.4/32         2.2.2.2         131067U 131067   --            --
4.4.4.4/32         3.3.3.3         131067U 131067   --            --
4.4.4.4/32         4.4.4.4           --    131067 1/1/3         10.14.1.4
```

The “show router ldp bindings” command displays the contents of the LIB and should contain all labels locally generated and those received from any LDP neighbors, whether they are in use or not.

The local router PE1 has three peers and has generated label 131071 for FEC 1.1.1.1/32. This label is seen as the Ingress Label for the FEC. It will be sent to all peers except the downstream peer in the LSP, as a label for a FEC should never be advertised to the next-hop of the LSP for that FEC. As a result the label will be flagged with an 'N' for that peer.

Only one entry per prefix shows an egress interface and next-hop since this is the label provided by the same router that provided the active IGP route, and hence is the active label binding, except for PE1’s own system interface, since this router is the eLER for this prefix. For FEC 1.1.1.1/32 the router has also generated a label, 131071. Since the FEC is the system address of the local router, the label will be advertised to all peers and flagged with a 'U', but none of the entries have any Egress information since the local router is the destination. When the local router receives a packet labeled with 131071 it will POP the label.

LFIB: Show Router LDP Bindings Active

```
PE1# show router ldp bindings active
===============================================================================
Legend:  (S) - Static
===============================================================================
LDP Prefix Bindings (Active)
===============================================================================
Prefix                Op   IngLbl    EgrLbl    EgrIntf        EgrNextHop
-------------------------------------------------------------------------------
1.1.1.1/32            Pop  131071      --        --             --
2.2.2.2/32            Push   --      131071    1/1/2          10.12.1.2
2.2.2.2/32            Swap 131069    131071    1/1/2          10.12.1.2
3.3.3.3/32            Push   --      131071    1/1/4          10.13.1.3
3.3.3.3/32            Swap 131070    131071    1/1/4          10.13.1.3
4.4.4.4/32            Push   --      131067    1/1/3          10.14.1.4
4.4.4.4/32            Swap 131067    131067    1/1/3          10.14.1.4
-output omitted-
===============================================================================
PE1#
```

The "show router ldp bindings active" command displays the contents of the LFIB and should contain all labels used for label switching packets and the associated label action.

For FEC 2.2.2.2/32, two entries exist. The first is associated to a PUSH action, used in the case that an unlabeled packet for this FEC is received at this router, or is locally generated. The unlabeled packet will have label 131071 PUSHed onto it and it will be forwarded via interface 1/1/2 to next-hop 10.12.1.2.

The second entry shows the action performed by this router when forwarding packets to the FEC 2.2.2.2 on behalf of other routers. In this case, router 1.1.1.1 is acting as an LSR when forwarding packets to 2.2.2.2 on behalf of routers 3.3.3.3 and 4.4.4.4.

For FEC 1.1.1.1/32, the local system address, any packets destined for this FEC will have the label stack POPed and forwarded as unlabeled packets.

LDP: Verifying Default Settings

LDP defaults are:
- Downstream Unsolicited label distribution mode
- Liberal Label Retention mode
- Ordered Control mode

```
PE1# show router ldp parameters
===============================================================================
LDP Parameters (LSR ID 1.1.1.1)
===============================================================================
-------------------------------------------------------------------------------
Interface Parameters
-------------------------------------------------------------------------------
Keepalive Timeout  : 30 sec                Keepalive Factor : 3
Hold Time          : 15 sec                HELLO Factor     : 3
Propagate Policy   : system                Transport Address: system
Deaggregate FECs   : False                 Route Preference : 9
Label Distribution : downstreamUnsolicited Label Retention  : liberal
Control Mode       : ordered               Loop Detection   : none
-------------------------------------------------------------------------------
PE1#
```

The "show router ldp parameters" command displays the default settings for label distribution and retention.

LDP defaults are Downstream Unsolicited label distribution mode, Liberal Label Retention mode and Ordered Control mode.

RSVP-TE operates in Downstream on Demand (DoD) label distribution mode, Conservative label retention mode and Ordered Control mode.

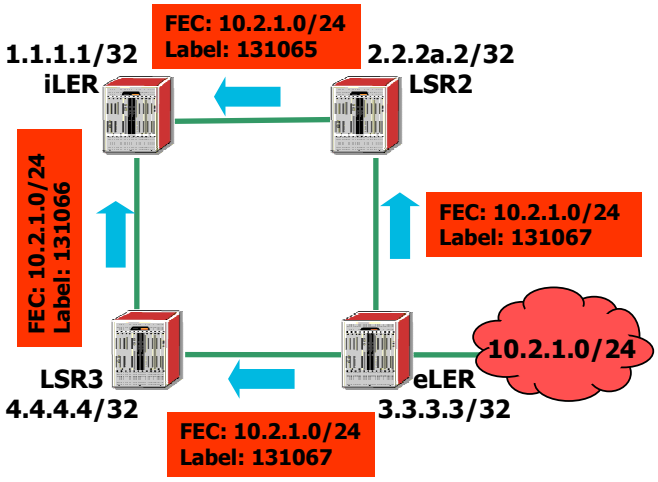
The following slides explain these settings more in detail.

LDP: Downstream Unsolicited Distribution Mode

Downstream:
Advertising labels from downstream to the upstream direction.

Unsolicited:
Each LSR will originate a label for its system address by default. Label mappings are provided to all peers for which the local LSR might be a next-hop, even when not explicitly requested.

iLER LIB	Prefix	Next-hop	Label
	10.2.1.0/24	LSR 2	131065
	10.2.1.0/24	LSR 3	131066



The MPLS architecture allows an LSR to distribute label bindings to other LSRs that have not explicitly requested them.

Downstream Unsolicited label distribution is not dependent on the data plane. Labels will be generated and distributed strictly in the control plane prior to the arrival of any user originated traffic.

In Downstream Unsolicited mode, label mappings are provided to all peers for which the local LSR might be a next-hop for a given FEC. This would typically be done at least once during the lifetime of a peer relationship between adjacent LSRs. The label received from the router providing the active IGP route for the FEC is used from the best next-hop.

This technique allows the IGP routing topology to provide some level of redundancy should there be any network issues. For example, should the router providing the active IGP route fail or the route via that router become unavailable, once the IGP converged to a new active route (from another router), the label for the FEC received from that peer will be immediately used.

LDP: Liberal Retention & Ordered Control Mode

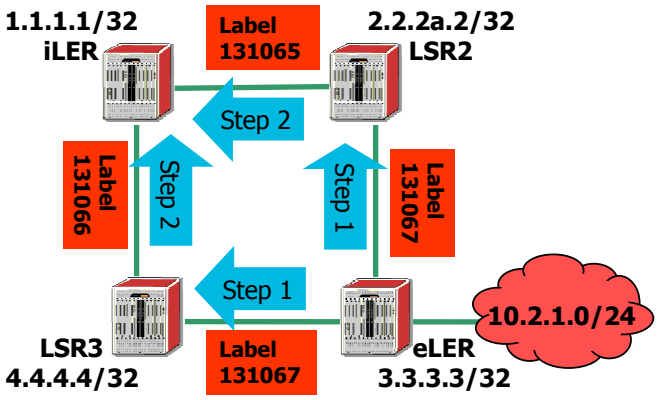
Liberal Retention:
The label received from the router providing the active IGP route for the FEC is used and the other labels are kept

Ordered Control:
The LSP will not pass data until the setup messages have propagated from end to end

iLER LIB	Prefix	Next-hop	Label
	10.2.1.0/24	LSR 2	131065
	10.2.1.0/24	LSR 3	131066

iLER FIB	Prefix	Next-hop
	10.2.1.0/24	LSR 3

iLER LFIB	Prefix	Next-hop	Label
	10.2.1.0/24	LSR 3	131066



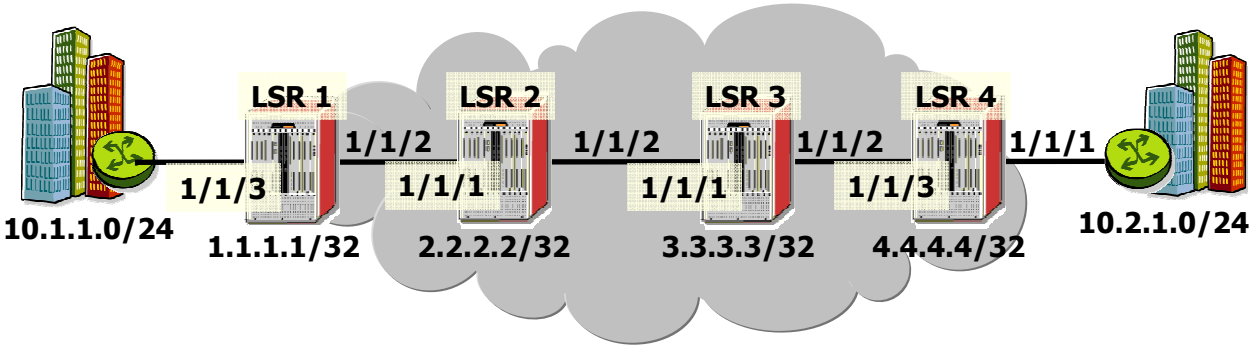
Using Liberal Label Retention mode, all label mappings received from all peer LSRs are saved. This approach consumes more memory on the LSR, but has the benefit of faster convergence. If the used label is lost, a label for the same FEC may have been previously received from another peer and is already present on the router, without the need for signaling.

In Ordered Control mode, LSP setup is initiated at one LSR and propagates from the eLER toward the iLER. A feature of Ordered Control mode is that an LSP is not completely set up until the associated control setup messages have propagated from end to end. As a consequence, data is not sent on the LSP until it is known to be loop free.

LDP Signaling

Each LSR will originate a label for its system address by default

Each LSR may originate a label for a FEC for which it has a next-hop that is external to the MPLS domain: an export policy is needed !



LSR 1 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.2.1.0/24	-	131068	1/1/2	LSR2
	4.4.4.4/32	-	131071	1/1/2	LSR2

LSR 4 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	-	131065	1/1/3	LSR3
	1.1.1.1/32	-	131070	1/1/3	LSR3

How are labels bound to a FEC?

There are two ways defined in the industry to bind a label to a FEC.

Advertise a label for each prefix in the routing table of the advertising router. This implies a large number of labels in the MPLS network and may cause scalability issues. The second way is to advertise a label per system address.

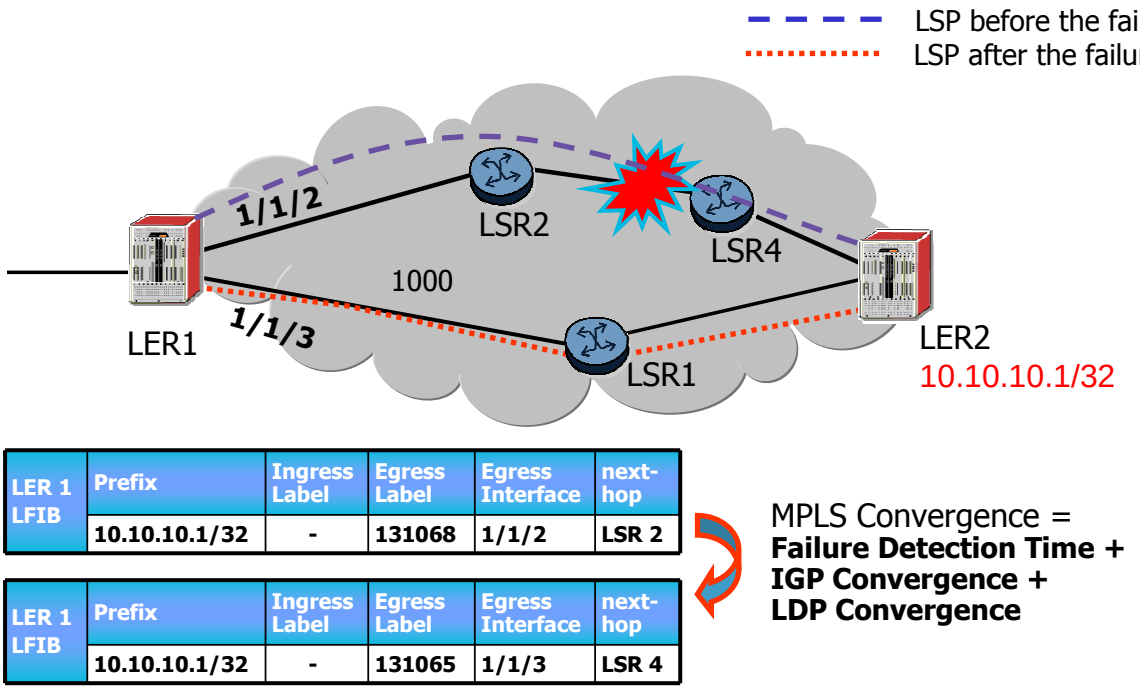
Alcatel-Lucent products advertise a label per system address as the default, resulting in fewer labels being used.

Once LDP neighbors have been established, each LSR will originate a label for its system address, and may originate a label for each FEC for which it has a next-hop that is external to the MPLS domain. To do this, an export policy is required.

In the example shown above, LSR 4 has an external next-hop for FEC 10.2.1.0/24 and has created a local label binding for this FEC and will propagate it into the MPLS domain to its LDP peer LSR 3.

Similarly, LSR 1 will originate a label for FEC 10.1.1.0/24.

LDP Convergence



LER 1 LFIB	Prefix	Ingress Label	Egress Label	Egress Interface	next-hop
	10.10.10.1/32	-	131068	1/1/2	LSR 2

LER 1 LFIB	Prefix	Ingress Label	Egress Label	Egress Interface	next-hop
	10.10.10.1/32	-	131065	1/1/3	LSR 4

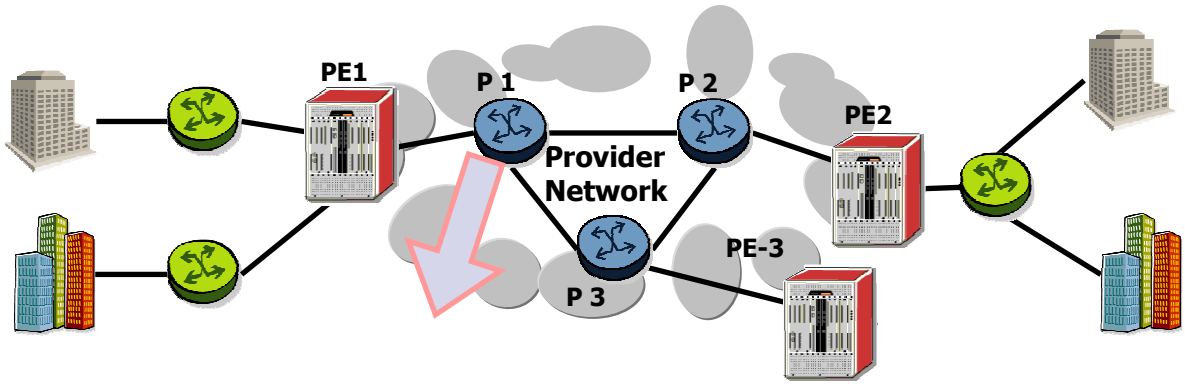
MPLS Convergence = **Failure Detection Time + IGP Convergence + LDP Convergence**

IGP convergence requires a recalculation of the SPF algorithm and a new selection of the best route. The IGP best route must then be offered to the Routing Table Manager or RTM to determine if this route is to be the active route and subsequently installed in the FIB.

In this case, the new best and active route to the destination 10.10.10.1/32 is the route via LSR1. LDP convergence may then follow. One benefit of liberal label retention mode is that the label previously received from the new best route is already installed in the LIB. It has been present all along, but until IGP routing converges, the LFIB may not be populated.

Now that IGP routing has converged, the LFIB may use the label from LSR2 since the next-hop now belongs to the active route.

LDP: Minimum Configuration




```
P1# configure router
P1>config>router# ldp
P1>config>router>ldp# interface-parameters
P1>config>router>ldp>if-params$ interface
    "P1-PE1"
P1>config>router>ldp>if-params>if$ exit
P1>config>router>ldp>if-params$ interface
    "P1-P2"
P1>config>router>ldp>if-params>if$ exit
P1>config>router>ldp>if-params$ interface
    "P1-P3"
P1>config>router>ldp>if-params>if$ exit all
```

All provider core facing interfaces must have LDP enabled. LDP must be enabled on each router's interface, to allow direct LDP sessions to be established between adjacent routers. Note that LDP is not enabled on the router's system interface.

RSVP-TE

Section 4: RSVP-TE.

Dynamic Label Distribution: RSVP-TE

RSVP-TE is used for establishing LSPs in MPLS networks.

RSVP-TE operates in ***downstream-on-demand (DOD)*** label advertisement mode with ***ordered LSP control***.

- A request to bind labels to a specific LSP tunnel is initiated by an ingress node through the RSVP Path message
- Labels are requested downstream and distributed (propagated) upstream by means of the RSVP Resv message

Advantage of using RSVP to establish LSP tunnels is that it enables the allocation of resources along the path.

- For example, bandwidth can be allocated to an LSP tunnel using standard RSVP reservations and Integrated Services service classes

RSVP-TE is an MPLS signaling protocol based on the resource reservation protocol originally used for signaling IP quality of service connections.

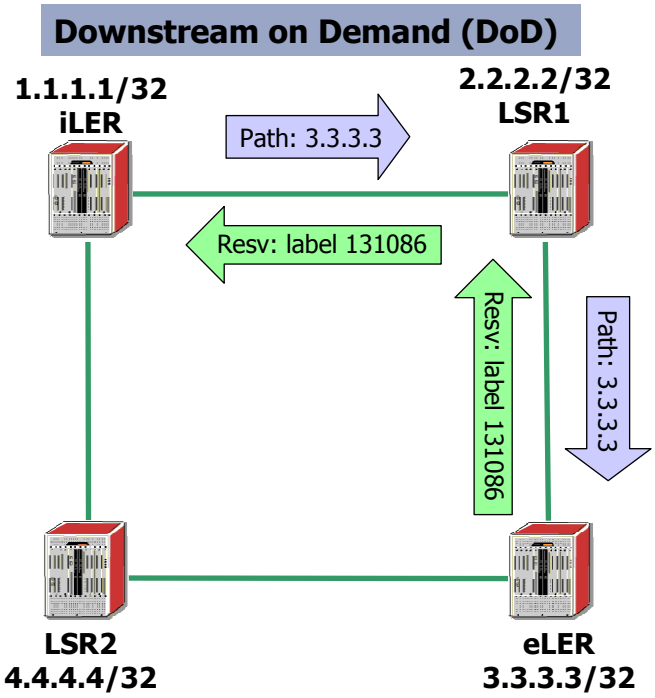
RSVP-TE defines a set of traffic engineering extensions to the Resource Reservation Protocol standard. RSVP-TE extensions provide a method by which RSVP may be used for traffic engineering in MPLS environments. These extensions add support for assigning MPLS labels and specifying explicit paths as a sequence of loose and strict hops. These extensions are supported by providing a Label Request field and an Explicit Route Objects field in the path message. The destination LSR responds to a label request by providing a label object in its RESV message. Labels are then assigned at each intermediate node which processes the "reserve" message. RSVP-TE operates in downstream-on-demand (DoD) label advertisement mode with ordered LSP control.

Since the flow along an LSP is identified by the label applied at the ingress node of the path, these paths may be treated as tunnels. A key application of LSP tunnels is traffic engineering with MPLS. The resulting label switched tunnels can be automatically routed away from network failures, congestion, and bottlenecks.

RSVP-TE: Message Types

RSVP-TE is a network protocol used to request and deliver quality of service (QoS) information for set up.

- PATH message sent towards tunnel destination
- Receiver sends RESV message back towards sender
- eLER sends label binding info in RESV message
- Path Refresh and RESV Refresh messages are sent periodically



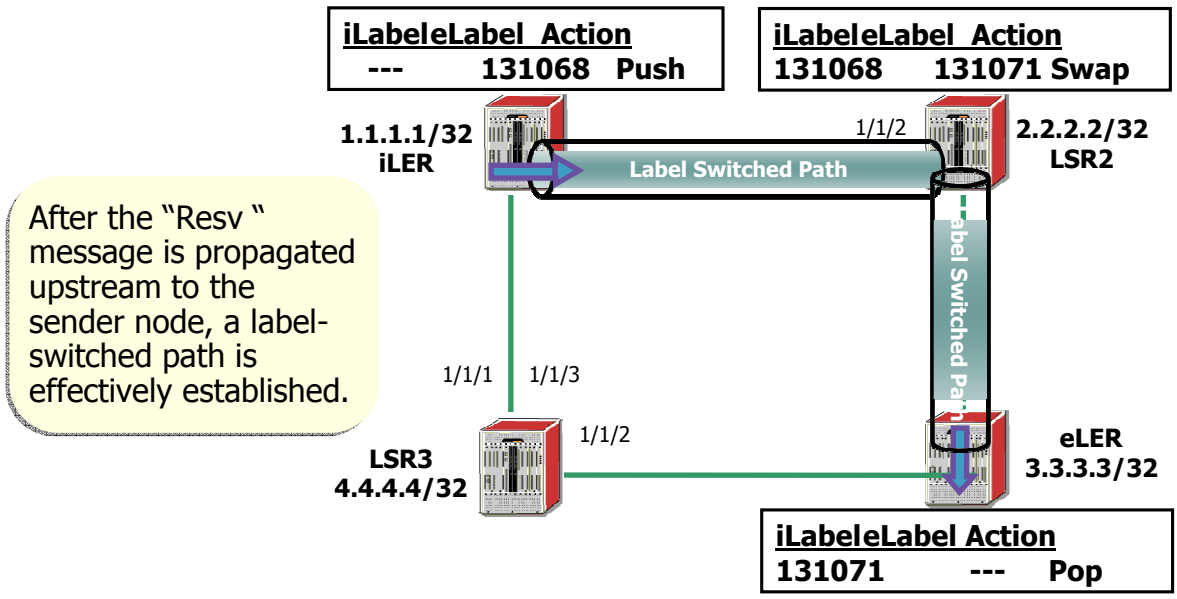
RSVP is a network control protocol used by a host to request specific qualities of service from the network for particular application data streams or flows. RSVP is also used by routers to deliver quality of service requests to all nodes along the path of the flows and to establish and maintain state to provide the requested service. RSVP requests generally result in resources reserved in each node along the data path. MPLS leverages this RSVP mechanism to set up traffic engineered LSPs. RSVP requests resources for simplex flows. It requests resources in only one direction.

RSVP treats a sender as logically distinct from a receiver, although the same application process may act as both a sender and a receiver at the same time. Duplex flows require two LSPs, one to carry traffic in each direction. RSVP is not a routing protocol, it operates with unicast and multicast routing protocols. Routing protocols determine where packets are forwarded and RSVP consults local routing tables to relay RSVP messages.

The sender, the ingress LER, sends PATH messages toward the receiver, the egress LER, to indicate the FEC for which label bindings are desired. PATH messages are used to signal and request label bindings required to establish the LSP from ingress to egress. Each router along the path observes the traffic type. PATH messages facilitate the routers along the path to make the necessary bandwidth reservations and distribute the label binding to the router upstream. The egress LER sends label binding information in the RESV messages in response to PATH messages received. The LSP is considered operational when the ingress LER receives the label binding information.

RSVP-TE: Label allocation

iLER makes use of the shortest IGP path from tunnel head to tunnel destination.



RSVP Path messages use a label request attribute and wait for a label reply in the RSVP Resv message. Once the Resv message is received, a label mapping is created in the LIB which provides a POP, SWAP, or Push action. The terminology of ingress or egress label is used with respect to the data plane of the packet flow.

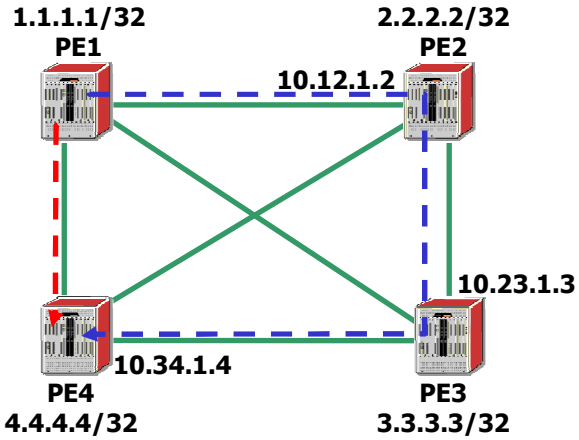
Once an LSP is established, the traffic through the path is defined by the label applied at the ingress node of the LSP. The set of packets that are assigned the same label value by a specific node are considered to belong to the same FEC which defines the RSVP flow.

RSVP-TE: Strict versus Loose hops

Example of strict and loose path

Blue path => ERO (Explicit Route Object) defines strict hops

- ERO explicitly defines the path PE2 => PE3 => PE4



Red Path => ERO defines only loose hops

- ERO defines only 4.4.4.4 as loose

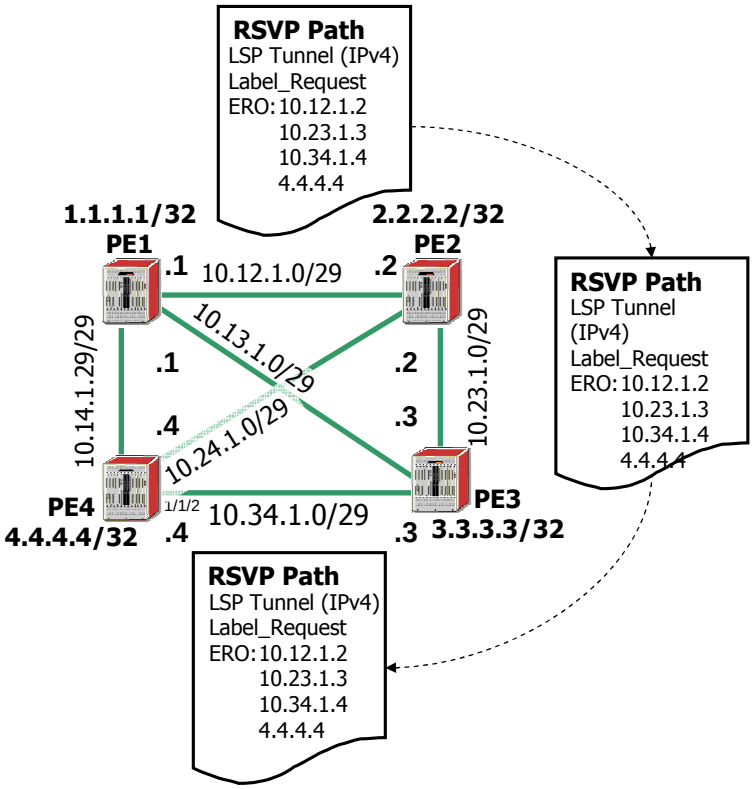
Each LSP can have a path configured as primary and one or multiples paths configured as secondary. The maximum number of primary and secondary paths is 8, configured as as 1 primary and 7 secondary or 8 secondary paths. Each path can be defined with strict or loose hops.

When defined with strict hops, the next hop address of each hop must be directly connected.

When defined as loose, the next hop does not need to be directly connected and IGP is used to reach this next hop.

RSVP-TE: Optional object – Explicit Route Object (ERO)

- ERO provides specific path information for the RSVP Path message to follow
- If ERO is not present then IGP is used to follow the path
- ERO can be manually provided or computed based on RSVP requirements such as bandwidth, hop limit, link coloring



The Explicit Route Object or ERO is used to specify the route RSVP Path messages take for setting up the LSP.

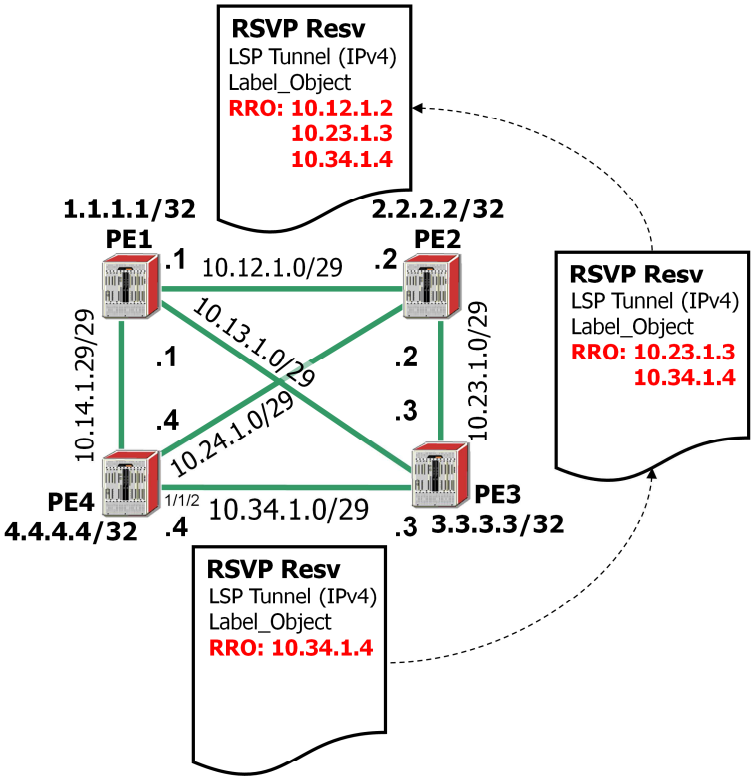
When the EXPLICIT_ROUTE Object is present, the Path message is forwarded towards its destination along a path specified by the ERO.

Each node along the path records the ERO in its path state block. A path state block is an internal structure which holds the information necessary to identify a particular RSVP session. Nodes may also modify the ERO before forwarding the Path message. In this case the modified ERO is stored in the path state block in addition to the received ERO.

The LABEL_REQUEST object requests intermediate routers and receiver nodes to provide a label binding for the session. If a node is incapable of providing a label binding, it sends a PathError message with an "unknown object class" error. If the LABEL_REQUEST object is not supported end to end, the sender node will be notified by the first node which does not provide support.

RSVP-TE: Optional object – Record Route Object (RRO)

- Record Route Object (RRO) of RSVP-TE is used for route recording purpose
  - RRO records the actual route a packet traversed
- Recording the path allows the iLER to know, on a hop-by-hop basis, which LSRs the path traverses.



By adding a RECORD_ROUTE object or RR0 to the Path message, the sender node can receive information about the actual route that the LSP tunnel traverses. The sender node can also use this object to request notification from the network concerning changes to the routing path. The RRO is analogous to a path vector, and hence can be used for loop detection.

RSVP-TE: CSPF (Constraint Shortest Path First) Implementation

The CSPF functionality provided by OSPF and IS-IS provides the capability to traffic engineer LSPs based on the following constraints:

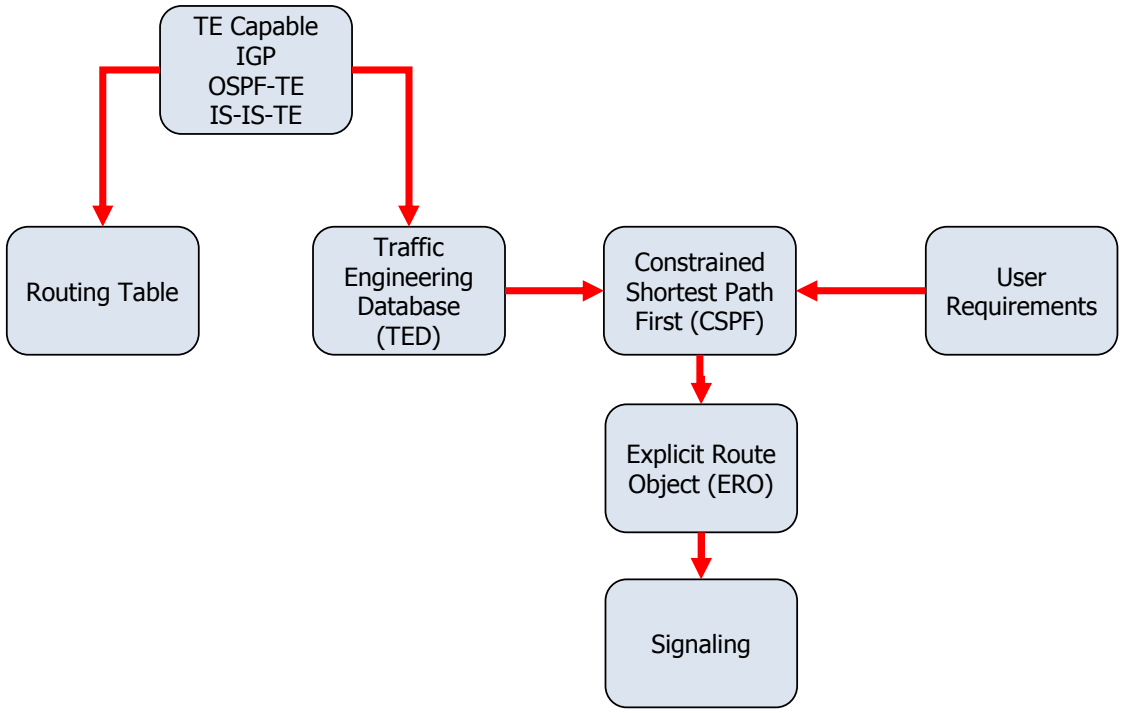
- Link constraints (include/exclude)
- Bandwidth requirements
- Hop count limitations

Enable Traffic-engineering on OSPF/IS-IS:

```
config>router>ospf# traffic-engineering

config>router>isis# traffic-engineering
```

The today's Link State Routing Protocols have traffic-engineering extensions, OSPF-TE and ISIS-TE. These extensions enable a special algorithm to take into account extra constraints when performing the Shortest Path First calculation. This algorithm is called the Constraint SPF (CSPF) algorithm which can be applied on an LSP when configured by the administrator.

RSVP-TE: Signaled LSP's with CSPF



The Constraint-based Shortest Path First process is an extension to the SPF process performed by link state routing protocols. The CSPF calculation uses constraints, obtained from the traffic engineering database or TED and local input, to compute the shortest path through the network that matches the configured requirements. Once a path is found by CSPF, the explicit route object in the RSVP message is updated with the path requirement information and RSVP uses the CSPF path to request the LSP set up.

Constraints taken into account by CSPF include:

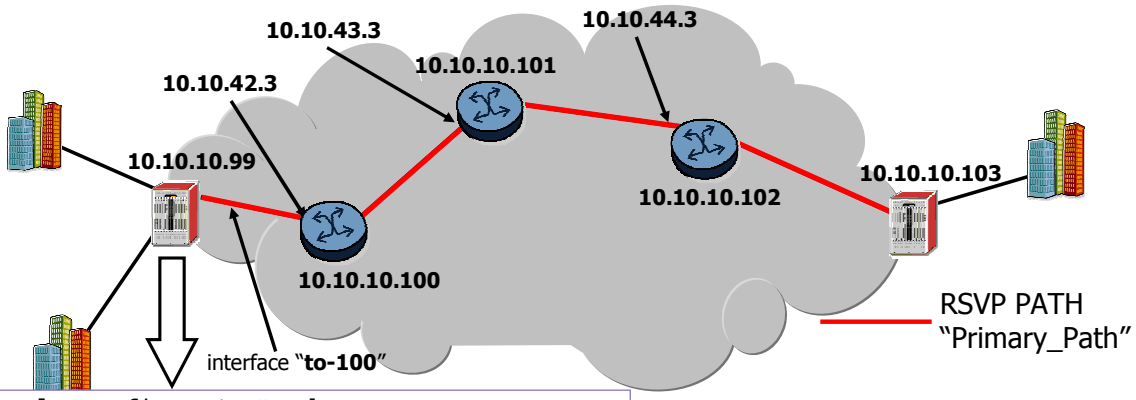
Admin groups, including link colors and resource classes

Bandwidth and

Hop limits

CSPF is also used to calculate the detour routes when fast-reroute is enabled. Avoiding a particular link is a constraint as well. Fast Reroute is covered further in the course.

RSVP-TE: LSP Configuration




```
Example: config>router# mpls
config>router>mpls# path "Primary_Path"
config>router>mpls>path$ hop 1 10.10.42.3  strict
config>router>mpls>path$ hop 2 10.10.43.3  strict
config>router>mpls>path# hop 3 10.10.44.3 strict
config>router>mpls>path# hop 4 10.10.10.103 loose
config>router>mpls>path# no shutdown
config>router>mpls>path# exit
```

```
config>router# mpls
config>router>mpls#lsp "LSP_99_103"
config>router>mpls>lsp# to 10.10.10.103
config>router>mpls>lsp# cspf
config>router>mpls>lsp# primary "Primary_Path"
config>router>mpls>lsp>primary# hop-limit 4
config>router>mpls>lsp>primary# bandwidth 256
config>router>mpls>lsp>primary# no shutdown
```

As shown above, the ERO is built up hop by hop in the "path" instance of the "mpls" context. After the creation of a certain path, the LSP can be created using this path as a primary or secondary path. Other constraints can be configured such as a hop limit, or a bandwidth reservation. For these extra constraints to work properly, CSPF must be enabled on the LSP which enables the IGP to perform a constraint SPF calculation.

RSVP-TE: LSP Protection

Path Protection

- Primary LSP with Secondary LSP
- Primary LSP with Secondary Standby LSP

Fast Reroute

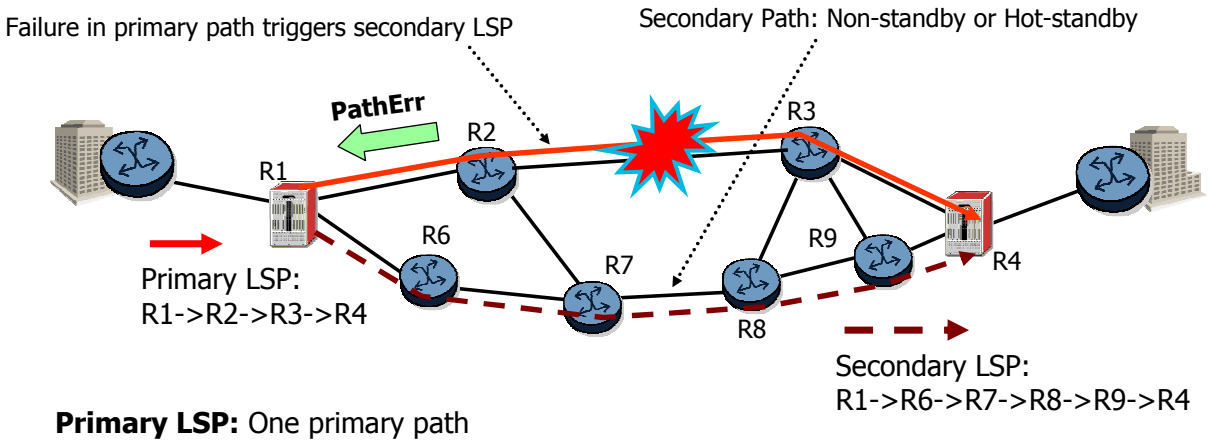
- One-to-One Backup
- Facilities Backup

An important concept of RSVP-TE LSPs is the ability for fast recovery, unlike LDP that is dependent upon the re-convergence time of the underlying IGP.

You can have up to 8 paths. This can be 1 primary path and 7 secondary paths or 8 secondary paths with no primary path. Once the iLER notices a problem further along the primary path, it will switch the traffic to a secondary path that can be configured, non-standby, or configured and signaled in advance, called hot standby.

For even faster, sub 50 ms, convergence, Fast Reroute can be activated. Fast Reroute provides a backup path for every link or node on the path so when a link or node failure occurs, every node can instantly switch to this backup temporarily until the iLER decides to use the secondary path.

RSVP-TE: Backup LSP – LSPs with Secondary Path



Primary LSP: One primary path

Secondary LSP

- Alternative path that is used if the primary path is not available.
- Non-Standby → needs to be signaled first (after primary path failure detection
- Hot-Standby → will be signaled upon creation
- Continuously tries to revert back to the primary path.
- Up to 8 secondary paths can be specified (or up to 7 if primary path is defined)
  - All the secondary paths are considered equal and the first available path is used
  - The software will not switch back among secondary paths

There can be only one Primary Path defined for a Label Switched Path. This primary path is the main path for the LSP and will be used in normal conditions.

If for any reason this primary path has become unavailable, up to 7 additional secondary paths can be defined to take over. Secondary paths all default as equal, but a preference can be assigned If a secondary path is configured as "hot-standby" the labels for this path have been signaled and maintained upon creation, which results in faster convergence but fewer labels left. This is a trade-off the network designer has to decide on. When the secondary path of the LSP is not in standby it will be signaled the moment a network failure causes the primary path to fail, and the Head-end node is made aware of the failure. After the switch over from the primary to the secondary, the software continuously tries to revert to the primary path. Up to 8 paths can be specified and are considered equal with the first available path being used. The software will not switch back among secondary paths. The Primary and Secondary paths can be configured with strict or loose hops, or with no hops specified.

RSVP-TE: LSP Protection with a Secondary Path

Pros

- Deterministic data flow during any point in primary path
- Multiple failures along the primary path can be handled by the same secondary path
- When statically configured, no nodes or links should be shared by the Primary and Secondary paths (otherwise if that link or node goes down, both are lost)
- Entire path is protected

Cons

- Notification of a link or node failure might take a while to reach head of tunnel
- Full path resources are reserved over both Primary and Secondary paths, therefore "double booking"
- Selective protection of link or node is not possible, only end-to-end

Primary and Secondary paths, when statically configured, are deterministic and will allow specific path protection as configured. The benefit to configuring Primary and Secondary paths is that path protection is guaranteed regardless of the location of the failure or if multiple failures exist. When statically configuring the Primary and Secondary paths, it is not advisable to configure common links since the failure of the common link will result in the failure of both LSPs. However, if the Primary and Secondary LSP are created using a Loose path, or if portions of the LSPs are making use of Loose path then it is entirely possible that the IGP path chosen will cause both LSPs to make use of a common link.

Primary and Secondary paths protect the entire path. However, a failure anywhere along the path requires the LSP head end to be notified that a failure has occurred. It's only the LSP head end that can signal the activation of the Secondary LSP. It might take some time for the head of the LSP to be notified that a failure has occurred downstream. Furthermore, by creating Primary and Secondary paths, network resources are being reserved for each LSP. This results in "double booking" the resources and overstating the actual resource utilization for the network, especially with hot standby secondary LSPs.

FRR - Fast Re-Route

Section 5: FRR or Fast Re-Route.

RSVP-TE: Fast Reroute (FRR) Overview

MPLS Fast Reroute (FRR):

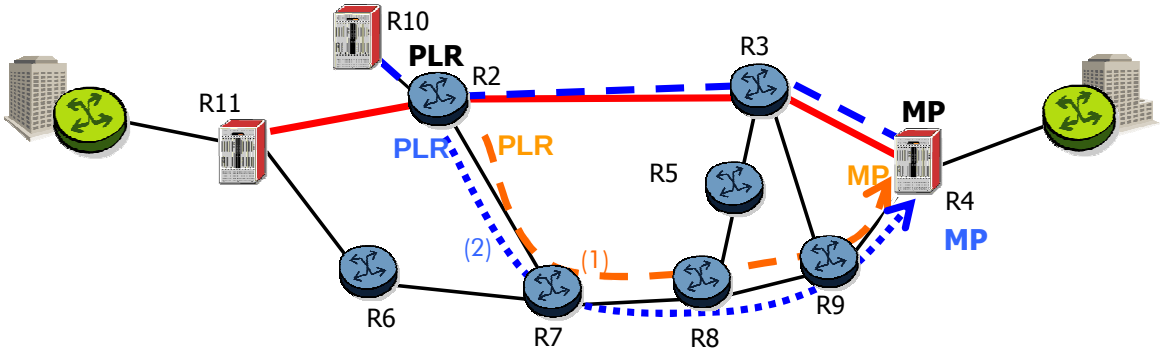
- Defines ways of pre-configuring and signaling backup paths before a failure.
- Allows traffic to flow almost continuously (in the 10s of ms timeframe)
- Uses LSPs established using RSVP-TE
- Allows protection to be applied as close to point of failure as possible

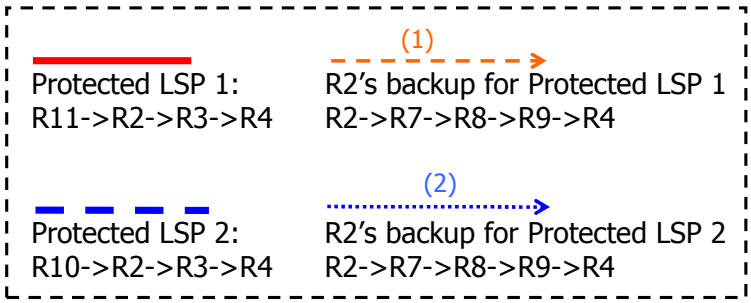
A common problem with IP networks is the restoration time after a link or node failure. With traditional routing protocols, it can take anywhere from several seconds up to nearly a minute for a failure to be bypassed.

MPLS Fast Reroute addresses these issues by defining ways of pre-computing and signaling backup paths before a failure, so that traffic can immediately be switched over to the backup path by the nearest node upon a failure. This allows traffic to flow almost continuously, without waiting for routing protocol convergence and signaling overhead.

MPLS Fast Reroute depends on LSPs being established using the RSVP-TE. Using RSVP-TE, it is possible to predetermine the path a LSP should take by specifying an explicit path for the LSP. This allows for the creation of alternative LSPs that do not depend on the same link or node as the LSP being protected.

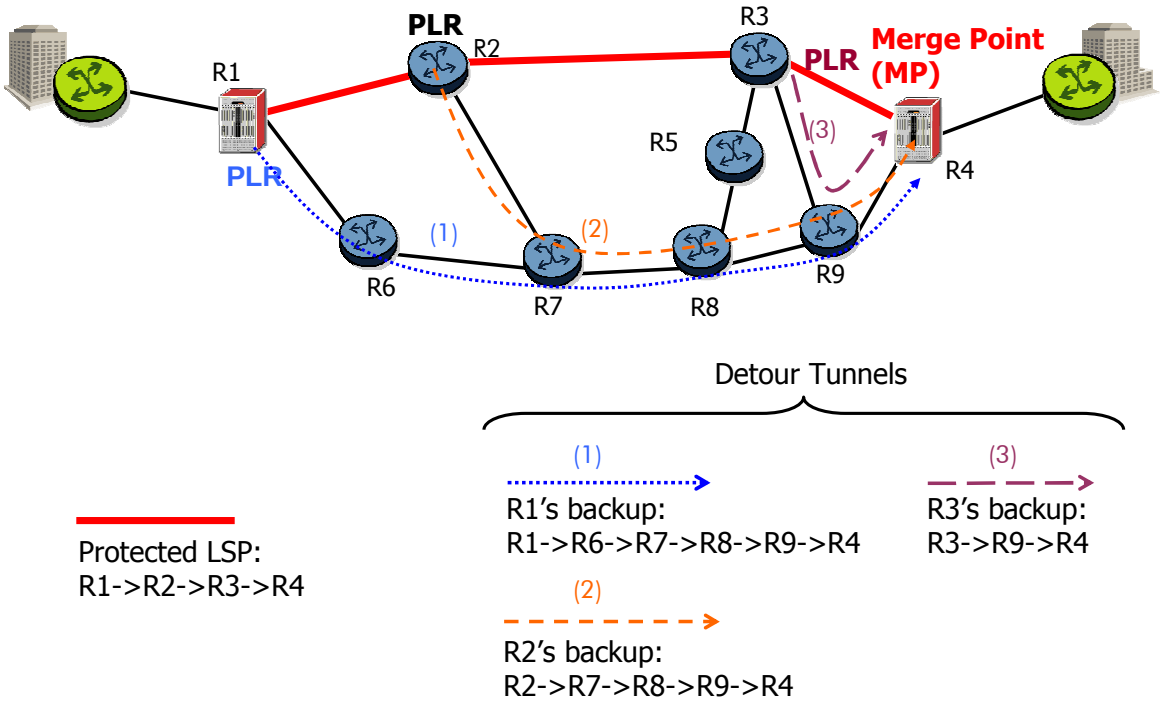
FRR: One-to-one Backup Method - Multiple LSPs





In the case of a one-to-one backup every LSP will create its own detour LSPs to protect itself. As shown above, only the Point of Local Repair or PLR of R2 is shown. R11 and R3 will also create detour LSPs. Two different LSPs with their own labels will be created and have the same Merge Point, the point where the detour merges with the original path, R4 in this example. The Facility Backup method, explained further in this course, will only create one bypass and therefore use less labels then the One-to-one Backup method.

FRR: One-to-one Backup Method - Path Setup

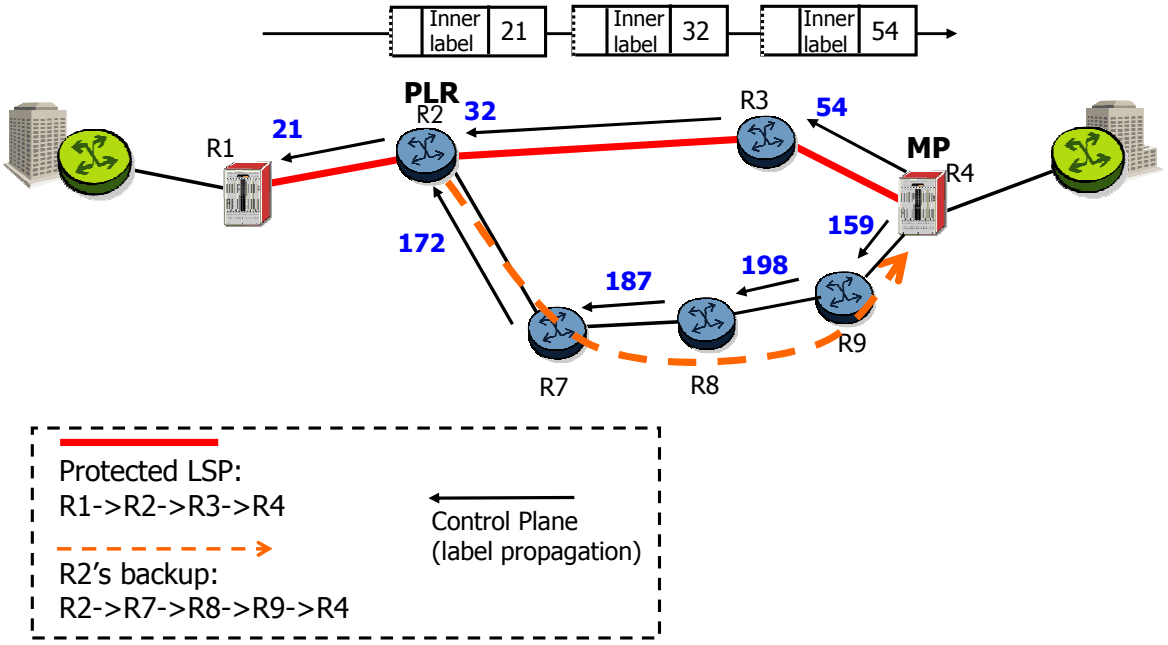


In One-to-one backup, each router along the LSP path establishes a detour LSP that is the best path to the destination node that avoids the node and/or link at the point of failure. For each LSP which is backed up, a separate backup LSP is established.

In the example above, the protected LSP runs from R1 to R4. Router R2 can provide user traffic protection by creating a partial backup LSP which is the best path to R4 that avoids the link between R2 & R3 (in case of link protection) and the node R3 (in case of node protection). The partial one-to-one backup LSP [R2->R7->R8->R9->R4] is referred as a detour.

To fully protect an LSP that traverses N nodes, there could be as many as (N - 1) detours. The paths for the detours necessary to fully protect the LSP in the above example are shown.

FRR: One-to one Backup Method - Label Exchange

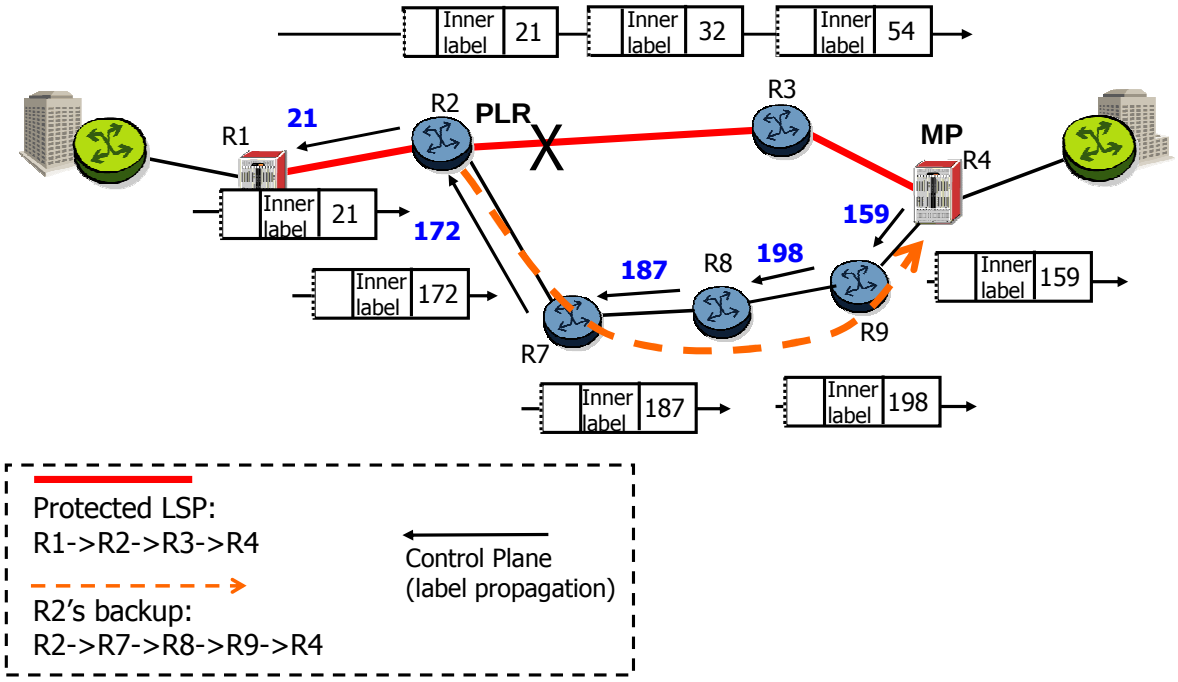


Labels along the Protected LSP path are advertised along the control plane according to common label propagation rules. Similarly, labels are propagated along the backup path in the same manner. CSPF is needed to perform this operation and is enabled automatically on the transit routers.

Therefore, the backup path is signaled, established, and prepared for the eventuality of a failed link. The only difference between the primary LSP and the backup LSP is that the primary LSP is being used in the data plane while the backup LSP is not.

If the link R2-R3 fails or if the node R3 fails, the PLR (R2) will swap incoming packets that have label 21 with label 172 and send it out the interface to R7. The MP (R4) will recognize that the packets arriving on its interface to R9 with label 159 must be POPped since it is the termination point of the protected LSP.

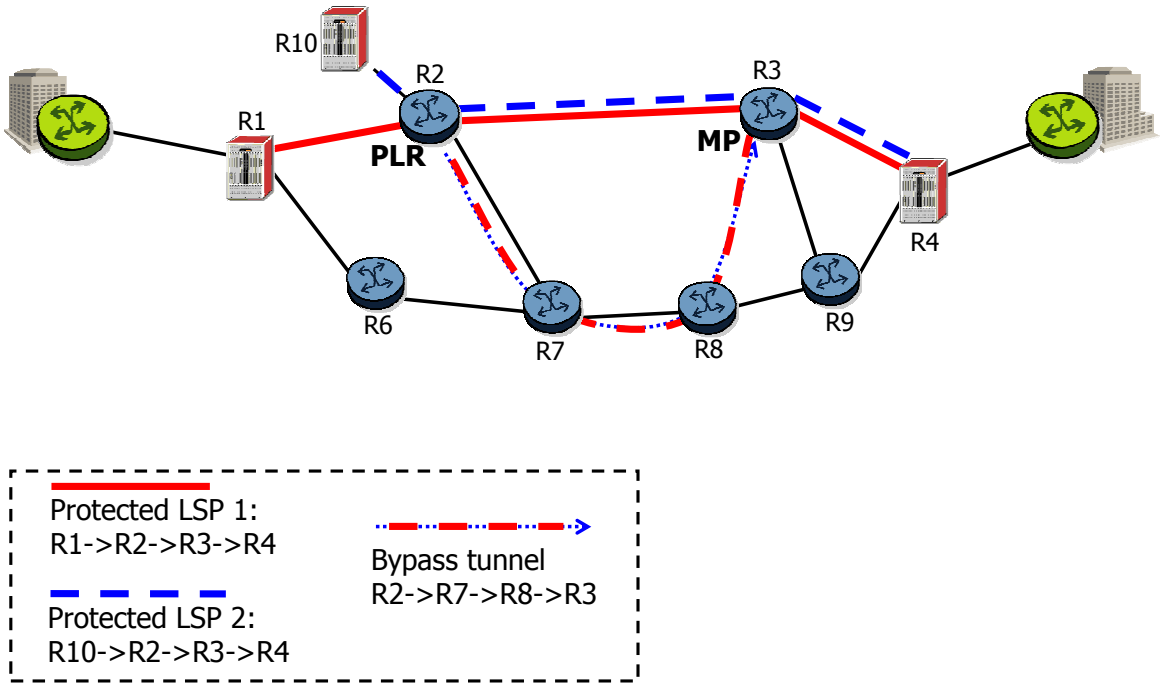
FRR: One-to-one Backup Method - Link/Node Failure



If the link R2-R3 should fail, the PLR will warn the upstream iLER about this failure and will swap the traffic over the detour LSP by swapping the outer label to the label received from R7 when the detour was created earlier. The label stack depth is unchanged, only the outer label will have a different value and a different egress interface.

The Merge Point router R4 will receive the traffic with a different transport label over a different ingress interface as before the link failure.

FRR: Facility Backup Method - Link Protection



The facility backup technique uses a different approach. Instead of creating a separate LSP for every backed-up LSP, a single LSP is created which serves to backup up a number of LSPs. We call such an LSP tunnel a bypass tunnel. The advantage of this method is the efficient use of labels in comparison with the One-to-one Backup method.

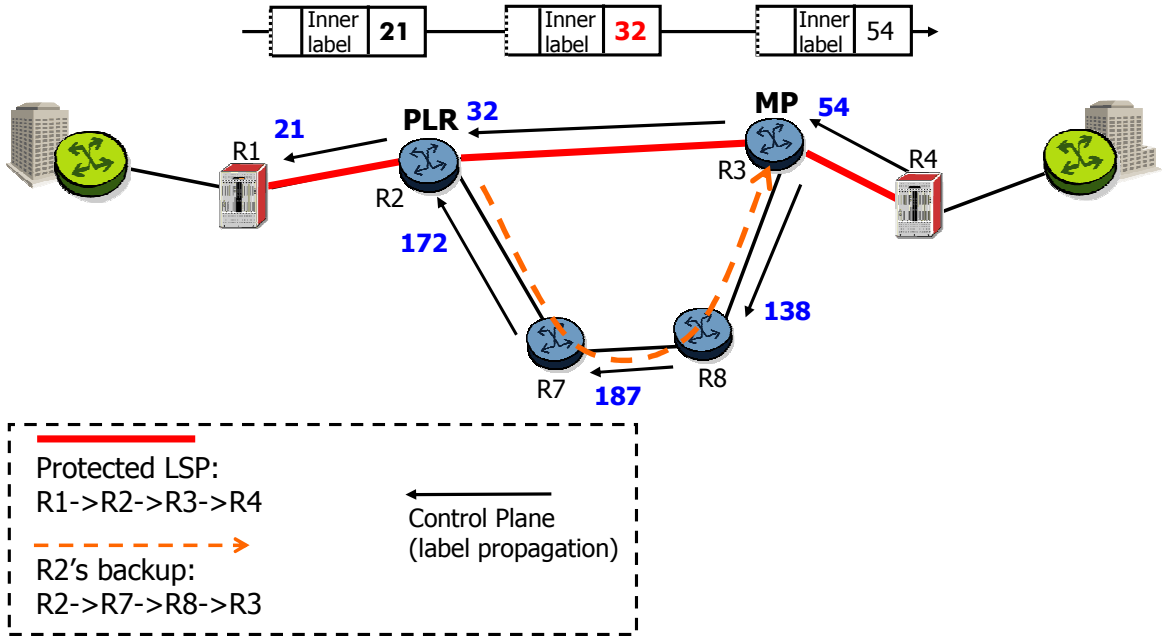
When such a bypass tunnel is created from a PLR to a downstream MP all the original LSPs using this partial path are candidates to use this bypass tunnel to protect themselves.

In the above example, R2 has built a bypass tunnel which protects against the failure of link R2-R3. The doubled lines represent this tunnel. Both the protected LSPs, regardless of their source and destination endpoints, can use this bypass tunnel, which provides a scalability improvement.

As with the one-to-one technique, to fully protect an LSP that traverses N nodes, there could be as many as (N-1) bypass tunnels. However, each of those bypass tunnels could protect a set of LSPs.

FRR: Facility Backup - Link Protection Label Exchange

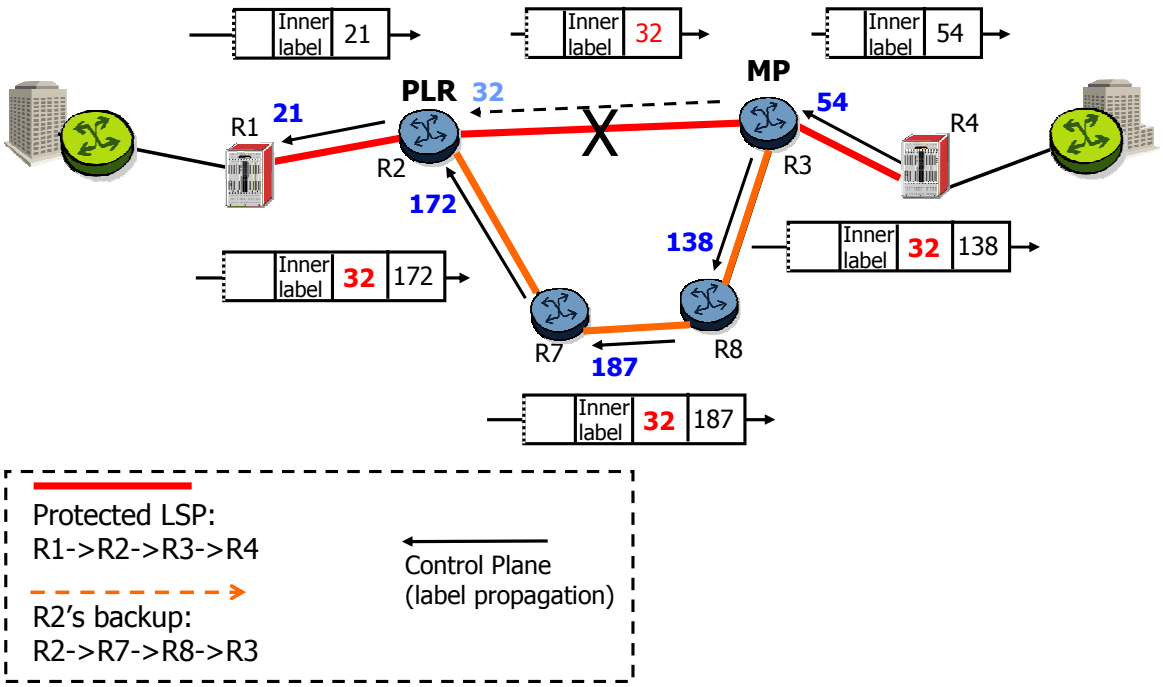
Works the same as One-to-One Backup under normal operating conditions



The data plane functions the same way in the stable LSP scenario as it did during the One-to-One backup method. The control plane prepares a bypass tunnel by signaling available labels from the MP to the PLR as a response to the PLR's RSVP-TE DoD request.

FRR: Facility Backup - Link Failure

MP receives same label from backup link as it would from Primary LSP



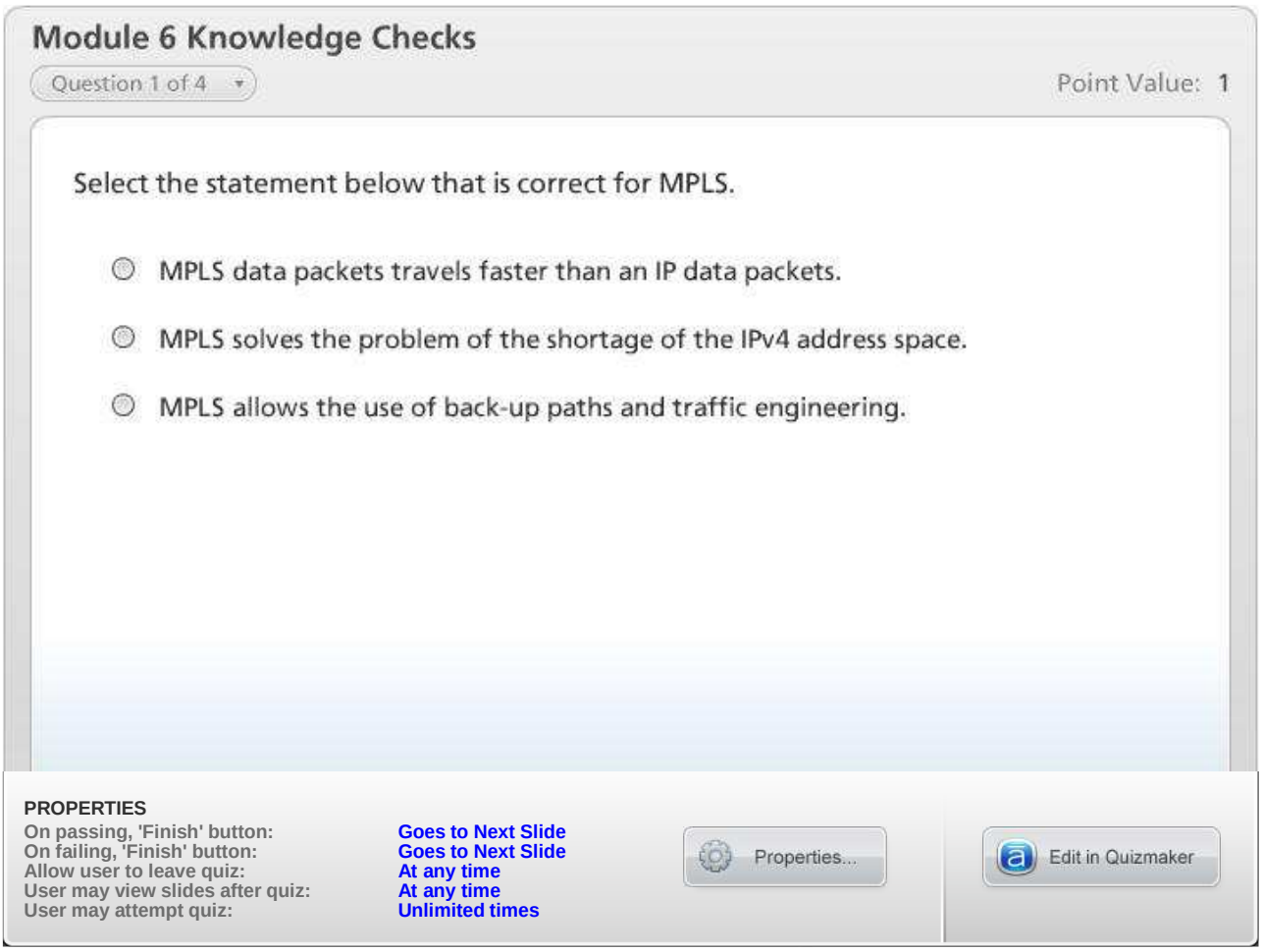
The example above demonstrates the Facility Backup method. If the link R2-R3 should fail, the PLR will first SWAP the outer label to label 32 because this is the label R3 would expect. On top of that label, the label 172 is PUSHed and the label stack has increased with one extra label. This packet will be sent over the backup egress interface and all the intermediate routers of the bypass tunnel (R7 and R8) will perform the normal swap operations on this extra outer label. It is the MP that will POP this label, investigate the 32 label and perform the SWAP like in normal operation.

This method allows the "tunneling" of multiple LSPs through the bypass and improves the scalability of the Protected paths.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 6

This completes module 6.

SR-OS Fundamentals
Module 7: SR-OS service module

IPD Development

Available as PDF

Welcome to the 7th module of the SR-OS fundamentals course.

Module 7 introduces the SR-OS services..

Table of Contents

Module 7 is divided into five sections.

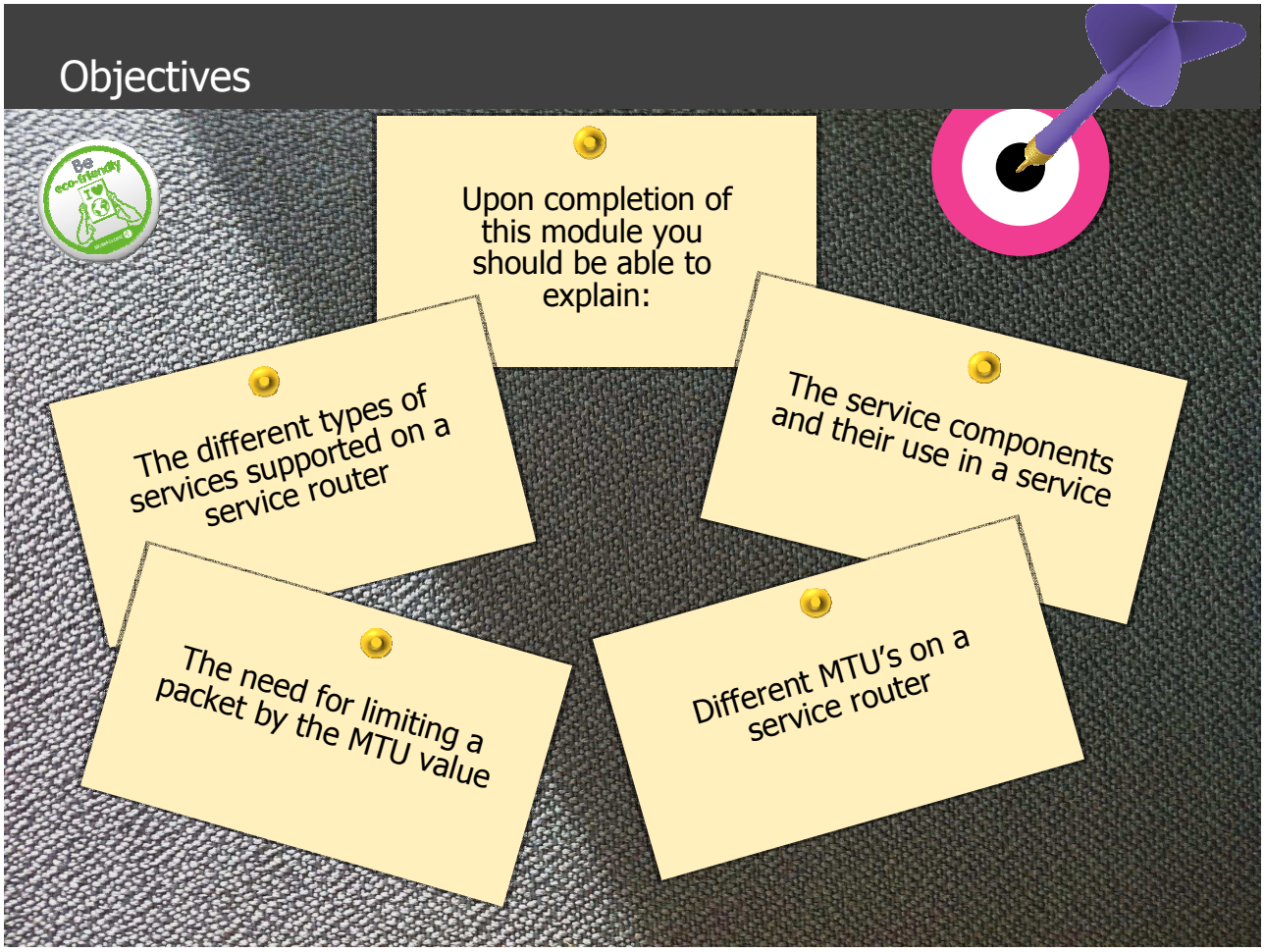
Section 1 explains what is so special about a service router compared to a traditional IP router

Section 2 provides an overview of the different supported services

Section 3 explains the principles of the service model

Section 4 describes the components that make a service

And section 5 talks about the maximum transmission unit or MTU



By the end of this module you will be able to explain:

- The different types of services supported on a service router
- The service components and their use in a service
- The need for limiting a packet by the MTU value
- Different MTU's on a service router

Introduction to service routers

Section1: Introduction to the concept of a service router.

What is a service router?

Traditional IP routers have limitations to support today's applications. New concepts were needed:

- Service oriented! Not port/interface oriented
- Uses MPLS (Multi-Protocol Label Switching) or GRE (Generic Routing Encapsulation) as tunneling mechanism which is well understood today
- Great for VPN type applications
  - QoS (Quality of Service)
  - Policies
  - Multicast
  - Switching
  - Routing
- Service concept found back in different market segments:
  - Mobile operator
  - Residential operator
  - Business operator
  - Strategic industries
  - Wireline and wireless carriers

Traditional IP routers have existed for quite some time. Initially they were mainly used in enterprises to route the data traffic. The higher capacity IP routers were mainly found in the core of the service operator. The service operators were running typical WAN technologies like frame relay and ATM in the edge to connect the enterprises and other types of customers to the IP core network.

The result of converging all the legacy networks towards one common technology introduced the Ethernet IP/MPLS network.

This strategy resulted in a cost efficient, scalable and bandwidth optimized network.

The IP/MPLS is the base of the service model as we know it today. The service model is oriented around a service and not around interfaces and ports like a traditional router.

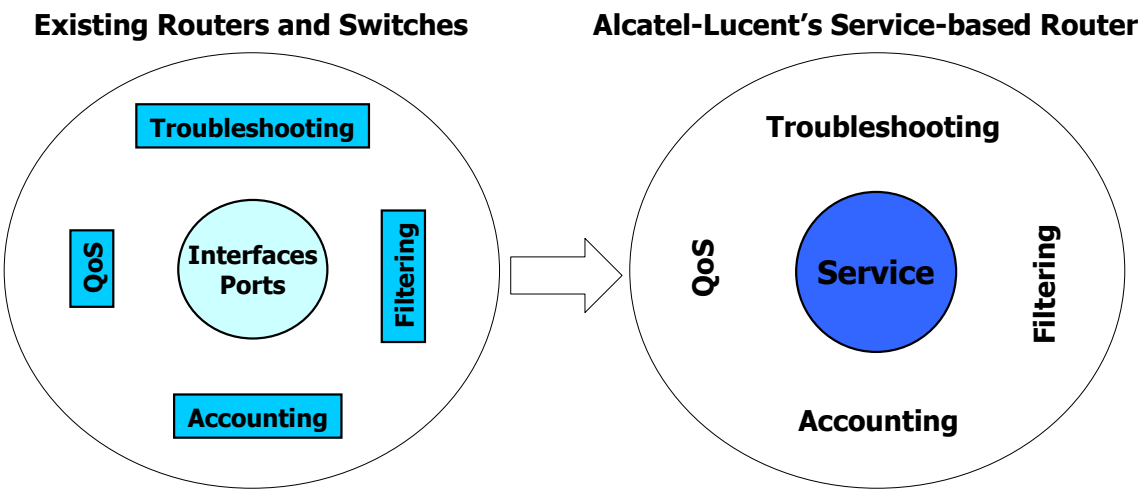
It also supports the tunneling concept, a well understood technology for VPN's.

All the features of QoS, policies, multicast, switching and routing are built on the basis of services.

The complete range of SR-OS routers were built from scratch with the service concept in mind.

These services do serve a wide range of application and are found in the networks of different kind of network operators. Services are the common base line for the mobile, residential, business, utility, wireline or wireless operator.

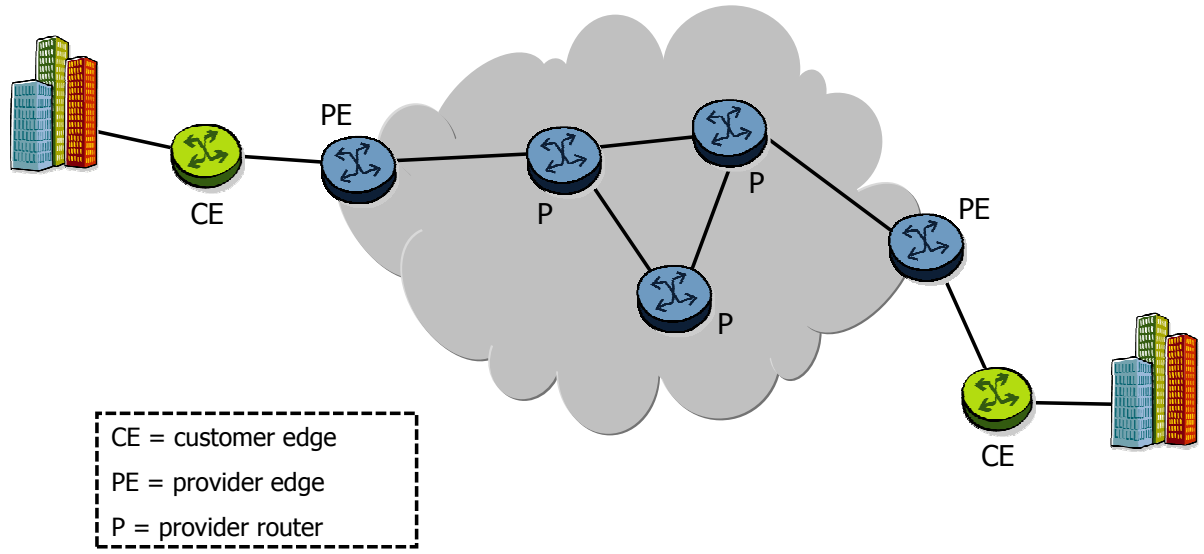
What is a service router?



The traditional routers and switches were designed for best effort forwarding of packets and frames. QoS, troubleshooting, filtering and accounting features were built around the interfaces and ports. There was no customer differentiation possible as all customers running traffic over a certain interface were getting the same treatment.

The Alcatel-Lucent Service router was designed from the beginning as a service router. The Services Router is designed to deliver Service Level Agreement -based services. This requires designing all the features around the concept of a service.

Network Component Naming Conventions



Before explaining how a service is built, let name some of the most important devices in a service enabled network.

A customer edge or CE device provides customer access to the service provider network over a data link to one or more provider edge or PE routers. The end-user typically owns and operates these devices. The CE devices run the routing protocol or protocols of the end-user and support the IP address scheme implemented by the end-user. They are unaware of the existence of the MPLS protocol or the VPNs inside the service network. Although there are some topologies where the CE device can run MPLS and are service aware.

CE devices used in Layer 2 VPNs may be Ethernet switches, in which case they do not need to participate in routing protocols. They must only be aware of VLANs running in the customer network.

A PE router is directly connected to the customer edge devices. In an MPLS network, PE routers are the Label Edge Routers.

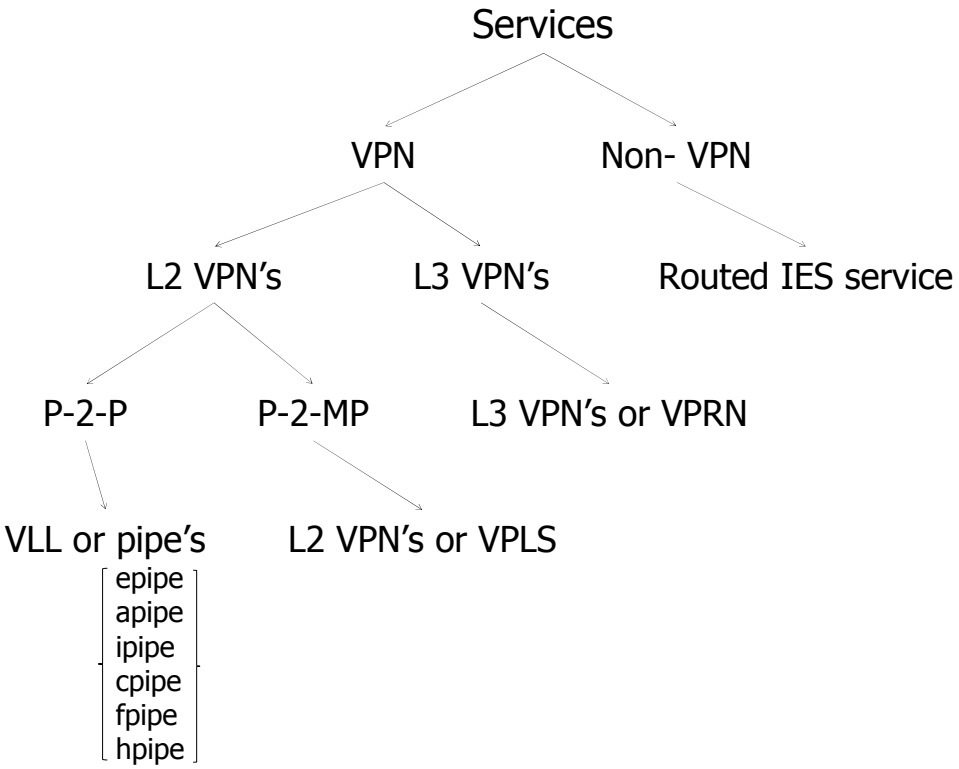
The routers in the provider core network with no direct connection to the CE devices are called P routers. In an MPLS provider network, these routers are the Label Switched Routers.

As explained later in more detail, services are set up from PE to PE and the P routers are unaware of the service.

Different supported services

Section 2: The different supported services.

Different services



Services can be divided into two main categories, VPN and non-VPN services. The majority of available services are VPN's services, either L2 or L3 VPN's. The non-VPN service of the SR-OS product family is the Internet Enhanced Service, a routed service that offers IP connectivity.

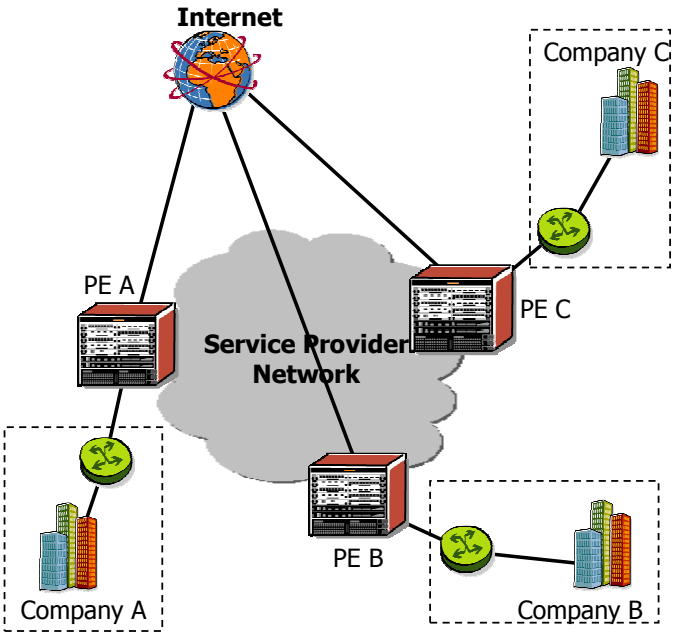
L2 VPN's can be further split up into point-to-point and point-to-mutipoint. The point-to-point services, pseudowires, virtual leased lines or also called pipe services come into different variant. The variant depends on what access technology need to be carried inside the VPN. In the point to multipoint service, the virtual private LAN service is found which can be best compared to an ethernet switch network.

Virtual private routed network is an L3 point to multipoint type of service VPN which offers routed services to customers.

Internet Enhanced Service (IES)

Internet Enhanced Service or IES provides direct internet access for the customer with the following features:

- From the customer's perspective it provides a direct connection to the Internet
- The Service provider can apply all billing, ingress/egress shaping and policing to the customer
- Customer connection is a SAP and supports:
  - Ethernet null
  - dot1q and q-in-q
  - SONET/SDH-IPCP
  - BCP-null
  - BCP-dot1q
  - ATM



The non-VPN routed IES service.

An Internet Enhanced Service or IES is a routed connectivity service where the subscriber communicates with an IP router interface to send and receive routed Internet traffic. The "I" in the name IES can be misleading. The IES service does not always have to be connected to the Internet; it can also be connected to a routed private network.

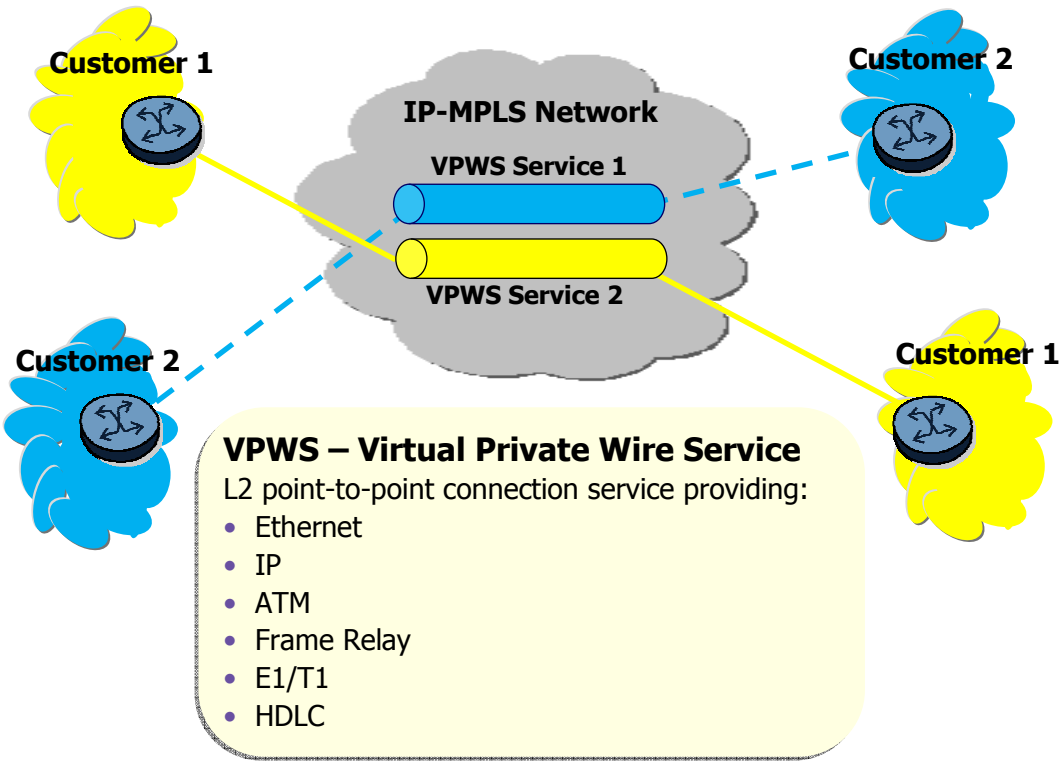
From a customer's perspective it provides a direct connection to a routed service, just like an IP interface. But because this is service and not a regular IP interface on a service router, the service provider can apply billing, ingress/egress shaping or apply any service policy.

The PE devices buffers service traffic and shapes it to conform to SLA parameters. Buffer allocation is programmable per-service to accommodate different maximum burst sizes. Each service can use multiple queues to enable shaping, policing and marking of different flows. The PE device can also shape and police on service egress so customers can purchase sub-rate services with asymmetric SLAs.

The connection to the customer is SAP based and not interface based. The SAP supports multiple encapsulation types including Ethernet null, dot1q and q-in-q, SONET/SDH-IPCP, BCP-null, BCP-dot1q and ATM. As with any SAP on a service, it can contain filter, accounting, QoS policies.

The IES service interface does support all the relevant IP routing protocols like RIP, OSPF, IS-IS, BGP, but is does support IPv6. Although this is a non-VPN service, it has the ability to connect to a pseudowire and do all the L3 termination.

VPWS Services

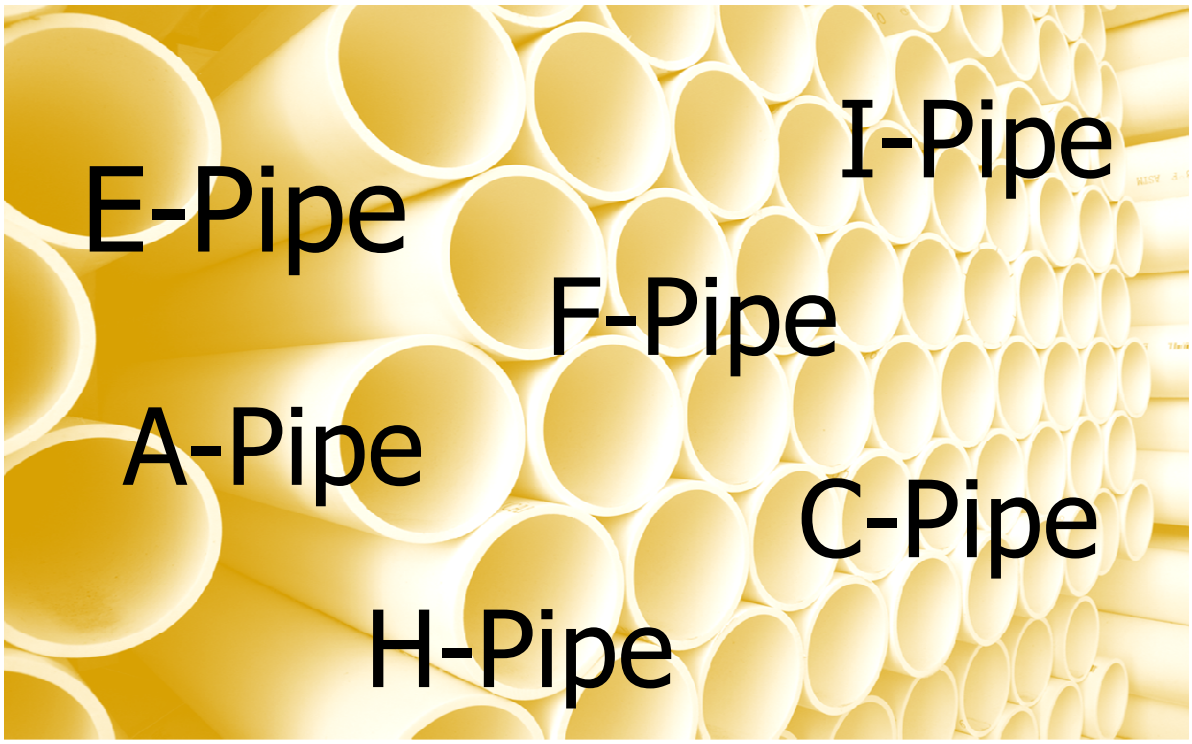


The L2 VPN point to point service or VPWS service.

A Virtual Private Wire Service is a layer 2 point-to-point connection service providing an emulation of layer 2 technologies such as Ethernet, IP, ATM, Frame Relay, E1/T1 and HDLC.

A VPWS service is also known under different names like a pseudowire service, VLL service, PWE3 service, martini service and pipe service.

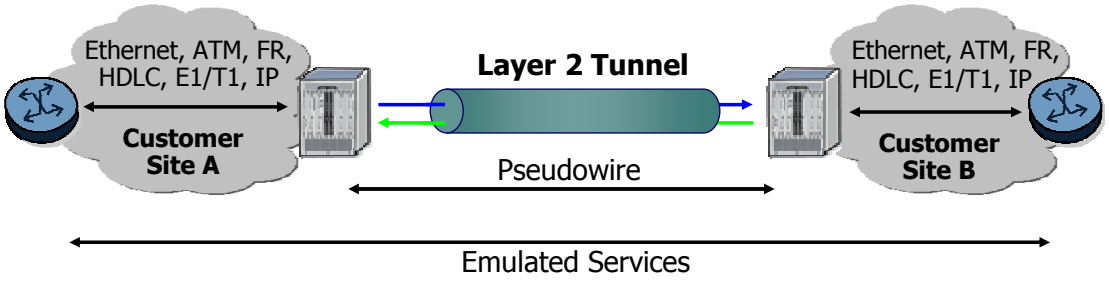
VPWS Services



Alcatel-Lucent's implementation of VPWS is known as E-pipe, I-pipe, A-pipe, F-pipe, H-pipe and C-pipe. These VPWS Services are based on the IETF "Martini Drafts" and the IETF Ethernet Pseudo-wire Drafts.

These services are layer 2 point-to-point services where the customer data is encapsulated and transported across a service provider's IP or MPLS network. The VPWS service is completely transparent to the subscriber's data and protocols. The VPWS service does not perform any MAC learning.

Why VPWS?

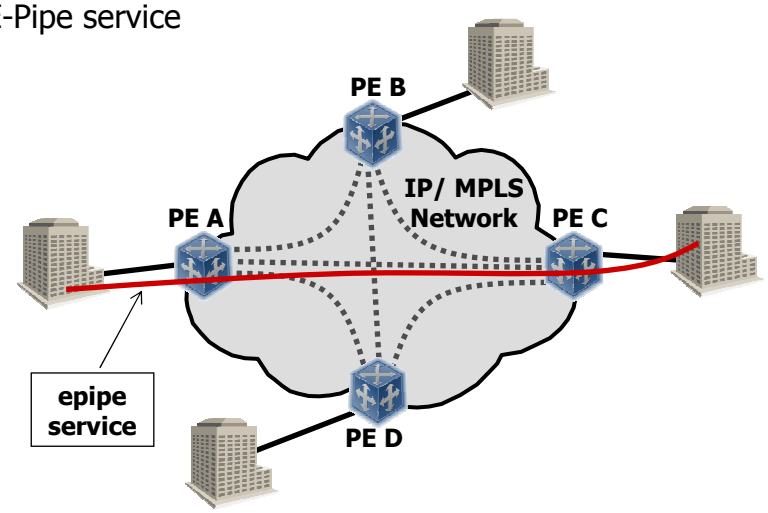


Service providers are seeking to offer multiple services across a common packet switched network.
Pseudowires enable you to emulate or to carry legacy Layer 2 services over a cost-effective and predominant IP/MPLS network. It decouples the services protocols and applications from the underlying facilities that carry them.

L2 Point-to-Point VLL Service

Alcatel-Lucent Ethernet Pipe (E-Pipe)

- Encapsulates customer Ethernet data
  - Transport data over IP or MPLS
  - Use GRE or MPLS tunnel
- Access to service provider's network via PE router SAP
  - Local service: Two SAPs connected on same node
  - Remote service: Two SAPs connected on different nodes
- Customer unaware of E-Pipe service
- Policies can be applied
- No MAC learing



As an example, let us have a look to an ethernet VPWS service.

This service is a Layer 2 point-to-point VLL service. The Alcatel-Lucent implementation of an Ethernet VLL is called an ePipe. The ePipe service encapsulates customer ethernet data and transports it across a service provider's IP or MPLS network in a GRE or MPLS tunnel.

Customer access to the service provider's network is through a Service Access Point or SAP on a PE router. An e-Pipe service connects two of those service access points on the same node, called a local service, or two SAPs on different nodes, called a remote service, through two uni-directional tunnels.

The customer is unaware of the e-pipe service and it looks like to him as a direct ethernet connection, while the two sides of the e-pipe could be geographically dispersed.

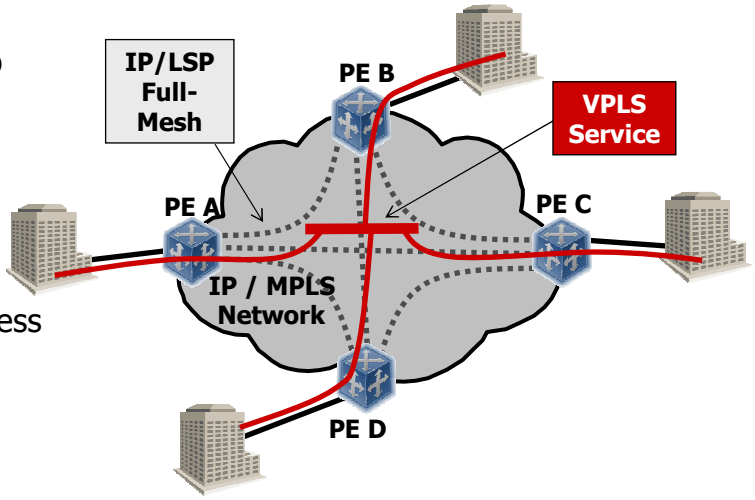
QoS and filter ingress and egress policies can be applied to ePipe services.

Whatever is sent to the SAP, exits on the other PE side on a SAP. There is no MAC learning required. Although some MAC address will not be passed due to a filter policy applied to the SAP.

Virtual Private LAN Service (VPLS)

VPLS is a class of VPN that allows the connection of multiple sites in a single bridged domain over a provider managed IP/MPLS network

- From the customer's perspective it looks as if all sites are connected to a single switched VLAN
- Service provider can reuse the IP/MPLS infrastructure to offer multiple services
- The Service provider can apply billing, ingress/egress shaping and policing



The Alcatel-Lucent point to multi-point L2 VPN is called a VPLS or Virtual Private LAN Service.

A VPLS is a multipoint Layer 2 service that allows multiple customer sites to be connected in a single bridged domain contained within a provider-managed IP/MPLS network. Customer sites in the VPLS appear to be on the same LAN, even if the sites are geographically dispersed.

A VPLS uses only Ethernet interfaces on the customer access side.

It enables customers to control and simplify routing strategies, as all routers in the VPLS are part of the same LAN. Each customer connected to the PE will be in the same subnet.

A VPLS can span a single node or multiple nodes. On a VPLS that spans a single node, subscriber data is distributed through multiple service access points on the node. A VPLS on a single node does not require service MPLS or GRE tunnels.

On a VPLS that spans multiple sites customer data enters the service using at least one SAP on each node. Data is transported among the nodes through service tunnels over an IP/MPLS provider core network. A VPLS that spans multiple nodes requires at least one tunnel at each node.

VPLS services switch traffic based on MAC addresses.

Although a VPLS is a Layer-2 VPN service and allows the use of Layer-2 switches as the CE device, most customers use routers at the LAN/WAN boundary.

Using a router as the CE device means that the PE device only has to learn one MAC address per-site per-service.

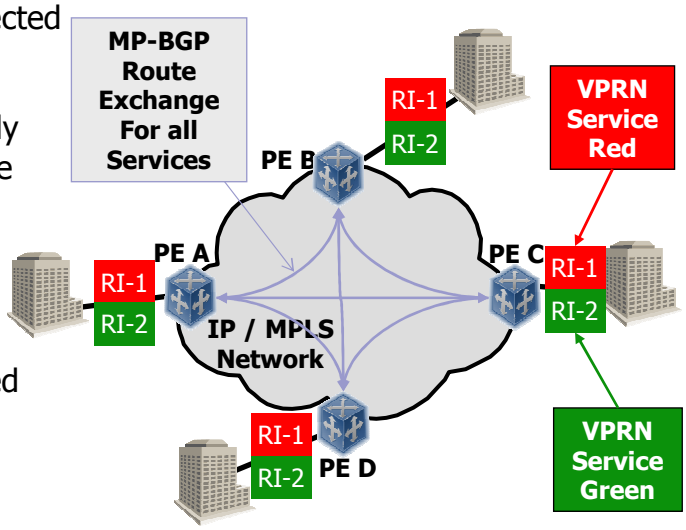
Using a Layer-2 switch as the CE device means that the PE device has to learn potentially hundreds of MAC addresses per-site per-service. The number of MAC addresses that the PE device has to learn can be limited through the use of MAC filters and/or by limiting the maximum number of MAC addresses accepted by the PE device.

As with the SAP on the e-pipe service, the same type of policies can be applied on the SAP of a VPLS service.

Virtual Private Routed Network (VPRN)

VPRN is a class of VPN that allows the connection of multiple sites in a routed domain over a provider managed IP/MPLS network:

- From the customer's perspective it looks like all sites are connected to a private routed network administered by the service provider for that customer only
- The service provider can reuse the IP/MPLS infrastructure to offer multiple services
- Each VPRN appears like an additional routing instance, routes for a service between the various PE's are exchanged using MP-BGP



In Feb. 2006, Internet Draft RFC2547bis was moved to 'standard' status, as **RFC 4364**.

The L3 VPN or Virtual Private Routed Network details a method of distributing routing information and forwarding data to provide a Layer 3 Virtual Private Network service to end customers.

Each Virtual Private Routed Network or VPRN consists of a set of customer sites connected to one or more PE routers. Each associated PE router maintains a separate IP forwarding table for each VPRN. Additionally, the PE routers exchange the routing information configured or learned from all customer sites via Multi-Protocol BGP peering.

Each route within a VPN is assigned a MPLS label. When BGP distributes a VPN route, it also distributes a MPLS label for that route.

Before a customer data packet travels across the service provider's backbone, it is encapsulated with the MPLS label that corresponds, in the customer's VPN, to the route which best matches the packet's destination address. The MPLS packet is further encapsulated with either another MPLS label or GRE tunnel header, so that it gets tunneled across the backbone to the proper PE router. Each route exchanged by the MP-BGP protocol includes a route distinguisher, which identifies the VPRN association and therefore, the backbone core routers do not need to know the VPN routes.

Service model

Section 3: The service model.

SR Service Configuration Model

In order to provision a service you must configure a:

- Service – define the type of service that you want
- Service Access Point (SAP) – a logical entity that serves as the customer access to the service.

If you want to provision a VPN service (VPRN, VPLS or VLL/ePipe) you also need a way of telling the router which other routers this VPN will connect to and how. This requires:

- Service Distribution Paths (SDP) – An SDP defines which other routers a service is connected to. They also tell the router what sort of tunnel encapsulation the service will be using RSVP-TE, LDP, or Generic Router Encapsulation (GRE).
- Auto-Bind - This feature is used with MPLS/LDP tunnels, and effectively auto-provisions the tunnels.

A service consists of a couple of components that makes the service.

Firstly, the service and the type of service need to be defined.

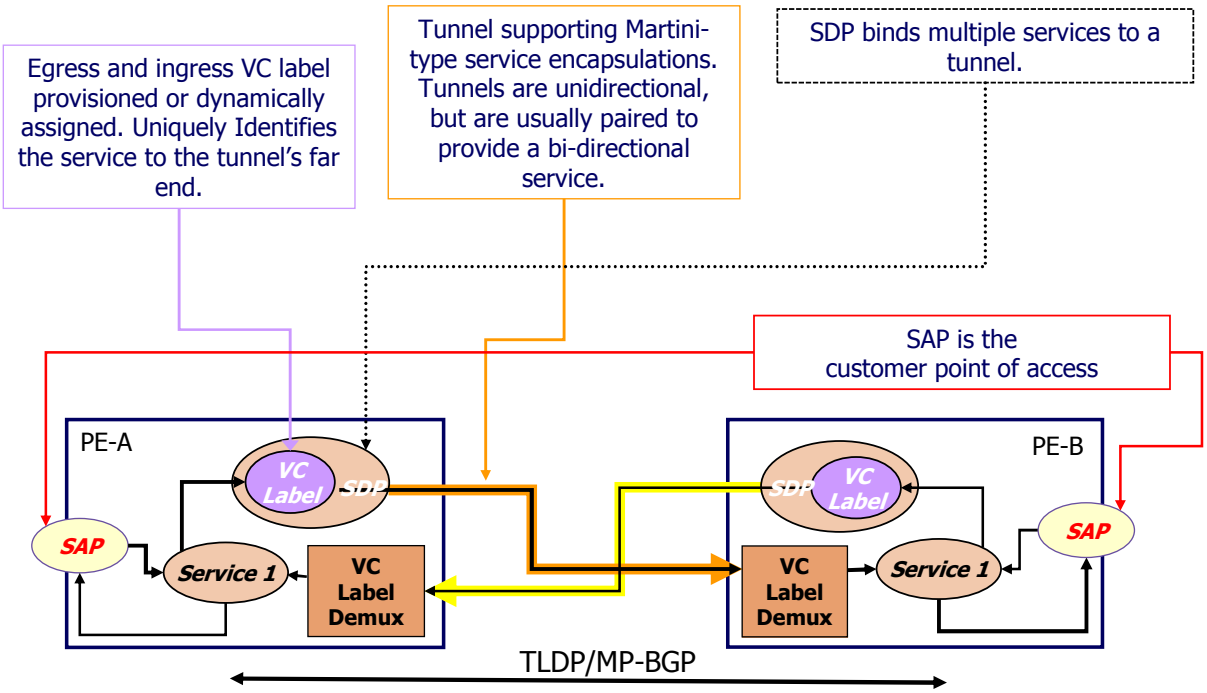
Secondly, one or more logical entities that serve as the customer access to the service, the SAP's, need to be configured.

If you want to provision a VPN service, VPRN, VPLS or VLL, you also need a way of telling the router which other routers this VPN will connect to and how. The SDP or service distribution point defines which other routers a service is connected to. They also define what sort of tunnel encapsulation is used.

Often there is some confusion about the abbreviation SDP. The "P" in the abbreviation is standing sometimes for path or sometimes for point. Although this is just a naming, "point" better describes the SDP as is discussed later.

The auto-bind feature is an option that can be used to auto provision the tunnels.

Logical Service Level Connectivity



The SAP is the customer access point to the service. A SAP can only be provisioned on ports configured as access or hybrid. A port is either access, network or hybrid. A port has a default of "network" for the 7450 or 7750 products and must be changed when SAPs are assigned. The port default is "access" for the 7705 products. The encapsulation type must also be specified when configuring the port.

The SAP is connected by configuration to only one service. But a service can contain multiple SAPs.

When the service is a distributed service, meaning it involves two or more PE sites, SDP's need to be connected to the service. An SDP, represented by a locally unique ID number, representing a GRE or MPLS tunnel to transport the data customer traffic to the other end.

As the SDP is a logical representation of a GRE or MPLS tunnel, it is uni-directional, but almost always paired to another SDP to provide a bi-directional service.

Before the data traffic is tunneled and sent to the other side, an additional label next to the MPLS transport label is added. This label, called the inner, service label or VC label, uniquely identifies the service to the tunnel's far end. The far end has a de-multiplexer mechanism so that after popping the MPLS transport label, the VC label decides to which service the traffic belongs to. A service router can contain multiple services all making use of the same SDP or service tunnel.

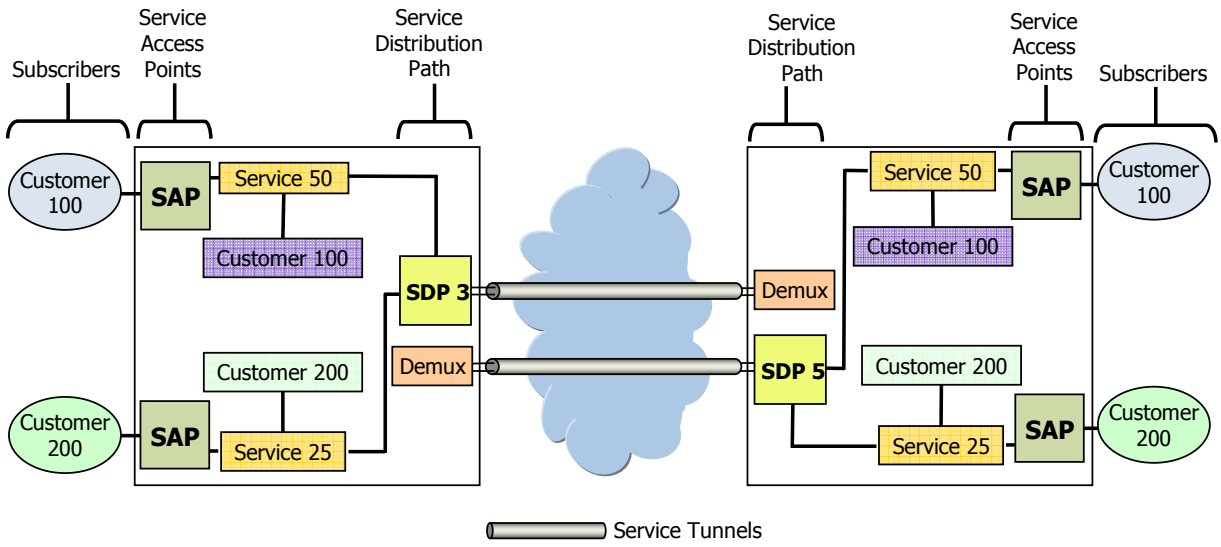
VC labels are only significant at both sides of the PE router. The intermediate P routers are unaware of this inner label and leave the label untouched.

The VC-label negotiating or distributing is performed by T-LDP or MP-BGP. T-LDP or targeted LDP is an adaptation the known LDP MPLS label distribution. LDP distributes label between directly connected peers and T-LDP distributes service labels between service PE routers. T-LDP

Service components

Section 4: Service components.

Service Components



Let us review the major components of a service.

A service uses a SDP service tunnel to direct traffic from one router to another. The service tunnel is provisioned with an encapsulation type of GRE, MPLS, or LDP and services are mapped to the service tunnel that most appropriately supports the service requirements.

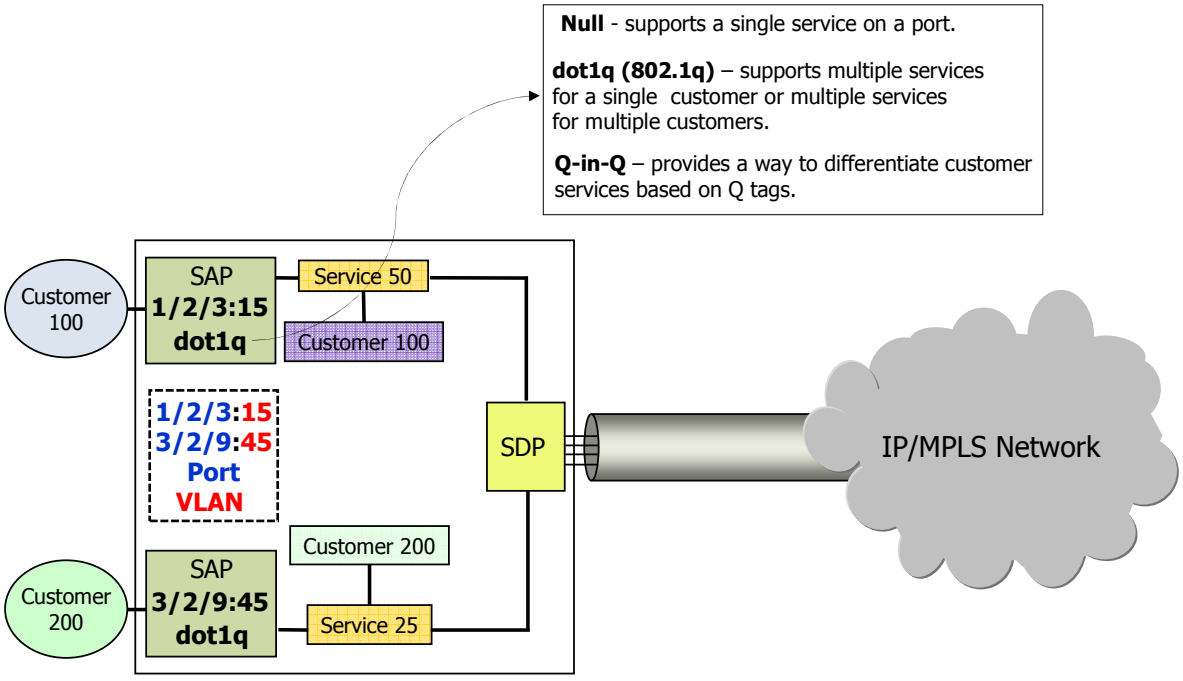
A SDP is configured before a service and different services can use this tunnel. A SDP is not a protocol, but a logical entity represented by a number that is linked to for example a LSP in case of a MPLS RSVP-TE SDP.

A SAP is a logical entity that serves as the customer's point of access into a service. Each subscriber service is configured with at least one SAP. A SAP can only be configured on a port that has been configured as an access port. The default configuration for a port is "network" or "access" or "hybrid" which means that you may need to reconfigure a port before you can configure it as a SAP. SAPs for IES and VPRN services are configured on IP interfaces within the service.

A SDP acts as a logical way of directing traffic from one router to another through a uni-directional service tunnel. An SDP originating on one node terminates at a destination node, which then directs incoming packets to the correct service egress SAPs on that node. A multi-node service needs at least one SAP and one SDP on each node. For a service to be bi-directional, a SDP must be provisioned on each node participating in the service.

The customer ID is an ID that is mandatory at the time the service is created. It is a local significant number that is mainly used to link customer credentials to a service. These credentials can be a phone number, name and address. When not needed, the default customer ID of 1 can be used.

Service Access Point (SAP)

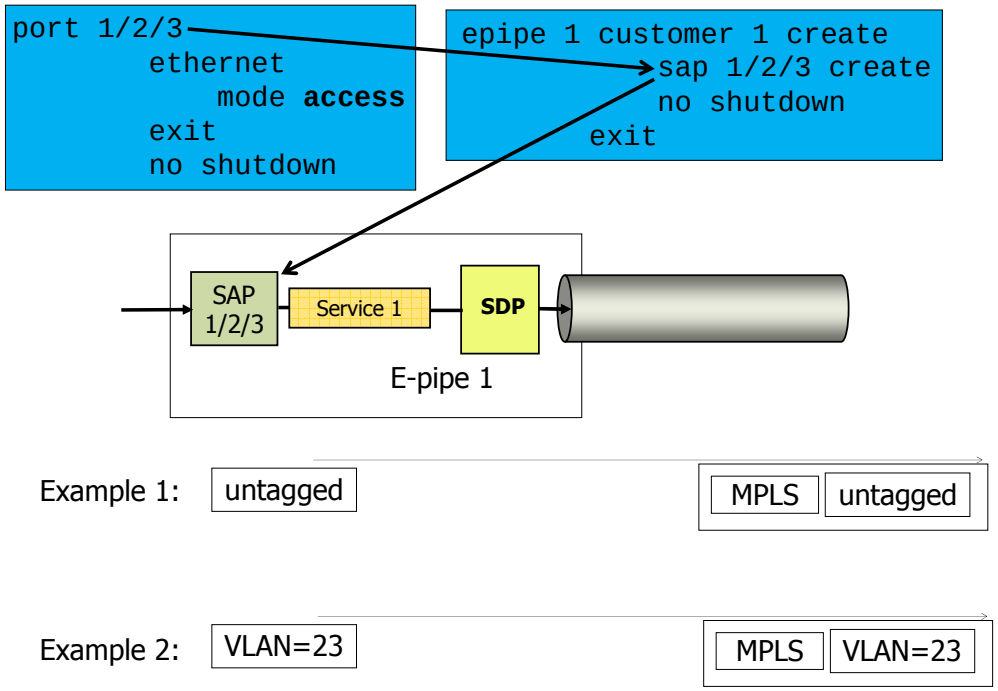


A SAP is identified by its port and by its encapsulation type. In case of Ethernet, three options of encapsulation are possible, null, dot1q and Q-in-Q.

In this example, each service on this service router contains one SAP with dot1Q encapsulation. The SAP for service 50 is 1/2/3:15 and the other service 25 has SAP 3/2/9:45. 15 and 45 are examples of VLAN's that are expected on the port from the customer. They will be stripped off before the packet handed over the service instance and sent to the other side. Instead of a value, a star can be set as VLAN value. This SAP, called the default SAP, accepts any VLAN value from the customer. But in this case, the VLAN is preserved and send to the other service side.

A null encapsulated port is meant for untagged or priority tagged frames. However, when a VLAN value is set anyway by the customer, it will be treated as customer traffic and will not be stripped off.

SAP example: Null SAP



This slide illustrates an example of a service with null encapsulation.

In the port context, port 1/2/3 has been adapted to mode access. The encapsulation type has not been changed which results in the default of null encapsulation.

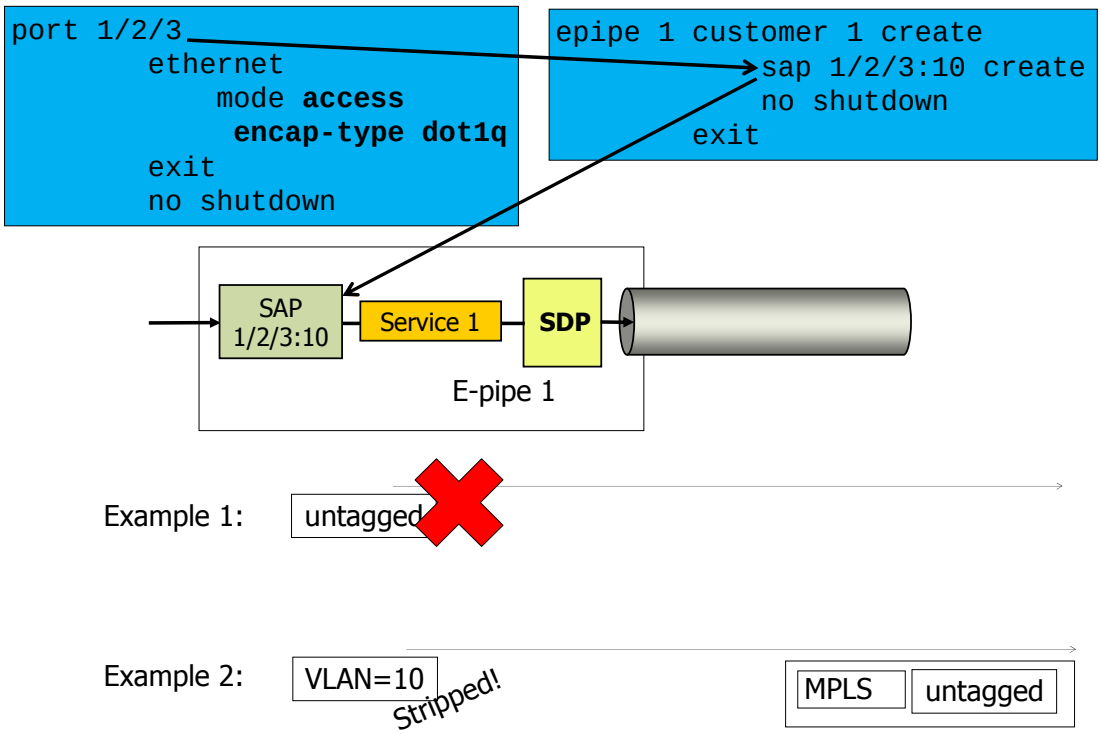
E-pipe with service id of one has SAP 1/2/3 connected.

Example 1 contains a frame of a customer on port 1/2/3 with no VLAN present. The customer data is encapsulated into MPLS or GRE and sent to the other far end of the service.

Example 2 contains a customer frame with a VLAN of 23. Because this traffic arrives on ports with null encapsulation, this VLAN tag is seen as customer traffic and not stripped.

Null encapsulated SAP's can also accept tags with VLAN 0. The value of 0 is ignored, but the priority bits are used for classification in QoS.

SAP example: dot1q SAP



This slide illustrates an example of a service with dot1q encapsulation.

In the port context, port 1/2/3 has been adapted to mode access and the encapsulation type is set to dot1q.

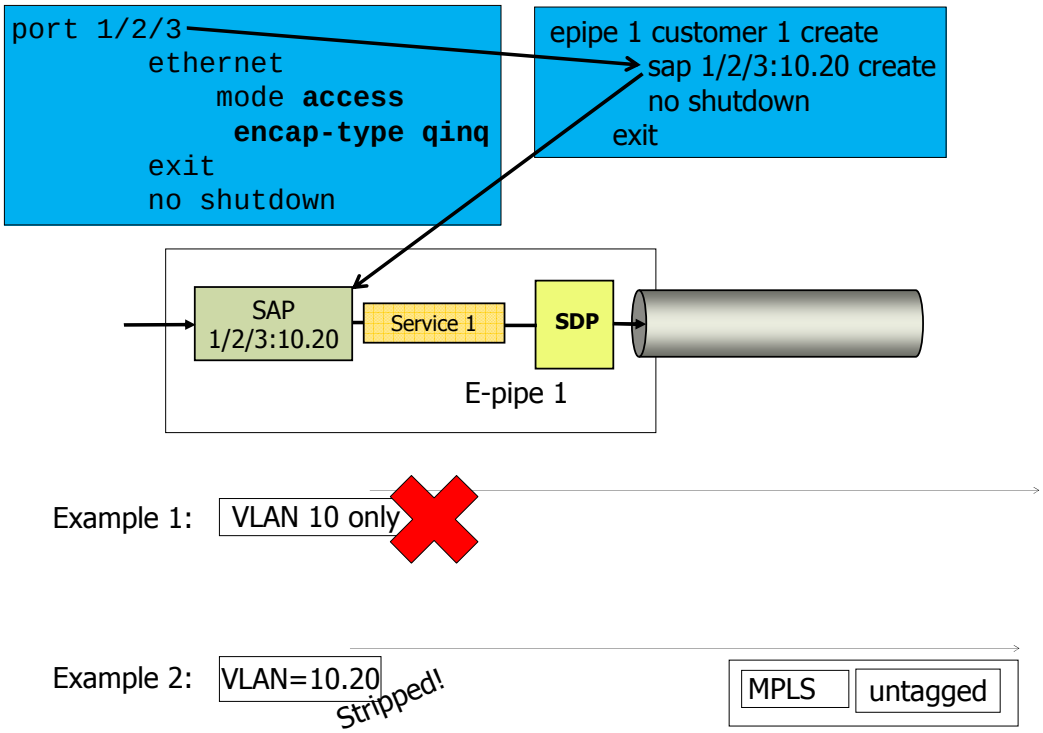
E-pipe with service id of one has SAP 1/2/3:10 connected. Trying to create a SAP without a VLAN tag specified would generate an error message, preventing the sap creation.

Example 1 contains a frame of a customer on port 1/2/3 with VLAN ten present. The customer data is discarded as no match was found.

Example 2 contains a customer frame with a VLAN of 10. VLAN 10 matches the configured SAP value and is therefore stripped and sent as untagged customer traffic into the network.

In case of the default SAP 1/2/3:* would have been configured, both the untagged and tagged frame would have been accepted. Nothing would have been stripped.

SAP example: q-in-q SAP



This slide illustrates an example of a service with q-in-q encapsulation.

In the port context, port 1/2/3 has been adapted to mode access and the encapsulation type is set to q-in-q.

E-pipe with service id of one has SAP 1/2/3:10.20 connected. Trying to create a SAP without VLAN tags specified will be impossible and result in an error message.

Example 1 contains a frame of a customer on port 1/2/3 with VLAN ten present. The customer data is discarded as no exact match was found. VLAN 10 was there, but second QTag was missing. The frame would have been accepted if SAP 1/2/3:10.* was configured. VLAN 10 would have been stripped and VLAN 20 would have been preserved.

Example 2 contains a customer frame with a VLAN of 10 and 20. This matches exactly the SAP value and hence all VLANs are stripped.

As a general rule we can say, a match of a VLAN or VLANS's means stripping the VLAN values.

SAP service characteristics

ePipe:

- SAPs can be defined on Ethernet ports, SONET/SDH or TDM channels
- Two SAPs are identified for a service that originates and terminates on the same device; one SAP and one SDP are defined for each service device in a two-node ePipe service

IES:

- SAPs can be defined on Ethernet ports, SONET/SDH or TDM channels
- SAPs are associated with an IP service; one SAP per logical interface

VPLS:

SAPs can be defined on Ethernet or POS (BCP) interfaces

VPRN:

SAPs can be defined on Ethernet ports or SONET/SDH (IPCP) interfaces

What are the service characteristics of a SAP per type of service.

E-pipe:

SAPs can be defined on Ethernet ports, SONET/SDH or TDM channels.

Two SAPs are identified for a service that originates and terminates on the same device; one SAP and one SDP are defined for each service device in a two-node ePipe service.

IES:

SAPs can be defined on Ethernet ports, SONET/SDH or TDM channels.

SAPs are associated with an IP service; one SAP per logical interface.

VPLS:

SAPs can be defined on Ethernet or POS interfaces.

VPRN:

SAPs can be defined on Ethernet ports or SONET/SDH (IPCP) interfaces.

SAP configuration considerations

Consider the following when configuring a SAP:

- A SAP is locally unique, in other words the same SAP ID value may be used on another device
- A SAP is associated with a single service and can only be configured on an access port
- A port or channel can have more than one SAP configured on it
- All SAPs must be explicitly created and are administratively enabled at the time of creation (no default SAPs)
- VLAN IDs have local significance
- A SAP can be configured with an:
  - Ingress and egress filter policy
  - Ingress and egress QoS policy
  - Ingress and egress scheduler policy
  - Accounting policy

When configuring a SAP the following considerations need to be taken into account.

A SAP is locally unique, in other words the same SAP ID value may be used on another device.

A SAP is associated with a single service and can only be configured on an access port.

A port or channel can have more than one SAP configured on it.

All SAPs must be explicitly created and are administratively enabled at the time of creation. There are no default SAP's.

VLAN IDs have local significance.

A SAP can be configured with an:

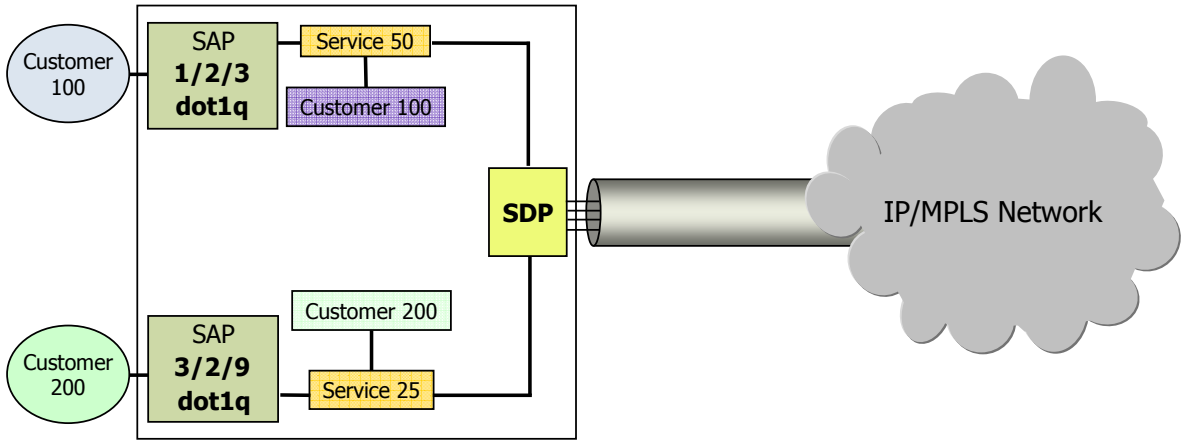
•Ingress and egress filter policy

•Ingress and egress QoS policy

•Ingress and egress scheduler policy

•And an accounting policy

Service Distribution Path (SDP) characteristics



- SDPs are locally unique; the same SDP ID can be used on another device
- SDPs use the System IP address to identify far end destinations
- SDP is not specific to one service; many services can use the same SDP
- All services bound to an SDP must use the same type of encapsulation (GRE, MPLS, or LDP)
- Operations on an SDP will affect all services that are bound to that SDP

A service distribution path or point acts as a logical way of directing traffic from one device to another through a uni-directional service tunnel.

The SDP terminates at the far end on a demulitiplexer where a forwarding decision is made based on the second label of the MPLS stack.

A node with a service that is connected to two remote service sites needs two SDPs, one SDP per far end.

The SDP must be configured before a VPN service can be assigned to it. The SDP must then be bound to the service in order for service traffic to be directed to the far-end of the service tunnel.

It is possible to collect ingress octets forwarded/dropped and egress octets forwarded/dropped statistics for every SDP on a per-service basis. Statistics are available via local show commands or management system.

An accounting file can also be created to collect SDP statistics and provide a way to bill customers or partners on a per-SDP per-byte basis.

SDP Encapsulation Types

GRE

- Encapsulates traffic in an IP/GRE header; appears like an IP packet.
- Low control plane overhead.
- Uses normal IP routing to find a path.

MPLS

- Uses LDP or RSVP for label signaling.
- LDP auto-bind available to simplify configuration.
- LDP uses an IGP or static configuration to find its path
- RSVP
  - Requires manual configuration
  - Can be loose or strict
  - May reserve bandwidth
  - Can use Fast ReRoute to speed convergence

SDP represent a GRE or MPLS tunnel.

GRE appears like an IP packet and has low control plane overhead. Because it relies on the normal IP routing to find a path through the network it misses all the great resiliency features of MPLS.

The recommended encapsulation method for SDP's is MPLS. The SDP's bound to an MPLS based tunnel are either LDP or RSVP-TE signaled. LDP auto-bind feature is available to simplify the configuration. RSVP-TE based SDP can be more resilient than an LDP based one since LDP relies on the IGP conversion time.

RSVP-TE based SDP can have loose or strict hops, reserve bandwidth or use the sub 50 ms fast re-route features.

SDP configuration considerations

Consider the following SDP characteristics:

- SDPs can be created as either GRE or MPLS
- Each multi-site service must have an SDP defined for every remote PE to provide Epipe, VPLS, and VPRN services
- A service must be bound to an SDP. By default, no SDP is associated with a service. Once an SDP is created, services can be associated with that SDP
- An SDP is not specific or exclusive to any one service or any type of service
- An SDP can have more than one service bound to it
- In order to configure an MPLS SDP, an MPLS infrastructure must be in place
- In the SDP configuration, automatic ingress and egress labeling (targeted LDP) is enabled by default.
- Ingress and egress VC labels are signaled over a TLDP connection between two PE routers

Let's review the typical SDP characteristics.

SDPs can be created as either GRE or MPLS.

Each multi-site service must have an SDP defined for every remote PE to provide Epipe, VPLS, and VPRN services.

A service must be bound to an SDP. By default, no SDP is associated with a service. Once an SDP is created, services can be associated with that SDP.

An SDP is not specific or exclusive to any one service or any type of service. An SDP can have more than one service bound to it.

In order to configure an MPLS SDP, an MPLS infrastructure must be in place.

In the SDP configuration, automatic ingress and egress labeling (targeted LDP) is enabled by default. Ingress and egress VC labels are signaled over a TLDP connection between two PE routers.

Configuring SDPs

```
config>service>sdp sdp-id [gre | mpls]

    description description-string
    far-end ip-addr
    keep-alive
    ldp (only for MPLS SDPs)
    lsp lsp-name [lsp-name] (only for MPLS SDPs)
    signaling {off|tldp}
    no shutdown
```

Creating an SDP for MPLS (RSVP-TE)

```
config>service# sdp 8 mpls create
config>service>sdp# far-end 10.10.10.104
config>service>sdp# lsp "to-104"
config>service>sdp# no shutdown
```

ing an SDP with GRE

```
ig>service# sdp 122 gre create
config>service>sdp$ far-end 10.10.10.122
ig>service>sdp$ no shutdown
```

Creating an SDP with MPLS LDP

```
config>service# sdp 104 mpls create
config>service>sdp# far-end 10.10.10.94
config>service>sdp# ldp
config>service>sdp# no shutdown
config>service>sdp# exit
```

These CLI examples show the three options used for creating SDP's with MPLS RSVP-TE, MPLS LDP and GRE.

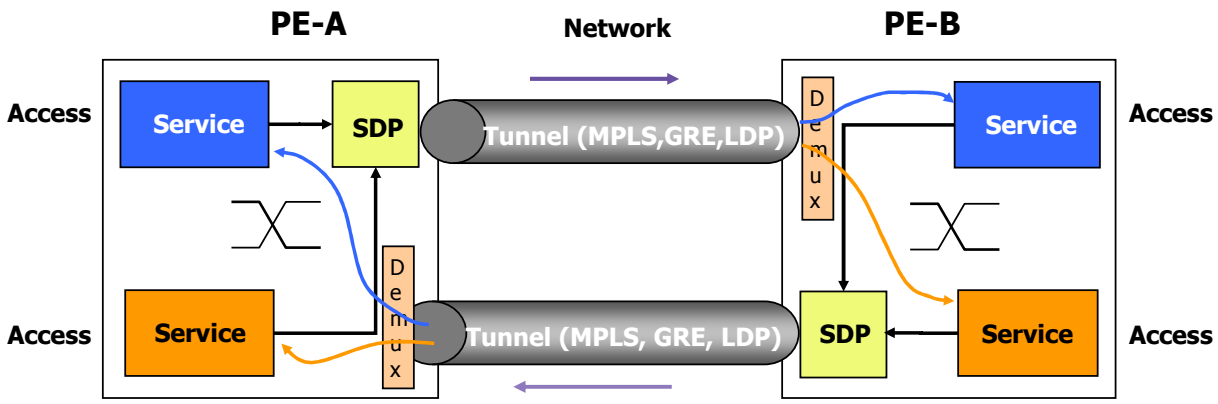
When nothing is specified during the creation of the SDP, the SDP uses the default of GRE.

SDP 8 uses RSVP-TE. Beside the far end node system address, the MPLS LSP is assigned.

SDP 104 uses LDP and therefore only the keyword LDP has to be mentioned. If LDP is enabled on the appropriate interfaces, LDP will create it paths automatically to the far-end.

SDP 122 uses GRE and will find its far end path through the IGP.

Service Tunnels and VPN Services

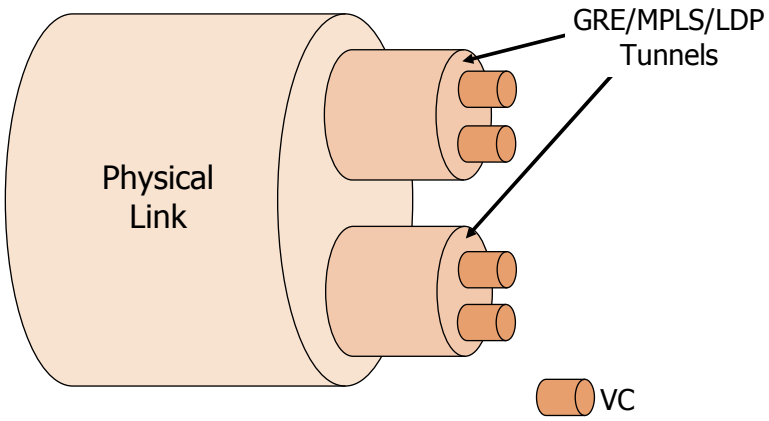


- Once a tunnel is created, multiple services can be carried in it.
- Operations on the tunnel affect all the services that are associated with the tunnel.
- A tunnel uses the system IP address to identify the far-end edge router.

This example illustrates two services making use of the same SDP.

Traffic from both services at PE-A is inserted into the same SDP and hence the tunnel. The destination node directs packets to the correct service egress interfaces on that device. This is the responsibility of the demultiplexer.

Physical links, tunnel LSPs and VCs



A physical link between a service router has one or multiple GRE or MPLS tunnels inside. Each of the tunnels can carry traffic from different services.

MTU - Maximum Transmission Unit

Section 5: MTU.

Maximum Transfer Unit (MTU) considerations

An MTU defines the maximum size of a packet in bytes. Limiting the MTU size of a packet is done for different reasons.

Alcatel-Lucent SR-OS service routers support five different MTU settings:

- Port MTU: Determines maximum packet size on a given physical wire
- Service MTU: Associated with a service and determines the maximum packet size that can be sent from the customer across the service.
- SDP Path MTU: This is the MTU of the SDP between the service endpoints; determines the maximum packet size that can be sent over the SDP
- VC-MTU: Negotiated by T-LDP; the maximum IP payload size that can be carried inside the tunnel. Derived from Service MTU
- IP-MTU: Used by IES and VPRN interfaces for spoke-SDP terminations; should be equal to the VC-MTU of the spoked Epipe/VPLS

First 4 MTU values are important to get Services in the Operational UP state

The Maximum Transmission Unit or MTU defines the maximum size of a packet in bytes. Limiting the MTU size of a packet is done for different reasons. Some servers might not be able to handle large packets. Routers in the network might also create too much delay when processing long packets.

Traditional routers can only change the MTU on the level of the port. Modern service routers have more options to set MTU values. This is needed, because on a port, there is a mix of paths, packets and services each may have different MTU requirements.

The Alcatel-Lucent SR-OS service routers support five different MTU settings: the port MTU, the service, SDP path, VC-MTU and IP MTU.

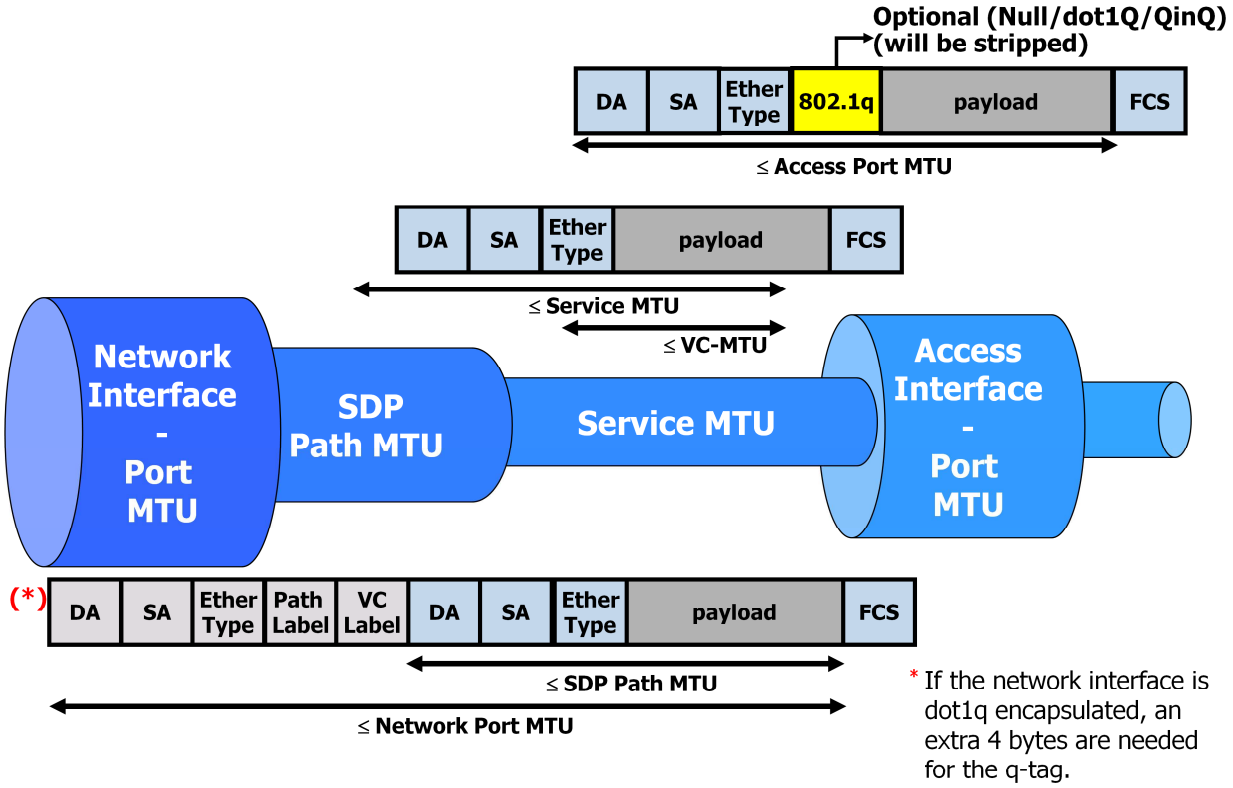
The Port MTU determines the maximum packet size on a given physical wire. It should be the highest overall MTU value set.

The service MTU is associated with a service and determines the maximum packet size that can be sent from the customer across the service. This value will be or can be different for each service. For example, two services configured, one VPLS service allowing customer packets of 1500 bytes and an e-pipe service accepting customer packets of maximum 1600 bytes.

The SDP path MTU should then be set to a minimum of 1600 bytes if it carries those two services. A SDP-path MTU defines the MTU of packets carried within the path. The VC-MTU is related to the service MTU. It is the service MTU minus the ethernet overhead in case of L2 VPN's. This MTU should match at both sides of the service. A mismatch will bring the service down.

The equivalent to the VC-MTU in L2 VPN's is the IP-MTU for L3 based services, IES and VPRN.

Layer 2 [L2] Service MTUs (Ethernet)



This picture graphically shows the different MTU sizes in relation to each other. When designing a network, you need to start with the service MTU as this not only the smallest value to begin with but also defines the customer packet size as the limiting factor. An operator typically wants to control the size of the packet a customer will send through the network.

For L2 VPN's, the VC-MTU is derived from the service MTU and negotiated by the TLDP protocol. A difference on the VC-MTU will bring the service down. If not manually specified, the VC-MTU is derived and therefore the service-MTU should be set the same at both sides. There are default values of service MTU's per service type. So if both sides of a service are of the same service type, the service MTU will be the same.

Next exercise is to look to the SDP path MTU. It should be greater or at least the same as the highest available service MTU on the path. The path MTU is not derived from the service MTU as this could mean a change every time there is a new service deployed, but it is taken from the network port MTU. The path MTU is the network port MTU minus the Ethernet DLC overhead minus the MPLS labels.

The access MTU is a port MTU but at the side of the SAP. It needs to be equal to the service MTU + any VLAN encapsulation.

Access Port and Service MTU

Access Port MTU:

- configure port x/x/x ethernet mtu [512..9212] bytes
- Default values: 1514 for Null, 1518 for Dot1Q, 1522 for Q-in-Q
- show port

Service MTU:

- configure service [vpls|[e|a|i|f]pipe] service-mtu [1..9194] bytes
- default value: 1514
- show service id x base
- not used in L3 services

VC MTU:

- Derived from the Service MTU (Service MTU – 14 Ethernet)
- show router ldp bindings (service-id <service-id>)

SDP Path MTU:

- Should be equal or smaller than (Network Port MTU – 2 labels – Ethernet header)
- configure service sdp x path-mtu [576..9194] bytes
- show service id x base

Here you see the most important configuration and show commands.

L2 – Minimum Physical MTU on Network Interfaces (MPLS)

If a router needs to support services offering a 1514 byte service payload over MPLS tunnels:

Ethernet	POS	Overhead
1514	1514	Service Payload
4	4	MPLS tag used as service ID
4	4	MPLS tag used for egress LSP
(4)	(4)	MPLS tag (if FRR bypass is used)
n/a	2	PPP MPLSCP Header
14	n/a	DLC Header
1536 (1540)	1524 (1528)	Total

An example:

If we start from the given that the network I/F port is a 10/100 Ehernet port, the default MTU value is 1514 and that the service size that needs to be supported is 1514 bytes on the I/F, the minimum physical port MTU should be 1536 bytes. Two times four bytes of MPLS tags and 14 bytes of DLC added. In case of fast re-reroute, one to one backup method, an additional 4 bytes needs to be added." Type of Ethernet port must be specified as if a Gig E was used, the default would be 9212.

On a POS interface, the overall port MTU is less as an on Ethernet interface. The DLC header is here replaced by the PPP MPLSCP header.

Minimum Physical MTU on Network Interface (GRE)

If a router needs to support services offering a 1514 byte service payload over GRE tunnels:

Ethernet	POS	Overhead
1514	1514	Service Payload
4	4	MPLS tag used as service ID
8	8	GRE Header
20	20	IP Header
n/a	2	PPP IPCP Header
14	n/a	DLC Header
1560	1548	Total

Same exercise as before, but now on a GRE based SDP. The result is a higher port MTU as the GRE header and IP header need to be taken into account.

show service id "x" all - command

This command can be used to verify MTU

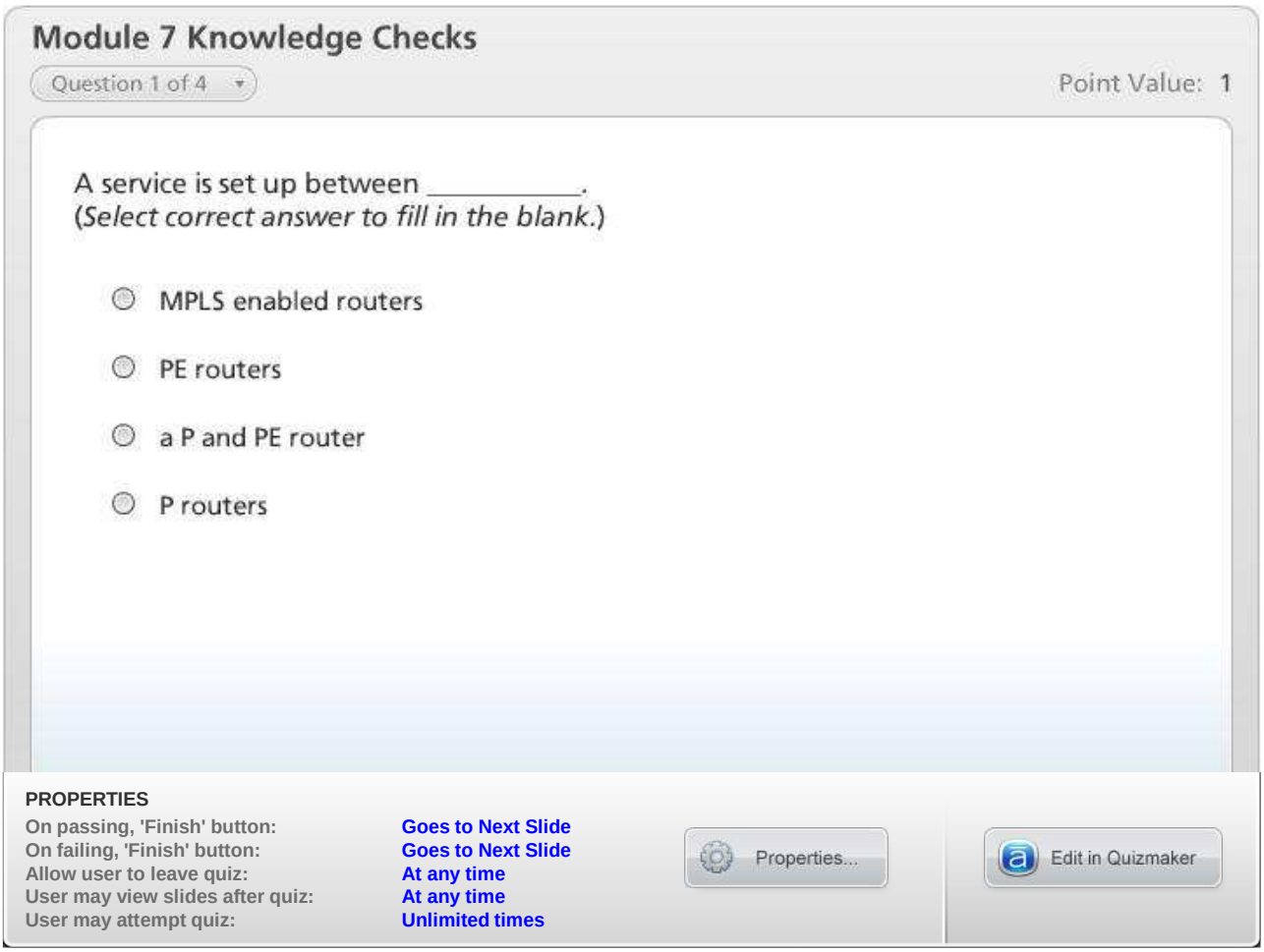
```
SR12# show service id 5500 all
===============================================================================
Service Detailed Information
===============================================================================
Service Id        : 5500              Vpn Id            : 5500
Service Type      : Epipe
Description       : Distributed epipe service to east coast
Customer Id       : 5 Last            Mgmt Change       : 07/14/2003 03:26:46
Adm               : Up                Oper              : Down
MTU               : 1514
SAP Count         : 1                 SDP Bind Count    : 1
-------------------------------------------------------------------------------
Service Destination Points(SDPs)
-------------------------------------------------------------------------------
Sdp Id 2:123 -(10.10.10.104)
-------------------------------------------------------------------------------
Description       : GRE-10.10.10.104
SDP Id            : 2:123             Type              : Spoke
Admin Path MTU    : 4462              Oper Path MTU     : 4462
Far End           : 10.10.10.104      Delivery          : GRE
Admin State       : Up                Oper State        : TLDP Down
Ingress Label     : 6600              Egress Label      : 5500
Ingress Filter    : n/a               Egress Filter     : n/a
Last Changed      : 07/14/2003 03:29:21 Signaling       : TLDP
```

The *show service id* command allows you to check the MTU that is configured and the actual MTU that will be supported across the entire network connection. The slide shows part of the information provided by this command.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



WRAP-UP
Module Summary

Module summary.

Module Summary

- IP/MPLS is the base ground of the service model as we know today
- A SAP, SDP, service ID, customer ID are the main components to configure a service
- T-LDP and MP-BGP are the two service label distribution mechanisms
- GRE and MPLS are the two encapsulation options for service tunnels
- Service tunnel or SDP is not a protocol but a logical representation of a tunnel
- SDP has to be created upfront before bound to a service
- Different MTU values are available to set the maximum size of a packet
- VC-MTU mismatch will bring down the service

The above slide provides a summary of this mod. Please take the time to review it.

End of Module 7

This completes module 7.

SR-OS Fundamentals

Module 8: IES

IPD Development

Welcome to the eighth module of the SR-OS fundamentals course.
Module 8 provides more details on the IES or Internet Enhanced Service.

Agenda

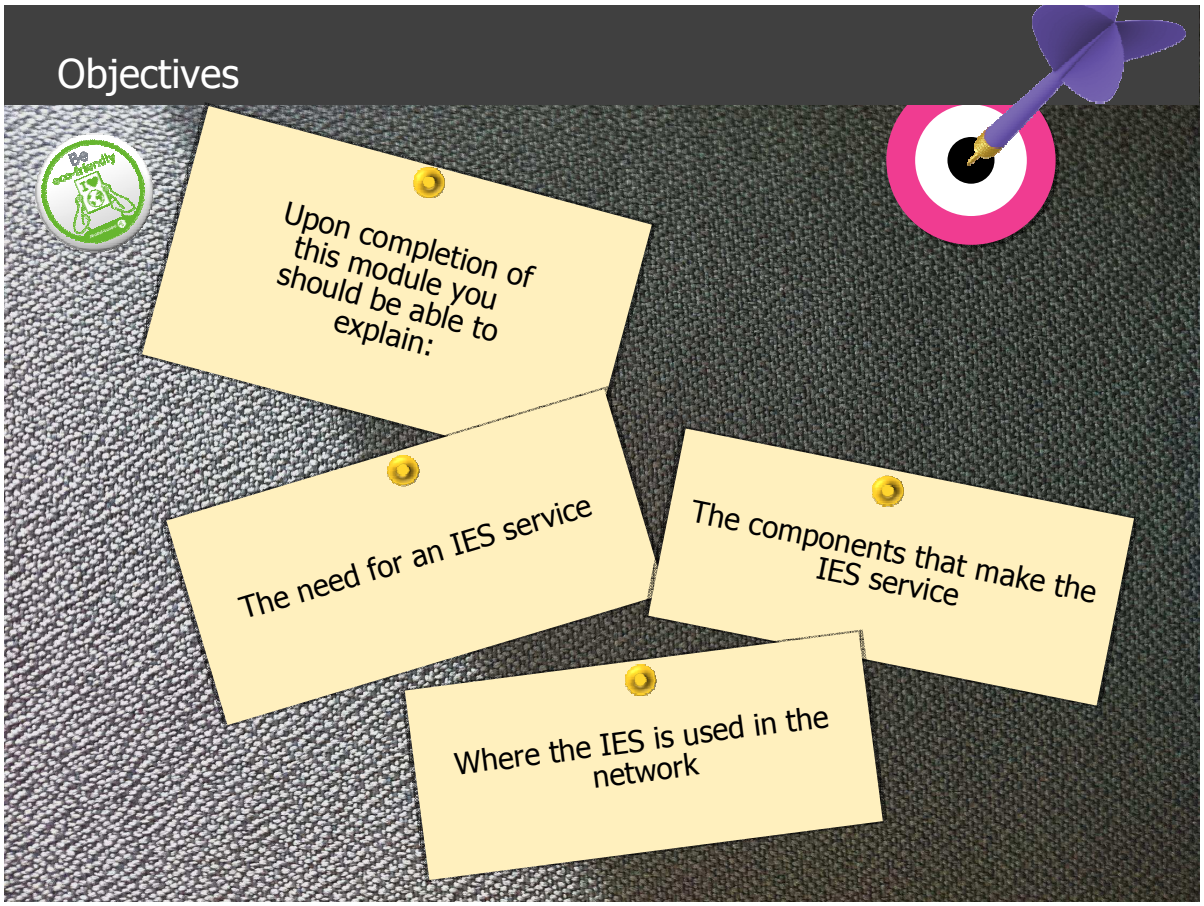
- Module 8:
  - Section 1:
    - Introduction to IES service
  - Section 2:
    - IES service model
  - Section 3:
    - Application

Module 8 is divided into three sections.

Section 1 introduces the IES services

Section 2 explains the IES service model

Section 3 provides information on where the IES service is applied in today's networks



By the end of this module you will be able to explain:

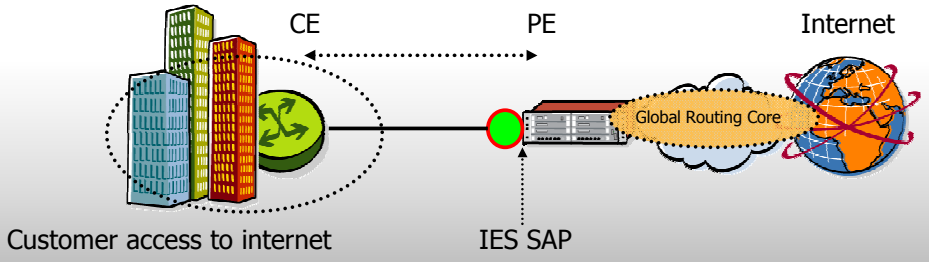
- The need for an IES service
- The components that make the IES service
- Where the IES is used in the network

SECTION 1: INTRODUCTION TO IES SERVICE

Section1: Introduction to IES service

What is an IES service?

- What is an IES?
  - IES provides direct internet access to customers
  - IES is a layer 3 service
  - IES may participate in public IP scheme
  - Global Routing table is used to route traffic from the customer to the destination
  - An IES has one or more logical IP routing interface each with a SAP which acts as the access point to the subscriber's network
  - IES allows customer facing IP interfaces to participate in the same routing instance used for network core routing connectivity



What is an IES?

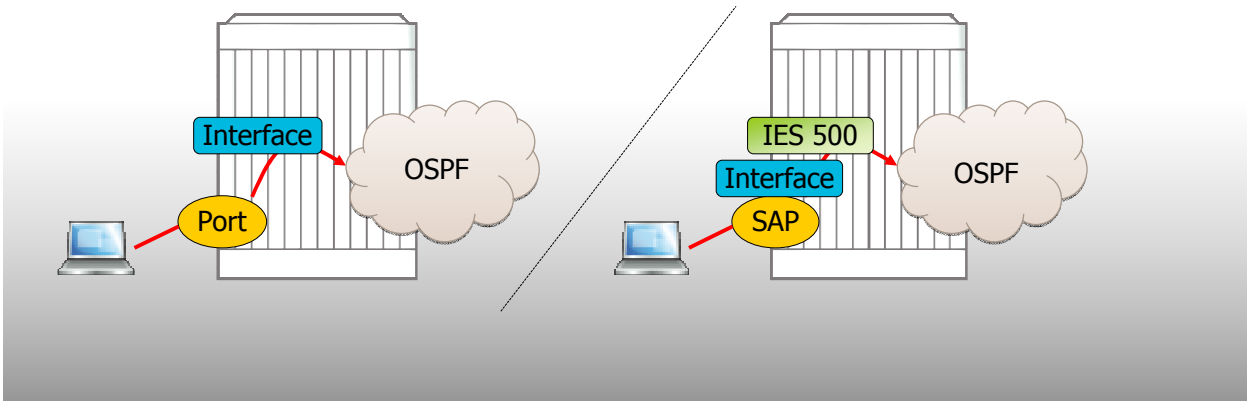
IES or Internet Enhanced Service is a non-VPN type of service. Although there are options to do L3 pseudowire termination as discussed later in this module. But basically, this service provides direct Internet access to customers OR any other layer 3 private network. The "I" in the abbreviation might be misleading here, an IES might participate in the public IP scheme, but might also participate in a private IP network.

The IES is a layer 3 services and uses the global routing table to route traffic from the customer to the destination. The access point for the customer towards the subscriber's network is a SAP which is connected to an IP interface.

The IES allows customer facing IP interfaces to participate in the same routing instance used for network core routing connectivity.

IES interface vs IP interface

- An IES is a routed, layer 3 connectivity service:
  - Interface name
  - SAP
  - IP Address
- A regular IP router interface:
  - Interface name
  - Port
  - IP Address



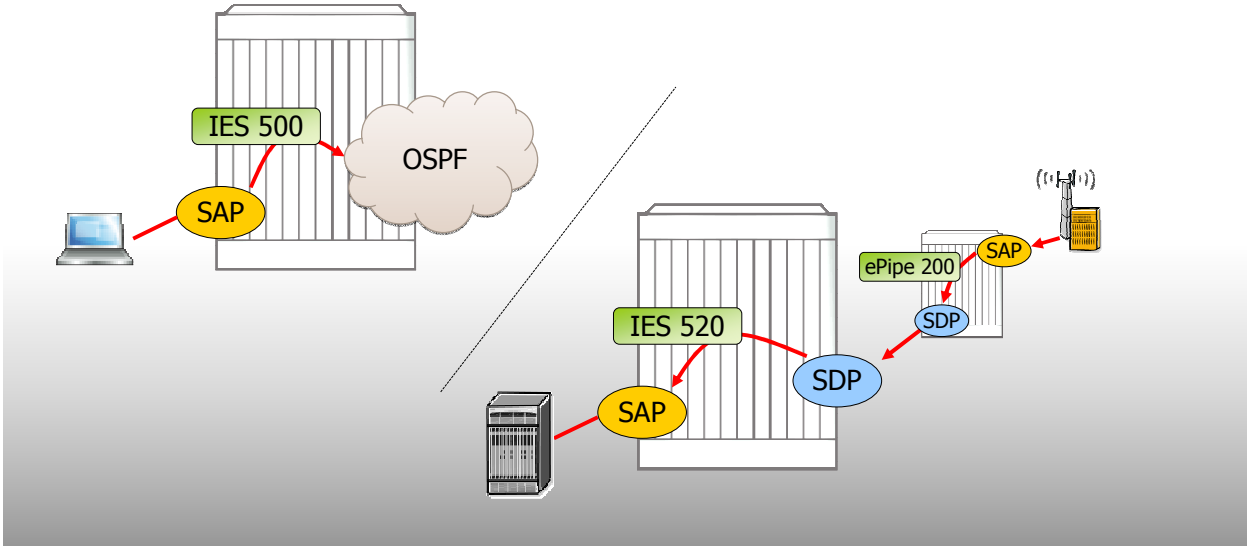
Internet Enhanced Service is a routed connectivity service where the subscriber communicates with an IP router interface to send and receive Internet traffic. It is comparable to a regular router interface and on a first sight it behaves exactly the same as any IP interface on a service router. You could think of offering L3 non-VPN type of services with a regular IP interface. But an IP interface is connected to a port and an IP interface within an IES service is connected to a SAP. This is a fundamental difference as most of the policies are service oriented. SAP ingress and egress QoS, filter and accounting policies work on a basis of a SAP.

Like a regular IP interface, an IES interface can run a routing protocol like OSPF in this example.

An IES can have more than one logical IP routing interfaces each with a SAP which acts as the access point to a subscriber's network.

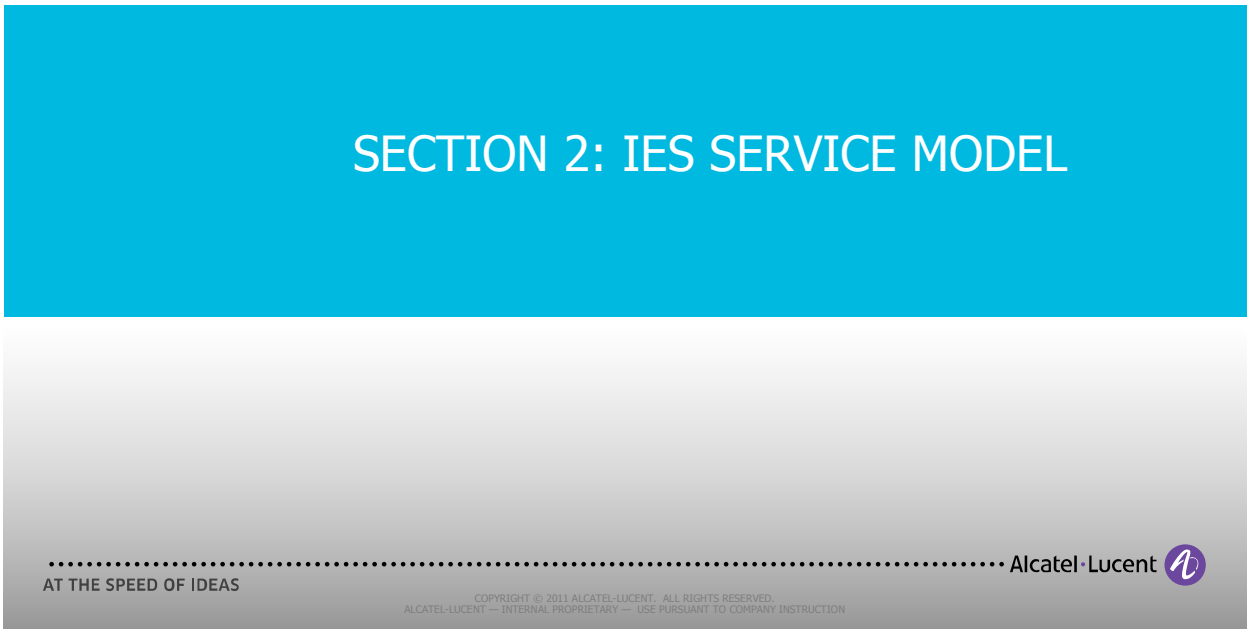
Internet Enhanced Service (IES)

- IES Interfaces can be included in the following routing protocols:
  - Static, RIP, OSPF, ISIS, and BGP
- Other service can be bound to an IES
- IES can terminate SDPs such as an EPipe, IPipe, or H-VPLS service



The interface of an IES service can be included in Static, RIP, OSPF, ISIS, and BGP routing protocols. It runs the routing protocol signaling with the customer as any other routed IP interface.

There is an option not to make the connection directly to a routed network, but use the IES to connect to a SDP. The SDP pseudowire on an PE is bound to the IES service and on the other PE bound to another non-IES service. Typically a VPLS, I-pipe or an e-pipe type of service.

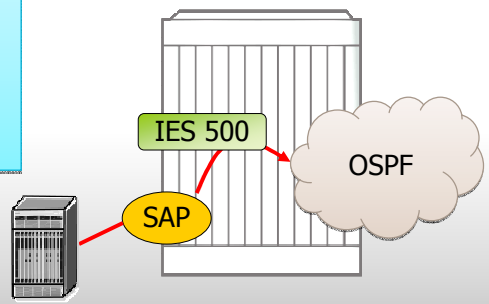


Section 2: IES service model

IES Routed Connectivity Service Example

- Since the traffic in an IES service communicates using an IP interface for the core routing instance, there is no need for the concept of tunneling traffic to a remote router
  - A basic IES does not require the configuration of any SDPs

```
A:SARf-133# configure service ies 1000 customer 1 create
A:SARf-133>config>service>ies# info
----------------------------------------------
        description "IES training"
        interface "to_CE" create
            address 192.168.100.1/30
            sap 1/2/8 create
            exit
        exit
        no shutdown
```



A basic IES, an IES without the connection of SDP's, doesn't need to have the concept of tunneling. After the customer ID and service ID have been defined, three components have to be added as a minimum to bring the service up; an interface name, IP address and a SAP.

After configuration, never forget to do a "no shutdown" to bring the service in the "up" state.

At this moment the IP subnet will appear in the routing table.

IES Verification

- Verify IES interface has been added to the router interfaces:

```
A:SARf-133# show router interface
===============================================================================
Interface Table (Router: Base)
===============================================================================
Interface-Name                   Adm       Opr       Mode    Port/SapId
   IP-Address                                          PfxState
-------------------------------------------------------------------------------
system                           Up        Up        Network system
   10.10.10.133/32                                     n/a
to_CE                            Up        Up        IES     1/2/8
   192.168.100.1/30                                    n/a
to_SAR114                        Up        Up        Network 1/2/1
   192.168.2.1/30                                      n/a
-------------------------------------------------------------------------------
Interfaces : 2
===============================================================================
```

First verification is to see if the interface is up. This can be checked with the same command to see if the other router interfaces are up; "show router interface".

The IES interface in our example is admin and operational up and of mode IES.

IES Verification

- Verify IES interface has been added to the router route table:

```
*A:SARf-133>config>router# show router route-table

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix                       Type    Proto    Age        Pref
      Next Hop[Interface Name]                                 Metric
-------------------------------------------------------------------------------
10.10.10.133/32                    Local   Local    03h07m04s   0
      system                                         0
192.168.2.0/30                     Local   Local    03h05m52s   0
      to_SAR114                                      0
192.168.100.0/30                   Local   Local    03h05m50s   0
      to_CE                                          0
-------------------------------------------------------------------------------
No. of Routes: 3
===============================================================================
```

The subnet appears in the global routing table with the type and protocol of local.

IES Verification

- Activate IES
  - Verify IES is operational

```
A:SARf-133# show service id 1000 base

===============================================================================
Service Basic Information
===============================================================================
Service Id        : 1000
Service Type      : IES
Description       : IES training
Customer Id       : 1
Last Status Change: 05/07/2010 15:39:32
Last Mgmt Change  : 05/07/2010 15:29:51
Admin State       : Up                 Oper State       : Up
SAP Count         : 1

-------------------------------------------------------------------------------
Service Access & Destination Points
-------------------------------------------------------------------------------
Identifier                  Type        AdmMTU  OprMTU  Adm  Opr

sap:1/2/8                   null        1514    1514    Up   Up
===============================================================================
```

The base command to see if any service is admin and operational up is "show service id base". Besides the status, it shows the major components and their status of up or down.

IES Configuration: PE - CE Routing

- Enable CE-PE routing:
  - Routes to customer prefixes reachable from the CE should be configured as static on the PE (optional bfd-enable)

```
A:SARf-133>config# router
A:SARf-133>config>router# static-route 1.1.1.132/32 next-hop 192.168.100.2
```

  - Also, other routes across the network need to be reachable from the CE, thus a default-route may be configured on the CE to reach all other routes via the PE

```
A:SAR132>config>router# static-route 0.0.0.0/0 next-hop 192.168.100.1
```

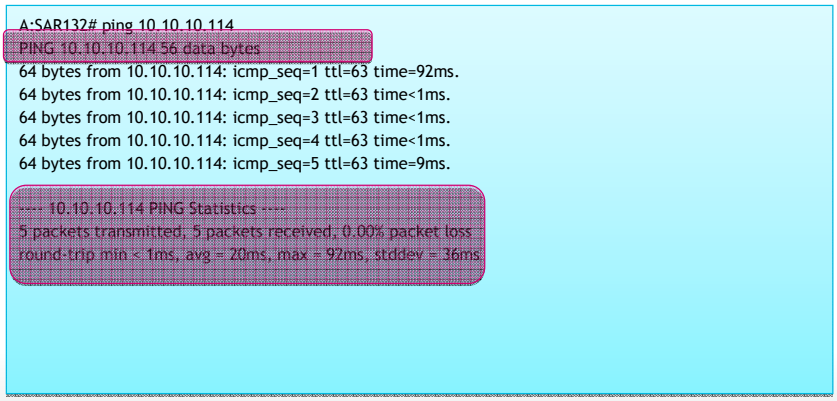
To enable CE-PE routing without the enabling of a routing protocol, the PE and CE should each have a static route. On the PE, a static route is configured with destination prefix the system address of the CE router and the far-end host address as next-hop.

In this way, if a ping is launched from the CE, the PE will know a way back to send the reply.

On the CE, a default route is added towards the PE as all traffic should be sent to the PE.

IES Verification

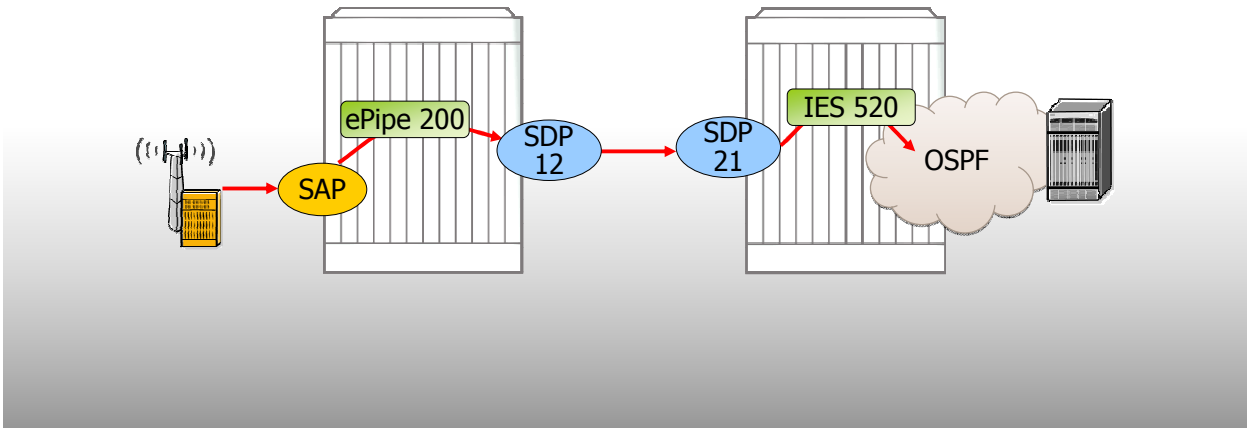
- Activate IES
  - Verify that the far end PE is reachable from the CE



A ping is sent from the CE towards the PE. If the ping was sent to a non directly connected subnet, the system address of the CE is used. If properly configured, the PE sends the reply back to the CE and the ping is successful.

Spoke termination for IES

- Spoke termination:
  - Introduces the ability to cross-connect traffic entering on a spoke SDP to an IES or VPRN service
    - A spoke SDP could be tied to an EPipe, an H-VPLS service, or an IPipe
    - Traffic entering the 7750 via the spoke SDP is classified and queued based on network QoS policies, not access QoS policies



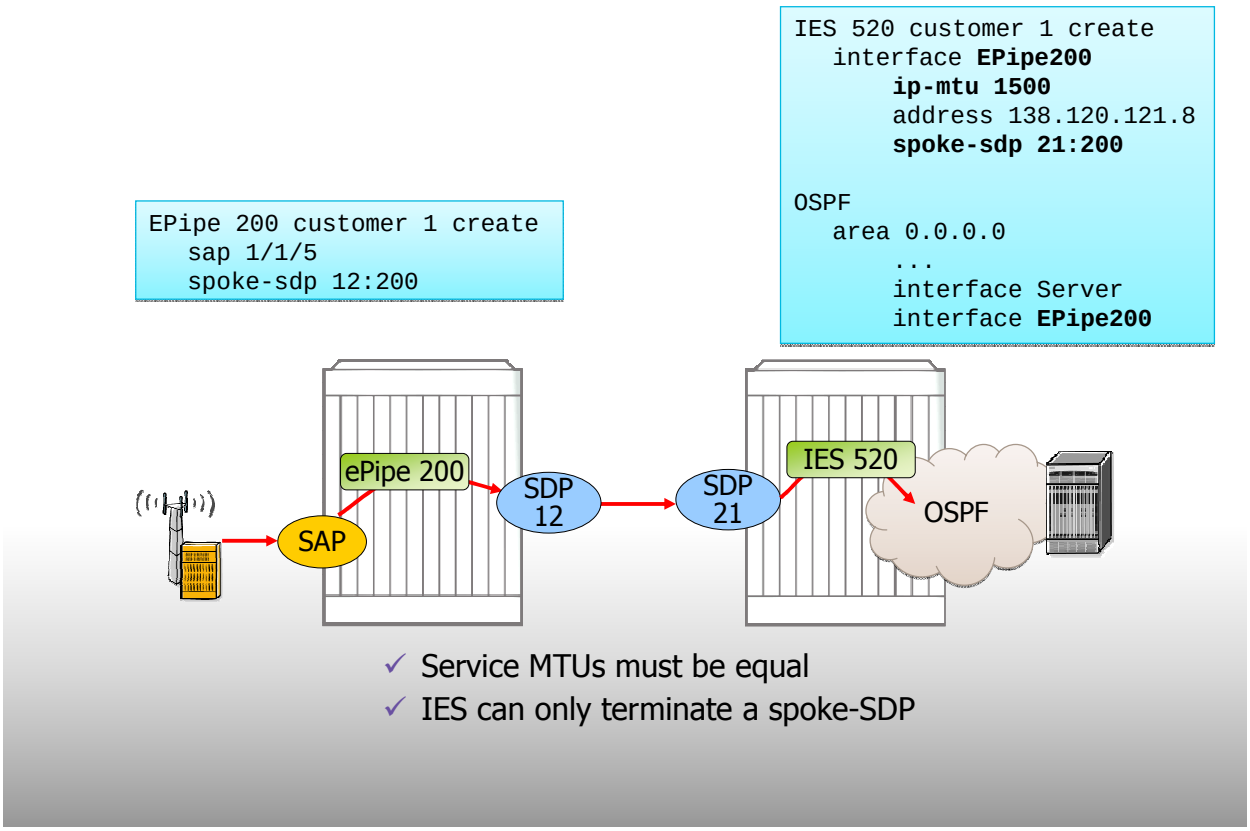
This feature provides the ability to cross-connect traffic entering on a spoke SDP, used for Layer 2 VLL or VPLS type of services, on to an IES service. From a logical point of view the spoke SDP entering on a network port is cross-connected to the layer 3 service as if it had entered via a service SAP. The traffic entering the Layer 3 service via a spoke SDP is handled via network QoS policies rather than an access QoS policies.

In this example, the traffic from the mobile station is sent to the SAP and further directed to the e-pipe SDP 12. The pseudowire or SDP on the other PE terminates on an IES service, rather than on an e-pipe service. This is also called a composite service. An IP address look-up at the global routing table is performed and traffic is further routed over the OSPF network in this case.

Traffic to be terminated on a given IES service is identified by the VC label present in the service packet.

The VC labels or service labels on both sides have to match. They are distributed by T-LDP.

Spoke termination for IES



The main difference in the IES service configuration compared to a SAP based IES interface configuration is the spoke-SDP binding to the interface instead of a SAP.

There are two types of SDP's, spoke and mesh, as explained in a later module. The type on an IES service is always spoke.

Special care on the MTU settings should be taken as these are composite services. And composite service do come with different default MTU settings.

The VC-MTU, derived from the service MTU in case of an e-pipe service has to match the VC-MTU of the IES. And in an IES, the IP-MTU is used as the negotiated VC-MTU value. The service MTU for an e-pipe is 1514. The VC-MTU is the service MTU minus the DLC overhead which results in 1500 bytes used at the negotiation phase. As the IES uses the IP-MTU derived from the SDP path MTU, this value has to be manually adapted to 1500.

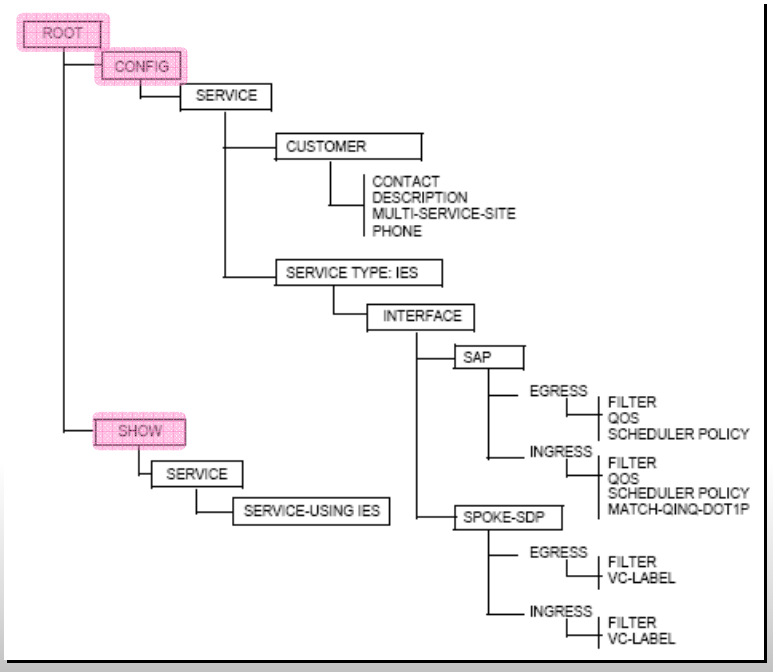
In this configuration it is important to note that during the VC Label signaling process, if the VC MTU does not match, the service will be operationally down.

"show service id 520 base" displays the Operational MTU for the IES

"show service id 200 base" displays the Operational MTU for the EPipe

An alternative to changing the maximum MTU for the IES interface, the SDP maximum MTU of 1514 could have been changed but this would impact all services carried by the SDP. Any service requiring a larger MTU than 1514 would go operationally down.

IES CLI Structure



This example shows the CLI tree structure from the root with the "config" and "show" context for the IES service.

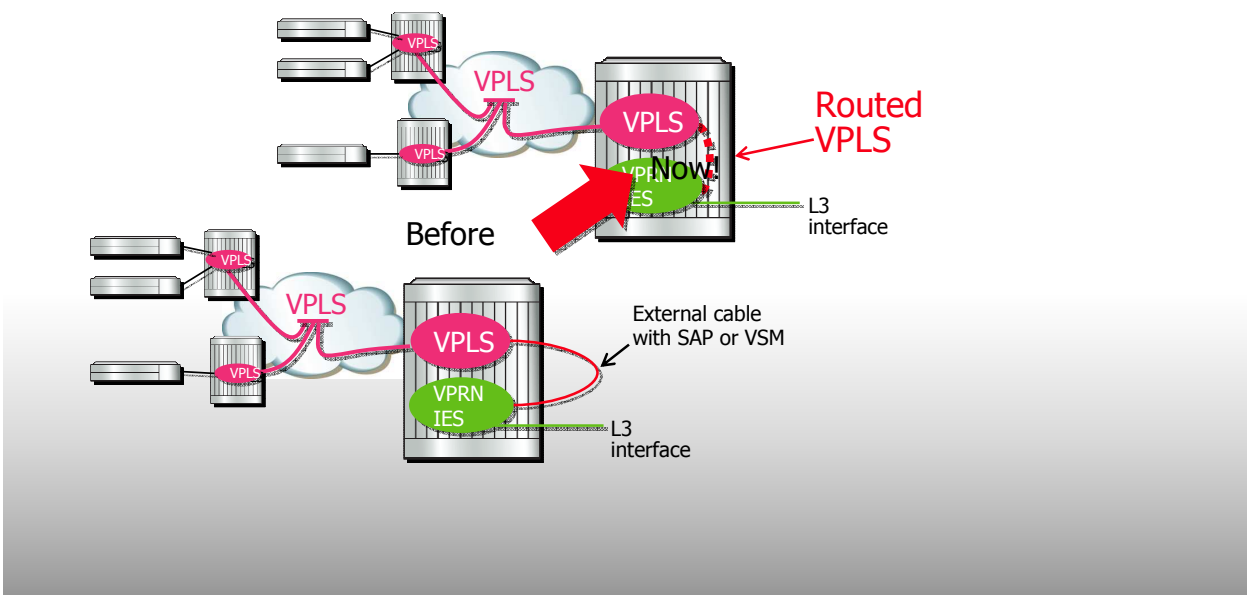
Configuring an IES

1. Create an IES service and associate it with a customer ID.
2. Create an interface within the IES.
3. Define SAP parameters on the interface.
   - Select node(s) and port(s)
   - Optional - select QoS policies other than the default (configured in the config>qos context)
   - Optional - select filter policies (configured in the config>filter context)
   - Optional - select accounting policy (configured in the config>log context)
4. Enable the service.

In conclusion. What do we need to configure an IES service?
First step is to create the service instance itself and associate it with a customer ID.
Secondly create an interface within the IES.
Add a SAP or SDP, IP address and optionally select the appropriate QoS, Filter and accounting policy.
Enable the service.

Routed VPLS

- Binding an IES (or VPRN) IP interface to a VPLS service
  also known as Integrated Routing and Bridging (IRB) or Bridge Virtual Interface (BVI)



VPLS is handled in a separated module but mentioned here as it interworks with an IES service. A VPLS is an Ethernet L2 point-to-multipoint VPN type service. The focus is not to explain the VPLS in detail, but focus on the connection between an IES and VPLS service on one service router.

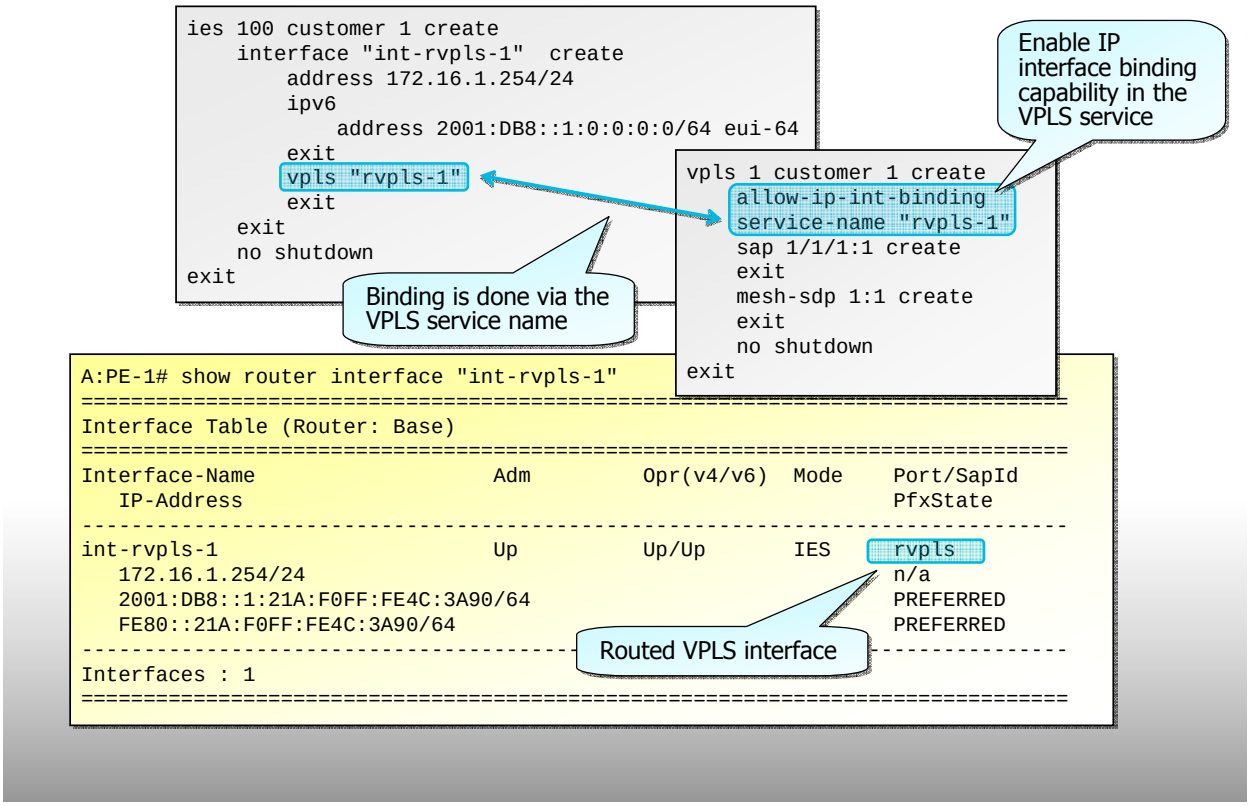
A Routed VPLS or R-VPLS allows a VPLS instance to be bound to an IES or VPRN IP interface.

Within an R-VPLS service, traffic with a destination MAC matching that of the associated IP interface will be routed based on the IP forwarding table, all other traffic will be forwarded based on the VPLS forwarding table.

The R-VPLS service can be associated with either an IPv4 or IPv6 interface and can run routing protocols over the R-VPLS service including OSPF, IS-IS, RIP and BGP. Note that BGP is supported in 7450 ESS in mixed-mode only. The R-VPLS requires that all network interfaces, all SAPs within the same routing domain as the R-VPLS and all SAP interfaces associated with the R-VPLS instance to be located on IOM3-XP or IMM cards. Note that routed VPLS is only a feature supported on the 7450 and 7750 type of products.

This features was possible in previous releases, but needed an external cable to interconnect two SAP's or making use of the additional VSM MDA module.

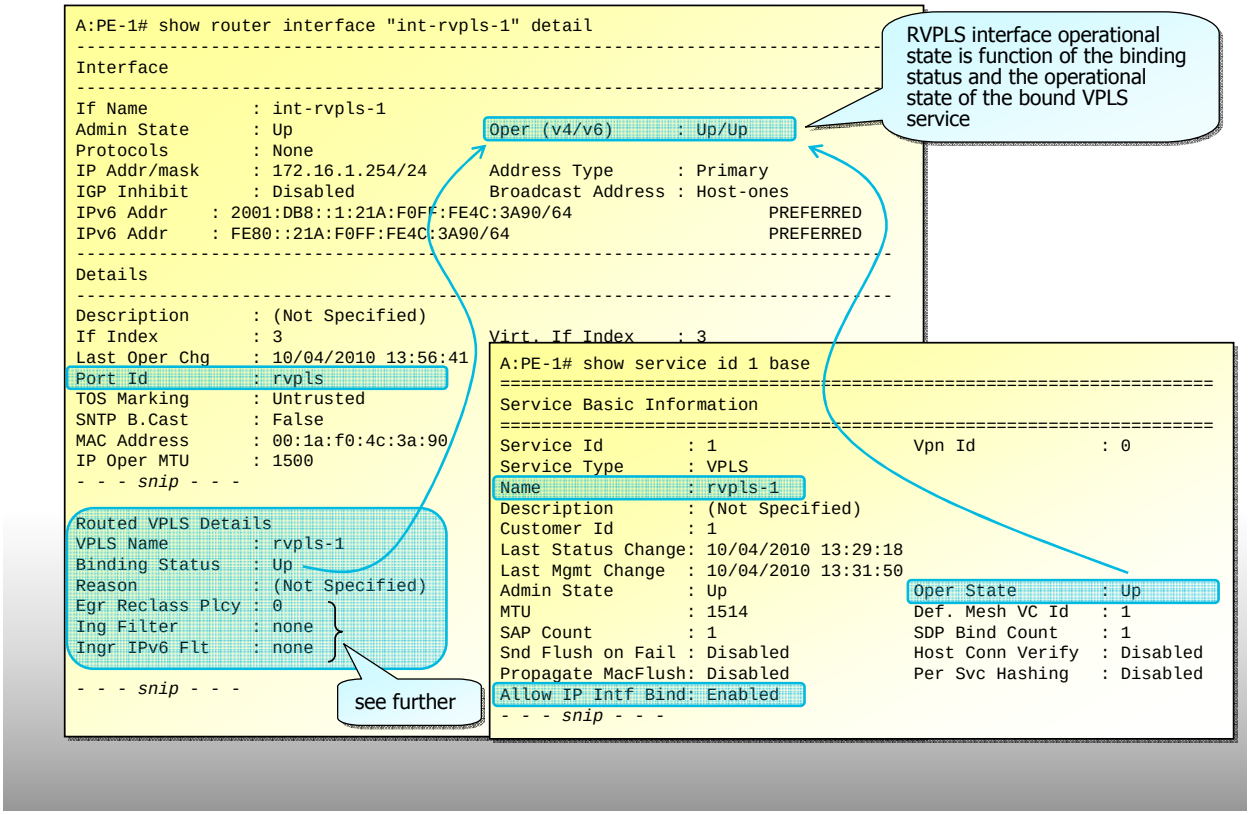
Routed VPLS Configuration



The binding between the IES and VPLS service on the same router is achieved by stating inside the IES service the name of the VPLS service. The VPLS service name is something that can be additionally configured to a VPLS service instance. In the VPLS service the command ”allow-ip-int-binding” has to be set.

When checking the IES interface status, “rvpls” is found at the place where you would normally see the port or sap binding.

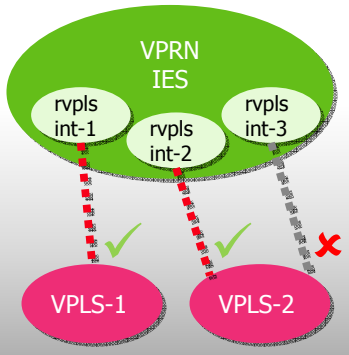
Routed VPLS configuration



The RVPLS interface operational state is a function of the binding status and the operational state of the bound VPLS service.

Routed VPLS facts

- A VPLS can be bound to a single IP interface
  - Service binding through the VPLS service name
- The service context containing the IP interface (IES or VPRN) may have other IP interfaces bound to other VPLS service contexts
- The IP interface is operationally down until the bound VPLS is created or until the service name is assigned to the VPLS
- When the IP interface is successfully bound to a VPLS service, the operational state of the IP interface will be dependent upon the operational state of the VPLS service



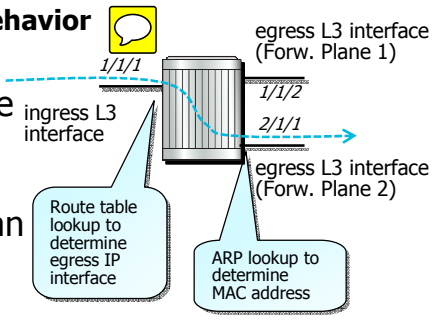
A VPLS can be bound to a single IP interface. But an IES interface might have several other IP interfaces bound to other VPLS services. The IP interface is operationally down until the bound VPLS is created or until the service name is assigned to the VPLS.

When the IP interface is successfully bound to a VPLS service, the operational state of the IP interface will be dependent upon the operational state of the VPLS service.

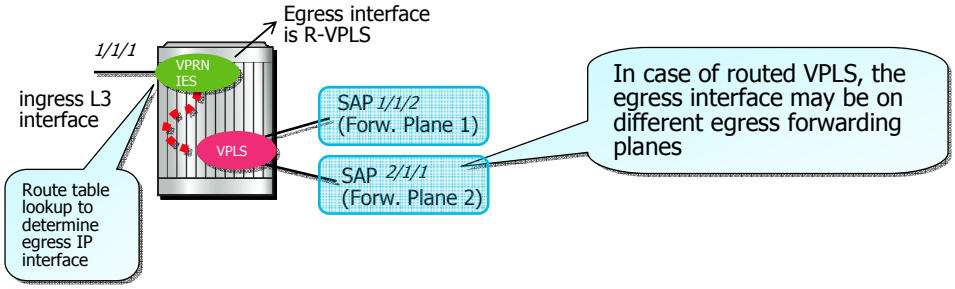
Routed VPLS L3 Forwarding behavior

Normal IP forwarding behavior

- At ingress forwarding plane, determine the egress interface from the routing table and send the packet to the corresponding egress forwarding plane
- At the egress forwarding plane, perform an ARP lookup to determine the MAC address



R-VPLS IP forwarding behavior



In a scenario where all IP interfaces are directly connected to the service router, a routing table look-up is performed to the destination IP address of an incoming packet. If an entry is found in the routing table, the next-hop IP address and the egress interface are resolved. An ARP lookup is performed to determine the MAC address of this next hop. Then the packet can be further send on the egress interface.

In case of R-VPLS, the behavior is slightly different. The destination address of an incoming IP packet results in an egress interface and a next-hop after routing table look-up which can be on different forwarding planes because of the VPLS. SAP 1/1/2 could be forwarding plane 1 and SAP 2/1/1 could be forwarding plane 2. Besides SAP's also SDP's can be seen as different forwarding planes.

A lookup in the ARP table is done to determine the MAC address to be used by an IP next-hop. If no entry is found, an ARP request is triggered.

Routed VPLS L3 forwarding in case of Routed VPLS

- At the ingress forwarding plane, determine the egress interface from the routing table, then perform an ARP lookup and an FDB lookup:

Enhanced ingress processing mandates the requirement for FP2 based hardware (IOM3/IMM)

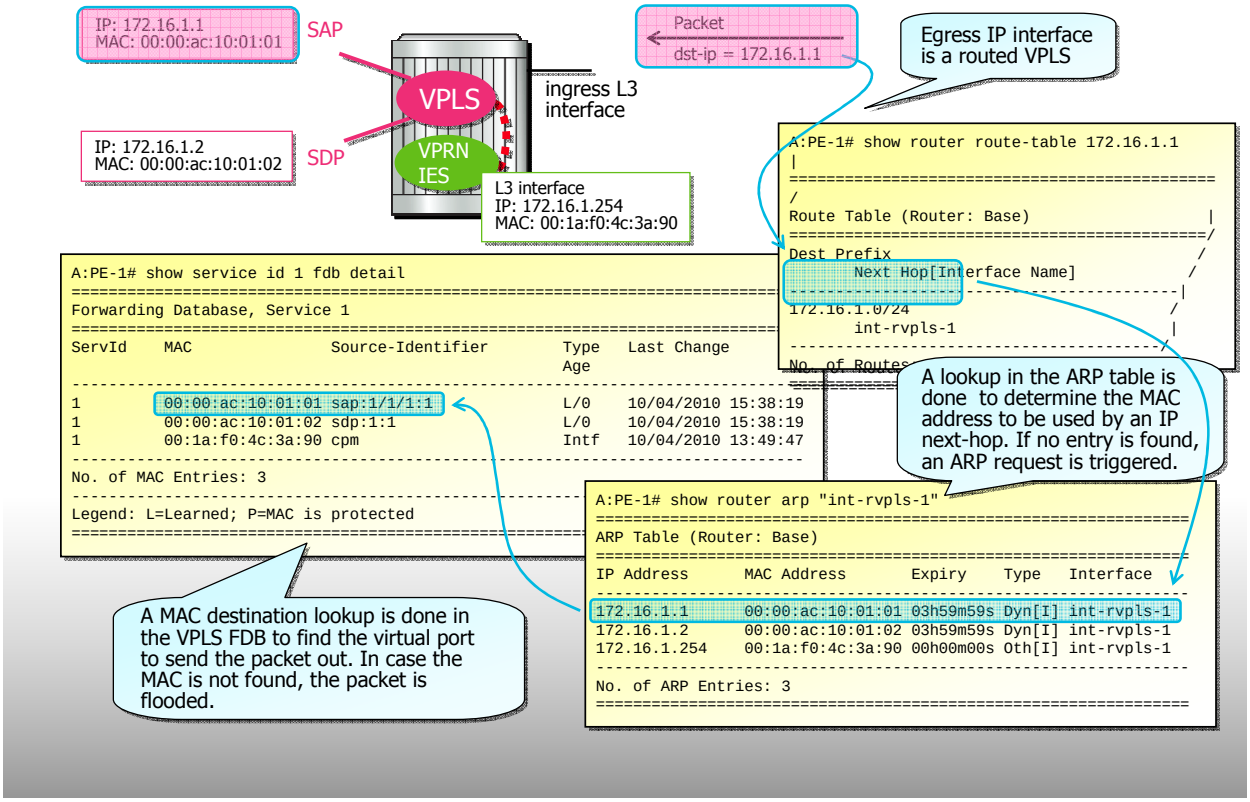
Next-hop ARP cache entry	Next-hop MAC FDB entry	Ingress forwarding plane behavior
No entry	known or unknown	Flood to all egress forwarding planes associated with the VPLS (ARP request triggered at egress) IP routing stops until ARP is resolved
Entry found	known	Forward to the egress forwarding plane associated with the VPLS virtual port (SAP/SDP)
Entry found	unknown	Flood to all egress forwarding planes associated with the VPLS

The egress forwarding plane uses the result from the ingress lookup to forward the packets.

When a packet enters the ingress forwarding plane, the egress interface from the route table is determined and an ARP look-up and FDB look-up are performed. If the ARP table hasn't an entry for this next-hop, an ARP flood is done. If an ARP entry is found and the MAC address is in the FDB MAC table of the VPLS, the ARP is sent to one of the egress forwarding planes associated to the VPLS virtual port; a SAP or an SDP. If an ARP entry was found but no MAC entry in the FDB, an ARP flood is done as well.

The egress forwarding plane uses the results from the ingress lookup to forward the packet.

Routed VPLS L3 forwarding in case of Routed VPLS



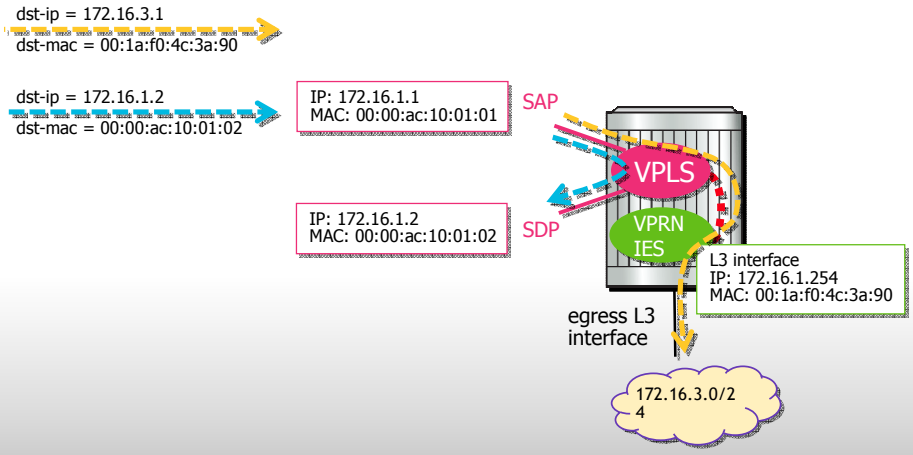
An example.

An IP packets enters the router with address of 172.16.1.1. The routing table identifies as next hop the IES interface. The destination address is connected to one of the VPLS SAP interfaces.

Based on this info, a lookup in the ARP table is done to determine the MAC address to be used by an IP next-hop. If no entry is found, an ARP request is triggered and send to all forwarding planes of the VPLS; in this case one SAP and one SDP. But in this example, the ARP table is populated, so the MAC address is resolved. A second look-up is now performed against the fdb table of the VPLS. The entry exists in this case so the packet is forwarding on SAP 1/1/1:1. If the entry was in the ARP table BUT not in the fdb table, the packet is flooded.

Routed VPLS Forwarding in the Routed VPLS

- Within an R-VPLS service, traffic with a destination MAC matching that of the associated IP interface will be routed based on the IP forwarding table; all other traffic will be forwarded based on the VPLS forwarding table.



When traffic is coming from the VPLS service itself, the destination MAC address matching that of the associated IP interface will be routed based on the IP forwarding table; all other traffic will be forwarded based on the VPLS forwarding table.

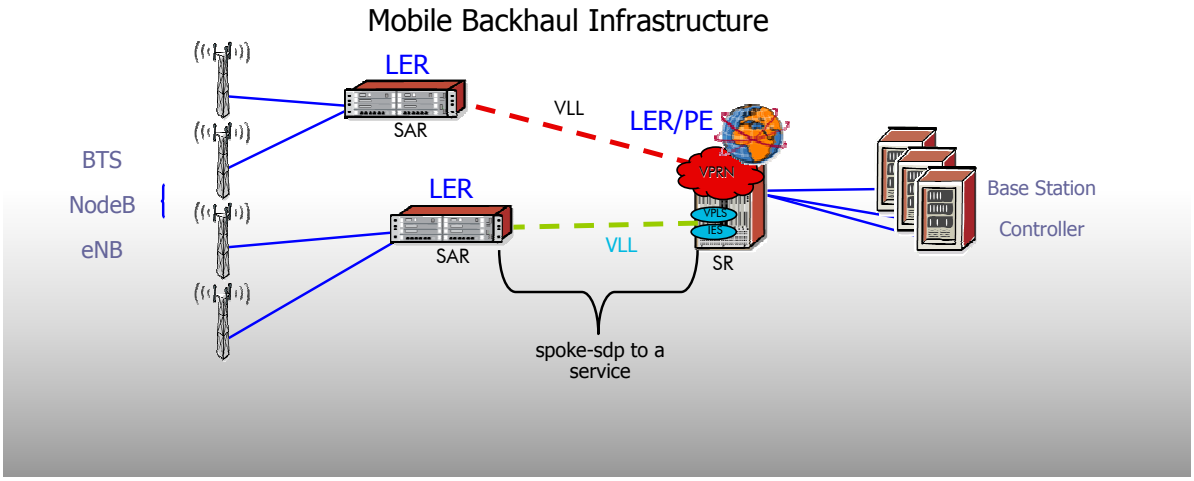
SECTION 3: IES APPLICATIONS

Section 3 provides an overview of applications and scenario's where an IES service is deployed.

IES in mobile backhaul

Presently, customer traffic is switched from the SAR to an aggregation SR via pseudo-wires (PWs)

- Requires less expertise and management at SAR node
- Potentially non-optimal delivery of packets between CEs
- Opens an improvement opportunity: Local routing on SAR would conserve NW bandwidth and hop reduction (delay and jitter advantages)



Let us have a look where IES could be used in a mobile backhaul network.

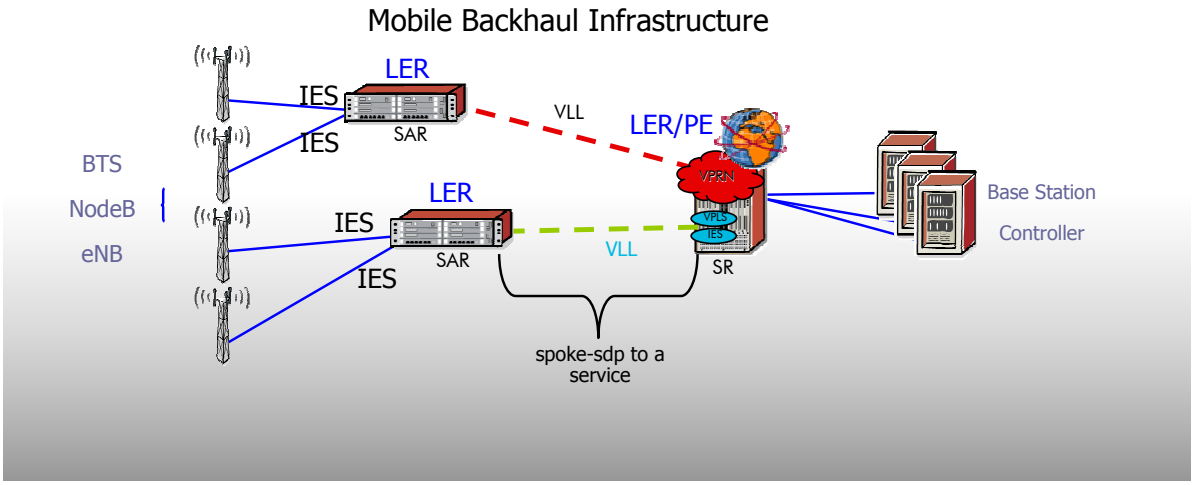
Currently, all traffic from the BTS, NodeB or eNodeB is directly connected to a L2 service and backhauled to the SR over a VLL. This requires less expertise and management on the SAR node as no table look-up has to be done and all traffic is sent into the tunnel. An IES is only used at the PE site. The disadvantage of this solution is the potential of non-optimal forwarding of packets for communication that happens between CE devices. In this case, all traffic is sent to the SR and then routed back to the same SAR router.

Local routing on the SAR would conserve network bandwidth and hop reduction resulting in less delay and jitter.

IES in mobile backhaul

Internet Enhanced Service (IES) provides IP routing on the SAR.

- Currently SAR-initiated PWs are terminated in an IES, VPRN or VPLS at an aggregation SR
- IP Routing allows traffic to be forwarded to another SAR or SR based on the destination IP address
- Seamless migration to IP Services for fixed backhaul

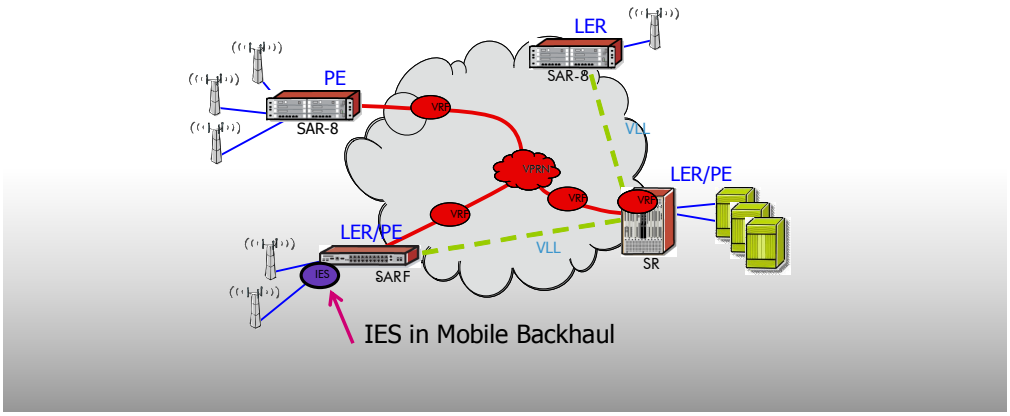


The traffic from mobile sites is terminated into an IES service right at the SAR node. With this termination, SAR effectively routes customer IP traffic to its destination based on destination IP addresses. Unlike the point-to-point nature of PW services, with IP services, customer traffic can be forwarded to different destinations solely based on the destination IP address. In other words, traffic from a given customer interface can be forwarded to another SAR or to a SR node within the network. To sum up, IP Routing Services offer any to any IP forwarding capabilities on SAR.

IES in mobile backhaul

Mobile Traffic can be backhauled with IES:

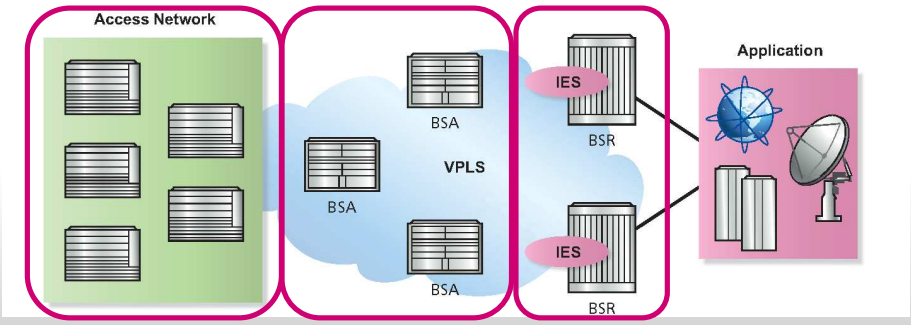
- eNB is bound to an IES interface on SAR
- For traffic destined to another eNB, SAR forwards the IP packets based on the destination address (it refers to the global routing table)



The eNodeB is bound to an IES on the SAR. For traffic destined to another eNodeB, SAR forwards the IP packets based on the destination address to the appropriate SAR. Traffic not destined to another eNodeB is sent to the SR.

IES in the residential service delivery network

- Uses VPLS to connect the Access network to the Application Network through VPLS services (L2 VPN)
- A single unicast VPLS per BSA per subnet
- A single multicast VPLS
- An IES service on the BSR to terminate each VPLS

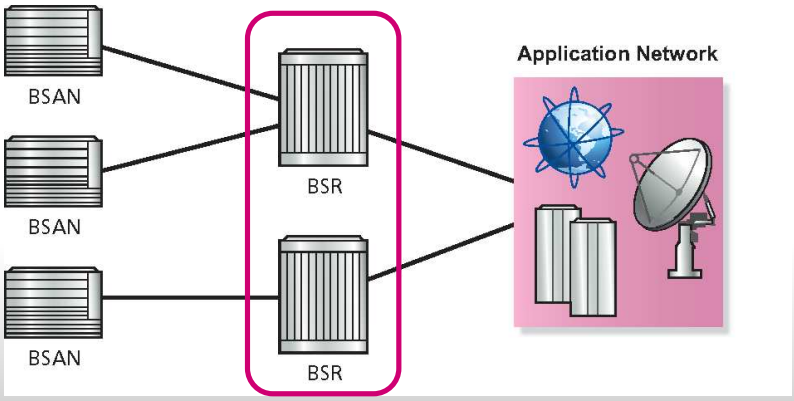


A residential service delivery network consists of an access network which DSLAM's in case of DSL technology in the first mile. An aggregation network with BSA's or Broadband Service Aggregators and at the edge of the aggregation network the BSR's or Broadband Service Routers.

The access and aggregation network is entirely layer 2. The L3 terminations happens on the BSR and is achieved by termination the VPLS SDP's into an IES service. So a VPLS on a BSA is connected to an IES on the BSR. A BSA can also be dual homed to two BSR IES interfaces for redundancy. This redundancy between BSR is called SRRP or Service Routing Redundancy Protocol.

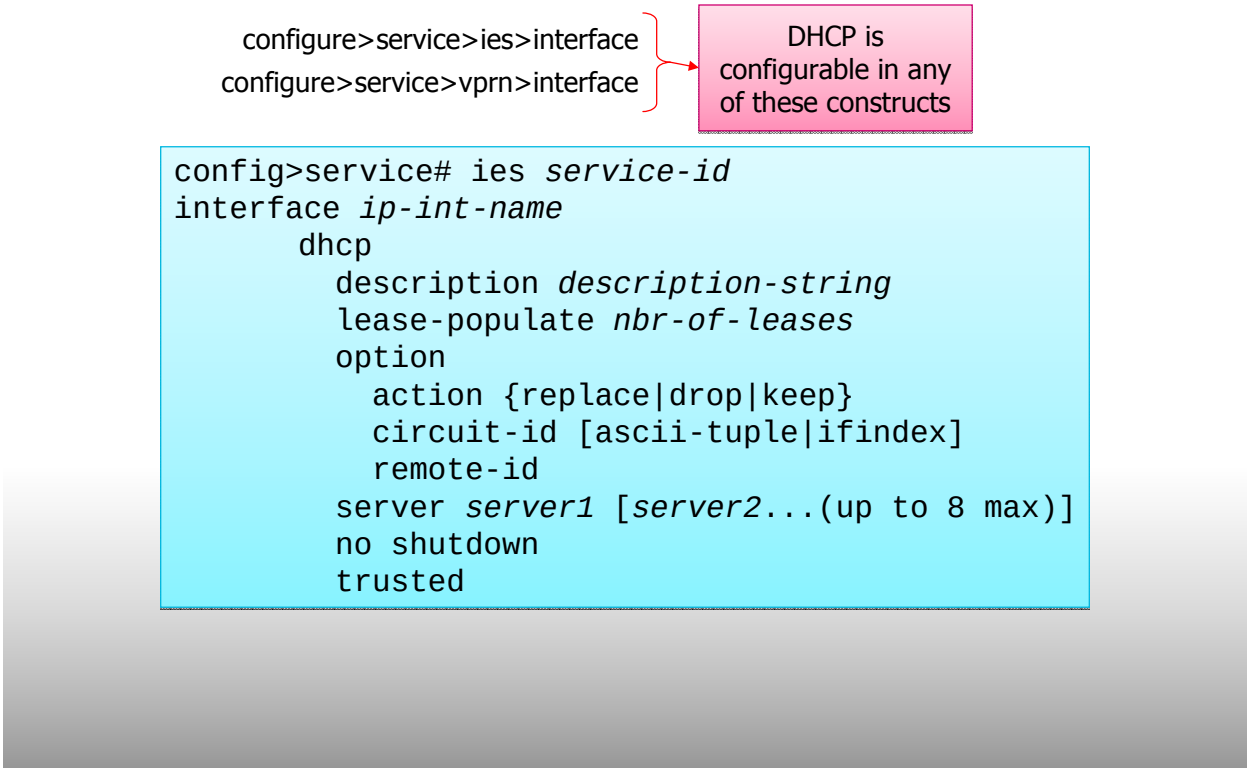
IES in the residential service delivery network

- BSANs in the Access Network are connected directly to a BSR
- Removes the requirement for VPLS based network and BSAs



In the Routed CO architecture the BSAN devices in the Access Network are connected directly to the BSRs. No aggregation network is in between and this solution is therefore more suited for smaller networks. The DSL BSAN is directly connected with an IES Gigabit connection to the BSR's IES service. So the IES is found here at the BSR.

DHCP Configuration



In a service delivery network, the receiving of a DHCP discover message can trigger the DHCP interaction with a DHCP server. This relay of DHCP is configured under the IES interface.

Module Summary

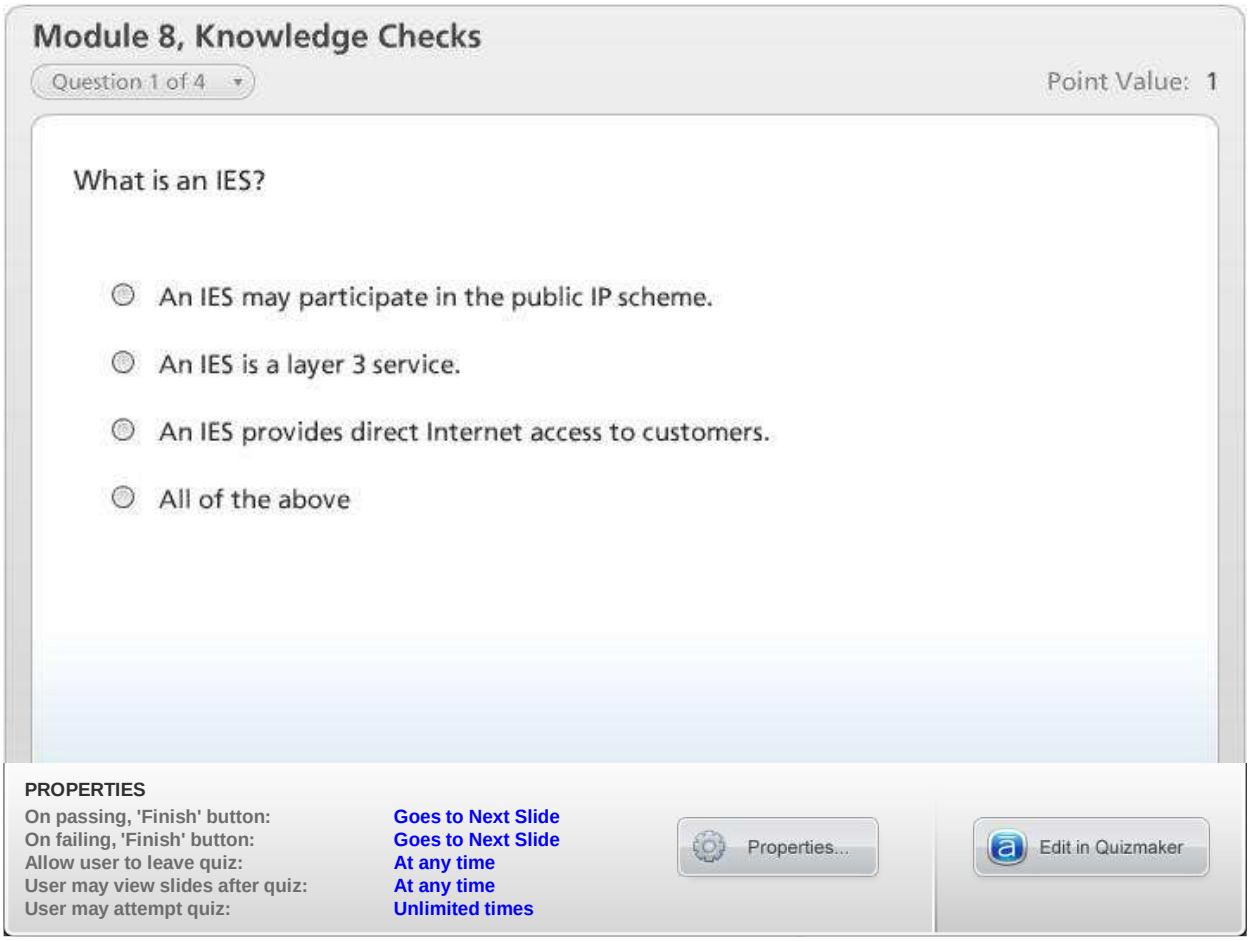
- IES or Internet Enhanced Service is a non-VPN type of service.
- An IES service contains IP interface(s) with a SAP or SDP as entry point
- IES services are used to terminate L2 VPN services in an L3 routed network
- IES and VPLS can now be connected on one service router without additional cables or MDA's
- IES can be found in the residential service delivery network and in the mobile backhaul network

Module summary. An IES or Internet Enhanced Service is an L3 service and part of the non-VPN type of services. However, there is the option to termination an L2 VPN service on an IES service. In this case, the IES is not using a SAP, but an SDP under the IP interface. On the 7450 and 7750 products, an IES and VPLS can be connected without additional cables called routed VPLS. IES is typically found in a residential service delivery network or mobile backhaul network.

KNOWLEDGE CHECKS

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempt at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 8

This completes Module 8.

SR-OS Fundamentals
Module 9: VLL - Virtual Leased Line

IPD Development

Available as PDF

Welcome to the 9th module of the SR-OS fundamentals course.

Table of Contents

Module 9 is divided into nine sections:

Section 1: VLL Introduction

Section 2: E-pipe

Section 3: I-Pipe

Section 4: F-pipe

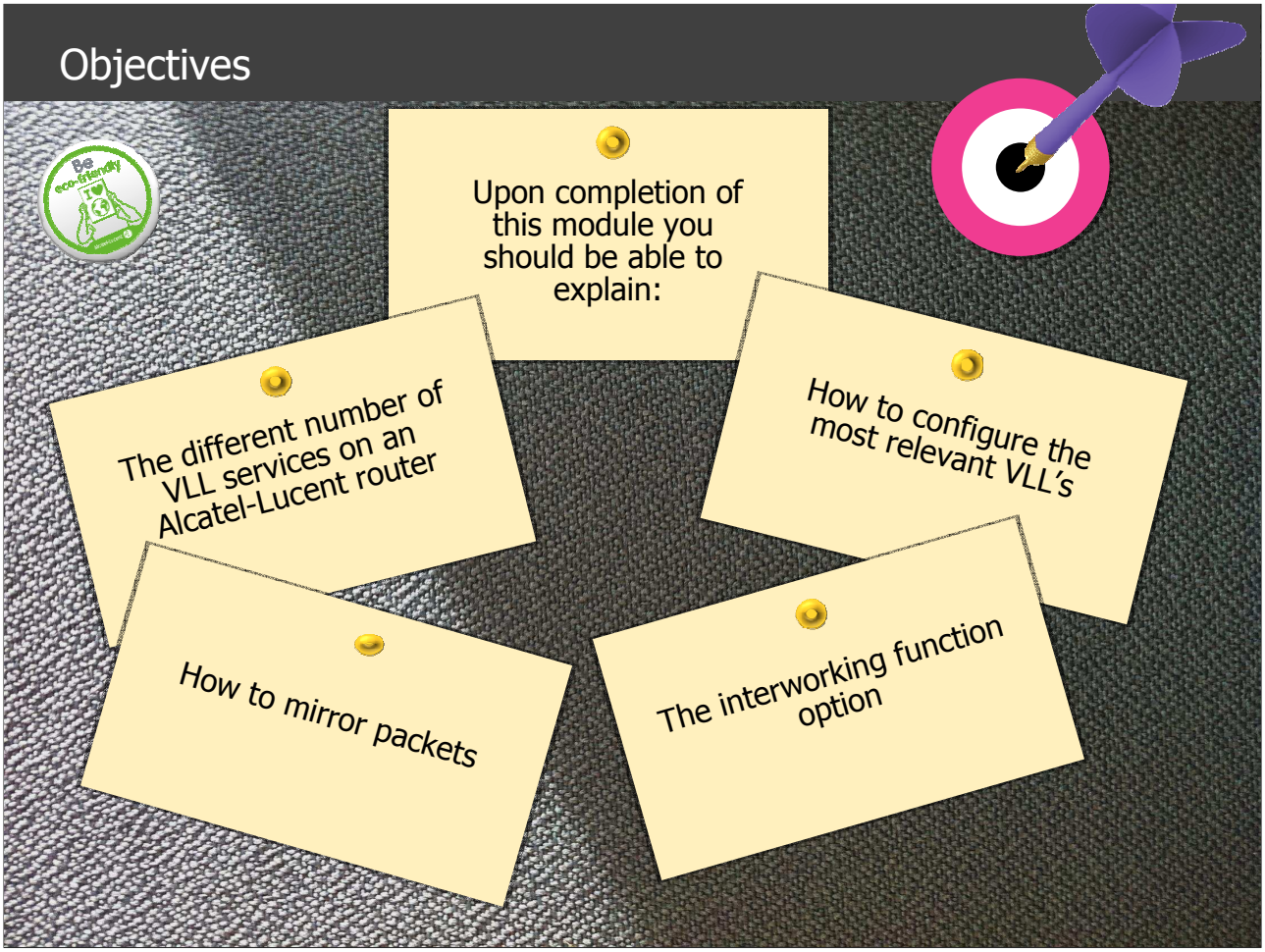
Section 5: A-Pipe

Section 6: C-pipe

Section 7: H-Pipe

Section 8: Interworking function

Section 9: Mirror-service



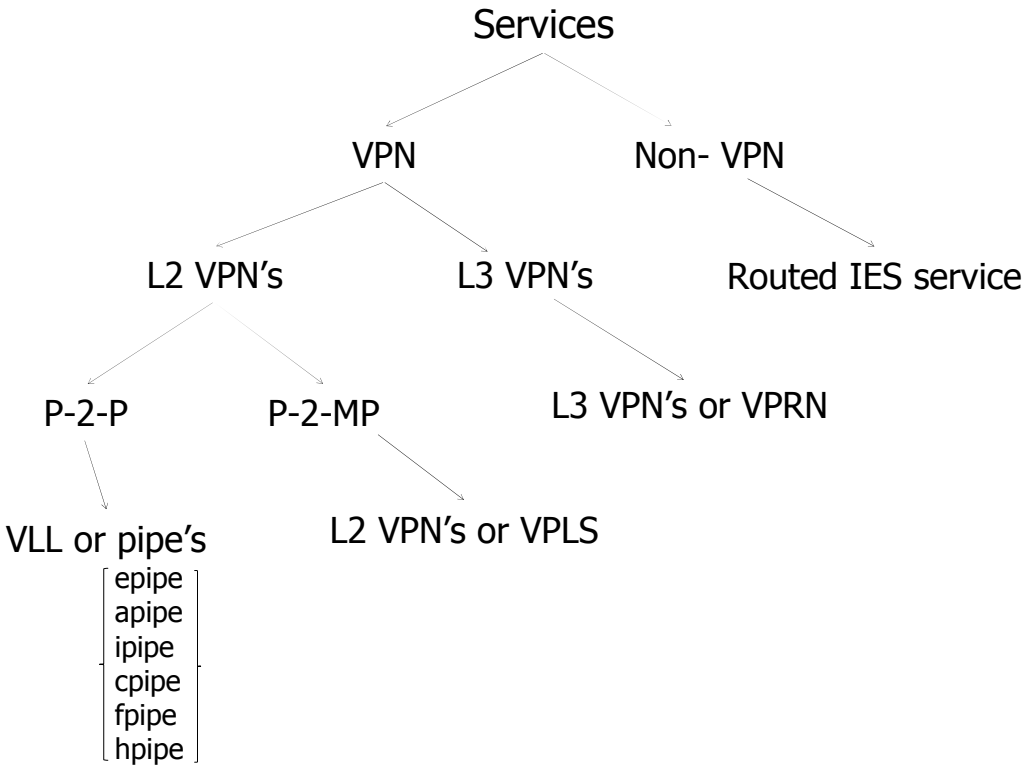
By the end of this module you will be able to explain:

- The different number of VLL services on an Alcatel-Lucent service router
- How to configure most relevant VLL's
- How to mirror packets
- The interworking function options

Introduction to VLL - Virtual Leased Line

Section 1: Introduction

Different services



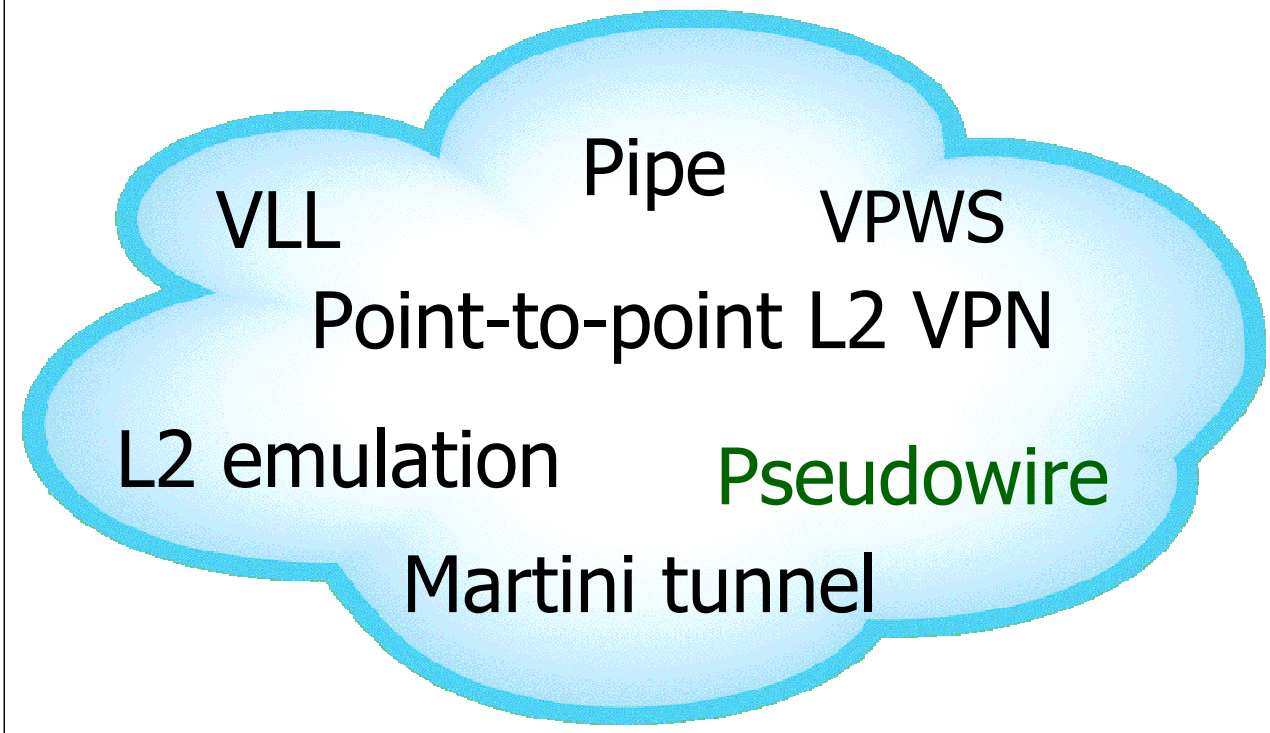
Services can be divided into two main categories, VPN and non-VPN services. The majority of available services are VPN type of services, either L2 or L3 VPN's. The non-VPN service of the SR-OS product family is the Internet Enhanced Service, a routed service that offers IP connectivity.

L2 VPN's can be further split up into point-to-point and point-to-multipoint. The point-to-point services, pseudowires and virtual leased lines are also called pipe services and come in different variants. The variant depends on the access technology that is carried inside the VPN. In the point-to-multipoint service contains the virtual private LAN service , which is best compared to an Ethernet switch network.

The Virtual Private Routed Network is an L3 point-to-multipoint type of service VPN. It offers routed services to customers.

This module explains the different point-to-point VLL services.

What's in a name



What's in a name?

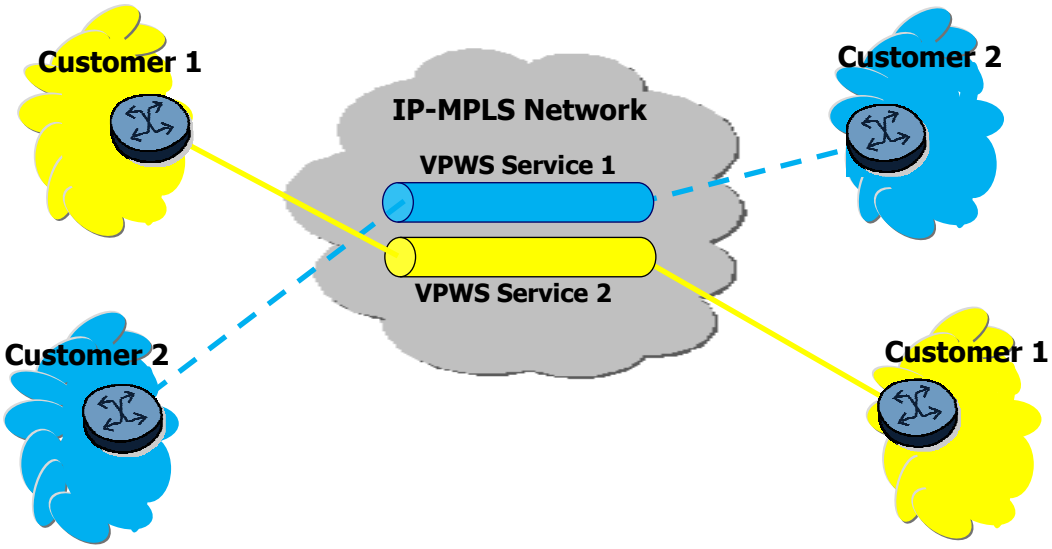
A pseudowire point-to-point service is known in the industry under different names, but all of them refer to the same type of service.

A Virtual Leased Line or VLL is a general well known industry standard name. Alcatel-Lucent calls it VLL's pipes.

Other names could be point-to-point L2 VPN or L2 emulated service or Martini tunnel.

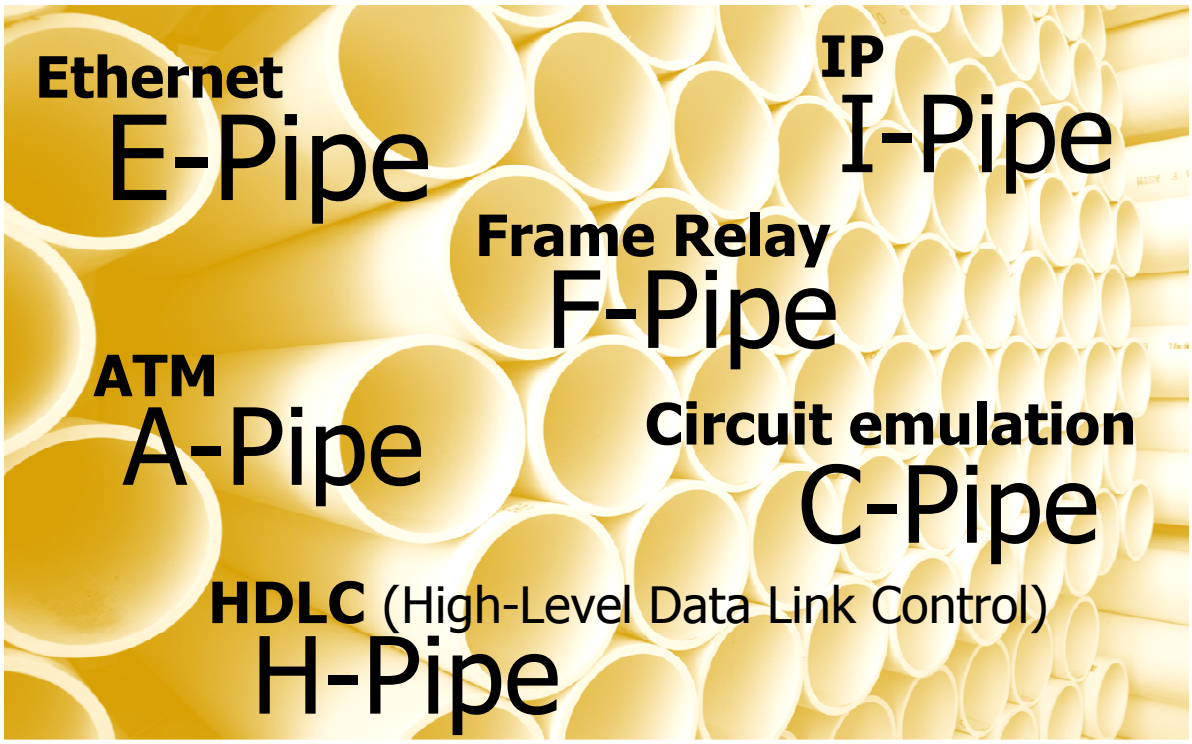
VPWS is also a common used name and stands for Virtual Private Wire Service.

VPWS services



The L2 VPN point-to-point service or VPWS (Virtual Private Wire Service) is a layer 2 point-to-point connection service between two customer sites. It provides an emulation of layer 2 technologies such as Ethernet, IP, ATM, Frame Relay, E1/T1 and HDLC.

VPWS services -> Pipes

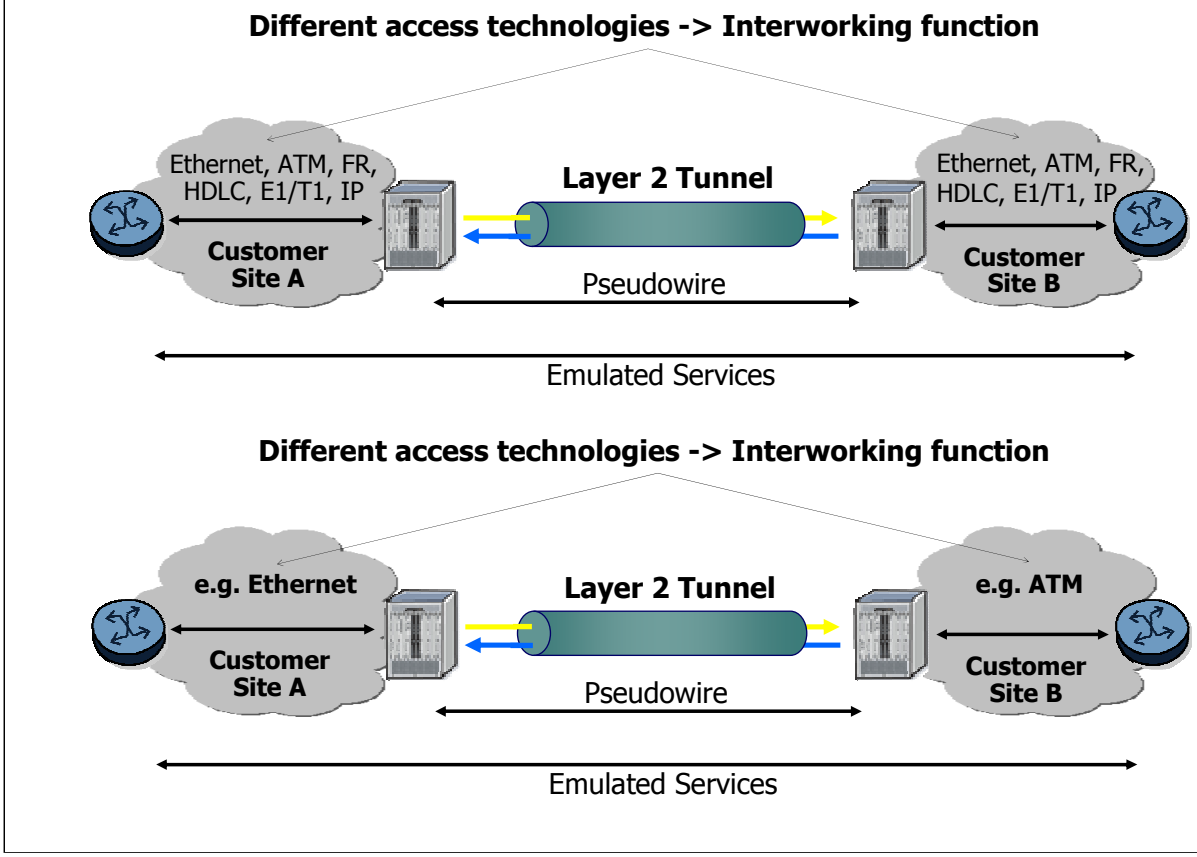


Alcatel-Lucent's implementation of VPWS is known as E-pipe, I-pipe, A-pipe, F-pipe, H-pipe and C-pipe. These VPWS Services are based on the IETF "Martini Drafts" and the IETF Ethernet Pseudowire Drafts.

These services are layer 2 point-to-point services where the customer data is encapsulated and transported across a service provider's IP or MPLS network. The VPWS service is completely transparent to the subscriber's data and protocols. The VPWS service does not perform any MAC learning.

The letter before the name of the pipe indicates the layer 2 technology that is encapsulated. E is for Ethernet, I for IP, A for ATM, F for frame relay, H for HDLC and C for circuit emulation.

Why VPWS?



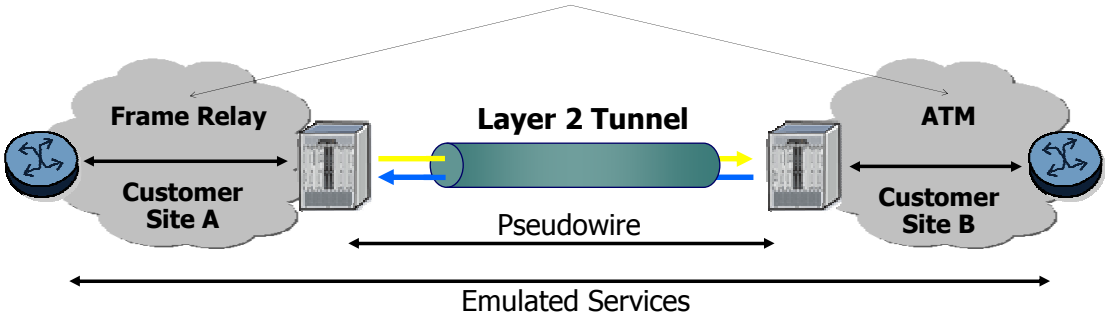
Service providers are seeking to offer multiple services across a common packet switched network.
Pseudowires enable you to emulate or to carry legacy Layer 2 services over a cost-effective and predominant IP/MPLS network. It decouples the services protocols and applications from the underlying facilities that carry them.

One of the major advantages is the use of different access technologies on each side of the VPWS. This is called the interworking function.

In this example Ethernet is the access technology for customer A and ATM is the access technology for customer B.

Interworking function

Different access technologies -> Interworking function



Customer Site A \ Customer Site B	**ATM**	Frame Relay	**Ethernet**
ATM	A-pipe	A-pipe	I-pipe E-pipe
Frame Relay	A-pipe	F-pipe	I-pipe E-pipe
Ethernet	Ipipe Epipe	I-pipe E-pipe	E-pipe

This diagram shows the different available interworking options and the type of pipe to use when interconnecting different access technologies.

For example if frame relay on one side connects to ATM on the other side, an A-pipe needs to be used.

Two options are available when connecting frame relay to Ethernet; Ipipe and Epipe.

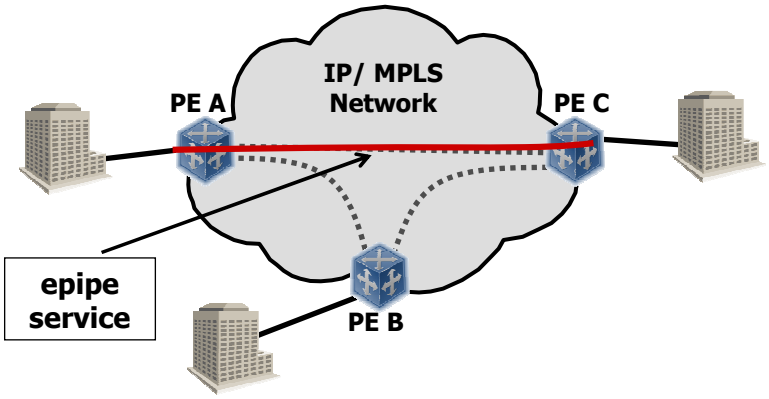
e-pipe

Section 2: E-pipe

E-Pipe service

An ePipe service provides a point-to-point connection between two nodes.

- From the customer's perspective it looks as if a leased link exists between the two locations.
- No MAC learning required
- The Service provider can apply billing, ingress/egress shaping and policing



Let's take a closer look at the e-pipe service.

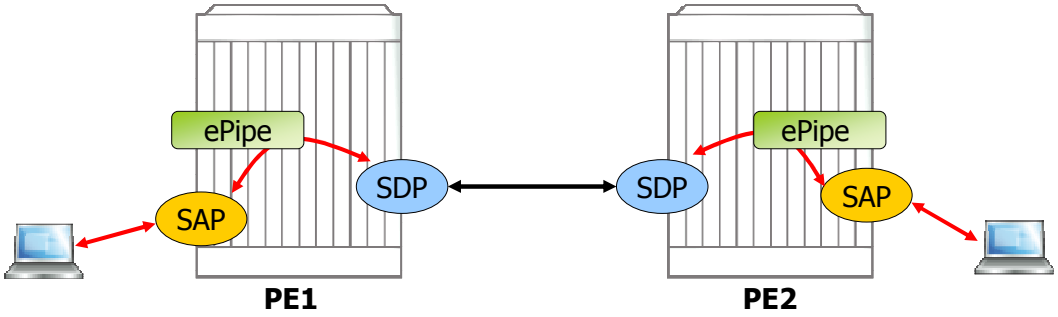
The e-pipe service provides a point-to-point connection of Ethernet between two locations. From the customer's perspective it operates as a leased line between the two locations. There is no MAC learning, so all traffic presented on the Ethernet SAP on one side will be sent over the pseudowire towards the other side's Ethernet SAP.

However, the operator has the option to apply billing, ingress and egress QoS and filter policies.

Distributed e-pipe

Distributed e-pipe

- 2 PE's
- 1 SAP/PE + 1 SDP/PE
- Traffic arriving on SAP directed to far-end via SDP -> No MAC look-up



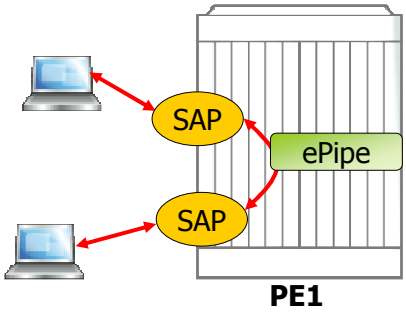
A distributed e-pipe service always involves two PE's. Each PE does have one e-pipe service instance configured and one SAP and one SDP connected.

Traffic coming from a SAP will be sent to the SDP and traffic coming from the SDP is sent to the SAP.

Local e-pipe

Local e-pipe

- 1 PE
- 2 SAP's/PE -> No SDP
- Traffic arriving on SAP directed to other SAP on the same PE

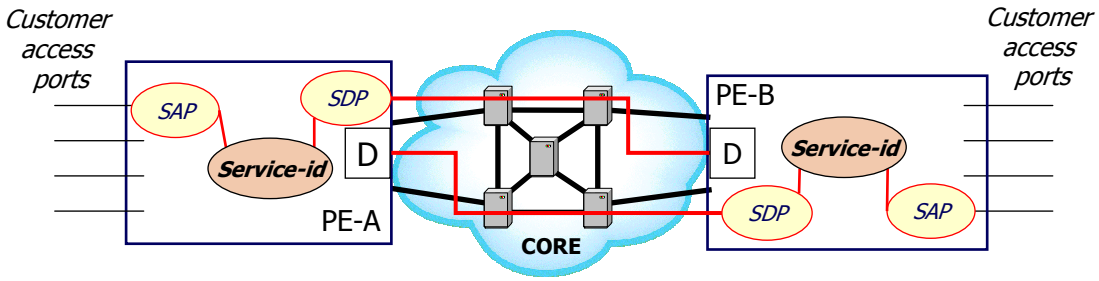


A local E-pipe service involves only one PE site. This PE has one e-pipe service instances configured with two connected SAP's. Traffic coming on one SAP is sent to the other SAP and visa versa.

This type of local service is mostly used to enforce policies or to shape the traffic.

Local point-to-point services are also used as a local interworking function in a migration scenario.

Creating a distributed e-pipe service



Service-id:	***ePipe***
	Customer (subscriber) Association
Service Access Point	***Select node and port***
	Select Accounting Policy (optional)
	Select Ingress/Egress QOS Policies (optional)
Service Distribution Path	***Binds a service to an SDP***

What are the key components that make the e-pipe service?

The first thing to do is the creation of the e-pipe service instance itself. During this creation, a customer ID has to be assigned. A customer ID refers to a name of a customer, address, phone, etc. However, if this is not wanted, the default customer ID of 1 can be used. This customer ID doesn't have any credentials.

Within the service context a SAP and/or a SDP needs to be configured. The configuration of the SAP and SDP under the service context automatically binds them to the service.

Optionally, different types of policies can be configured under the SAP or SDP.

E-Pipe configuration

Create a customer:

configure service epipe 5500 customer 100 create

config>service>cust$

```
[no] contact            - Configure contact Information
[no] description        - Description for this customer
[no] multi-service-*  + Configure a multi-service site for this customer
[no] phone              - Phone numbers for the contact
```

This example shows the creation of an e-pipe service with service ID of 5500 and customer ID of 100.

The customer assignment is only done during the first creation of the e-pipe.

Contact, description, multi-service site and phone number are the options to describe the customer. This customer ID is useful for trouble shooting. When a particular customer ID is applied multiple times to certain services, it is easier to retrieve all service information from a dedicated customer.

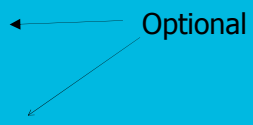
E-Pipe configuration

Creating an ePipe Service:

config>service# epipe 5500 customer 100 create
config>service>epipe$ description "epipe service"
config>service>epipe# no shutdown

Applying Ingress and Egress SAP Parameters:

Node>config>service# epipe 5500
config>service>epipe# sap 2/1/3:0
config>service>epipe>sap# ingress
config>service>epipe>sap>ingress# qos 555
config>service>epipe>sap>ingress# filter ip 1
config>service>epipe>sap>ingress# exit
config>service>epipe>sap# egress
config>service>epipe>sap>egress# qos 627
config>service>epipe>sap>egress# scheduler-policy alpha
config>service>epipe>sap>egress# exit
config>service>epipe>sap#



A description is always optional in CLI, but can be useful for troubleshooting and clarification.

The dot1q SAP 2/1/3:0 is added to epipe service 5500.

The ingress and egress CLI contexts are optional here. They house the policies.

For example, in the ingress direction, QoS policy 555 and IP filter policy 1 is applied. Whereas QoS policy 627 and scheduler policy "alpha" are applied in the egress direction.

SAP Encapsulation for ePipe services

Port Type	Encapsulation
Ethernet	NULL
Ethernet	dot1q
Ethernet	QinQ
SONET / SDH	IPCP
SONET / SDH	BCP-Null
SONET / SDH TDM	BCP-dot1q
SONET / SDH TDM	ATM
SONET / SDH FR	FR

Shown here are the different SAP encapsulation options available for an e-pipe.

Null – Supports a single service on the port. For example, a single customer edge device attached to the port

Dot1q – Supports multiple services on the port. For example, a customer edge device running Virtual LANs

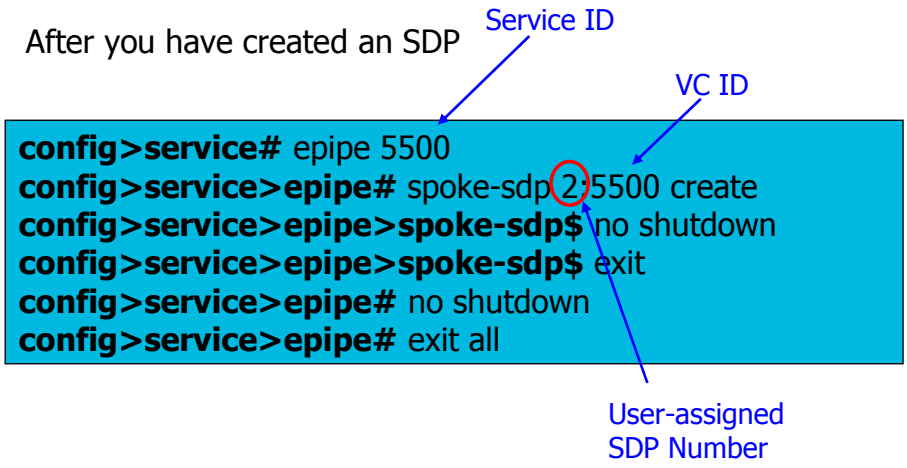
QinQ – Supports tags within tags

IPCP – Internet Protocol Control Protocol is typically used for interconnection using point-to-point protocol

Bridging Control Protocol is typically used for bridging a single service between two devices using PPP over SONET/SDH with an encap ID of 0.

Bridging Control Protocol supports multiple services on the SONET/SDH port/channel.

E-Pipe SDP Bindings



Under the e-pipe service 5500 SDP 2 is added with a VC ID of 5500. This VC ID is the ID that must match the other side's VC ID. This is the ID that is used during the TLDP negotiation. Spoke SDP is the only type of SDP possible as there can't be any loops in a point-to-point service.

I-pipe

AT THE SPEED OF IDEAS

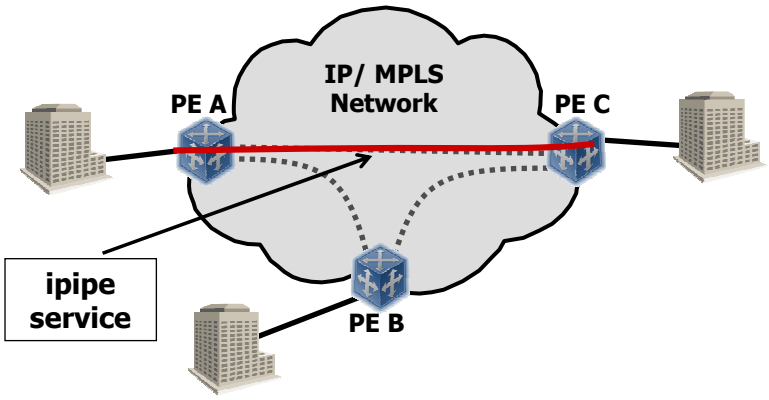
Alcatel·Lucent

Section 3: I-pipe

I-Pipe service

An I-Pipe service provides a point to point connection between two nodes.

- IP point to point VLL
- Layer 3 IGP/EGP routing protocol that runs over IP can be run over the IPipe
- OSPF can run over an I-Pipe
- IS-IS uses Layer 2 for its routing messages and is therefore not supported within an Ipipe
- **Mainly used as interworking function** -> see later
- The Service provider can apply billing, ingress/egress shaping and policing



The I-pipe provides a pure IP encapsulation VLL service between two nodes. The L2 is stripped at the SAP and the encapsulated payload is only IP.

The regular L3 IGP/EGP routing protocol that run over IP can be run over the I-pipe. OSPF can run over the I-pipe, not IS-IS as this protocol uses a layer 2 for its routing messages.

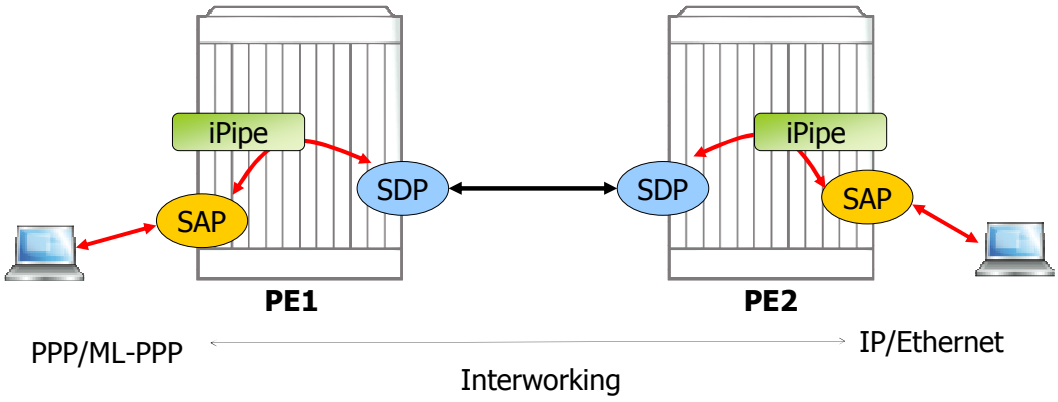
The I-pipe is mainly used for interworking as discussed later on in this module.

As with the e-pipe, the customer can apply the appropriate policies.

Distributed I-pipe

Distributed i-pipe

- 2 PE's
- 1 SAP/PE + 1 SDP/PE
- Typical example -> Interim step during ppp network to Ethernet network migration
  - IP connectivity between a CE attached to a point-to-point access circuit (PPP/IPCP) with routed encapsulation and a CE attached to an Ethernet interface
- Local i-pipe possible



A distributed I-pipe requires at least two PE's with one SAP and one SDP per PE

A typical example of an I-pipe would be in a migration scenario. IP encapsulated in PPP is migrated to IP encapsulated over Ethernet.

This interworking is also available as a local service.

E-pipe versus I-Pipe

E-pipe	I-Pipe
●Provides Ethernet VPWS	●Provides IP interworking VPWS
●Supports bridged encapsulation	●Supports routed encapsulation
●More network overhead because Ethernet header is transported	●Less network overhead because Ethernet header is taken out
●Both hosts appear to be the same Ethernet LAN	●Both hosts appear to be on the same IP subnet or network
●No Mac learning	●Mac learning
●Easier to configure	●Local & Remote CE IP address configuration is required

Let us look at the main differences between an e-pipe and an I-pipe.

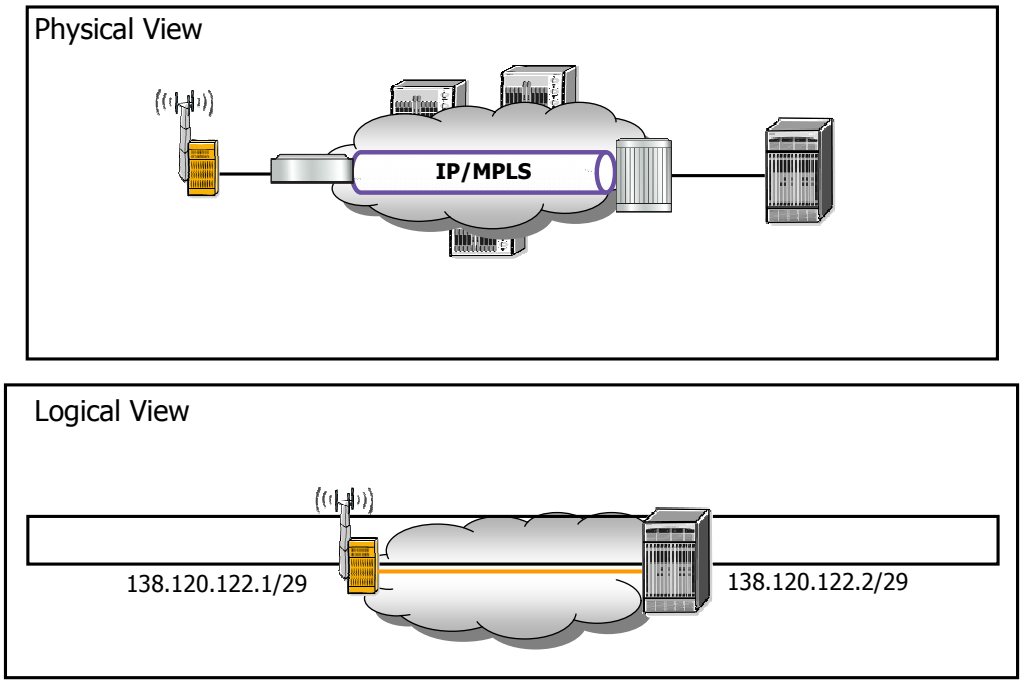
The fundamental difference is the encapsulation. An e-pipe supports bridge encapsulation resulting in more overhead. An I-pipe supports routed encapsulation with less overhead as the Ethernet header is taken out.

With an e-pipe, both hosts appear to be an on the same LAN while with an I-pipe both hosts appear to be on the same IP subnet.

To enable the interworking, the I-pipe needs to learn MAC addresses. E-pipe will never learn MAC addresses.

An I-pipe takes more to configure. The local and remote CE IP address configuration is required.

Distributed I-Pipe



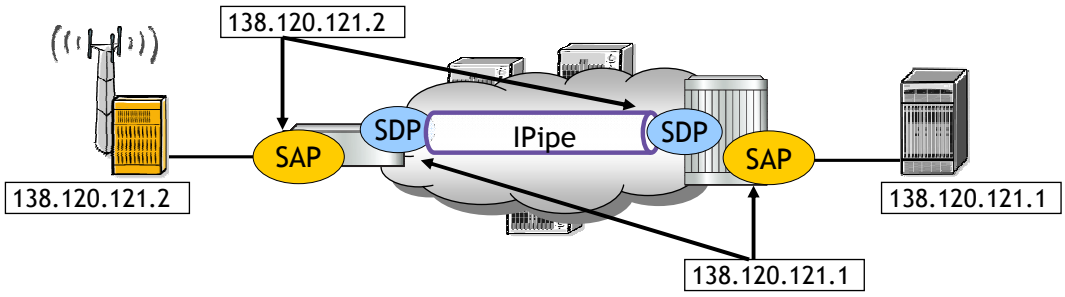
The physical view of an I-pipe involves PE's and/or P routers and an IP/MPLS network.

The customer sees only the logical view. Two IP subnets connected to each other. The customer is unaware of the IP/MPLS service cloud.

Configuration overview

The following steps are required to configure a distributed IPipe service:

- Configure an IPipe Service with Customer ID
- Configure a SAP and associate it with the IPipe service
- Assign the local IP address to the SAP
- Configure a spoke-SDP and associate it with IPipe service
- Assign a remote IP address to the spoke-SDP



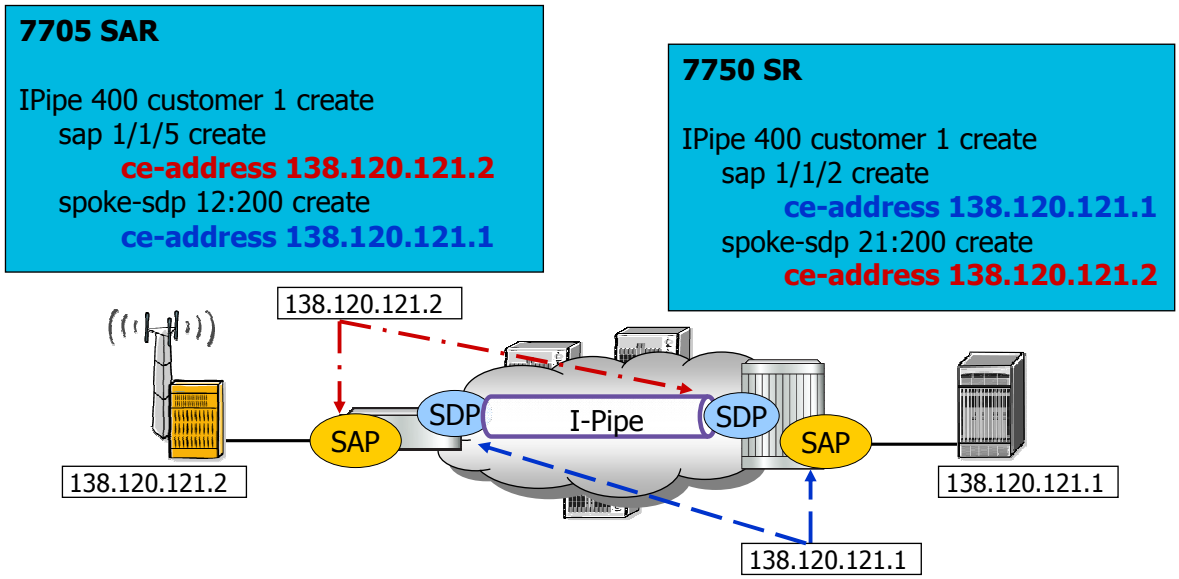
The creation of an I-pipe service is the same as e-pipe service, except that the service type is I-pipe. In addition, an I-pipe requires a local and a remote IP address.

The local IP address is the address of the locally connected CE.

The remote IP address is the IP address of the remotely connected CE.

I-Pipe service: configuration overview

CE address needed for ARP mediation or PPP negotiation -> see interworking section



When a CE sends an ARP for a MAC address of the far-end connected CE device, the ARP can't be transmitted over the pseudowire as the Ethernet is stripped off. Therefore the local PE router will answer on behalf of the far-end CE device with its own MAC address. This is only possible if the local PE router knows the IP address of the far-end CE device.

The local IP address, configured under the SAP, is also needed to discover the MAC address of the connected CE device for traffic that needs to be sent out towards the CE.

Native IP traffic that is coming from the SDP is encapsulated in Ethernet with a destination MAC address learnt from the ARP message. The ARP uses as input the local CE address configured under the SAP.

F-pipe

AT THE SPEED OF IDEAS

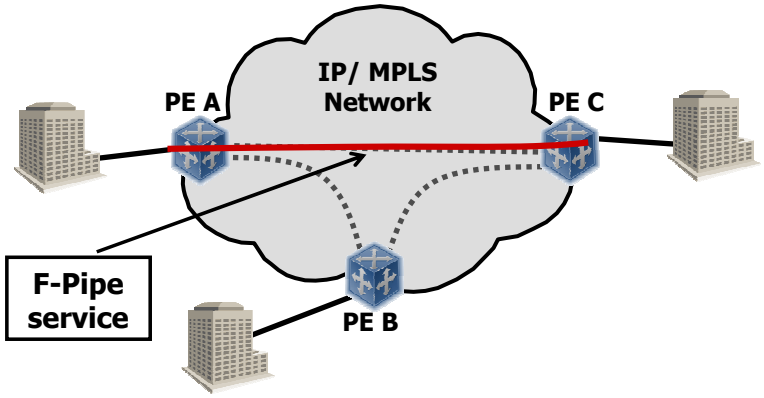
Alcatel·Lucent

Section 4: F-Pipe

F-Pipe service

An F-Pipe service provides a point-to-point connection between two nodes:

- Frame Relay point-to-point VLL
- Traffic is mapped into a SAP belonging to a defined DLCI value
  - DLCI (Data Link Connection Identifier) - Identifies virtual circuit



An F-Pipe service provides a point-to-point connection between two nodes transporting layer 2 frame relay.

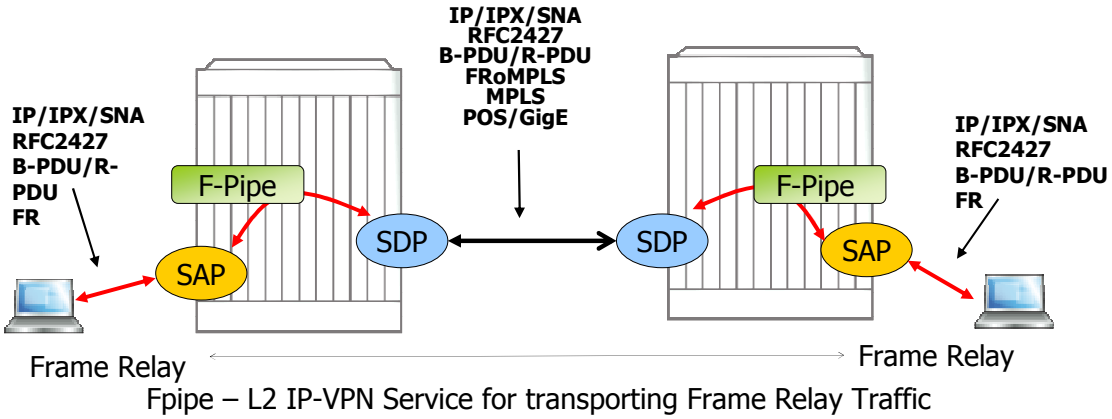
A SAP is characterized by a DLCI value indicating the frame relay.

The Data Link Connection Identifier or DLCI bits identify the virtual connection. Multiple DLCIs can be present on the same physical channel. DLCI values range from 0-1023. 0 – 15 are reserved, as are 1008-1023.

Distributed F-pipe

Distributed F-pipe

- 2 PE's
- 1 SAP/PE + 1 SDP/PE
- Frame relay at both SAP's but interworking function possible
- Local F-pipe also possible



Fpipe – L2 IP-VPN Service for transporting Frame Relay Traffic

This slide depicts an application of a Frame Relay pseudowire. Both PE's receive a standard frame on the Frame Relay SAP and encapsulate it into a pseudowire packet according to the 1-to-1 Frame Relay encapsulation mode defined in RFC 4619, Encapsulation Methods for Transport of Frame Relay over MPLS Networks.

The Frame Relay pseudowire feature supports local cross-connecting when the users are attached to the same PE node. The FR PW is initiated using Targeted LDP signaling.

F-pipe common configuration tasks

Associate an F-pipe service with a customer ID

- Define SAP parameters
  - Optional - select egress and ingress QoS and/or scheduler policies (configured in config>qos context)
  - Optional - select accounting policy (configured in config>log context)
- Define spoke SDP parameters
- Enable the service

Apart from the specific frame relay SAP, all other service settings are identical to the former VLL type of services.

Creating an F-pipe service

CLI Syntax: config>service# **fpipe *service-id*** [customer *customer-id*] [vpn *vpn-id*] [vc-type {fr-dlci}]
description *description-string*
service-mtu *octets*
sap *sap-id*
no shutdown

The service F-pipe needs a service ID because a service ID, in general, is unique per router, a different service ID from an already associated service ID should be chosen. This is different from the customer ID. A customer ID can be applied multiple times, even to different types of services.

The description is as always optional.

The "vc-type" specifies a 15 bit value that defines the type of the VC signaled to the peer. Its values are defined in *draft-ietf-pwe3-iana-allocation* and it defines both the signaled VC type as well as the resulting data path encapsulation over the f-pipe.

Creating an F-pipe service - example

```
Example:  R4>config>service# fpipe 1 customer 1 create
          R4>config>service>fpipe# description "F-pipe "
          R4>config>service>fpipe# service-mtu 1400
          R4>config>service>fpipe# sap 1/2/3:17 create
          R4>config>service>fpipe# exit
          R4>config>service>fpipe# no shutdown

Example:  R1>config>service# fpipe 1 customer 1 create
          R1>config>service>fpipe# description "F-pipe"
          R1>config>service>fpipe# service-mtu 1400
          R4>config>service>fpipe# sap 1/2/3:22 create
          R4>config>service>fpipe# exit
          R1>config>service>fpipe# no shutdown
```

This example shows an f-pipe configuration on either side of a pseudowire. Beside the SDP that is needed to bring the f-pipe up, traffic is mapped into a SAP belonging to a defined DLCI value. The Data Link Connection Identifier or DLCI bits are used to identify the virtual connection.

Configuring F-pipe SDP Bindings

CLI Syntax: config>service# **fpipe *service-id*** [customer *customer-id*]
[vpn *vpn-id*] [vc-type {fr-dlci}]
spoke-sdp *sdp-id*:*vc-id*
egress
filter ip *ip-filter-id*
vc-label *egress-vc-label*
ingress
filter ip *ip-filter-id*
vc-label *ingress-vc-label*
no shutdown

As with all VLL's, the SDP type is spoke.

The spoke-SDP can be configured with an IP filter on both the ingress and egress direction.

VC-labels can also be manually set if TLDP is disabled.

Configuring F-pipe SDP Bindings - example

PE router 4
Example: R4>config>service# **fpipe 1**
R4>config>service>fpipe# **spoke-sdp 1:1 create**
R4>config>service>spoke-sdp# **exit**

PE router 1:
Example: R1>config>service# **fpipe 1**
R1>config>service>fpipe# **spoke-sdp 1:1 create**
R1>config>service>spoke-sdp# **exit**

This example show the binding of the spoke-SDP's to the f-pipe's on both sides.

A-pipe

AT THE SPEED OF IDEAS

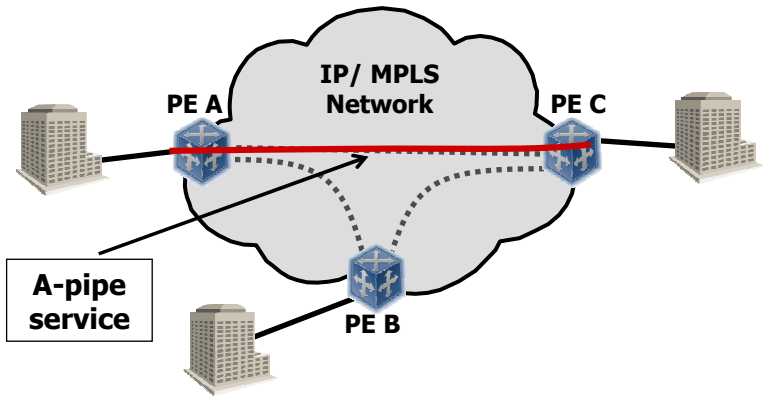
Alcatel·Lucent

Section 5: A-Pipe

A-Pipe service

An A-Pipe service provides a point to point connection between two nodes.

- A-pipe provides a bi-directional layer 2 connection of ATM
- Local or distributed A-pipe
- N:1 cell mode or AAL5 SDU mode supported
- The Service provider can apply billing, ingress/egress shaping and policing



An a-pipe or ATM pipe provides a bi-directional layer 2 connection of ATM, available as local and distributed service.

A-pipe supports two main ways of transporting ATM cells over the pseudowire, N:1 cell mode or AAL5 SDU frame mode as explained in the next slide.

Cell Mode Encapsulation

N:1 Cell Mode

- Allows a service provider to offer an ATM PVC or SVC based service across a network
- Used to transmit a single ATM cell or multiple ATM cells per Packet Switch Network (PSN) PDU
- Allows multiple ATM VCCs or VPCs to be carried within a single PSN tunnel
- Supports the binding of multiple VCCs/VPCs to a single pseudowire

AAL5 SDU Frame Encapsulation

- allows the transport of ATM AAL5 PDUs traveling on a particular ATM PVC across the network to another ATM PVC
- requires Segmentation and Reassembly on the ingress PE-CE ATM interface
- Once reassembled, the AAL5-SDU is carried via a pseudowire to the egress PE

N:1 Cell Mode:

In the simplest case, this encapsulation can be used to transmit a single ATM cell per PSN PDU. However, in order to provide better PSN bandwidth efficiency, several ATM cells may optionally be encapsulated in a single PSN PDU. This process is called cell concatenation.

According to RFC4717, in the N:1 Cell Mode, ATM cells are transported individually. The encapsulation of a single ATM cell is the only required encapsulation for ATM. The encapsulation of more than one ATM cell in a PSN frame is optional and supported by Alcatel-Lucent.

AAL5 SDU Frame Mode

The AAL5 payload VCC service defines a mapping between the payload of an AAL5 VCC and a single pseudowire. The AAL5 payload VCC service requires ATM segmentation and reassembly support on the PE.

Even the smallest TCP packet requires two ATM cells when sent over AAL5 on a native ATM device. It is desirable to avoid this padding on the pseudowire. Therefore, once the ingress PE reassembles the AAL5 PDU, the PE discards the PAD and PDU trailer and then the ingress PE inserts the resulting payload into a pseudowire PDU. The egress PE MUST regenerate the PAD and trailer before transmitting the AAL5 frame on the egress ATM port.

A-pipe configuration example

```
configure
    service
        apipe 100 customer 1 vc-type atm-vcc create
            service-mtu 1482
            sap 1/1/9.1:0/50 create
            exit
            spoke-sdp 104:100 create
            exit
            no shutdown
        exit
    exit
exit
```

This slide shows a typical a-pipe configuration in N:1 Cell Mode. ATM vc-type is set to atm-VCC.

The VCC cell transport service is characterized by the mapping of a single ATM VCC or VPI/VCI value to a pseudowire.

The alternative, the VPC service is defined by mapping a single VPC or VPI to a pseudowire. As such it emulates a Virtual Path cross-connect across the packet switched network. All VCCs belonging to the VPC are carried transparently by the VPC service.

The VPI/VCI value is defined under the SAP definition.

C-pipe

AT THE SPEED OF IDEAS

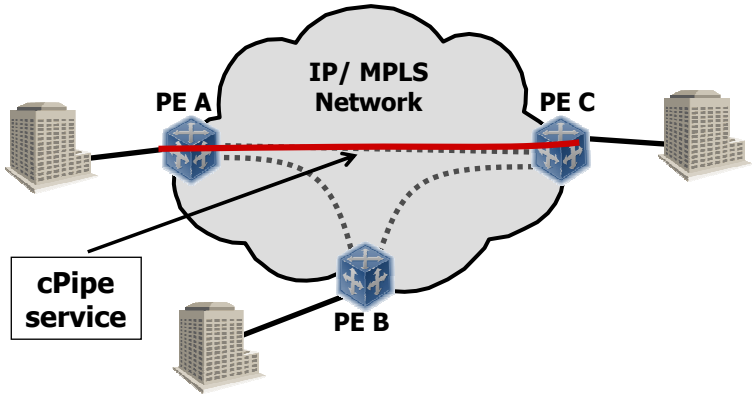
Alcatel·Lucent

Section 6: C-pipe

C-Pipe service

A C-Pipe service provides a point to point connection between two nodes:

- C-Pipe provides a bi-directional E1/T1 layer 2 connection
- TDM is defined as the transport protocol within existing 2G networks.
- C-Pipe PW capabilities will provide the transport service for 2G connectivity between the BTS and BSC.
- (CESop & SAToP)
- 1588v2 and Sync-E supported



A c-pipe provides a bidirectional E1/T1 layer 2 connection. TDM still plays an important role within the utility market or as transport protocol for existing 2G mobile networks. As the mobile operator evolves to an all IP/MPLS packet core, a migration scenario is needed to support this type of legacy services.

The C-pipe supporting both CESop and SAToP uses two supported Ethernet synchronization standards, 1588v2 and Sync-E.

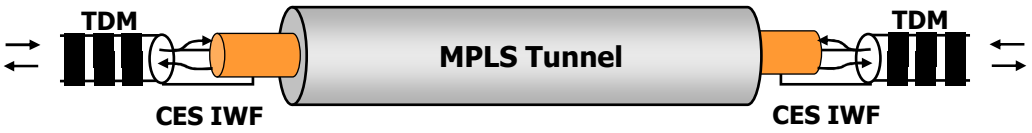
Transport of structured and unstructured connections

Two principal services used for structured and unstructured connections

- CESoP — Circuit Emulation Service over Packet → individual timeslots
  - Provides fractional services (nxDS0)
- SAToP — Structure Agnostic TDM over Packet → Full E1/T1
  - Provides structured T1/E1 services

Services referred to as:
Circuit Emulation Services (CES)

Services transported over:
MPLS enabled Metro Network using Pseudowire (PWE3) point-to-point tunnels



For private line and leased line applications a fundamental requirement is being able to transport TDM voice and data while maintaining timing and not introducing significant delay and jitter into the path.

Circuit Emulation Services over a packet switched network or CESoPSN and Structure-Agnostic TDM over Packet or SAToP are two types of TDM pseudowires.

CESoPSN is a structure-aware emulation for the transport of structured TDM traffic. Structured mode is supported for DS1, E1 and n times 64 kbps. Full DS1 or E1s can be transported by selecting all 24 DS0 timeslots, channels 1 to 24 and 32, channels 2-32, respectively. Framing bits are reproduced at the far end. When sending individual timeslots, they can be sent to different destinations.

SAToP is the Structure Agnostic emulation for the transport of Unframed TDM packets. This unstructured mode does not align to any framing.

SAToP C-Pipe configuration example

Create a SAToP C-pipe for E1:

```
A:SAR-154>config>service# cpipe 100 customer 1 vc-type satop-e1 create
A:SAR-154>config>service>cpipe# sap 1/1/1.1 create
A:SAR-154>config>service>cpipe>sap# back
A:SAR-154>config>service>cpipe# spoke-sdp 119:100 create
A:SAR-154>config>service>cpipe# no shutdown
```

This slide shows an example of a SAToP C-pipe configuration for transporting E1 type of traffic.

H-pipe

AT THE SPEED OF IDEAS

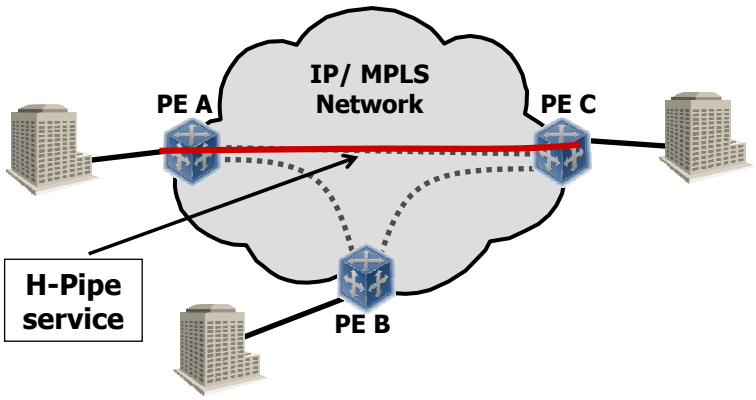
Alcatel·Lucent

Section 7: H-pipe.

H-Pipe service

A H-Pipe service provides a point to point connection between two nodes:

- H-pipe allows HDLC PDUs to be carried over MPLS network
- Provides IP-VPN services by emulating HDLC point-to-point connections
- HDLC Pseudo Wire implementation is described in RFC-4618



An H-pipe allows HDLC PDUs to be carried over MPLS network. It provides IP-VPN services by emulating HDLC point-to-point connections. HDLC Pseudo Wire implementation is described in RFC-4618

High-Level Data Link Control or HDLC, is a bit oriented that provides both connection oriented and connectionless service.

It is a data link control protocol, and falls within layer 2 of the OSI model

It supports both half duplex and full duplex communication lines, point-to-point and multi-point networks.

Cisco HDLC or cHDLC is an extension to HDLC protocol.

cHDLC frame contains an extra field that HDLC protocol doesn't have, which is called "Protocol Code" field.

Configuration example H-Pipe

```
*A:SAR-124>config>port# info
----------------------------------------------
        tdm
            ds1
                channel-group 1
                    encap-type hdlc
                    timeslots 1-24
                    no shutdown
                exit
                no shutdown
            exit
        exit
        no shutdown
----------------------------------------------
```

```
hpipe 1 customer 1 create
            sap 1/6/1.1 create
            exit
            spoke-sdp 1:1 create
            exit
            no shutdown
        exit
```

This slide shows a configuration example for H-Pipe where the encap-type is set to HDLC.

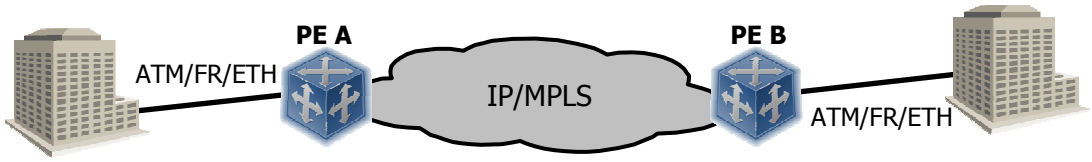
Interworking

AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 8: Interworking.

L2 Business network services => options



Bridged Encapsulation(1): IP traffic encapsulated in Ethernet frame transported via ATM port to an Ethernet port, an Epipe is used.
Routed Encapsulation(2): Native IP traffic transported via an ATM port to and Ethernet port, an Ipipe is used.

PE A \ PE B	ATM	Frame Relay	Ethernet
ATM	Apipe	Apipe	Ipipe **(2)** Epipe **(1)**
Frame Relay	Apipe	Fpipe	Ipipe Epipe
Ethernet	Ipipe Epipe	Ipipe Epipe	Epipe

Let's take a look to the different options for connecting business point-to-point L2 VPN's together. The access technology on one side does not have be the same on the other side of the point-to-point link. This is called in the interworking function.

The table lists the various interworking scenarios possible with ATM, Frame Relay and Ethernet ports at either end of an MPLS pseudowire.

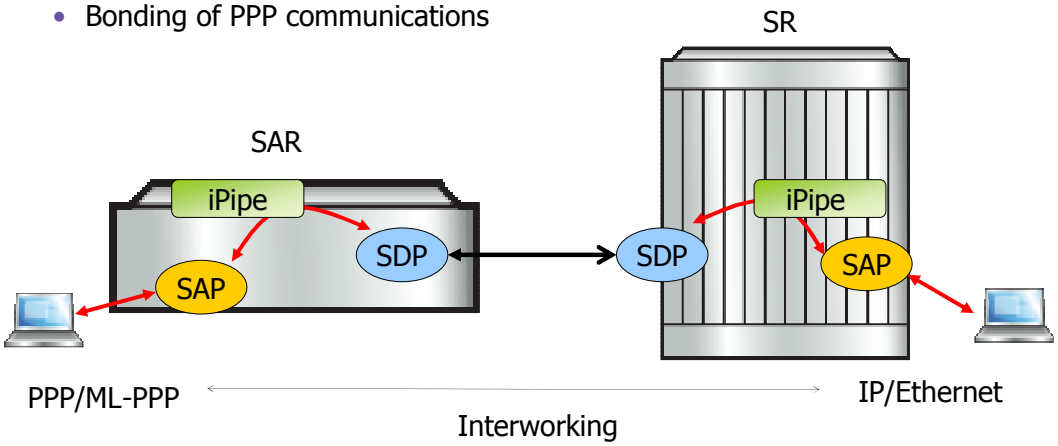
For example, when IP traffic, encapsulated inside an Ethernet frame, needs to be transported by a bridge/router over its ATM port to a bridge/router with an Ethernet port, bridged encapsulation is used. An E-pipe is required to transport this type of traffic. More details are given in the service module of this course.

Routed encapsulation is used when native IP traffic is transported across a bridge/router with an ATM port to a bridge/router with an Ethernet port. An Ipipe is required to transport this type of traffic. For further details, please refer to the Ipipe section in this module.

Example PPP/ML-PPP <-> IP/Ethernet interworking

Distributed i-pipe

- Distributed i-pipe
- SAR <-> SR
- IP/PPP <-> IP/Ethernet
- PPP -> Point to point data link protocol
  - Used here as encap. for SDH/Sonet
- ML-PPP -> extension to PPP
  - Bonding of PPP communications

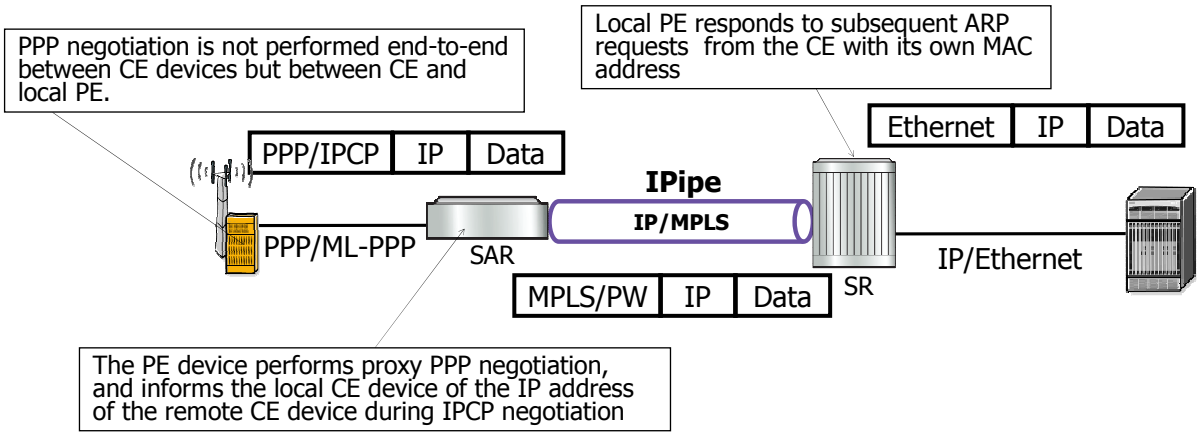


This is an example of a distributed I-pipe between a SAR and a SR within an interworking between PPP/ML-PPP and IP/Ethernet. IP encapsulated in PPP or ML-PPP is translated into an IP over Ethernet.

ML-PPP is an extension to PPP and binds several PPP communications.

Example PPP/ML-PPP <-> IP/Ethernet interworking

1. The SAR receives the PPP packet on the SAP
2. The SAR un-encapsulates the PPP or ML-PPP.
3. The SAR encapsulates the IP packet directly into the PW
4. The SR receives MPLS packet and removes the PW encapsulation
5. The SAR and SR need to have static IP address of local and remote CE for:
   a) ARP mediation
   b) Proxy PPP negotiation



The PPP frames received on the SAP of the SAR are sent to the SR, which re-encapsulates the frame again into an Ethernet frame.

To achieve this interworking, both the SAR and SR have a local and remote IP address configured.

At the side of the PPP, the PE device performs proxy PPP negotiation, and informs the local CE device of the IP address of the remote CE device during IPCP negotiation.

At the Ethernet side, the local PE responds to subsequent ARP requests from the CE with its own MAC address.

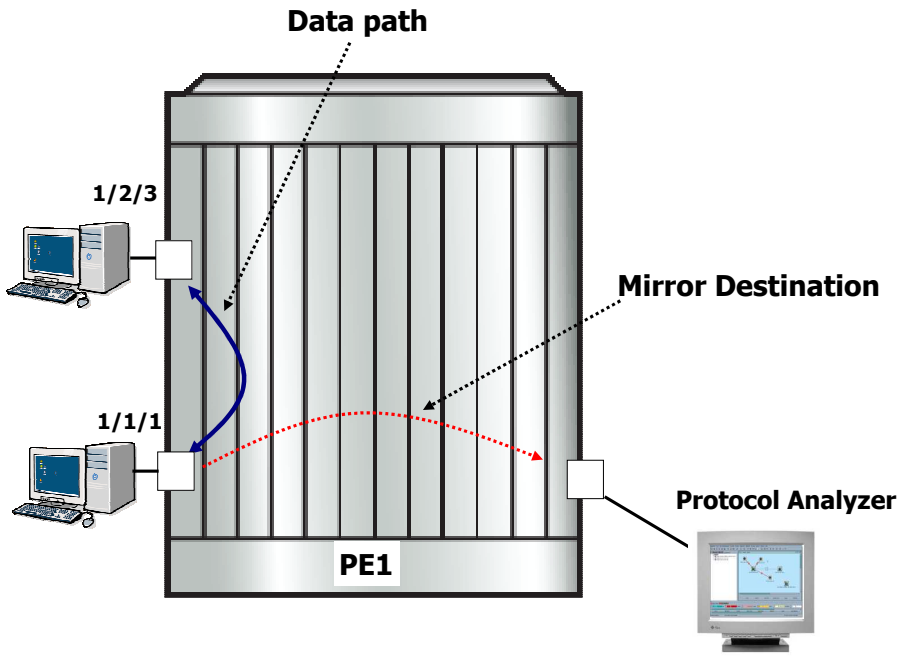
Mirror service

AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 9: Mirror service.

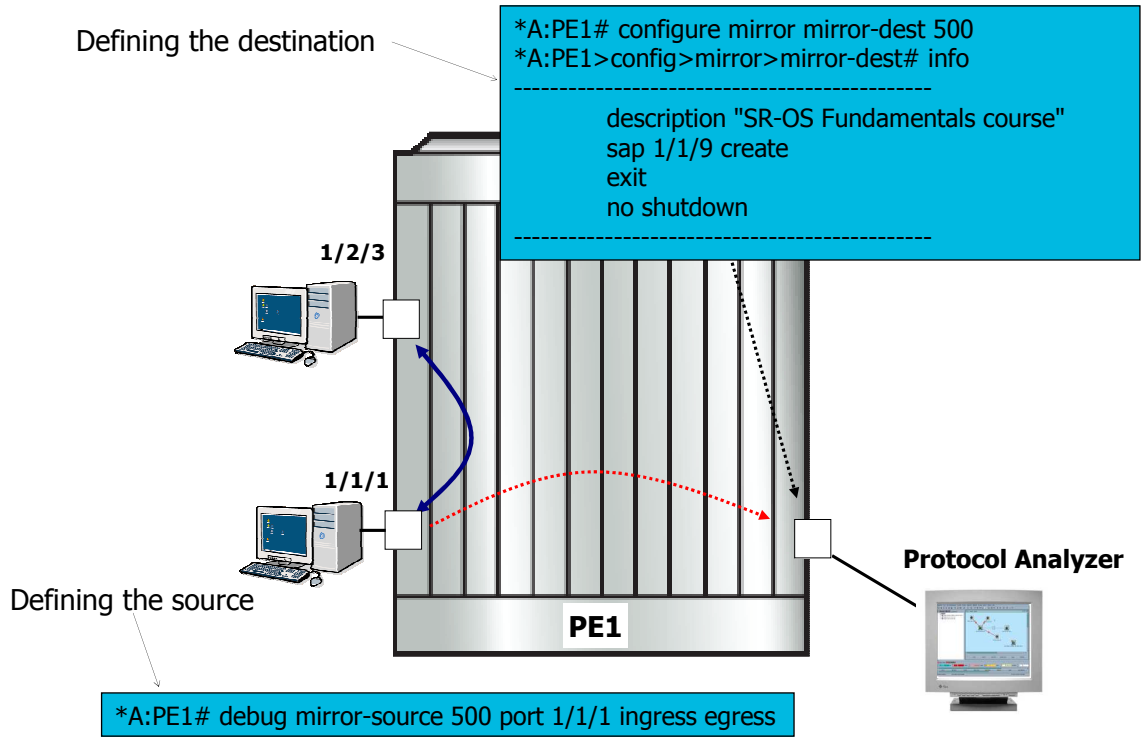
Local Mirror service



A mirror service is a type of service on a service router to mirror packets on a data link and to send it to a protocol analyzer.

In a local mirror service, the source and the destination of the mirror service are located on the same PE router.

Local Mirror service - configuration



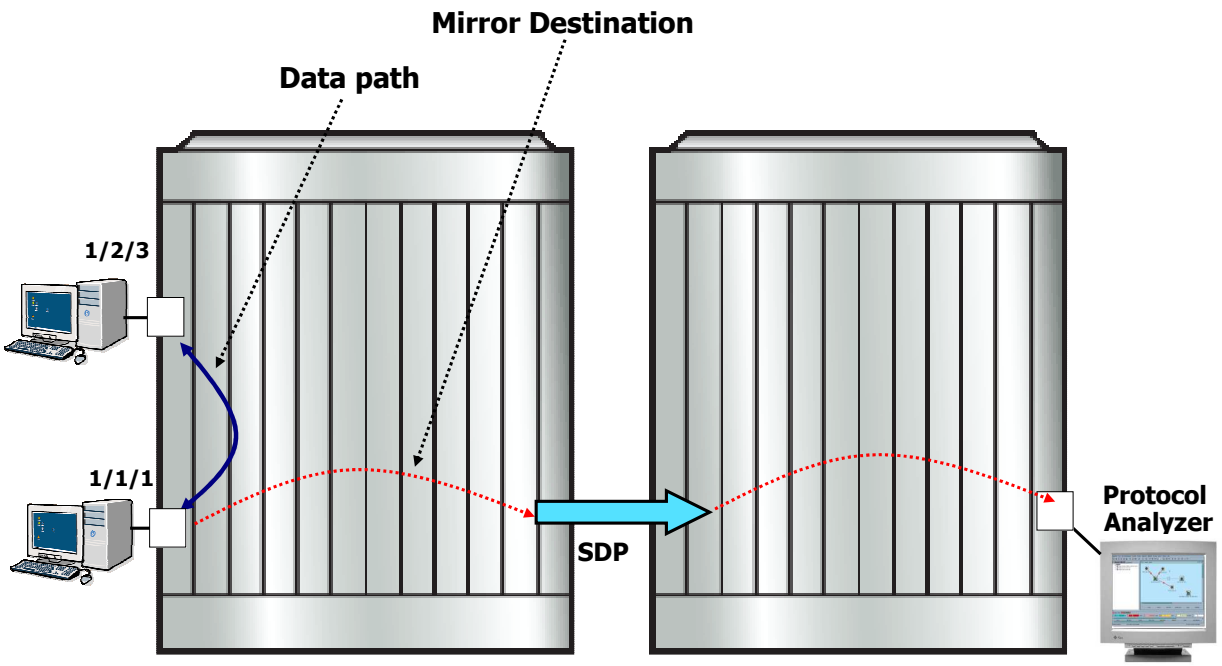
To configure a mirror service, a service ID needs to be configured within the context "configure mirror mirror-dest".

Only the destination needs to be defined here. For example SAP 1/1/9.

The source is defined under the debug context. It defines which port or SAP from which to mirror the traffic.

If not all packets needs to be mirrored, a filter can be optionally applied.

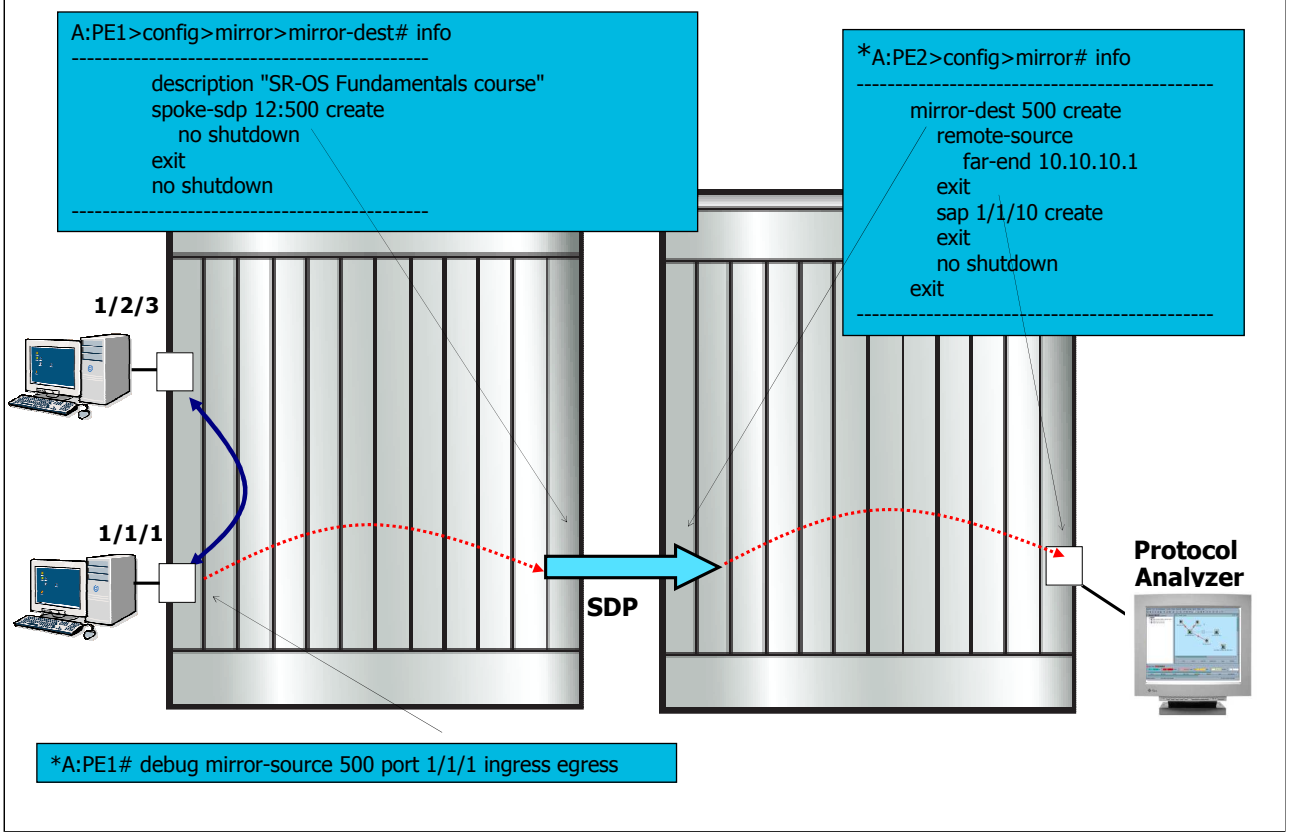
Remote Mirror service



A mirror service can also mirror packets on a remote service router. Instead of connecting the protocol analyzer to a SAP directly to the router, traffic is further forwarded to an SDP. On the destination PE, traffic is taken out of the SDP and forwarded to the connected protocol analyzer.

In this way, all traffic in the network can be seen for inspection.

Remote Mirror service - configuration



A mirror service is like a regular service and service labels are manually set or negotiated by TLDP. However, there are only labels negotiated for unidirectional traffic. There is no need to have traffic from the destination PE to the source PE.

On the PE where the source resides, a mirror service is created with the destination as the spoke-SDP.

The source is set in the debug command.

On the receiving PE, the same service is created with same ID, with the remote far-end defined of the source. SAP 1/1/10 is the location where the protocol analyzer is connected.

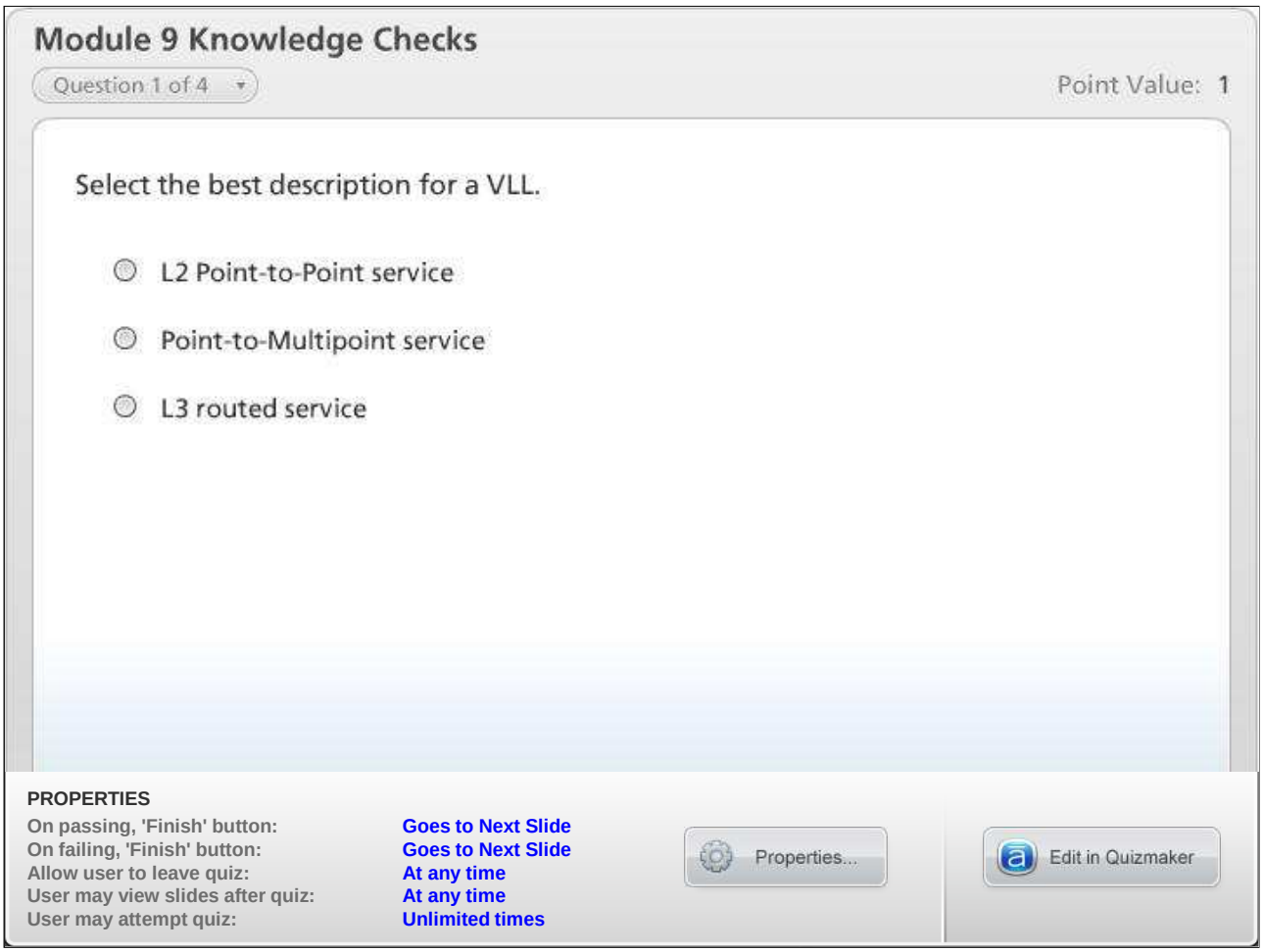
The service label is automatically negotiated and doesn't have to be specific.

The protocol analyzer can analyze traffic from different sources.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 9

Learning experience powered by Alcatel-Lucent University

Alcatel·Lucent

This completes module 9. Module 9 provided an overview of all VLL types available on the SR-OS products. At the end of the module, the local and remote mirror service was explained.

SR-OS Fundamentals

Module 10: VPLS – Virtual Private LAN Services

IPD Development

Available as PDF

Welcome to the 10th module of the SR-OS fundamentals course.

Table of Contents

Module 10 is divided into 7 sections.

Section 1: Introduction to Switching

Section 2: VPLS - Virtual Private LAN Service

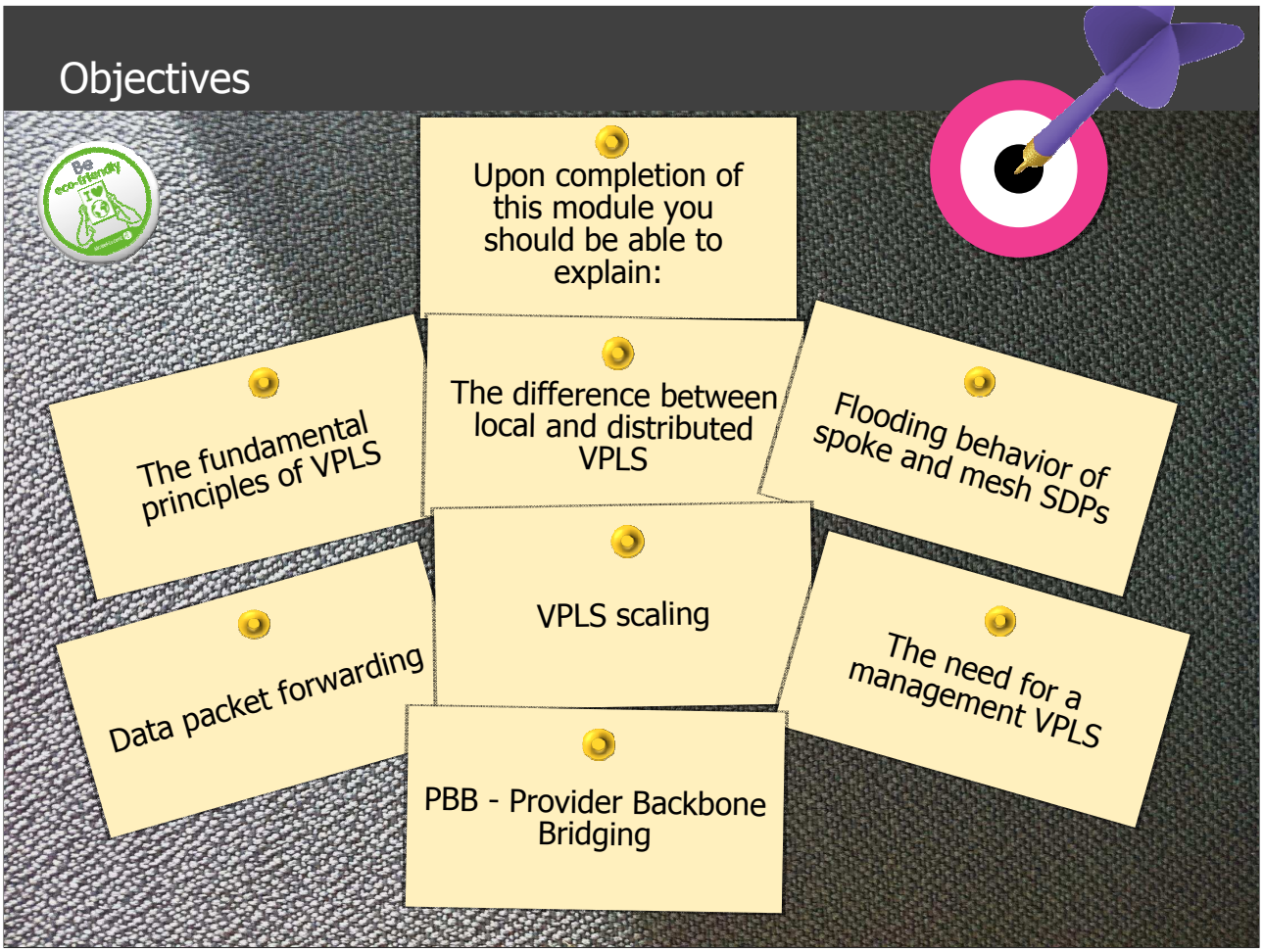
Section 3: Data forwarding

Section 4: VPLS Diagnostics

Section 5: VPLS scaling

Section 6: M-VPLS

Section 7: PBB - Provider Backbone Bridging



By the end of this module you will be able to explain:

The fundamental principles of VPLS

The difference between local and distributed VPLS

Flooding behavior of spoke and mesh SDP's

Data packet forwarding

VPLS scaling

The need for a management VPLS

PBB

1

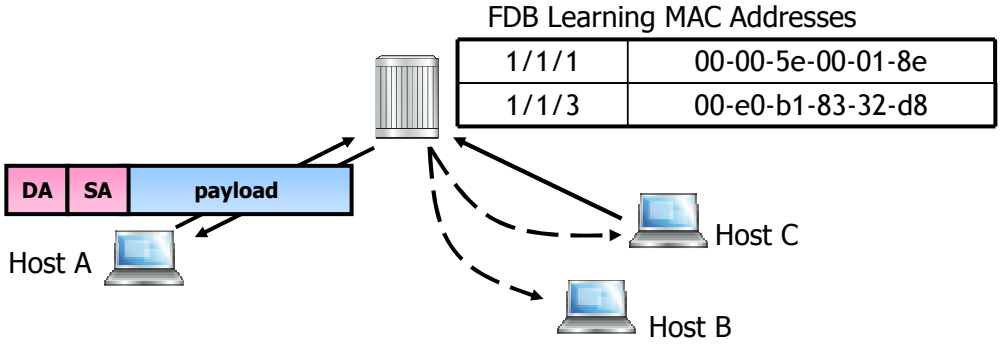
Introduction to switching

Section 1: Introduction to switching.

Populating a Forwarding Database (FDB)

Host A sends a frame to Host C

1. The switch receives a frame on port 1/1/1
   - Host A's MAC address is learned from the frame's Source Address field and is associated with port 1/1/1
2. If the Destination Address is not already in the FDB, the switch floods the frame out all ports except for the source port 1/1/1
3. Host C responds to Host A, and all other devices drop the flooded frame
   - From Host C's response frame, the switch associates Host C's MAC address with port 1/1/3



A VPLS is an ethernet L2 switch based service. So let us review the fundamentals of adaptive MAC learning on a switch.

When host A wants to send a packet to host C it sends its ethernet frame to its point-to-point physical switch connection. If the MAC forwarding table is empty on the switch, it doesn't know where to send the packet. It does not block the packet but floods it to all interfaces except back to host A. Host B and C receive the frame and C discovers that it is the one that is addressed and Host B will discard the frame. In the mean time, the switch has learned the MAC address of host A and on which port it is located. When host C replies, the switch can send the packet in a uni-directional way to host A and the switch will learn the MAC address of host C and its port.

No flooding is required for all further communication between A and C.

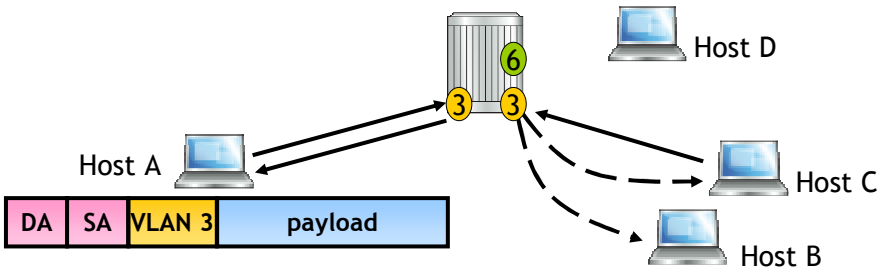
Although, in case of no refresh of the FDB table, the MAC addresses will be flushed and flooding behavior will be performed again.

Forwarding Databases Using Tagging

Virtual LANs (VLANs) are used to split a single switch into multiple virtual switches, each with the ability to service the hosts associated within that VLAN.

Graphic: Hosts A, B, and C are part of VLAN 3, Host D is part of VLAN 6.

1. The switch receives a frame on port 1/1/1:3 (:3 signifies VLAN 3)
   - Host A's MAC address is learned and is associated with VLAN 3's FDB
2. If the Destination Address is not already in the FDB, the switch floods the frame out **all ports associated with VLAN 3**
3. Host C responds to Host A, and all other devices drop the flooded frame
   - From Host C's response frame, the switch populates Host C's MAC address in VLAN 3's FDB



To limit the flooding, a switch network can be split into a couple of virtual switched networks or VLAN's. These VLAN's are all running on the same physical network but are using VLAN numbers to indicate which virtual switched network they belong to.

VLAN's are used to split a single switch into multiple virtual switches, each with the ability to service the hosts associated within that VLAN.

In the example, hosts A, B and C belong to one VLAN. Traffic from host A will only be flooded to host C and B and not to host D as he is part of a different VLAN.

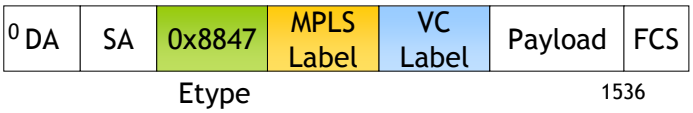
Note that in this example no communication between the two VLAN's is possible. A VPLS service does have this option as VLAN can be stripped off and replaced by another VLAN.

Ethertype (Etype)

The Ethertype field describes the payload of the ethernet frame

Ethertype for some common Protocols	Protocol
0x0800	Internet Protocol, Version 4 (IPv4)
0x0806	Address Resolution Protocol (ARP)
0x8100	VLAN – Tagging (Dot1Q)
0x9100	VLAN – Tagging (QinQ)
0x86DD	IPv6
0x8847	MPLS unicast
0x88a8	Provider Bridging

Etype 8847 indicates this packet is a Unicast MPLS frame



Ethertype - MPLS (8847)

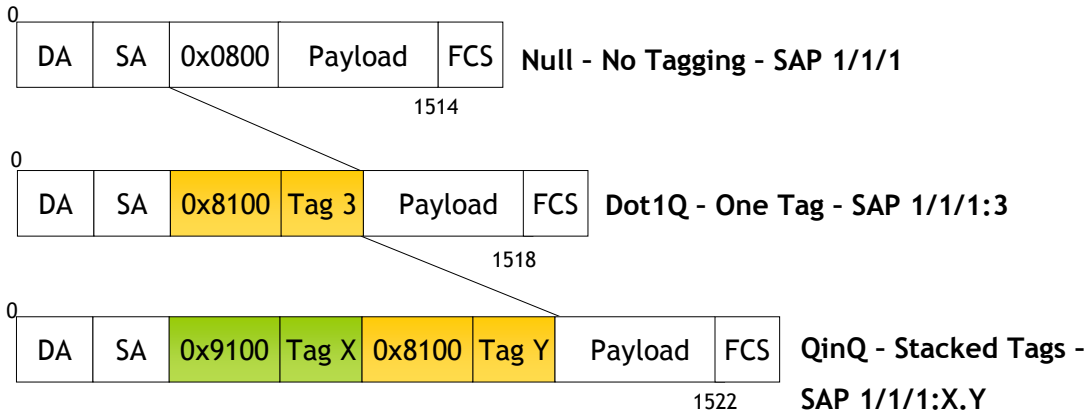
The Ethertype is a field found in an Ethernet version 2 frame and indicates what type of protocol is encapsulated.

Ethertype numbering generally starts from 0x0800. In the example shown, Etype 8847 indicates a unicast MPLS ethernet frame.

VLAN encapsulation type

Ethernet port encapsulation type can be configured as one of the following:

- Null Header (Raw Mode), all bits after the 'type' field are treated as data
- Dot1q Header (Single tag header), adds an additional 4 byte "TAG"
- QinQ Header (Two tags header), adds two TAGs after Source Address field



The Ethernet VLAN encapsulation type can be configured as one of the following:

•Null Header where all bits after the 'type' field are treated as data

•Dot1q Header that adds an additional 4 byte VLAN "TAG"

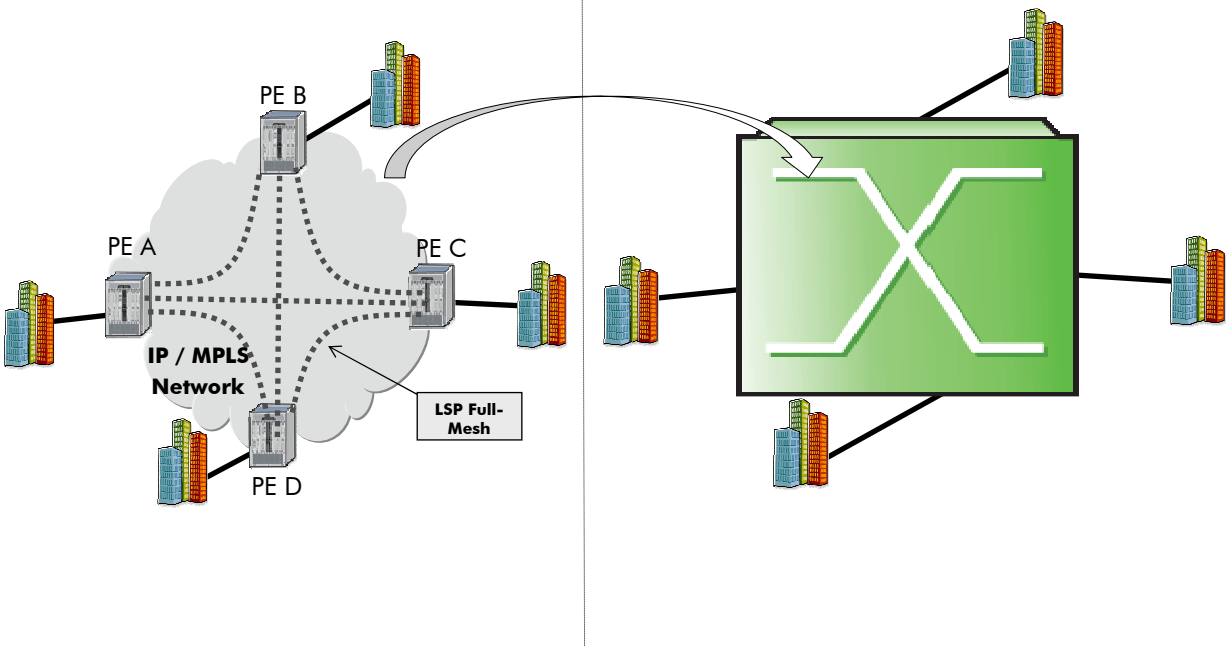
•QinQ Header that adds two VLAN TAGs after Source Address field

VPLS - Virtual Private LAN Service

Section 2: VPLS - Virtual Private LAN Service

VPLS <-> Switch

VPLS acts as an intelligent switch where the customers are connected to geographically spread PE's



A distributed VPLS service, a VPLS where more than one PE router is involved, can be seen as one switch.

All SAP's towards the CE's are then the ports of the switch, whereas all the SDP's are the internal traffic paths of a switch.

The FDB MAC table is spread amongst all PE routers. A separated FDB table is kept per VPLS service instance on a PE.

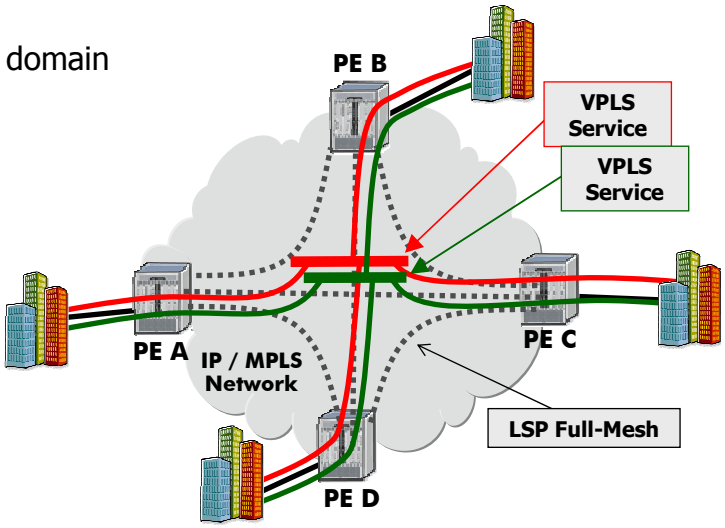
How is a VPLS provided over MPLS?

Bridging capable PE routers
Connected with a full mesh of MPLS LSP tunnels

Per-Service VC labels
Negotiated using draft-Martini

Unknown/broadcast
Traffic replicated in a service domain

MAC learning
- Over tunnel & access ports
- Separate FIB per VPLS



For each VPN at each site, a Customer Edge device connects to the Provider Edge router via a point-to-point access connection.

Ethernet serves as the framing technology between the CE device and the PE router in the provider's network. Frames can include IEEE 802.1Q Ethernet VLAN tags, which allow customers to segment their networks and assign quality of service priorities to LAN traffic. VPLS also supports "QinQ" encapsulation, where a second VLAN tag is added as a service delimiter. From the customer's perspective, the entire VPN looks like a single Ethernet switched LAN, with the PE acting as a switch that switches frames on the basis of their Layer-2 destination MAC addresses.

On the provider's side, however, PEs are interconnected with Generic Routing Encapsulation and/or Multiprotocol Label Switching (MPLS) tunnels. If PEs are connected using GRE tunnels traffic is encapsulated and routed through the core network using standard IP frame formats and addressing. If PEs are connected using MPLS tunnels, traffic is encapsulated in an MPLS frame and transmitted using MPLS labels.

In a basic configuration, all PE's should be fully meshed to connect all PE's to a VPLS service. However, a different SDP is available so that there is no need for a full mesh as discussed later.

The VC labels negotiation uses draft-martini/TLDP on a point to point basis. So for 4 PE's, 6 bi-directional point-to-point TLDP sessions are established.

Unknown data traffic or broadcast traffic is replicated in the VPLS service domain. Unknown data traffic arriving on a SAP will be flooded to all SDP's belonging to the VPLS service. At the same time, MAC addresses are learned over the SDP tunnels and SAP access ports.

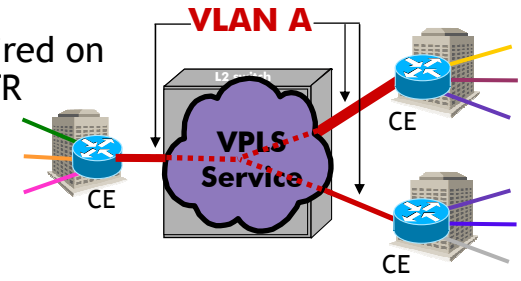
VPLS: Customer view

All locations appear to be on the same Ethernet LAN.

Entire provider network appears to be a Layer 2 switch/VLAN.

CE-PE interface

- Simple Ethernet interface
- Removes L2 protocol conversion between LAN and WAN
- No additional training required on WAN technologies such as FR



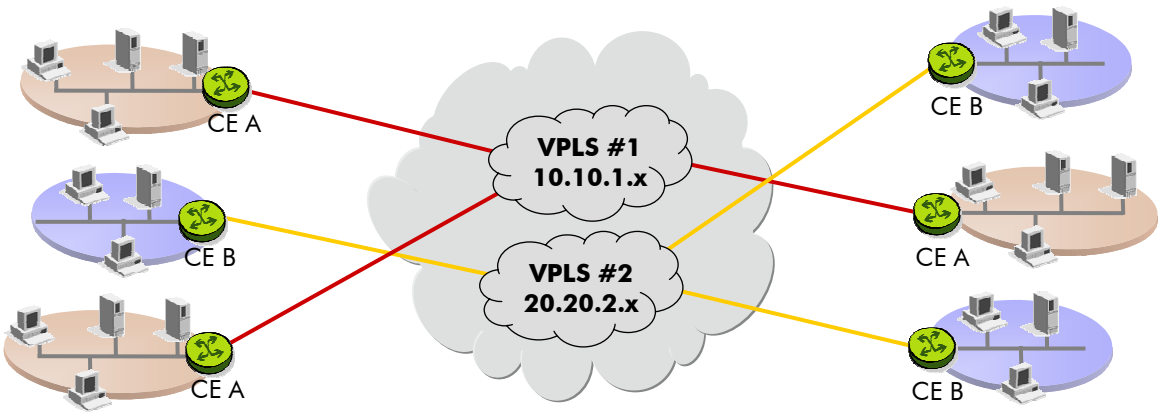
From the customer perspective, all geographically spread locations appear to be on the same ethernet LAN.

The entire provider network appears to be a Layer 2 switch.

The CE-PE connection is a simple ethernet interface. No additional WAN expertise is required here.

VPLS: Customer operation

- Customers maintain complete control over routing
- Adding new sites simplified: no re-configuration at existing sites



Customers maintain complete control over their routing. The provider is never interfering at the IP level.

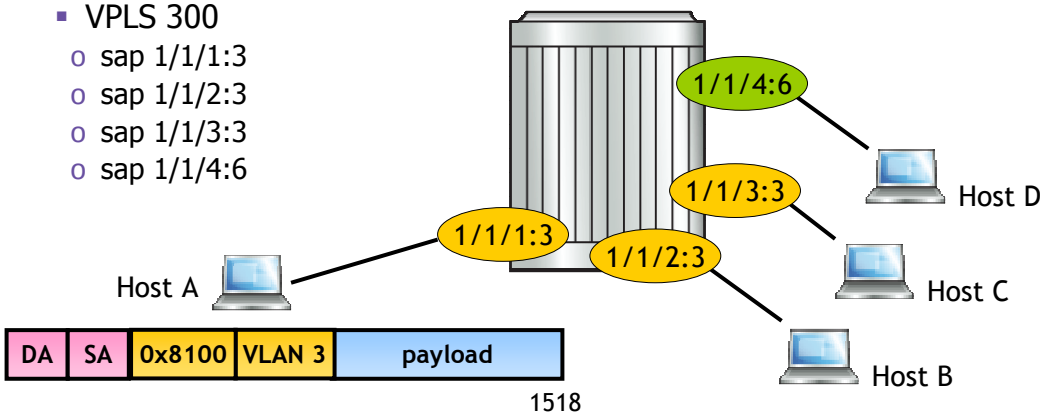
This clear demarcation of functionality between the service provider and customer makes troubleshooting easier.

Adding a new site requires no configuration of the service provider's equipment or the customer equipment at existing sites.

Building a Local Virtual Private LAN Service (VPLS)

Creating a Local VPLS Service, meaning a service that exists solely on a single node:

- Change the ports to port mode access
  - Ethernet mode access
- Enable tagging on the ports
  - Ethernet encap-type dot1q
- Create Service Access Points (SAPs) within the VPLS service using the desired port/tag
  - VPLS 300
    - sap 1/1/1:3
    - sap 1/1/2:3
    - sap 1/1/3:3
    - sap 1/1/4:6



A local VPLS is a service on 1 PE having SAP's and no SDP's.

All frames that ingress to this local VPLS service are compared to the FDB to determine which SAP the frame is to be forwarded out. Because the egressing port is a SAP, the frame will be regenerated with the appropriate tag's based on the SAP definition. For example, a dot1q SAP that is provisioned as 1/1/1:3 will have the dot1q tag 3 inserted into all frames that egress that port based on the FDB.

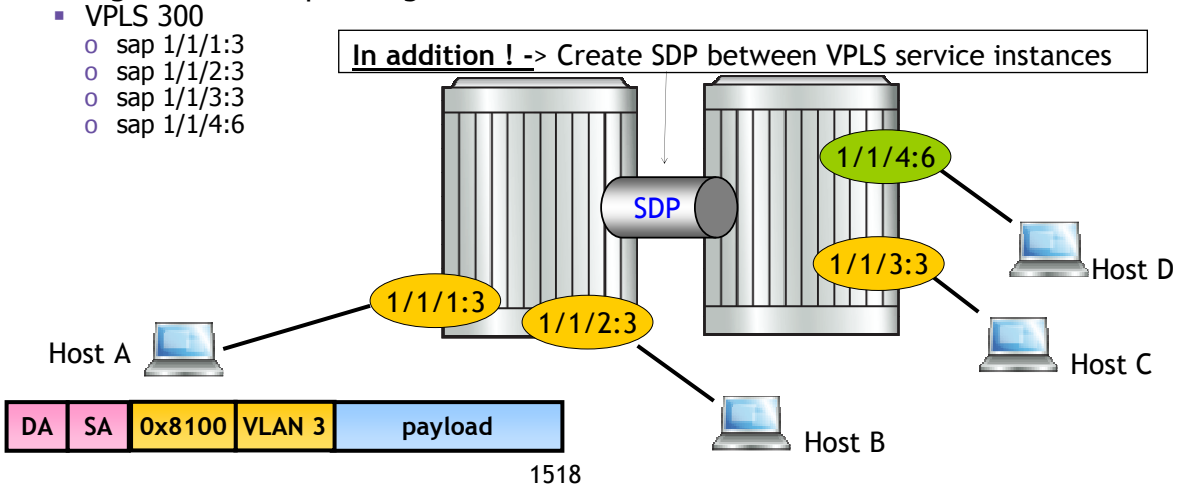
When a wildcard is used, for example 1/1/1:* although the SAP is defined as dot1Q, no additional tags are added on egress.

A local VPLS service requires only a few configuration steps. After the VPLS service instance is created, the SAP's can be added. A SAP is created on a port in access or hybrid mode. After the access mode is configured, the encapsulation method can be set.

Building a Distributed VPLS

Creating a Distributed VPLS Service, meaning a service that exists on a minimum of two nodes:

- Change the ports to port mode access on both PE's
  - Ethernet mode access
- Enable tagging on the ports
  - Ethernet encap-type dot1q
- Create Service Access Points (SAPs) within the two VPLS service instances using the desired port/tag
  - VPLS 300
    - sap 1/1/1:3
    - sap 1/1/2:3
    - sap 1/1/3:3
    - sap 1/1/4:6

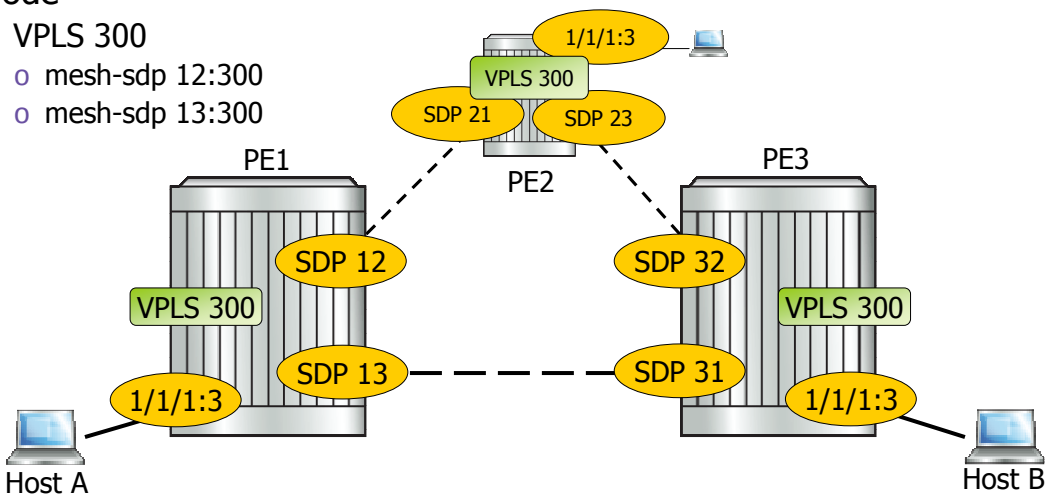


A distributed VPLS requires a minimum of two PE's running the same VPLS instance. Apart from the SAP that are added to each VPLS instance on the PE, an SDP service tunnel is needed to make the interconnection between the PE's.

Building a Distributed VPLS (cont.)

Creating a Distributed VPLS Service:

- Add Service Access Points (SAPs) to the VPLS service on each node
  - VPLS 300
    - sap 1/1/1:3
- Add Service Distribution Points (SDPs) to the VPLS service on each node
  - VPLS 300
    - mesh-sdp 12:300
    - mesh-sdp 13:300



What are the steps needed to build a distributed VPLS?

Before any service is provisioned the following tasks should be completed:

Build the IP or MPLS core network

Configure routing protocols

Configure MPLS LSPs if MPLS is used and not GRE

Construct the core SDP service tunnel mesh for the services

After this, services can be created. A distributed VPLS service is identified by a service id and customer id. It is suggested to have the same number on each PE. After the creation of the service instance, add the Service Access Points to the VPLS service on each node and add the Service Distribution Points to the VPLS service on each node.

VPLS - VC Lable

VC-label Signaling between PEs per VPLS service instance

- Each PE initiates a targeted LDP session to the far-end System IP address
- LDP protocol is used to communicate the VC label when sending packets for each service
- VC-IDs for the related services on different routers must match

PE1->PE2: For Svc-id 101 Use VC-label pe2-1

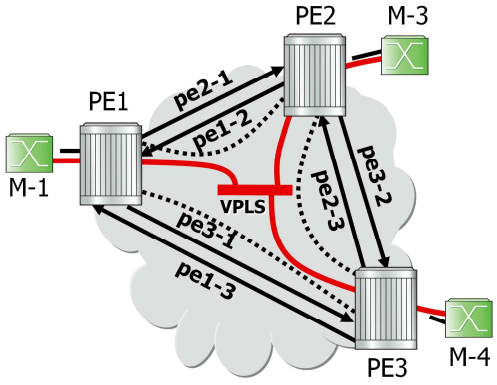
PE1->PE3: For Svc-id 101 Use VC-label pe3-1

PE2->PE1: For Svc-id 101 Use VC-label pe1-2

PE2->PE3: For Svc-id 101 Use VC-label pe3-2

PE3->PE1: For Svc-id 101 Use VC-label pe1-3

PE3->PE2: For Svc-id 101 Use VC-label pe2-3



Customer packets are transported either inside an IP packet (GRE) or inside an MPLS packet.

The packet carries an inner label that identifies the service the packet belongs to. This label is sometimes referred to as the Martini label which is an MPLS label that identifies the particular service that is being transported through the MPLS tunnel.

When a packet arrives at the destination, the outer IP address or MPLS label is stripped off. At this point the inner label is examined to determine which service the packet belongs to. There can be multiple services on one PE.

After determining which service the packet belongs to, the customer's Ethernet packet is examined and its MAC address is looked up in a table on the PE to determine which SAP the packet should go to.

VC labels can be assigned manually or automatically using targeted LDP (TLDP). The TLDP protocol is used to dynamically negotiate VC labels between PE's. This method is not error prone and scales much better then manually assigning labels.

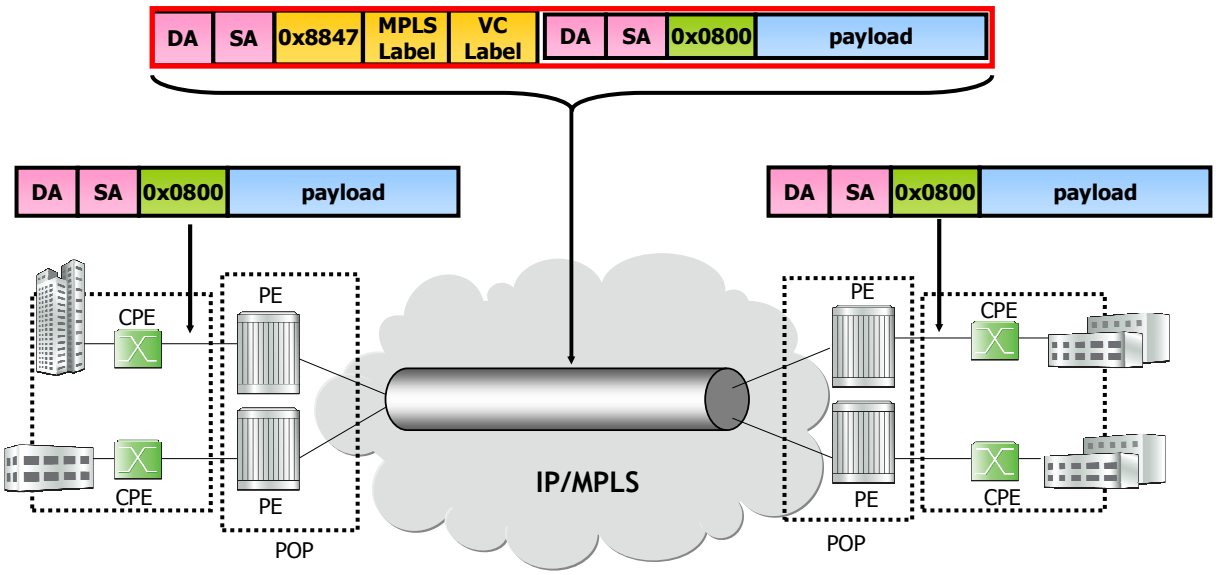
Data forwarding

AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 3: Data forwarding

Packet format



The customer frame becomes the payload of the MPLS frame

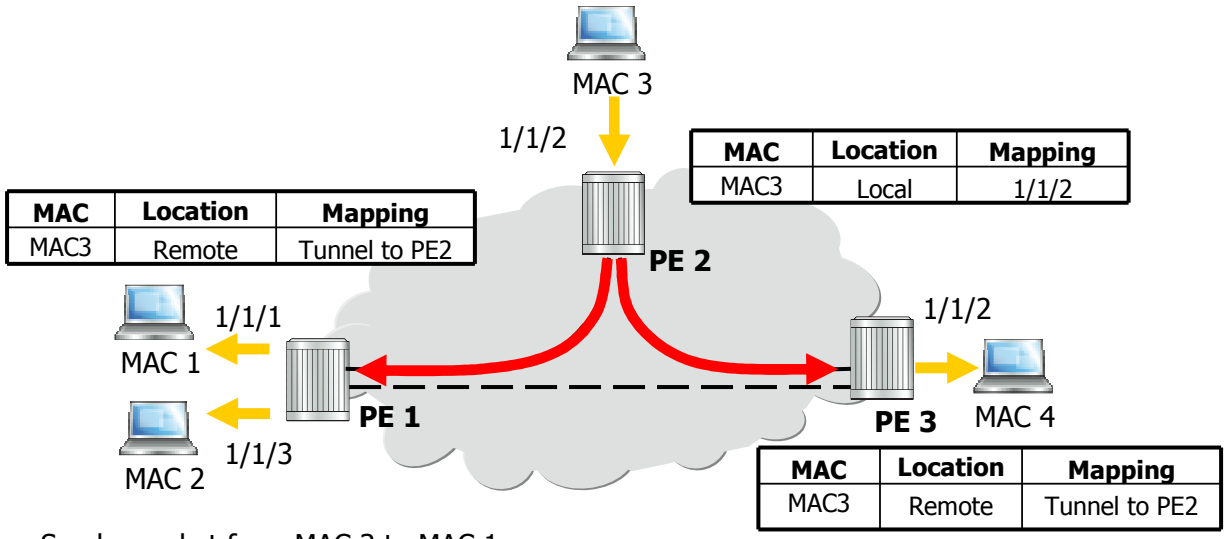
The ethernet data traffic coming from a customer is encapsulated into MPLS in case of a non-GRE SDP. The VC label is added first to identify the service this traffic belongs to. A second label is added that represents the transport tunnel.

Between the PE's there is also Ethernet framing with the Ethertype of 0x8847 indicating the MPLS frame.

This is why there are two Ethernet frames seen in the core provider's network.

Any dot1q or QinQ tags could be stripped off at the SAP before being encapsulated into MPLS.

VPLS Packet DATA walkthrough - VPLS learning



Send a packet from MAC 3 to MAC 1

- PE2 learns that MAC 3 is reached on Port 1/1/2
- PE2 floods to PE1 with VC-label pe2-1 and PE3 with VC-label pe2-3
- PE1 and PE3 learns that MAC 3 is behind PE2
- PE3 sends on Port 1/1/2
- PE1 sends on Port 1/1/1 & 1/1/3

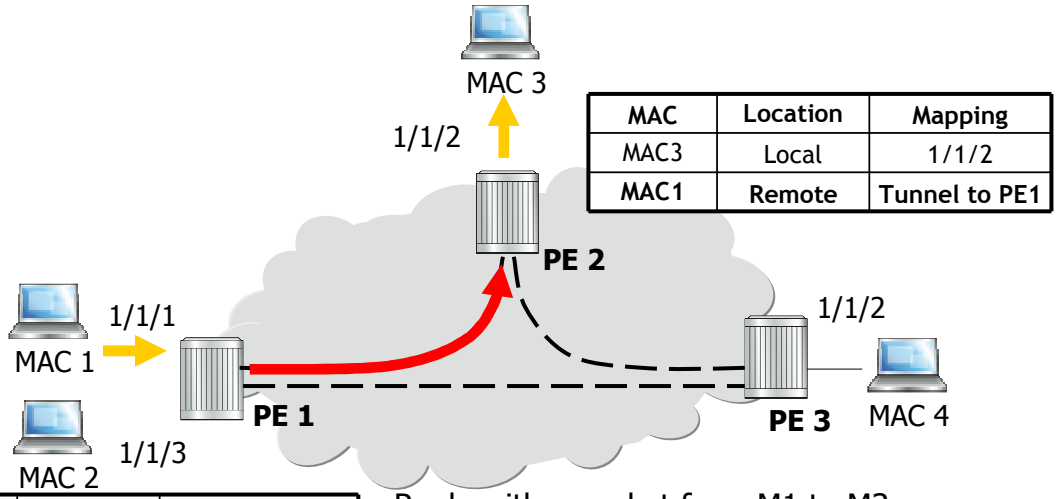
PE routers learn the source MAC addresses of the traffic arriving on their access and network ports.

Each PE router maintains a Forwarding Information Base or FIB for each VPLS service instance and learned MAC addresses are populated in the FIB table of the service.

All traffic is switched based on MAC addresses and forwarded between all participating PE routers using the LSP tunnels.

Unknown packets, i.e. destination MAC address has not been learned, are forwarded on all the LSPs to the participating PE routers for that service until the target station responds and the MAC address is learned by the PE routers associated with that service.

VPLS Packet Forwarding



MAC	Location	Mapping
MAC3	Local	1/1/2
MAC1	Remote	Tunnel to PE1

MAC	Location	Mapping
MAC3	Remote	Tunnel to PE2
MAC1	Local	1/1/1

Reply with a packet from M1 to M3

- PE1 learns M1 is on Port 1/1/1
- PE1 knows that M3 is reachable via PE2
- PE1 sends to PE2 using VC-label pe1-2
- PE2 learns MAC 1 is reachable via PE1
- PE2 knows that M3 is reachable on Port 1/1/2

All hosts that receive the frame but are not addressed by the destination MAC address will discard this frame. The host that is addressed will respond and the frame is sent, non-flooded, towards the originator based on the FIB forwarding tables on each node.

VPLS Forwarding Information Base Controls

Command: “disable learning”
- Prevents learning of new addresses for security reasons/debugging

Command: “discard unknown frames”
- Only allow packets from known customer MAC addresses for security

Configurable MAC FIB size to limit number of MAC per service
- High and low watermark

Can configure static MAC per SAP or SDP

Independent aging of MAC address:
- Local MAC Address (SAP)
- Remote MAC Address (SDP)
  - Remote MAC has longer expiry time to prevent unnecessary flooding

The learning of MAC addresses can be controlled by a couple of CLI commands.

There is the option to prevent the learning of new addresses, called the “disable learning” command.

“Discard Unknown Frames” allows only packets from known customer MAC addresses for security reasons.

A high and low watermark sign are available as threshold settings on the FDB table. Once the high watermark is exceeded, an alarm will be generated. Of the number of MAC entries decreases and passes the low watermark threshold, another alarm is raised.

MAC entries can also be manually added.

The age time per MAC is different when learned on a SAP or SDP. The default aging time on a SAP is llonger to prevent unnecessary flooding.

VPLS diagnostics

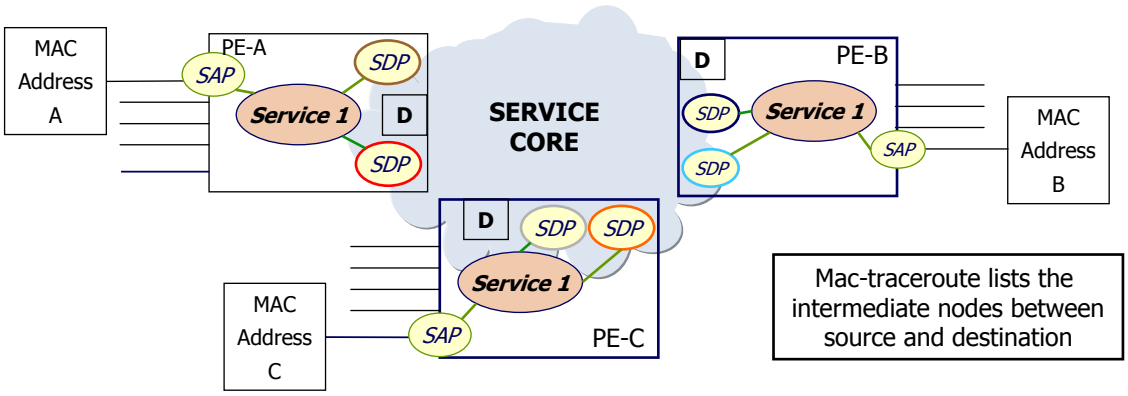
AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 4: VPLS diagnostics.

VPLS MAC diagnostics – MAC-PING

ACTION	RESULT
Mac-Ping Service 1 to destination MAC address B. PE-A will send a MAC-ping to PE-B (per the FIB).	PE-B has the MAC in its FIB, acknowledges the ping.
Mac-Ping Service 1 to destination D (unknown) Sends the MAC-ping to all nodes (per the FIB)	No node responds.
Mac-ping service 1 destination FF:FF:FF:FF:FF:FF	All nodes in the service respond.



Once set up there are mechanisms to test the service or to test the MAC table entries.

One of them is the MAC ping.

A MAC ping will poll the PE's to check if a certain MAC address is learned on a SAP. The PE that has the MAC address on a SAP will responds to the MAC ping with its system address and the SAP ID.

MAC Ping requests directed to the MAC broadcast address FF:FF:FF:FF:FF:FF are flooded throughout the service flooding domain and will receive a response from all operational SAPs. Note that SAPs that are operationally down do not reply.

MAC-PING CLI

```
SR1# oam mac-ping
  - mac-ping service <service-id> destination <ieee-address> [size <octets>]
    [ttl <vc-label-ttl>] [send-control] [return-control] [source <ieee-address>]
     [interval <interval>] [count <send-count>]

 <service-id>         : [1..2147483647]
 <ieee-address>       : xx:xx:xx:xx:xx:xx or xx-xx-xx-xx-xx-xx
                        All zero not allowed
 <octets>             : [1..65535]
 <vc-label-ttl>       : [1..255]
 <send-control>       : keyword - sends via the control plane
 <return-control>     : keyword - receives on the control plane
 <ieee-address>       : xx:xx:xx:xx:xx:xx or xx-xx-xx-xx-xx-xx
                        All zero and multicast not allowed
 <interval>           : [1..10]
 <send-count>         : [1..255]


SR1# oam mac-ping service 100 destination ff:ff:ff:ff:ff:ff

Seq Node-id                                                        Path     RTT
------------------------------------------------------------------------------
[Send request Seq. 1.]
1   10.250.1.4:sap1/2/2:100                  No FIB on Egress Self     0ms
1   10.250.1.5:sap1/2/2:100                  No FIB on Egress In-Band  0ms
------------------------------------------------------------------------------
```

This slide shows the MAC ping CLI options.

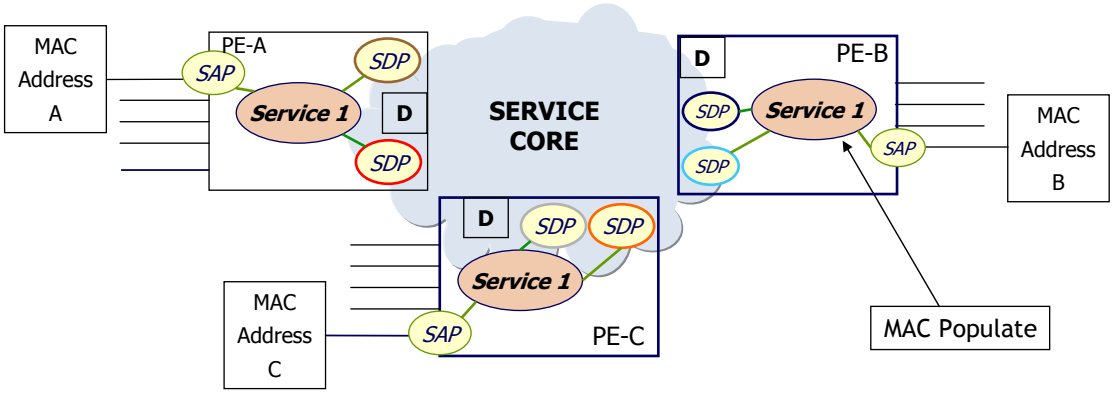
VPLS MAC diagnostics - MAC Populate

Scenario :

1.PE-A tries Mac-Ping Service 1 to Destination D (unknown) - no node responds.
2.Perform a “mac-populate” Service 1 destination D that will create a FIB entry and flood within VPLS service to PE-B.

- Example: **`SR1>show>service>id# oam mac-populate 1000 mac 12:34:56:78:90:ab flood`**

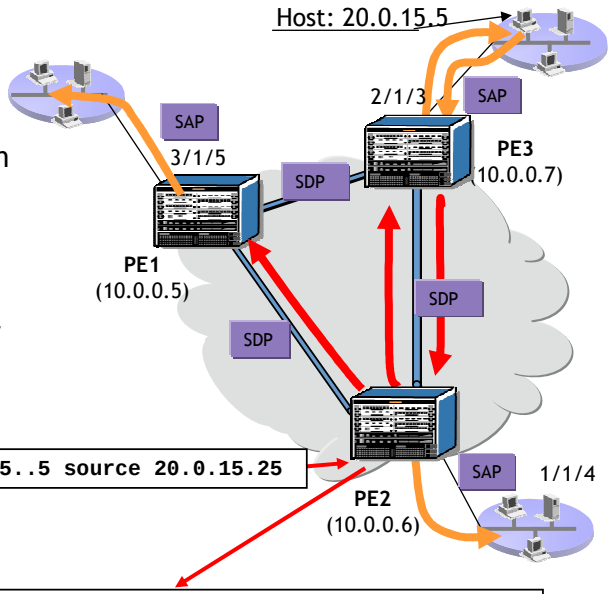
3.When PE-A retires a Mac-ping service 1 destination D - PE-B responds.



For testing purposes, the operator might manually add a MAC address. This MAC address can be added locally on 1 PE or added to all PE’s via flooding. This MAC ping can then be sent to this MAC address. The MAC addresses that have been populated will never age out and have to be removed by the MAC purge operation.

CPE-Ping

1. CPE-ping command is translated into a MAC-ping to a broadcast address.
   All destinations within the VPLS will be reached by the ping to the broadcast MAC address.
2. Each peer PE within the VPLS will issue an ARP request for the specific IP address.
   IP DA = configured value
   IP SA = configured "dummy" address
   MAC DA = all 1s
   MAC SA = CPM MAC address
3. The PE that receives a response will reply to the PE that issued the CPE-ping.




```
#oam cpe-ping service 10 destination 20.0.15..5 source 20.0.15.25
```

```
Seq  Sap-id                                    CPE Mac Address        RTT
------------------------------------------------------------------------------
[Send request Seq. 1.]
1   10.0.0.7:sap:2/1/3                  22:6a:01:01:00:02   10ms
[Echo replies received: 1]
------------------------------------------------------------------------------
```

A CPE ping is the only diagnostics tool that triggers the customer device. A MAC populates and MAC ping were never reaching out to the customer. A CPE ping will ping an IP address of a customer and the PE where the IP address resides on will responds.

But the communication from the PE towards the CE is not a ping, but a simple ARP message. The CE that's responds to the ARP is assumed to have the IP address configured.

The source address during the CPE-ping should be set to a "fake" address to avoid the CE from associating the MAC of chassis to a valid customer IP address. Any IP address will do.

CPE-Ping configuration

```
oam cpe-ping service <service-id> destination <ip-address> source <ip-address>
[ttl <vc-label-ttl>] [return-control] [source-mac <ieee-address>]
[interval <interval>] [count <send-count>] [send-control]
```

Source	Specifies the IP address to be used, must be in the same subnet as the destination IP address for the ARP command to work.
Destination	The destination IP address that you want to ping.
Return-control	Specifies the MAC OAM reply to a data plane MAC OAM request be sent using the control plane instead of the data plane. Default: MAC OAM reply sent using the data plane.
Send-control	Specifies the MAC OAM request be sent using the control plane instead of the data plane. Default: MAC OAM reply sent using the data plane.

This slide illustrates the different kind of options for a CPE ping.

VPLS scaling

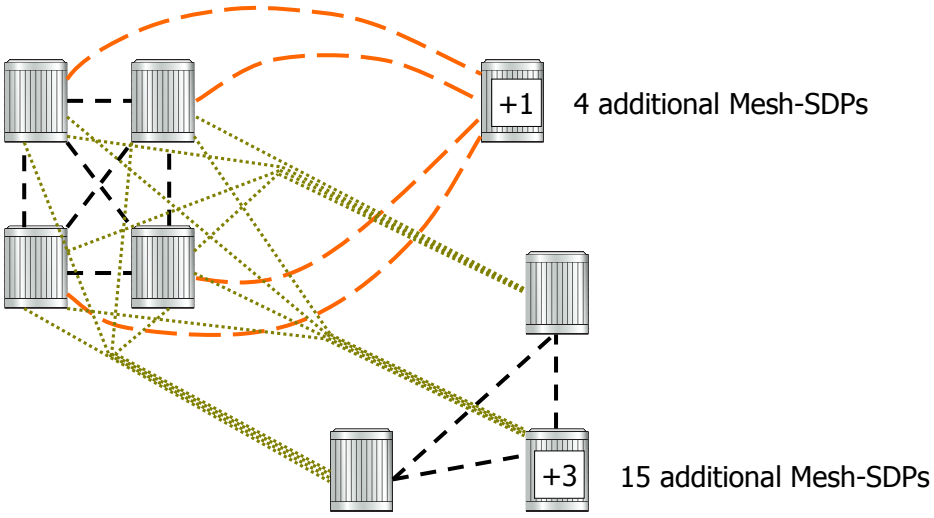
AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 5: VPLS scaling

VPLS Topologies - Scaling

When working with a large VPLS every addition to the existing mesh results extensive network touches, and an exponential growth of control sessions



In a basic VPLS set-up, all SDP's should be fully meshed.

The number of meshed SDP's equals n(n-1)/2, where n is the number of nodes.

4 nodes requires 6 meshed bi-directional SDP's and hence TLDP sessions.

An additional node adds four additional meshed SDP's

Adding three extra node would result in 15 additional meshed SDP's.

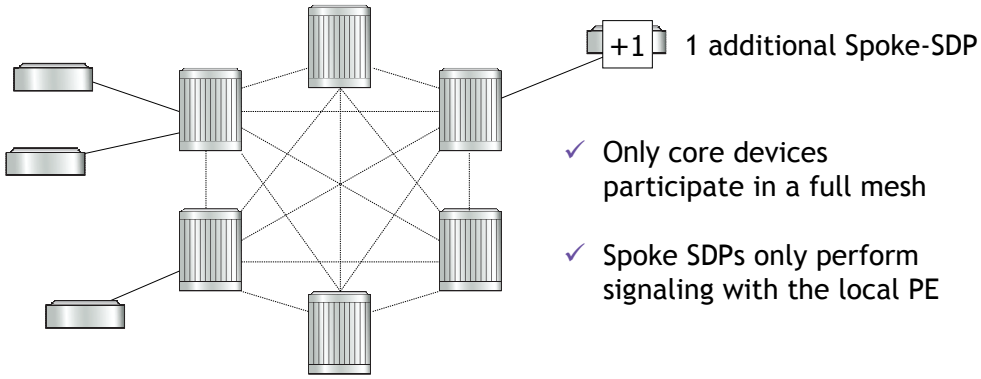
This is an exponential curve and not scalable for large networks.

Bringing a hierarchy would solve this issue.

Hierarchical VPLS - VPLS Scalability

Spoke-SDPs remove the requirement for full mesh connectivity to remote PEs

- Spoke-SDPs dramatically reduces the number of connections required
- Packet replication is reduced
- Simplifies edge devices, keeping cost down
- Minimizes signaling overhead since fewer SDPs are required for the VPLS

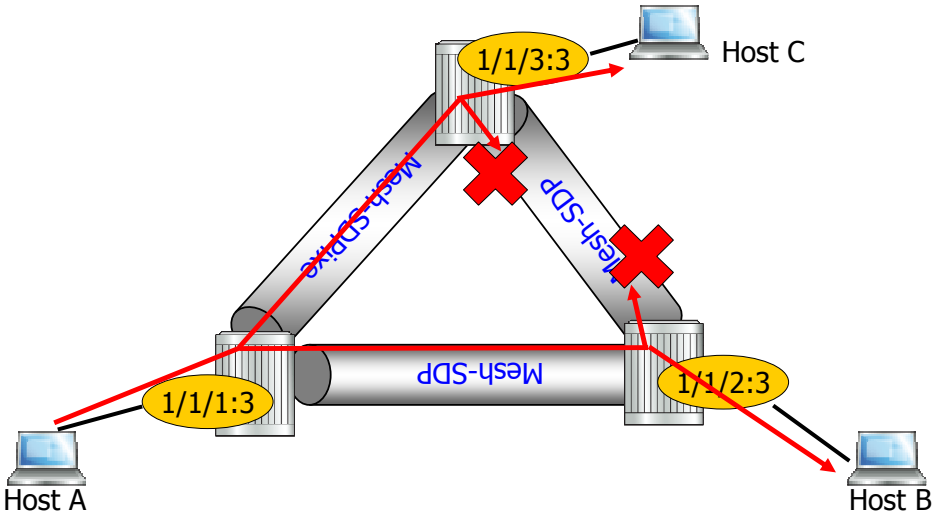


A special type of SDP, a spoke SDP, will create a hierarchical VPLS and decreases the number of SDP's.

This SDP, the spoke-SDP removes the requirement for full mesh connectivity to remote PE's.

Only the core devices participate in a full mesh and the spoke SDP's only perform signaling with the local PE.

Split Horizon and Mesh SDPs



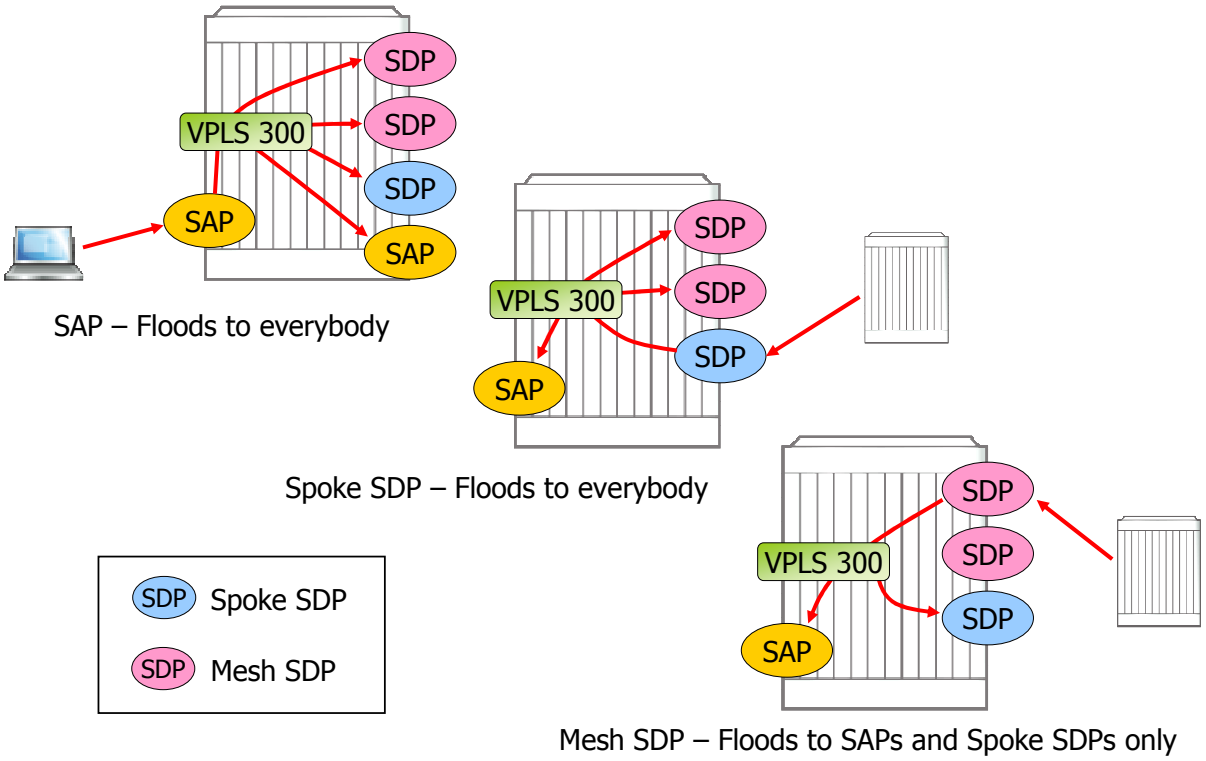
- Multipoint packets received from a mesh SDP are never forwarded to another mesh SDP within the same service.
- A mesh SDP cannot be used for redundancy purposes
- No loops for broadcast, unicast or unknown packets

The full meshed core of SDP's requires the SDP of the type "mesh". A Mesh SDP has the behavior that multipoint packets received from a mesh SDP are never forwarded to another mesh SDP within the same service.

A mesh SDP cannot be used for redundancy purposes but avoids loops for broadcast, unicast or unknown packets.

If all SDP's would have been from the type "spoke", broadcast, unicast or unknown packets might be flooding infinitely as spoke SDP's flood their traffic to other spoke-SDP's.

Flooding traffic within a VPLS



Flooded traffic received on:

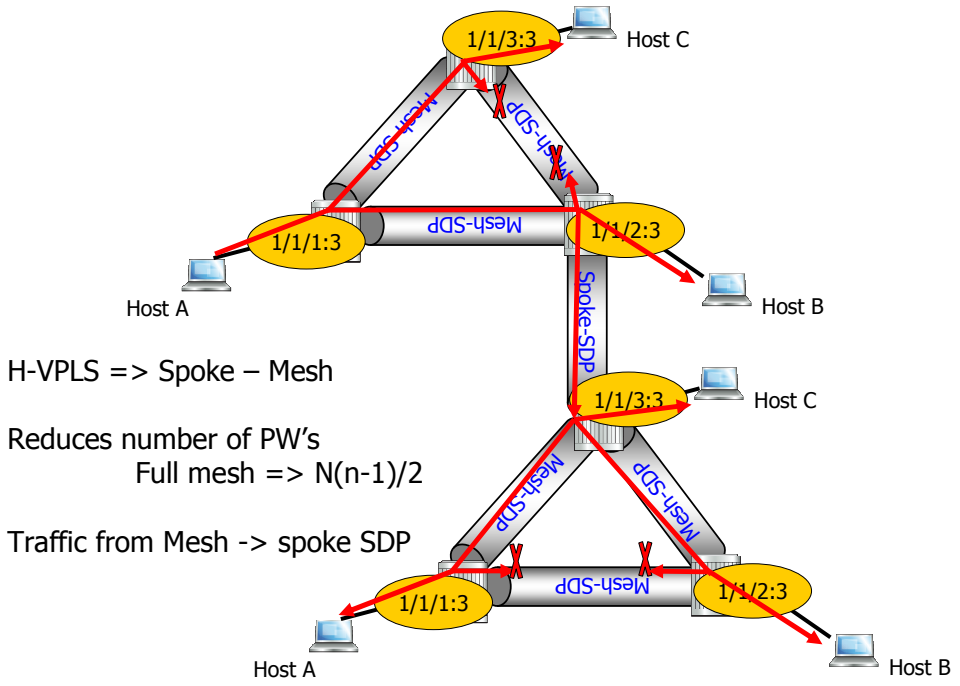
SAP – Flooded traffic received on a SAP is replicated to other SAPs, Spoke SDPs, and Mesh SDPs

Spoke-SDP – Flooded traffic received on a Spoke SDP is replicated to other SAPs, Spoke SDPs, and Mesh SDPs

Mesh-SDP – Flooded traffic received on a Mesh SDP is replicated to other SAPs, and Spoke SDPs but is not transmitted on other Mesh SDPs

All frames that ingress to a service are compared to the FDB to determine which SAP or SDP the frame is to be forwarded out. Flooding only occurs if there is no existing entry in the FDB. Egress tags are generated based on the SAP definition of the egress port in the FDB.

H-VPLS



An H-VPLS or hierarchical VPLS is nothing more than a combination of mesh and spoke SDP's.

Unknown traffic from host A is flooded to all mesh-SDP's. It is further flooded between mesh and spoke to the other meshed core. But never between mesh SDP's.

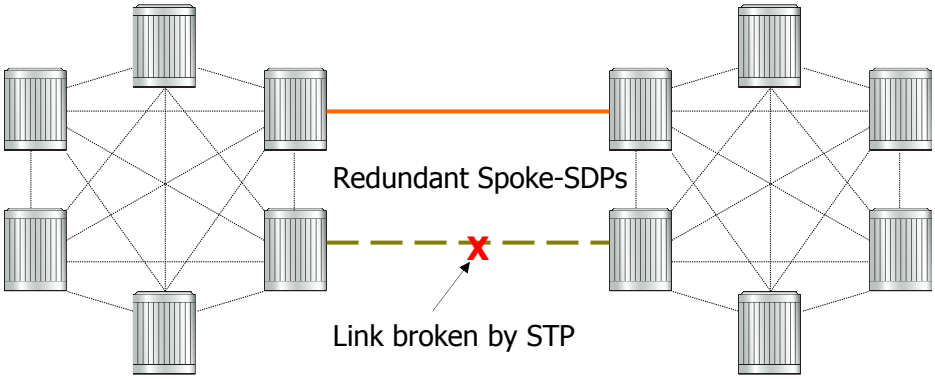
VPLS Topologies - Spoke Connectivity

Pros of using spokes to connect two large networks together

- Minimal connection requirements
- Reduced multicast replication

Cons

- A single spoke-sdp would become a single point of failure
- Two spoke-sdps requires loop protection due to the nature of spoke-sdp 'forward to everyone'



Connecting two meshed core together with a spoke SDP introduces a second problem. It creates a single point of failure. Two spoke SDP's connecting the two meshed core might be ok, but this results again in loops. To avoid loops here, Spanning Tree Protocol might be enabled to logically block one of the spoke SDP's.

Alcatel-Lucent's implementation of the Spanning Tree Protocol (STP) incorporates some modifications to make the operational characteristics of VPLS more effective.

A preferred STP variant can be configured per VPLS instance. The 7750 SR supports the following STP variants:

STP: Spanning Tree Protocol compliant with IEEE 802.1D-2004 (default mode)

RSTP: Rapid Spanning Tree Protocol compliant with IEEE 802.1w

MSTP: Multiple Spanning Tree Protocol compliant with IEEE 802.1s (this mode of operation is only supported in an mVPLS)

While the 7750 SR initially uses the mode configured for the VPLS, it will dynamically fall back to STP, IEEE 802.1D-1998, based on the detection of a BPDU of a different format. A trap or log entry is generated for every change in spanning tree variant. Some older 802.1W compliant RSTP implementations may have problems with some of the features added in the 802.1D-2004 standard. Interworking with these older systems is improved with the comp-dot1w mode.

M-VPLS

AT THE SPEED OF IDEAS

Alcatel·Lucent

Section 6: M-VPLS

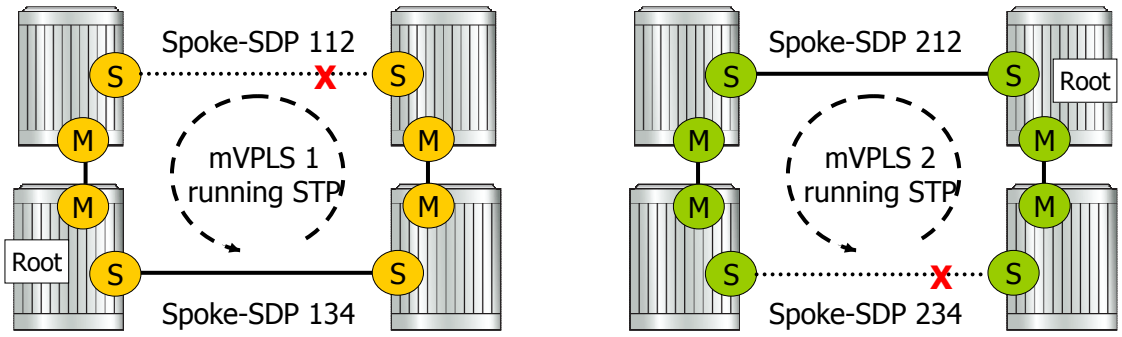
Management VPLS

To remove loops, STP can be configured within a Management VPLS.

- The mVPLS handles loop protection for all VPLS using the SDPs specified

Load sharing: Achieved in this topology with the addition of a second set of spoke-SDPs and a second mVPLS instance.

- Within the second mVPLS, priority is configured reverse to the first
  - mVPLS 1 forwards all traffic down spoke-SDP 134
  - mVPLS 2 forwards all traffic through spoke-SDP 212



To remove loops, STP can be configured within a Management VPLS.

The mVPLS handles loop protection for all VPLS using the SDPs specified. In this, only one STP instance has to run, controlling a couple of VPLS instances. This saves the processing power of BPDU's.

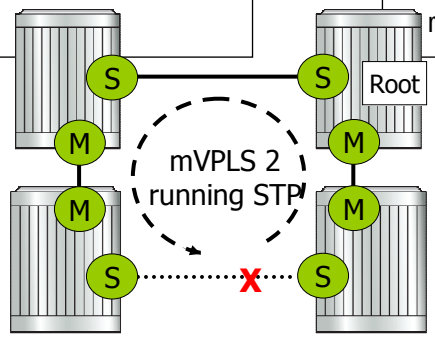
Because all the user VPLS instances prune there state from the management VPLS, blocking happens on the same point in the network for all user VPLS's. To avoid this, load sharing can be enabled.

Load sharing is achieved by creating another pair of redundant SDPs and a second mVPLS. Some of the uVPLSs can be associated with the first mVPLS and other uVPLSs can be associated with the second mVPLS. The path cost can be configured so that the first mVPLS prefers the top link and the second mVPLS prefers the bottom link.

Building a Management VPLS

```
vpls 2 customer 1 m-vpls create
  stp
    priority 32768
    no shutdown
  spoke-sdp 212:1
  mesh-sdp 300:1
```

```
vpls 2 customer 1 m-vpls create
  stp
    priority 0
    no shutdown
  spoke-sdp 221:1
  mesh-sdp 300:1
```




```
vpls 2 customer 1 m-vpls create
   stp
      priority 32768
      no shutdown
   spoke-sdp 234:1
   mesh-sdp 300:1
```

```
vpls 2 customer 1 m-vpls create
   stp
      priority 32768
      no shutdown
   spoke-sdp 243:1
   mesh-sdp 300:1
```

To configure protection for VPLS spoke-SDPs:

Define the SDPs between the redundant PEs

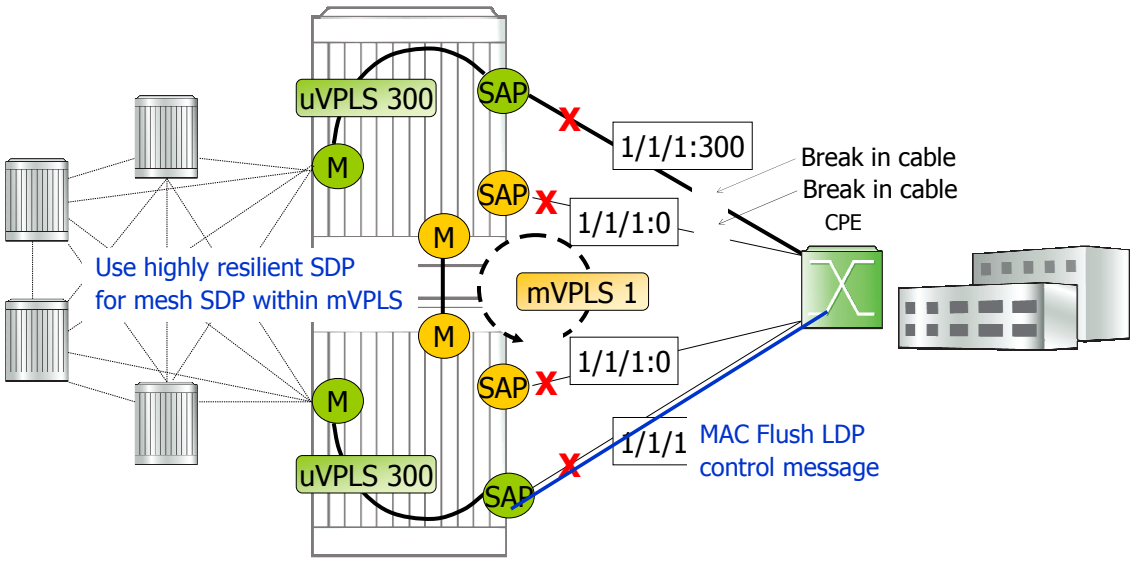
Create an mVPLS that will run the one RSTP instance to manage the redundancy. This mVPLS consists of the nodes with the redundant links.

Create the uVPLSs that will use the redundant SDPs

Management VPLS - Dual Homed Access Device

mVPLS is also used to remove loops in the network caused by the implementation of SAP redundancy

- Load balancing is achieved with two mVPLS instances, each managing multiple VLANs



mVPLS is also used to remove loops in the network caused by the implementation of SAP redundancy for a dual homed access device.

If there is a loop in the mVPLS domain, based on priority, one port will be placed into a blocking state to break the loop. If a SAP fails, the transition of traffic to the new sap is achieved when standby node broadcasts a MAC Flush LDP control message so the address of the newly active node can be relearned by all PEs in the VPLS.

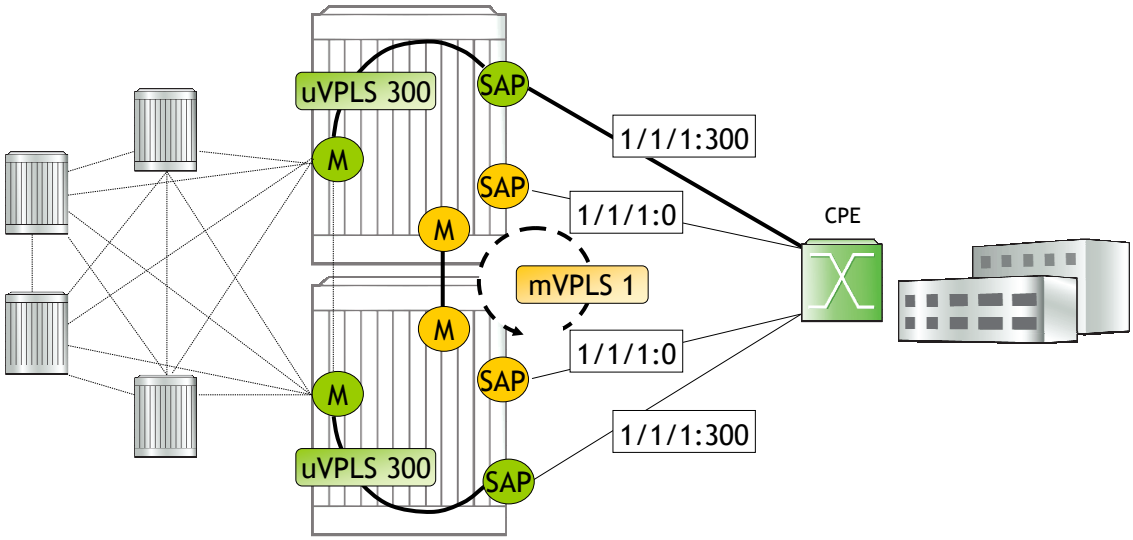
Always use a highly resilient SDP such as FRR for the mesh SDP within the mVPLS. The mesh SDP is used to close the loop and should it fail the loop is broken in the mVPLS and both SAPs will transition to a forwarding state causing a loop in the uVPLS.

Note: Like the redundant spoke scenario, the mVPLS configuration does not carry customer traffic but only STP BPDU's.

Management VPLS - Dual Homed Access Device (cont.)

```
vpls 1 customer 1 m-vpls create
  stp
    no shutdown
  sap 1/1/1:0 create
    managed-vlan-list
      range 300-600
  mesh-sdp 3612:1 create
```

```
vpls 300 customer 1 create
  stp
    shutdown
  sap 1/1/1:300 create
  mesh-sdp 12:300 create
  mesh-sdp 13:300 create
  mesh-sdp 14:300 create
```



The mVPLS is created on both PEs

The mesh SDP used in the mVPLS can be different than what is used in the uVPLS

A range of VLANs are defined under the SAP in the mVPLS

STP is enabled on the mVPLS and shutdown on the uVPLS

The uVPLS follows the state of the mVPLS. The show service id 300 stp displays the port state (1/1/1:300 Pruned/Discard), and that the decision to prune this port was managed by service 1.

Note: If the mVPLS is not created before the uVPLS, a loop may form in the uVPLS .

PBB - Provider Backbone Bridging

AT THE SPEED OF IDEAS

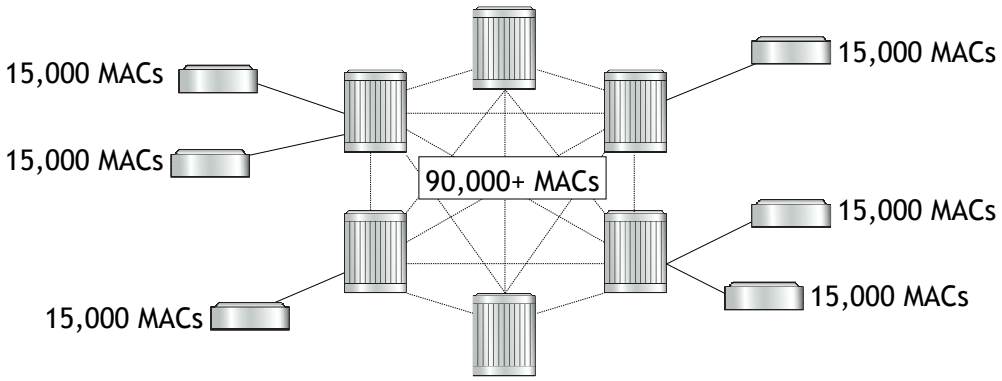
Alcatel·Lucent

Section 7: PBB

MAC Explosion in the Core

If a network is large enough delays in MAC learning and relearning after a topology change could potentially be service impacting

- The FDB of the 7750 SR can support over 100,000 MAC addresses
- If the FDB capacity is exceeded, new MAC addresses are not learned which results in unnecessary flooding
- Should a MAC flush occur it could take several seconds to relearn the MAC addresses resulting again in excessive flooding or loss of packets

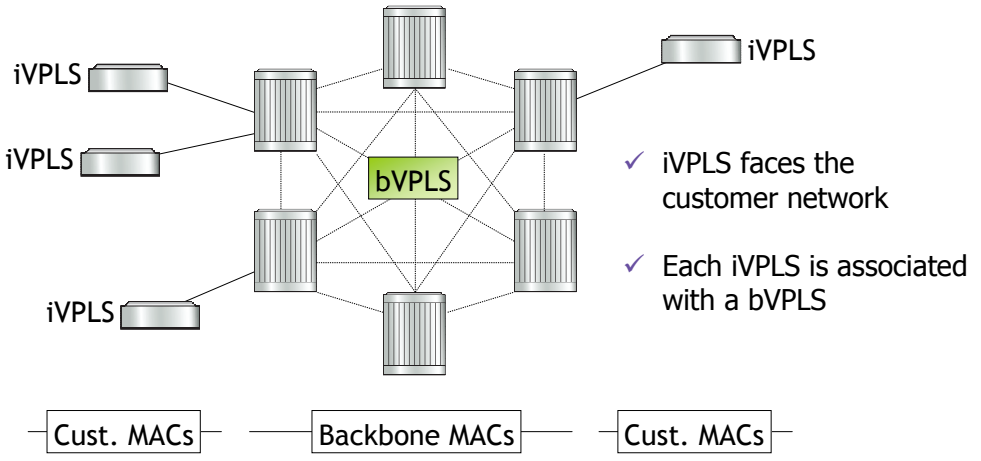


The use of spoke-SDPs in H-VPLS addresses many problems with VPLS scaling but it does not solve delays in learning and relearning large numbers of MAC addresses after a topology change. It is also possible that the number of MACs handled in the core exceed FDBs resulting in the flooding of unknown packets.

Provider Backbone Bridging (PBB)

PBB or MAC-in-MAC uses backbone MAC addresses in the core and hides the customer MAC addresses

- The core learns backbone MAC addresses
- Reduced requirements for MAC relearning



- ✓ iVPLS faces the customer network
- ✓ Each iVPLS is associated with a bVPLS

IEEE 802.1ad or Provider Backbone Bridging, is an amendment to 802.1Q or QinQ VLAN tagging. PBB was defined to resolve the issue of MAC explosion by offering separation of customer and provider domains. The customer MAC addresses are contained within the Instance of VPLS.

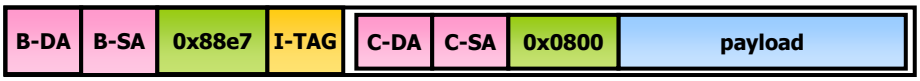
PBB

PBB defines a 48 bit B-DA and 48 bit B-SA to indicate the backbone source and destination MAC addresses.

The I-TAG contains a 12 bit B-VID (backbone VLAN ID)and 24 bit I-SID (Service Instance VLAN ID)

- I-SID distinguishes the services within a PBB domain

Addresses are learned based on the B-SA and ingress port value and hence is completely unaware of the customer MAC addresses



18 1532

- ✓ B-DA = 6 Bytes
- ✓ B-SA = 6 Bytes
- ✓ eType = 2 Bytes
- ✓ I-TAG = 4 Bytes

PBB defines a 48 bit B-DA and 48 bit B-SA to indicate the backbone source and destination MAC addresses.

The I-TAG contains a 12 bit B-VID or backbone VLAN ID and a 24 bit I-SID or Service Instance VLAN ID.

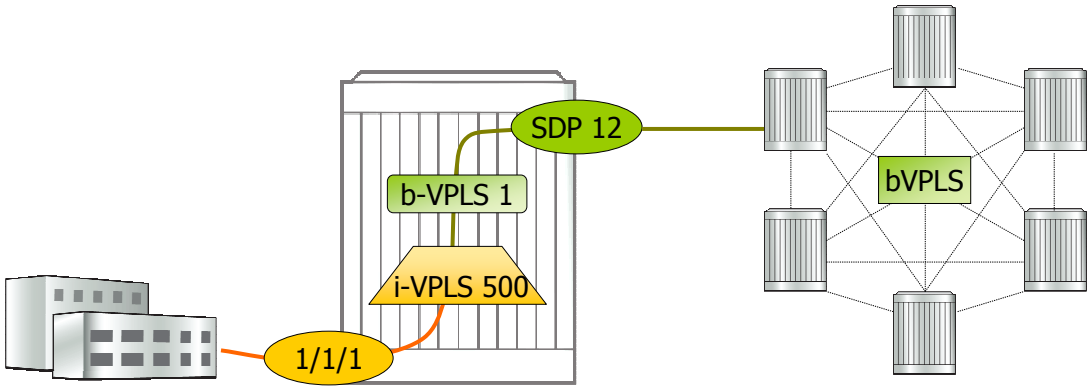
I-SID distinguishes the services within a PBB domain.

Addresses are learned based on the B-SA and ingress port value and hence is completely unaware of the customer MAC addresses.

Building a PBB

```
vpls 500 customer 1 i-vpls create
  backbone-vpls 1
  sap 1/1/1
```

```
vpls 1 customer 1 b-vpls create
  service-mtu 1532
  spoke-sdp 12:1
```



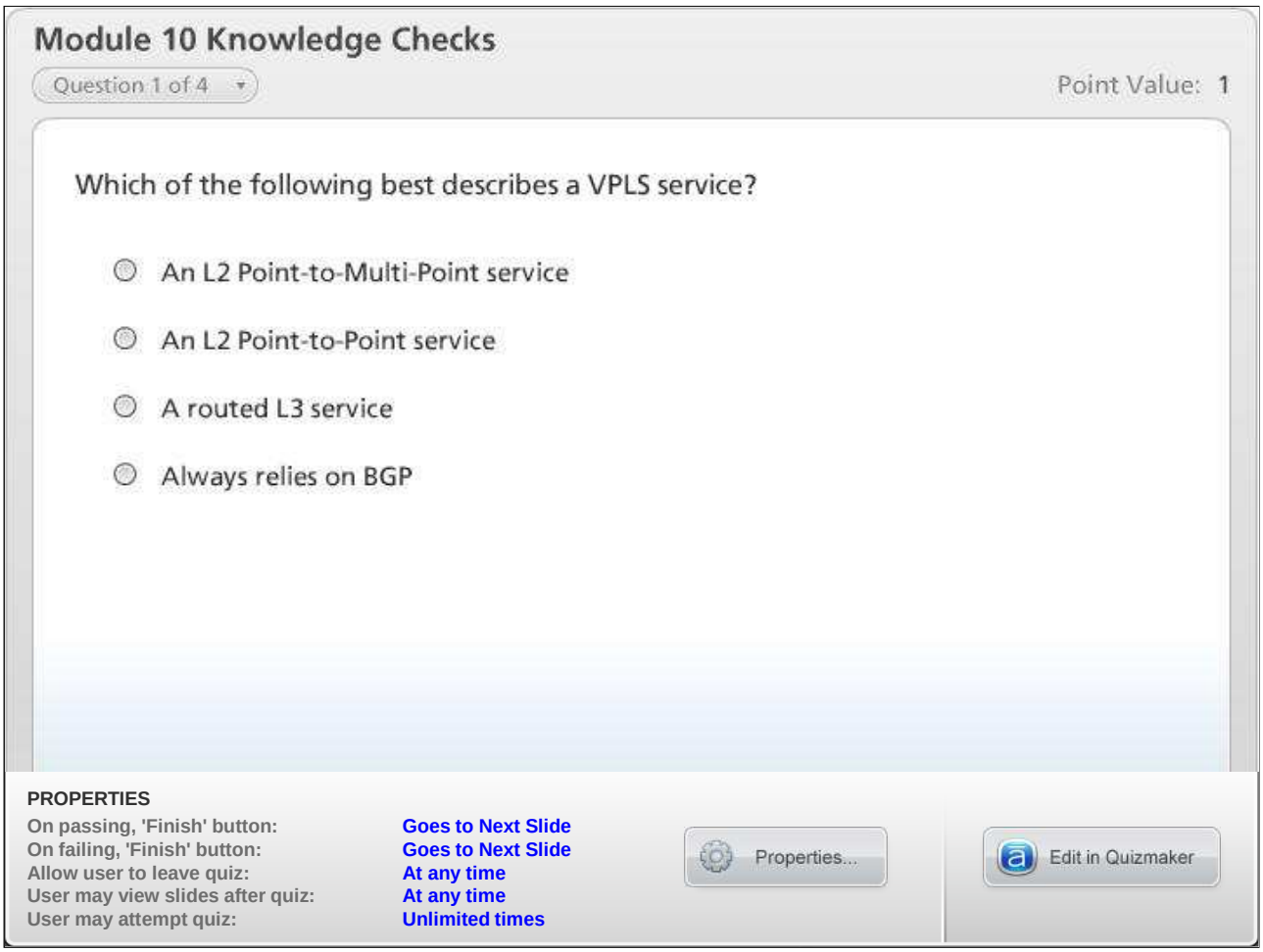
- ✓ b-VPLS supports multiple i-VPLS
- ✓ i-VPLS attaches to one b-VPLS
- ✓ b-VPLS and i-VPLS can coexist with regular VPLS instances on the same node

The above example involves traffic entering the iVPLS using SAP 1/1/1. Spoke-SDPs may also be used to bring traffic into the iVPLS. An MTU running VPLS can be plugged into the iVPLS instance via a spoke-SDP with the goal of keeping MTU complexity, cost and configuration to a minimum.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 10

Learning experience powered by Alcatel-Lucent University

Alcatel·Lucent

This completes module 10.

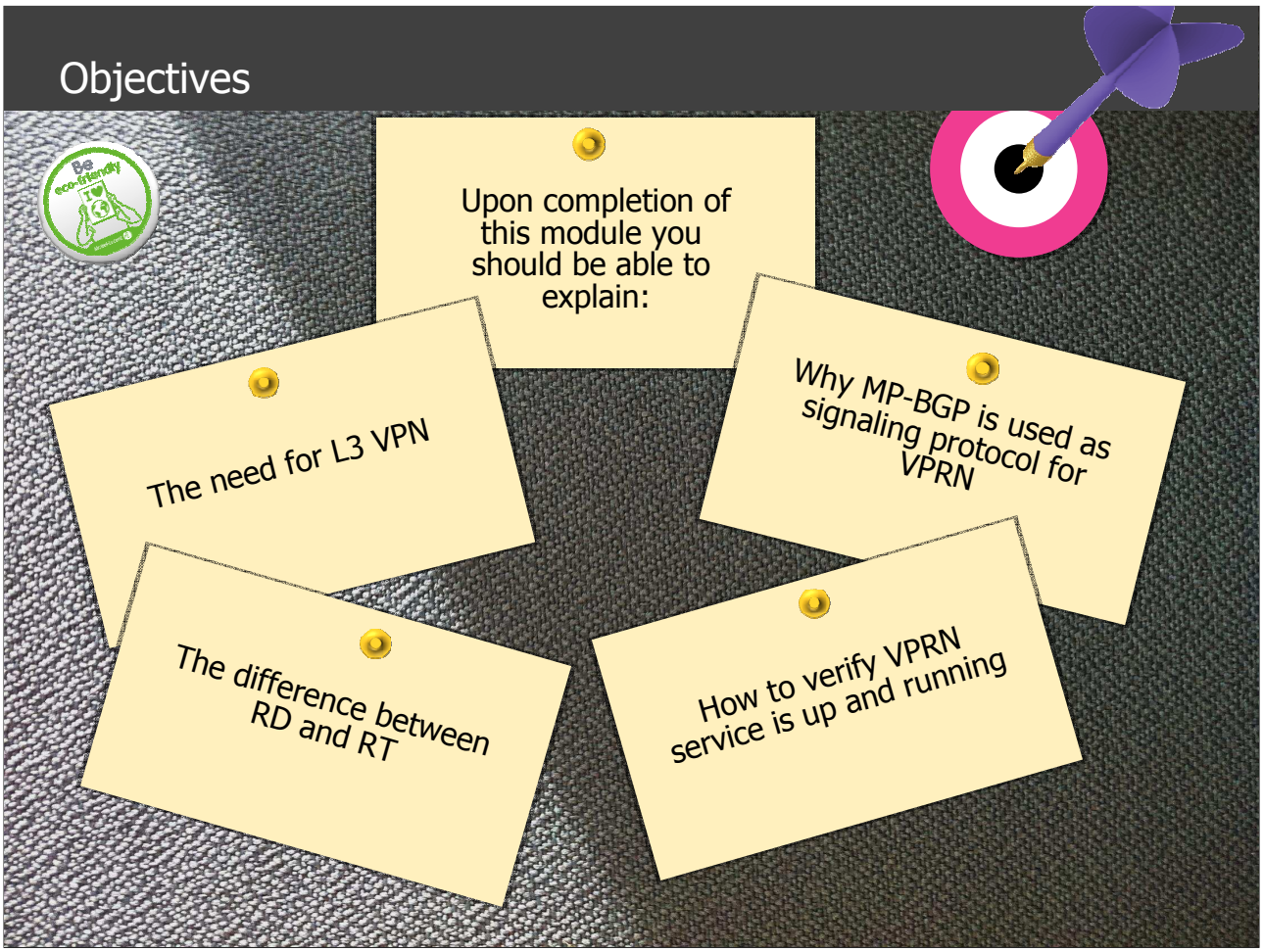
SR-OS Fundamentals

Module 11: VPRN – Virtual Private Routed Network

IPD Development

Available as PDF

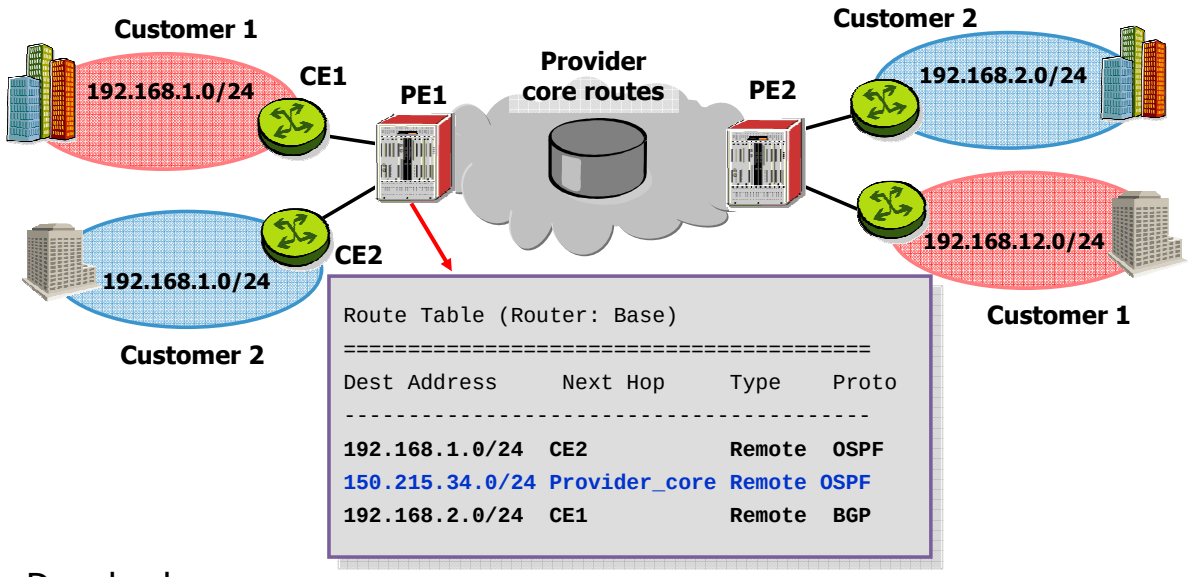
Welcome to the 11th module of the SR-OS fundamentals course.



By the end of this module you should be able to explain:

- The need for a L3 VPN
- Why MP-BGP is used as signaling protocol for VPRN
- The difference between RD and RT
- How to verify VPRN service is up and running

Need for VPRNs: Signal Routing Table



Drawbacks :

- All routes are mixed together in a single PE Routing table
- Large amount of information may be exchanged when routing changes.

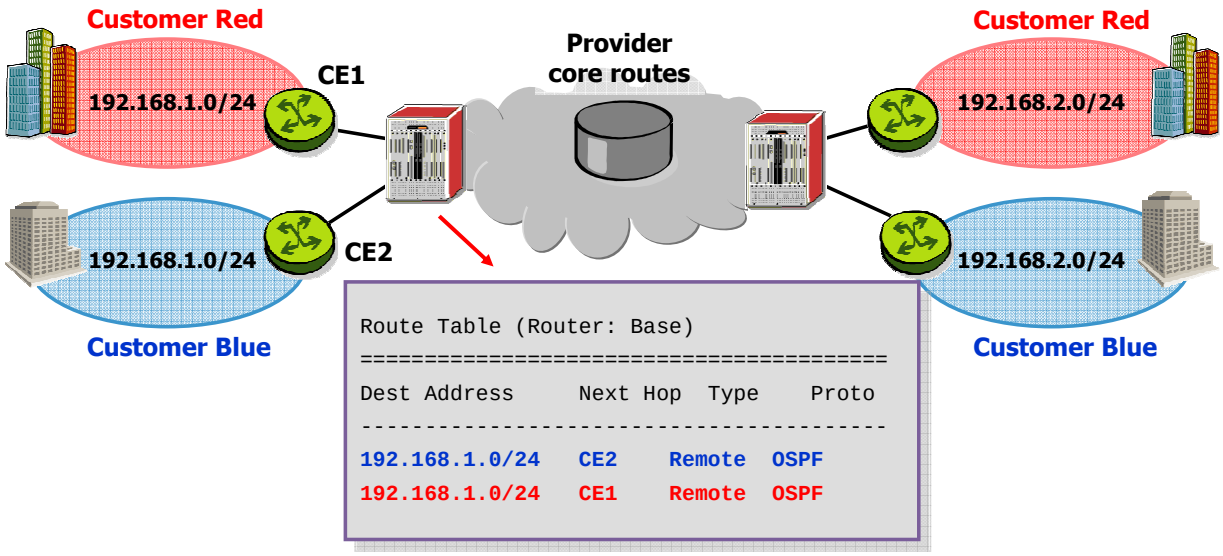
When a service operator interacts with the routing exchange of the customer the forwarding of data packets is based on IP addresses, the operator can give access via a regular IP interface and forwarding is based on the base routing table of the service router. All the routes of the different customers appear in a single routing table.

This solution sounds simple to maintain and to configure but has some major drawbacks.

All the routes are mixed together in a single PE routing table and customer to customer traffic might be possible. This is not desired and therefore the installation of a filter might be the solution, however this is time consuming and not scalable.

When all customers run the same routing protocol with the operator and are using a single routing table, a change in the routing topology could imply a large amount of customer information exchanged.

Need for VPRNs: Overlapping IP addresses



Drawbacks :

- Only the best route is selected in routing table or ECMP (Equal Cost Multipath) routing
- Communication between customers possible
- Routing problems

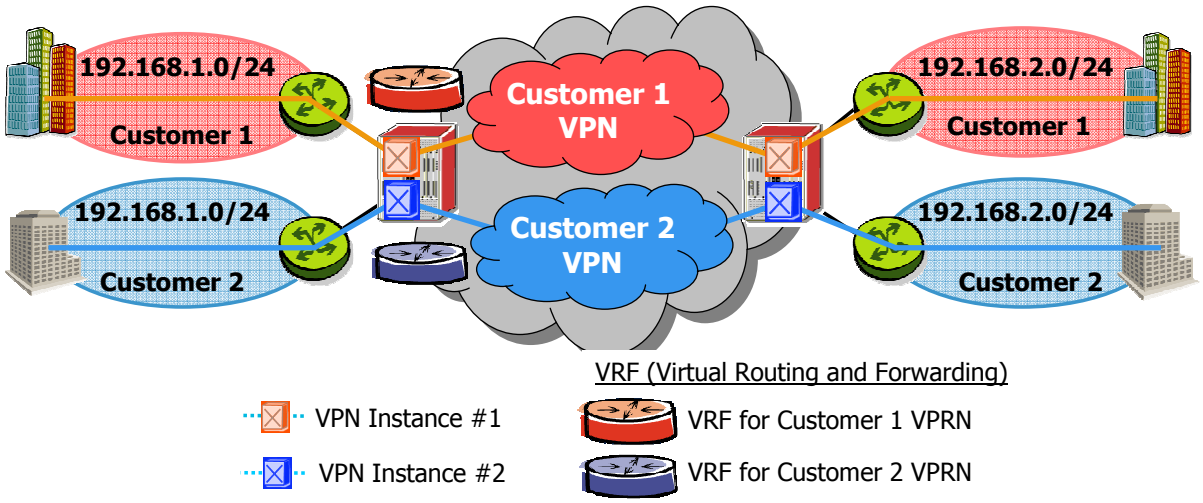
When each customer uses the same IP address range, for example IP addresses out of the private range, a single routing table will only install one of the subnets when no ECMP is enabled. Although this is technically possible, this results in serious routing problems and customer to customer communications is still possible.

ECMP or Equal Cost Multipaths allows the load sharing between two of the same routes with the same metric. When ECMP is enabled, the two subnets appear in a single routing table and traffic is load shared. This feature is not suited for this type of application as customer traffic might not end up with the right customer.

L3VPN: Introduction

L3 VPN = Alcatel-Lucent VPRN (Virtual Private Routed Network)
(RFC 4364 [Formerly RFC 2547bis])
Goals :

- Isolation of different VPN traffic
- Connectivity between customer sites
- Use of private address spaces in each site



A solid proven VPN solution where no customer to customer communication and overlapping IP addresses can be used is a L3 VPN.

A VPN service where the service operator forwards packets based on IP routing is called a L3 VPN or Alcatel-Lucent's name for a L3 VPN is called a VPRN or Virtual Private Routed Network.

Different routing tables are available per VPN instance. These routing tables are called VRF's, Virtual Routing and Forwarding instances. Traffic between VRF's of different customers is impossible, only traffic between the VRF's belonging to the same VPN instance or customer is possible.

L3 VPNs: Benefits

From a customer's perspective:

- Customer can choose their IP addressing scheme
  - Private addresses, overlapping addresses, etc.
- Transparency: customer is "unaware of VPN"
- Security: data separation
- QoS as defined in an SLA
- Different sites may use different access technologies

From a provider's perspective:

- Scalability:
  - Backbone network (signaling, state, etc.)
  - # VPNs/backbone, # sites/VPN, # routes/VPN, etc.
- Ease of provisioning: addition of a site to a VPN, creation of a new VPN, merging of VPNs, etc.

What are the benefits of a L3 VPN from the customer and provider's perspective?

Because of the private nature of a VPN, the customer can decide which IP addresses to use. It doesn't matter if they are private and/or overlapping.

Moreover, the customer is unaware of the VPN. The VPN starts at the PE and is fully managed by the service operator. An IP interface is the basic connection between the CE and PE, but the PE this interface is hooked up to a VPN and assures complete data separation.

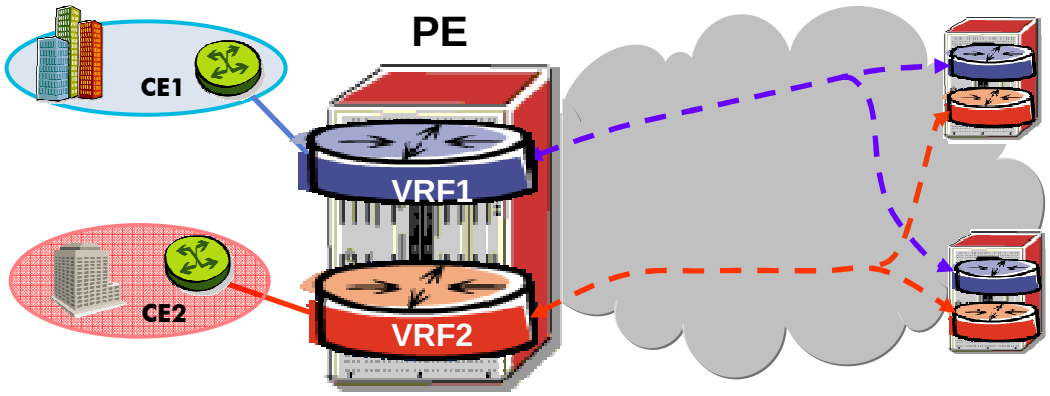
Each CE can have its own quality of service setting to meet his service level agreement.

And each customer might use its own access technology. For example, one customer uses IP over Ethernet and another customer uses IP over ATM.

The simplicity and scalability on the customer side is also found on the provider's side. The same and available MPLS or GRE service infrastructure can be used for VPRN services. The scalability and performance of Alcatel-Lucent service router results in a high number of VPN's per backbone, number of sites per VPN and a high number of routes per VPN.

Adding or merging new VPN VPRN site is easy to provision.

L3 VPN: Virtual Routing and Forwarding instances on the PE



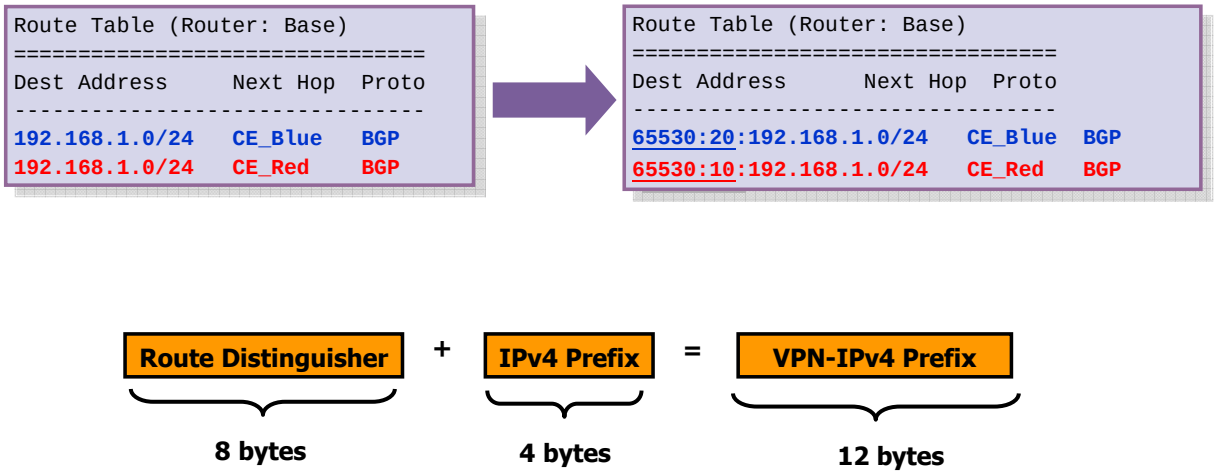
- Each PE router maintains a separate logical routing table for each VPRN
- This table is referred to as a Virtual Routing and Forwarding Table (VRF)
- Contains customer destination routes
  - local sites
  - remote sites.
- MP-BGP is used to carry the VPN routes

For each VPRN created on a service router, a VRF is created. This VRF or Virtual Routing and Forwarding Table within a PE device maintains the forwarding information on a per site basis. It contains customer destination routes from locally connected CE devices and routes from remotely connected CE devices.

The routes advertisements between PE service routers are handled by MP-BGP or Multi Protocol BGP. This is an extension to the well known BGP protocol.

L3VPN: Route Distinguisher

An identifier called the Route Distinguisher (RD) is added to all IPv4 prefixes

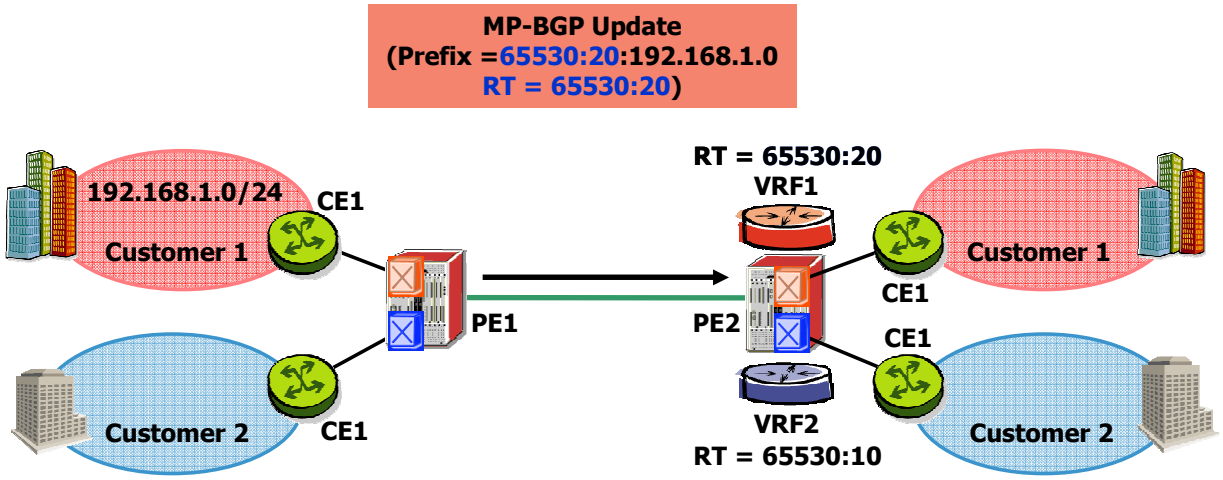


Because all VPN routes from different customers are carried in the same MP-BGP routing protocol, overlapping IP addresses would cause routing problems. BGP was initially not designed to carry the same addresses from different customers. To solve this, the address is extended with a unique number, the Route Distinguisher or RD. Each VPN has a specific RD to make the VPN routes unique.

This is achieved by pre-pending the 4-byte IPv4 address with an 8-byte Route Distinguisher to form a new address called the "VPN-IPv4 address".

L3VPN: Which routes go to which VRF?

- Mark a route by attaching a MP-BGP extended community attribute: Route Target (RT)
- Import/Export Route filtering based on the RT
- In many cases, the RT value chosen will be the same as the RD



Now that we know how to make IP addresses unique, the question is which routes go to which VRF?

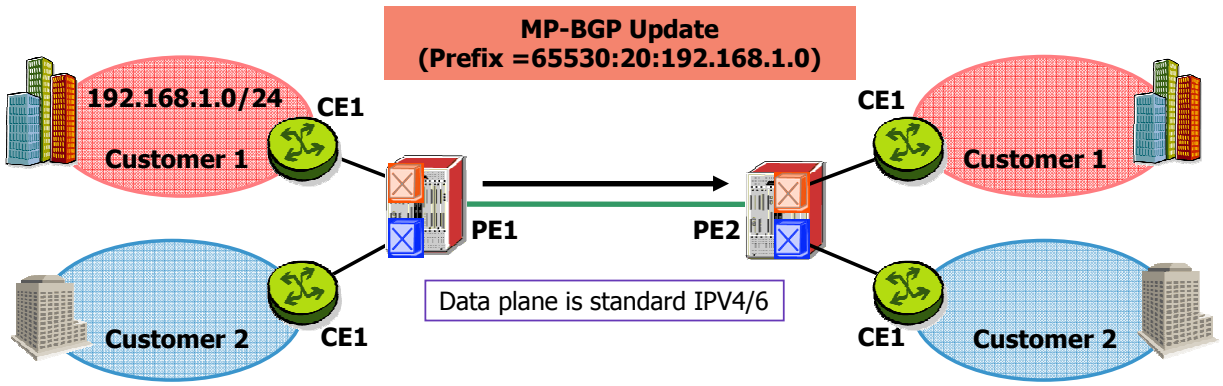
A way is needed to determine to which VRF a route belongs. A Route Target was defined to address this issue. The Route Target or RT is the closest approximation to a VPN membership identifier in the VPRN architecture, and identifies to the receiving PE the VRF table that a prefix is associated with.

Route Target is a MP-BGP extended community. One or more MP-BGP community attributes may be associated to any route, therefore one or more Route Targets may be associated to any route.

In simple VPN cases and for provisioning consistency, the Route Target value chosen is often the same as the Route Distinguisher value, however they should not be interpreted as meaning the same thing.

L3VPN: MP-BGP (Extended Community Attribute)

- MultiProtocol BGP (MP-BGP) : Allows BGP to carry VPN-IPv4 prefixes
- VPRN routes distributed in provider core network as 12 byte VPN-IPV4 routes
- VPN-IPv4 addresses:
- Only used in provider core control plane
- Used to exchange MP-BGP routing updates between PEs
- Only needs to be known by PE routers in the provider's network that are actually involved in exchanging routing information for VPN destinations



BGP is used with Multiprotocol BGP extensions to distribute VPRN routing information across the service provider's network.

In order to carry VPN-IPv4 prefixes in BGP, support for the additional address family must be enabled. Separate address families are viewed as different protocols in BGP, therefore, multi-protocol extensions to the BGP protocol are required. When BGP is configured or enabled in this fashion, it is referred to as Multiprotocol BGP or MP-BGP.

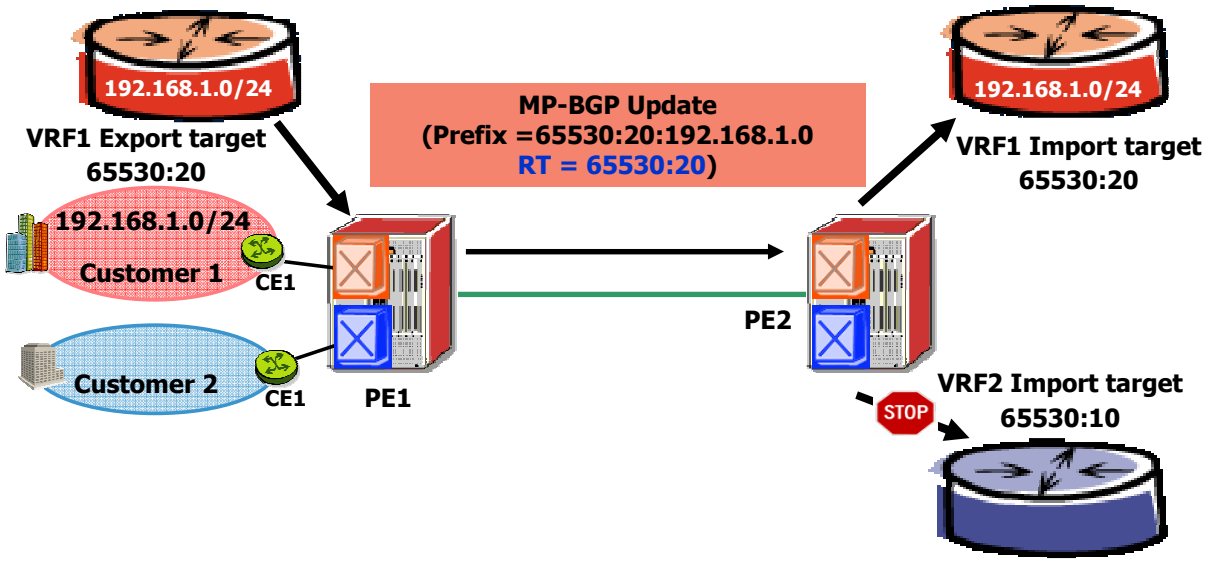
The Multiprotocol nature of MP-BGP allows the overlapping routing information to be transported across the provider core as VPN-IPv4 addresses. VPRN routes are not distributed within the provider core network as IPv4 routes, but as 12 byte VPN-IPv4 routes.

It is important to note that the VPN-IPv4 address family is used only in the provider core control plane when exchanging MP-BGP routing updates between PE's. The data plane remains as standard IPv4. All data traffic is carried in standard IPv4 packets.

The customer is unaware of the existence of VPN-IP addresses. The translation between customer IP routes in a particular VPN and VPN-IP routes distributed between provider routers is performed by the PE routers.

L3VPN: Export/Import Policy

- Export policies define the route target/s added to routes advertised to remote PEs
- Import policies (*Remote PEs*) are used to decide if route should be added to VRF
- Route isolation between VRFs is accomplished through careful administration



Export policies define the route target to be added to routes advertised to remote PE's. Upon receipt of a VPN-IPv4 route a PE router uses import policies to decide whether to add that route to a VRF.

Route isolation between VRFs is accomplished through careful policy administration. An administrator determines the appropriate export and import target relationships.

Since RTs are applied at the time the route is exported, they are called export-RT. In the above example, a single RT is attached to each route.

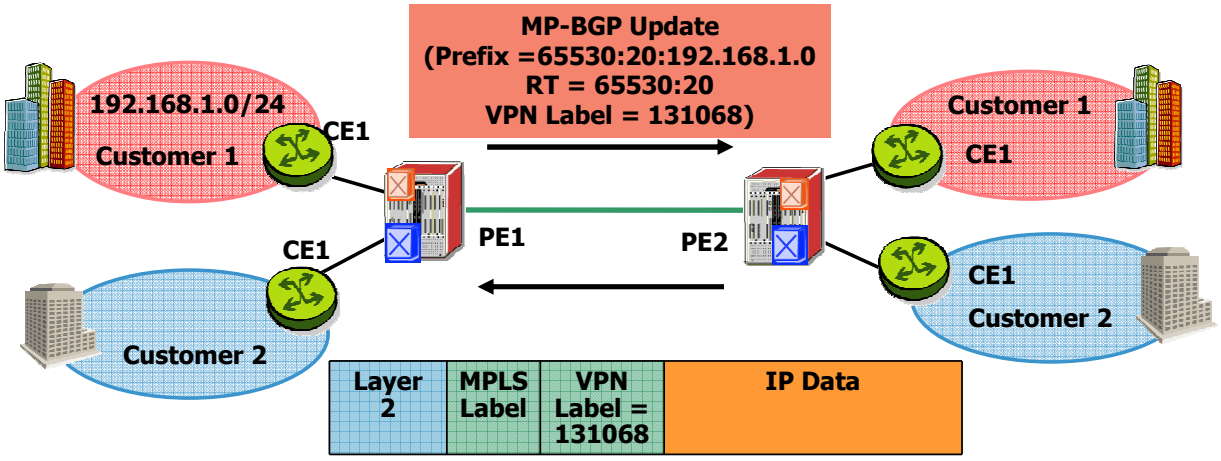
The decision to import the route into the VRF is done by matching the received routes against locally defined per-VRF import policies expressed in terms of RTs (import-RT). If the Router import RT matches the update RT it is stripped of the RD and imported into the VRF.

L3VPN: VPN Label

For data traffic at an egress PE, which VRF must be used to resolve the destination address?

- Associate a label (the VPN label) with a VPN route at the ingress PE
- Demux the traffic based on VPN label at egress PE

Distribute the VPN label with VPN Route information (MP-BGP)



The purpose of the VPN label is to de-multiplex VPN data traffic arriving at the PE.

The distribution of the VPN label is done using BGP along with the VPN route information.

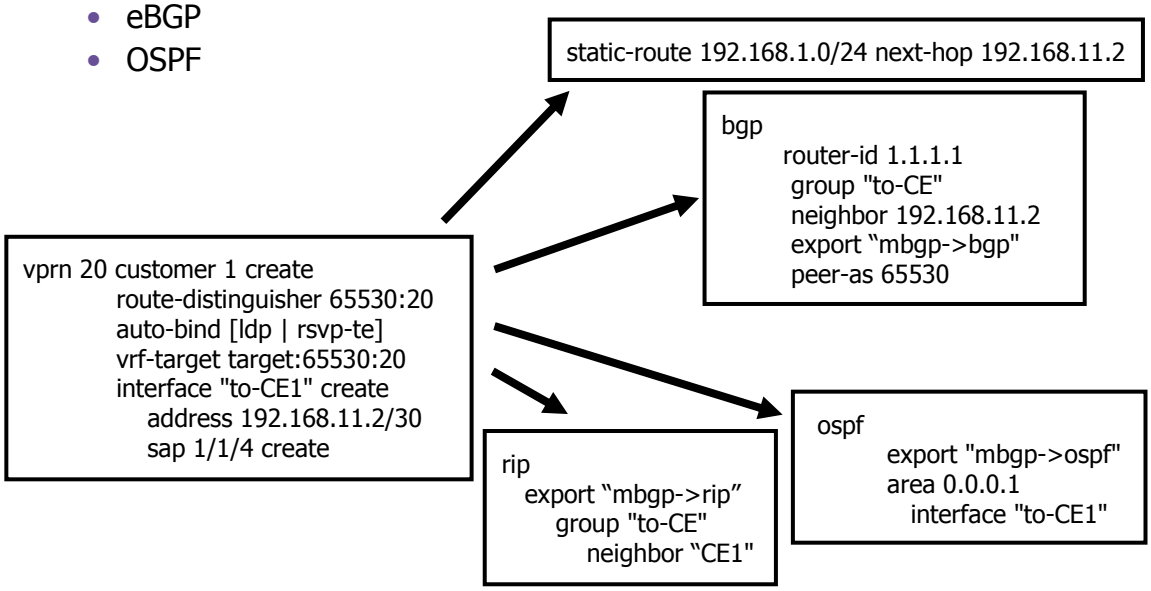
The distribution of the VPN tunnel information is automatic and does not require manual intervention.

The forwarding at the egress PE is based on an MPLS lookup on the VPN label to determine the appropriate VRF followed by an IP lookup in that VRF.

L3VPN: PE to DE route distribution

The following route distribution options are available on the 7750 SR:

- Static routes
- RIP
- eBGP
- OSPF



While MP-BGP is used between PE's, other routing protocols might be used between the CE and PE.

Static routes, RIP, eBGP and OSPF are the route distribution options available between the PE and the CE.

A static route might be configured inside the VPRN to point to a route on the CE network.

External BGP is used when the routing protocol between the CE and PE is BGP.

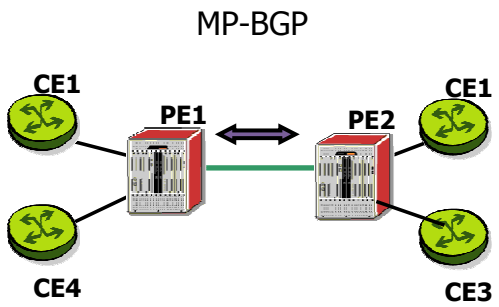
If no import route policy is specified, then all BGP routes advertised by the CE are accepted by the PE.

An export policy is needed for the PE to advertise the routes learned from other PE sites in the VPRN instance via MP-BGP to the CE router via eBGP.

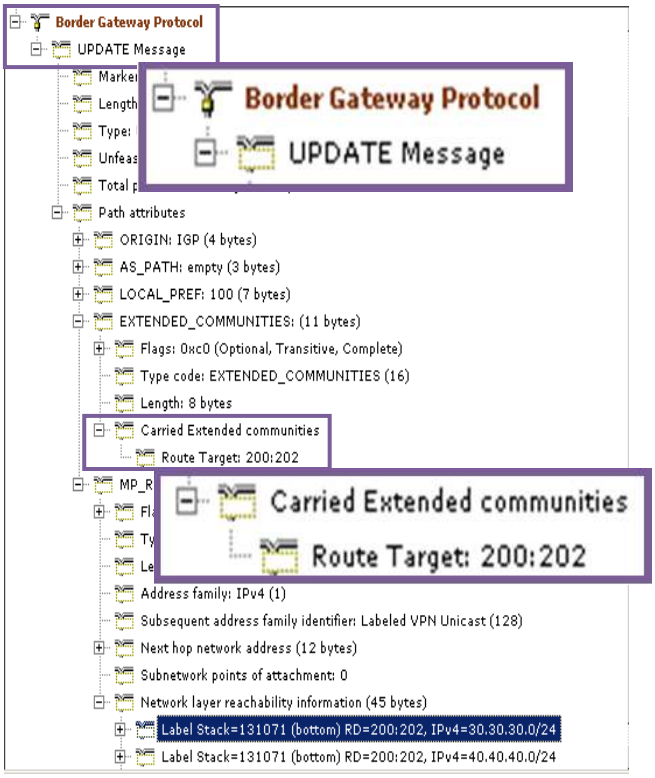
OSPF
OSPF LSA information is not transmitted natively across the IP-VPN. The OSPF routes are "imported" into MP-BGP as AS externals. As a result, other OSPF-attached VPRN sites on remote PEs will receive these via type 5 LSAs. This process is not automatic and requires the configuration of Route Policies.

RIP
By default RIP does not export routes it has learned to its neighbors. Therefore it is necessary to configure an export policy to enable MP-BGP routes learned from remote CEs belonging to the VPN to be redistributed into RIP.

L3VPN: PE to PE route distribution



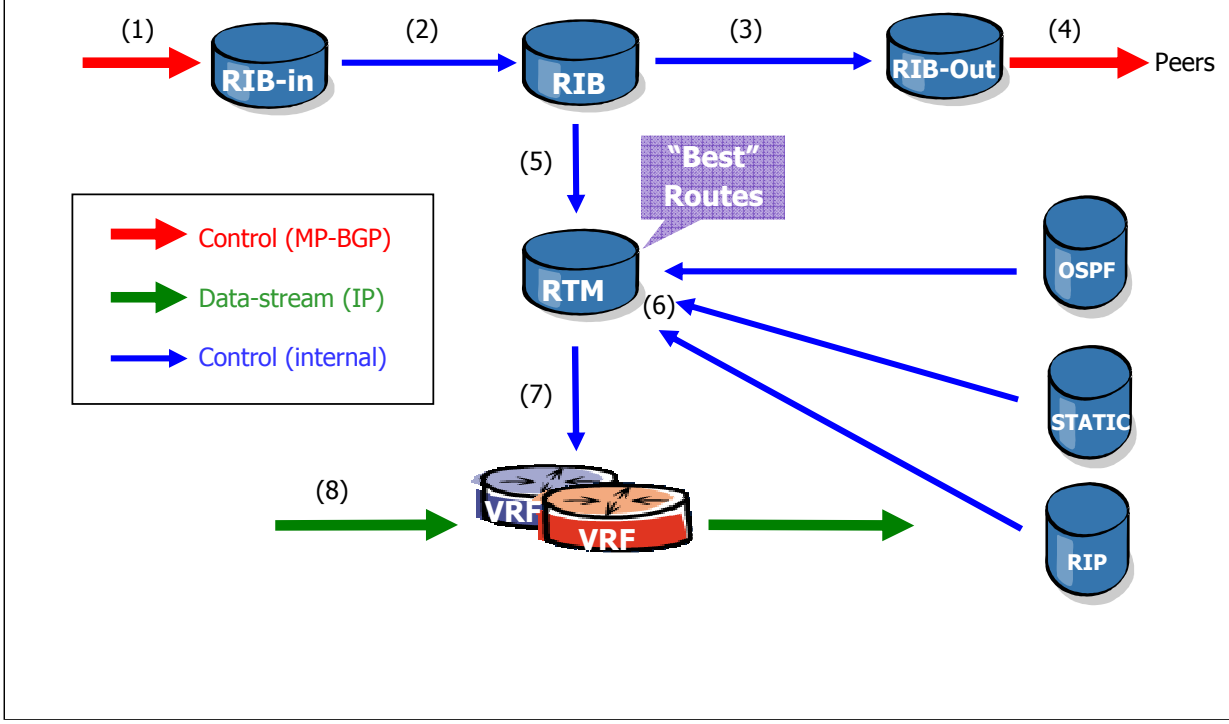

```
bgp
        family ipv4 vpn-ipv4
        group "bgp-core"
            peer-as 65530
            neighbor 2.2.2.2
```



When setting up BGP as MP-BPG between PE routers for VPRN, the family VPN-ipv4 needs to be supported in addition.

From this moment, MP-BGP can carry in his update message the extended community route target attribute, the RD and the service label. These are the three minimum parameters that MP-BPG must carry in a control update message.

L3VPN: Populating a VRF



A BGP process receives BGP updates and are collected in the RIB-IN or ingress Routing Information Base.

After applying import policies, if configured, the routes are sent to the RIB.

If routes are considered as being "best" and if they have to be advertised to one or more peers they are sent to the RIB-OUT, and export policies are applied.

BGP routes are sent out to the peers who need to receive the routes.

The RIB does the BGP selection process, and sends the "best" routes to the RTM or Route Traffic Manager.

The RTM collects all "best" routes from all protocols, and selects the best based on the preference.

The RTM sends all "best" routes to the corresponding VRF. The VRF's get downloaded on all forwarding complexes of the IOM.

IP traffic coming in will do a longest prefix match on the VRF to make a forwarding decision.

L3VPN: Monitoring a VRF

Verify the VPRN routing table of the PE

```
PE1# show router 20 route-table
===============================================================================
=
Route Table (Service: 20)
===============================================================================
=
Dest Prefix          Type    Proto    Age         Pref Metric
       Next Hop[Interface Name]
-------------------------------------------------------------------------------
-
192.168.2.0/24       Remote  BGP VPN  01d01h48m   170    0
       2.2.2.2
192.168.1.0/24       Remote  OSPF     01d01h48m   150    1
       192.168.11.1
192.168.11.0/29      Local   Local    01d01h48m   0  0
       to-CE1
-------------------------------------------------------------------------------
-
No. of Routes: 3
===============================================================================
=
PE1#
```

The show router <service-id> route-table command, in this case PE1, is used to verify the contents of the routing instance for the specified service. The service-id variable is actually the VRF id.

As shown here, this table should contain customer routes such as physical links to the PE devices, system interfaces of the CE devices and internal customer LAN networks.

The prefixes received from the local CE will be learned via the configured PE-CE protocol, in this case OSPF, and the prefixes received from the remote CE will be learned via the BGP VPN protocol.

Notice that in the VRF, the RD and RT have been removed before the prefix is imported into the VRF table.

L3VPN: Monitoring MP-BGP learned routes (RIB-In)

To find out the VPN label and the Route Target:

- No command like :
  - #show router bgp rib-in
- Two possible commands
  - # show router bgp neighbor 1.1.1.1 received routes
  - # show router bgp routes 65530:20:192.168.1.0/24 detail

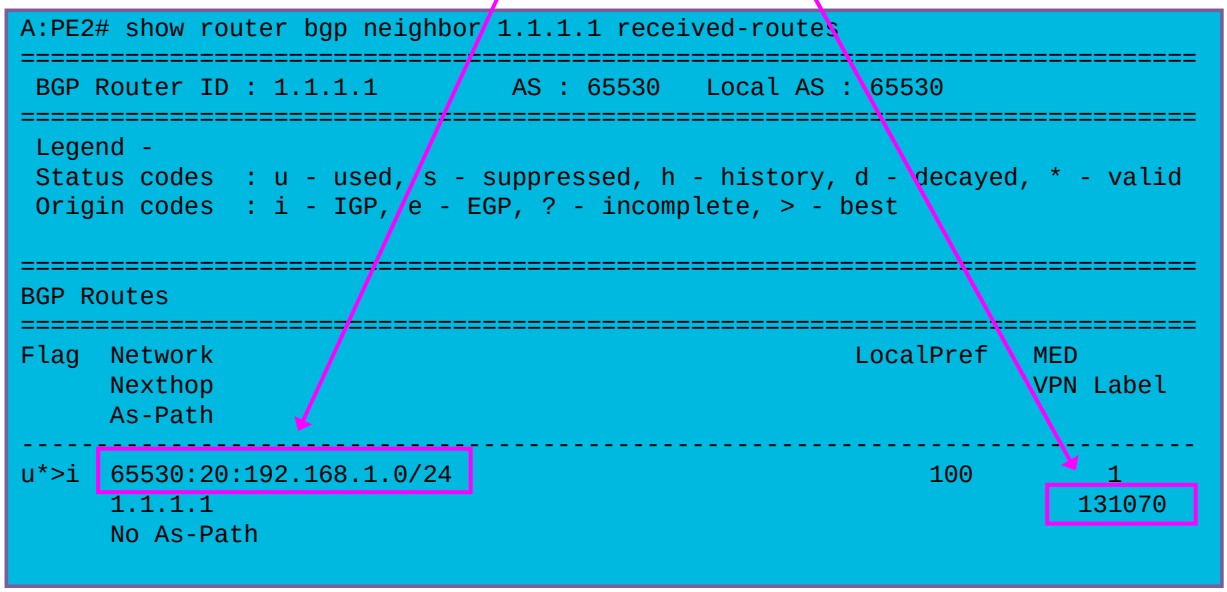
```
A:PE2# show router bgp routes 65530:20:192.168.1.0/24 detail
===============================================================
BGP Router ID : 2.2.2.2        AS : 65530   Local AS : 65530
===============================================================
Status codes  : u - used, s - suppressed, h - history, d - decayed,
* - valid  Origin codes  : i - IGP, e - EGP, ? - incomplete, > - best
===============================================================
BGP Routes
Network        : 192.168.1.0/24
Nexthop        : 1.1.1.1
Route Dist.    : 65530:20              VPN Label      : 131070
From           : 1.1.1.1
Res. Nexthop   : 192.168.12.1
Local Pref.    : 100                   Interface Name : pe2-1
Aggregator AS  : none                  Aggregator     : none
Atomic Aggr.   : Not Atomic            MED            : 1
Community      : target:65530:20
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 1.1.1.1
Flags          : Used  Valid  Best  IGP
AS-Path        : No As-Path
```

show router bgp routes 65530:20:192.168.1.0/24 details

This command shows all routes with this prefix that are received. Per route, there is an output on the original and modified path attributes. The original attributes display the route, as it was received from its peer, before applying the import policies. If no import policies are configured, the original and modified attributes are equal.

L3VPN: Monitoring MP-BGP learned routes (RIB-In)

BGP updates carry the Route Distinguisher and VPN Label, identifying the remote VRF



```
A:PE2# show router bgp neighbor 1.1.1.1 received-routes
===============================================================================
 BGP Router ID : 1.1.1.1        AS : 65530   Local AS : 65530
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
 Origin codes  : i - IGP, e - EGP, ? - incomplete, > - best

===============================================================================
BGP Routes
===============================================================================
Flag  Network                                          LocalPref   MED
      Nexthop                                                      VPN Label
      As-Path
-------------------------------------------------------------------------------
u*>i  65530:20:192.168.1.0/24                              100         1
      1.1.1.1                                                       131070
      No As-Path
```

show router bgp neighbor < . . . > received-routes

This command shows all routes that are received from this BGP-peer. There is no "detail" option available for this, so for more detailed information, a "per-route" display needs to be done.

L3VPN: Monitoring MP-BGP advertised routes (RIB-In)

No command like :

#show router bgp neighbor"

```
A:PE2# show router bgp neighbor 1.1.1.1 advertised-routes
===============================================================================
 BGP Router ID : 2.2.2.2         AS : 65530   Local AS : 65530
===============================================================================
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
 Origin codes  : i - IGP, e - EGP, ? - incomplete, > - best
===============================================================================
BGP Routes
===============================================================================
Flag  Network                                              LocalPref   MED
      Nexthop                                                          VPN Label
      As-Path
-------------------------------------------------------------------------------
i     65530:20:192.168.2.0/24                                  100         1
      10.1.1.22                                                        131057
      No As-Path
```

show router bgp neighbor < . . . > advertised-routes

This command shows all routes that are advertised to this BGP-peer. There's no "detail" option available for this, so for more detailed information, a "per-route" display needs to be done.

L3VPN: Monitoring MP-BGP advertised routes (RIB-In)

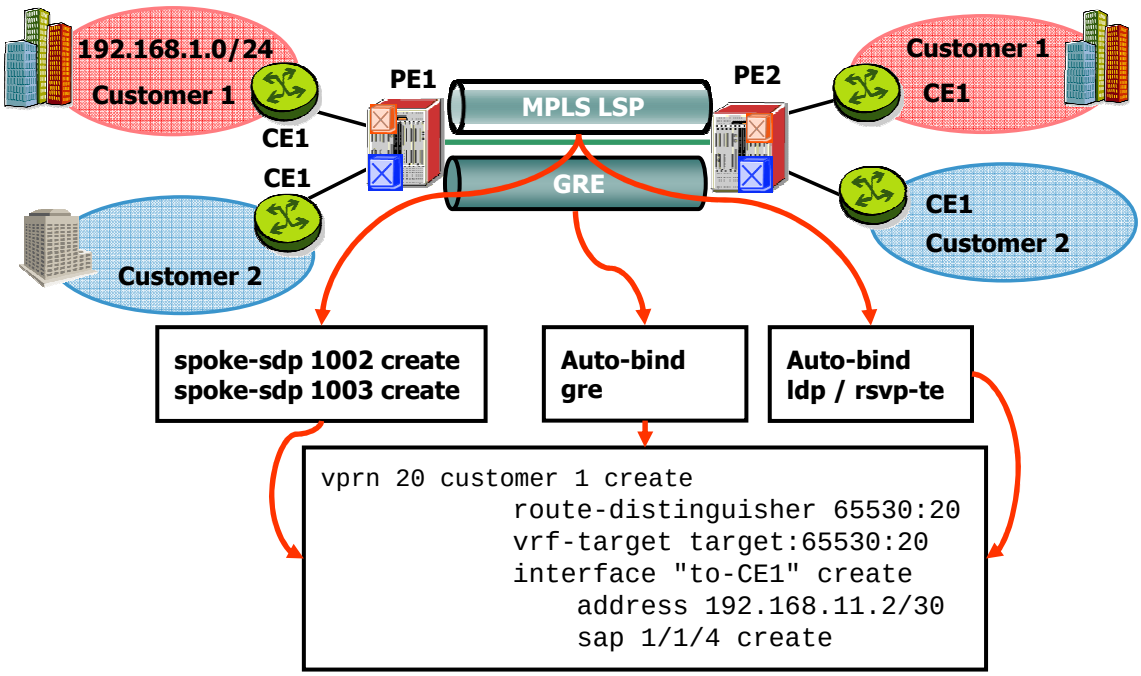
BGP updates carry the VPN label and the RT which identifies the remote VRF

```
PE1# show router bgp routes 200:20:192.168.1.0/24 hunt
===============================================================================
 BGP Router ID : 1.1.1.1        AS : 65530   Local AS : 65530
===============================================================================
 Legend -Status codes  : u - used, s - suppressed, h - history, d - decayed, * -
valid Origin codes  : i - IGP, e - EGP, ? - incomplete, > - best
===============================================================================
BGP Routes – RIB Out
===============================================================================
Network        : 192.168.1.0/24
Nexthop        : 2.2.2.2
Route Dist.    : 65530:20              VPN Label      : 131057
To             : 2.2.2.2
Res. Nexthop   : n/a
Local Pref.    : 100                   Interface Name : NotAvailable
Aggregator AS  : none                  Aggregator     : none
Atomic Aggr.   : Not Atomic            MED            : 1
Community      : target:65530:20
Cluster        : No Cluster Members
Originator Id  : None                  Peer Router Id : 2.2.2.2
Origin         : IGP
AS-Path        : No As-Path
===============================================================================
PE1#
```

show router bgp routes 200:20:192.168.1.0/24 hunt

This command shows all peers that have received this route. There is an output of the route-per-peer, after applying potential export-policies. Since export policies can be applied on a per-peer basis, the routes can have a different content, that's why they are displayed per-peer.

L3VPN: Data Plane - Transport Tunnel



After the establishment of the routing topology in the provider core, the network must be MPLS or GRE enabled to support services such as VPRN.

A full mesh of transport tunnels must be created between the PE's. The transport tunnel is either an MPLS LSP or a GRE point-to-point tunnel between PE's.

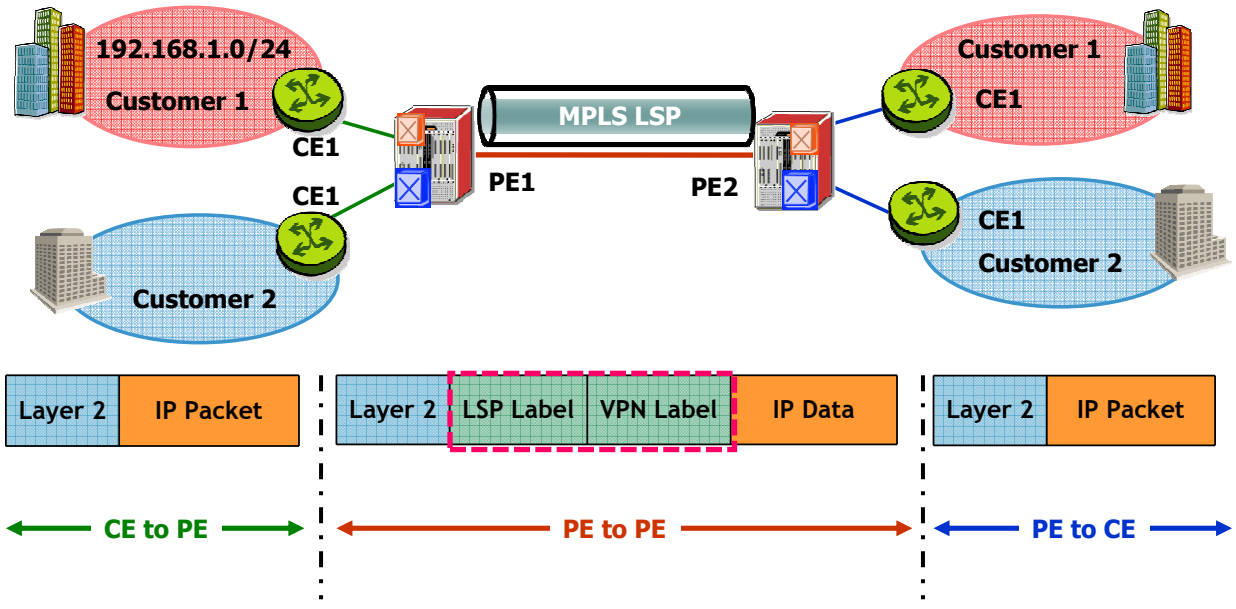
The tunnels serve as the label switched paths the customer packets will take as they cross the provider core network.

Each PE involved in a given VPRN service must be configured with a tunnel to every other PE participating in the same VPRN service in order to transport a customer's VPN traffic from one site to another.

The tunnel is created either through the configuration of a SDP or using the auto-bind option when creating a VPRN service instance.

If SDP tunnels are used, they must be created prior to the creation of the VPRN services.

L3VPN: Forwarding Traffic (1)



MPLS Labels:

LSP (Label Switched Path)- Used to propagate packet across provider core (PE-to-PE)

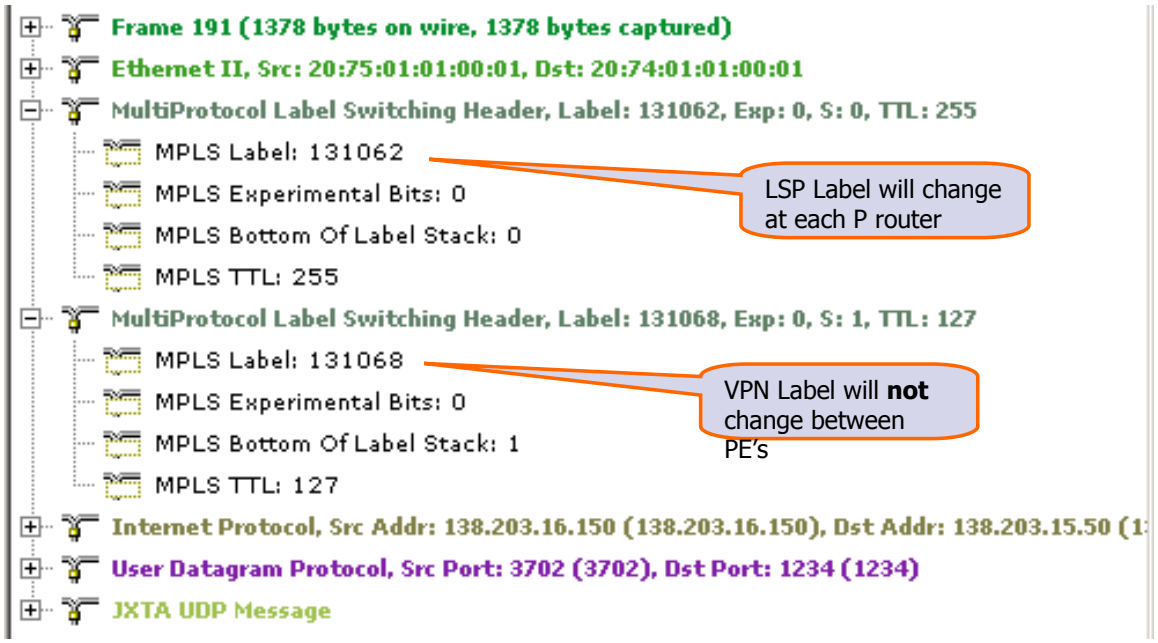
VPN (Virtual Private Network) - Identifies VPRN to egress PE

Shown here are the data packet formats visible in an MPLS network. Note that the packet is a standard IP packet to which MPLS labels have been inserted in the provider core. The Route Distinguisher and the Route Target are not used in the data plane.

Between the PE and the CE, only unlabeled data packets should be seen.

From PE to PE in the provider core, customer VPRN traffic will have a label stack consisting of 2 MPLS labels. The outer label is used to propagate the packet from PE to PE across the provider core. The inner label identifies the VPRN to the egress PE.

L3VPN: Forwarding Traffic (2) - Screenshot Analyzer SW



Shown here is the MPLS label stack with the two MPLS labels. The top label is the transport label that is changed or swapped at each P router. The second label in the stack is the VPN or service label and is only popped at the far-end service PE.

L3VPN: Debug command

"Debug router bgp update" results in…

```
7 2008/01/17 19:16:37.38 UTC MINOR: DEBUG #2001 - Peer 1: 1.1.1.1
"Peer 1: 1.1.1.1: UPDATE
Peer 1: 1.1.1.1 - Send BGP UPDATE:
    Withdrawn Length = 0
    Total Path Attr Length = 106
    Flag: 0x40 Type: 1 Len: 1 Origin: 0
    Flag: 0x40 Type: 2 Len: 0 AS Path:
    Flag: 0x80 Type: 4 Len: 4 MED: 1
    Flag: 0x40 Type: 5 Len: 4 Local Preference: 100
    Flag: 0xc0 Type: 16 Len: 8 Extended Community:
        target:65530:20
    Flag: 0x90 Type: 14 Len: 70 Multiprotocol Reachable NLRI:
        Address Family VPN-IPV4
        NextHop len 12 NextHop 10.1.1.22
        192.168.1.0/16 RD 65530:20 Label 131057
```

Here we can see that the route-target is translated into an extended community. The route distinguisher is part of the multi-protocol reachability NLRI.

L3VPN: OAM VPRN-Ping and VPRN-Trace

Use....	To Find.....
VPRN-Ping	Destination PE
	Outgoing SAP
VPRN-Trace	Service Label
	Route Targets
	Transport Mechanism (SDP/LDP/GRE
	Next IP Hop
	Remote Ifindex
	Metrics/Preferences

```
A:PE1# oam vprn-ping 20 source 192.168.11.1 destination 192.168.2.2

Seq Node-id                                    Reply-Path Size  RTT
------------------------------------------------------------------------------
[Send request Seq. 1.]
1   2.2.2.2:sap:1/1/4                      In-Band    84    0ms
------------------------------------------------------------------------------
A:PE1#
```

VPRN ping and VPRN trace are available as in-band packet based OAM tools to verify that a service is operational. The packets closely resemble customer packets to effectively test the customer's forwarding path, but they are distinguishable from customer packets so they are not forwarded to the customer.

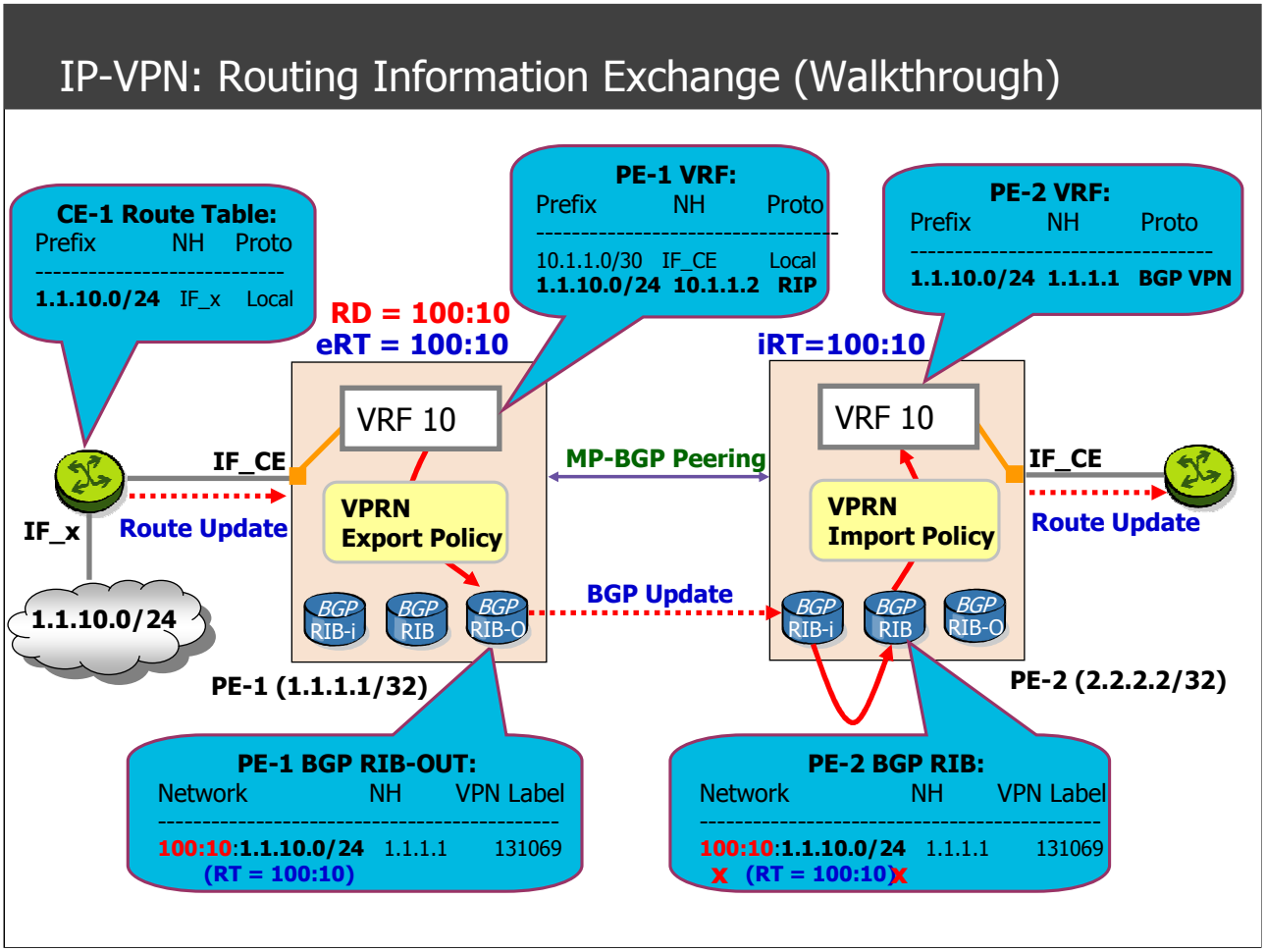
VPRN-Ping is similar to the traditional ping command, except that it sends the ping in-band through the customer's data forwarding path, allowing the administrator to verify that the path through the IP/MPLS core is working properly.

The minimum arguments to use the command are the VPRN service ID, source address and destination address (5, 5.0.0.1 and 5.1.1.1 in this case). Optionally specify the size of the packets, and the number of packets that are sent.

L3VPN: OAM VPRN-Trace

```
A:PE1# oam vprn-trace 20 source 192.168.11.1 destination 192.168.2.2
TTL Seq Node-id         Rcvd-on                   Reply-Path RTT
------------------------------------------------------------------------
----
[Send request TTL: 1, Seq. 1.]
1   1   2.2.2.2          cpm                        In-Band    0ms
  Requestor 1.1.1.1 Route: 192.168.2.0/24
    Vpn Label: 131048 Metrics 0 Pref 170 Owner bgpVpn
    Next Hops: [1] ldp tunnel
    Route Targets: [1]: target:65530:20
  Responder 2.2.2.2 Route: 192.168.2.0/24
    Vpn Label: 0 Metrics 1000 Pref 10 Owner ospf
    Next Hops: [1] ifIdx 2 nextHopIp 192.168.2.2
[Send request TTL: 2, Seq. 1.]
2   1   2.2.2.2      sap:1/1/4                 In-Band    0ms
```

VPRN-Trace is similar to traditional trace-route in that it will send back all of the hops that the trace packets traverse on the way to their final destination, but it also includes additional information such as the VPRN MPLS labels and the route-target.



1) CE-1 has a network 1.1.10.0/24 directly attached through the interface IF_x

2) CE-1 advertises this prefix to PE-1 via a route update.

3) The route update is received on PE-1 at the VPRN access interface "IF-CE" and thus populated into VRF 10.

4) Through the use of a "VPRN Export Policy", the entry in the local VRF is taken out to the BGP RIB-OUT on PE-1 to be advertised to PE-2

While performing this operation, PE-1 includes the following information:

- A RD value of 100:10 is prepended onto the IPv4 prefix, making it a unique VPN-IPv4 entry inside the BGP table. If another customer CE attached to PE-1 were to inject an entry with the same prefix value (1.1.10.0/24), conflict would have been avoided by choosing another unique RD tag for the other customer prefix.

- An export route target value of 100:10 is written into the Extended Community Attributes field. This is going to be used at the end (PE-2) to decide which VRF is going receive that route update.

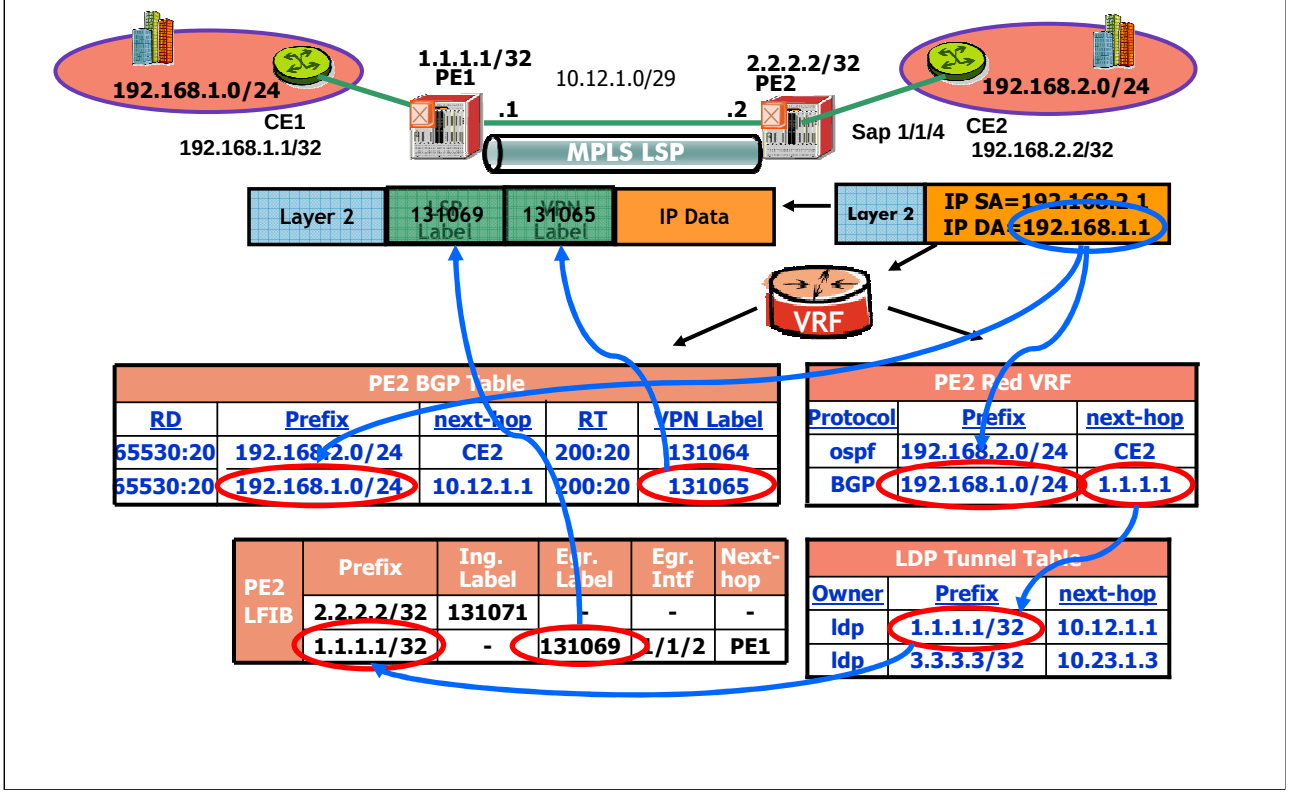
- Also the VPN Label associated with VRF 10 on PE-1 is communicated within the BGP Update message, later to be used on the data plane.

5) The BGP Update message is received on PE-2 and immediately populated into BGP RIB-in.

6) An optional "BGP import policy" could be used at this point to control the way, if and how, certain routes are accepted into the actual RIB. By default there is no BGP import policy and hence BGP RIB-IN = BGP RIB.

7) PE-2 checks the RT field of the incoming BGP update packet to see if there is a match with

Packet Walk: Data Plane (1)



Data Plane Packet Flow

An IP Packet destined for 192.168.1.1 is forwarded by the CE2 router via its routing table to the PE2 router.

At PE2, the packet is received on a sap associated with the Red VRF, so the PE knows that it will use the Red VRF for forwarding decisions.

In the Red VRF, the next-hop address for the destination is determined to be 1.1.1.1. The VPRN is configured with Auto-bind LDP which means that the next hop address will be resolved by using the LDP tunnel table. If the next hop address is located in the LDP tunnel binding table, a lookup is performed in LDP binding table to find out the egress MPLS label.

The packet will be labeled by PE2 with an LSP label of 131069, which defines the transport tunnel to the egress PE.

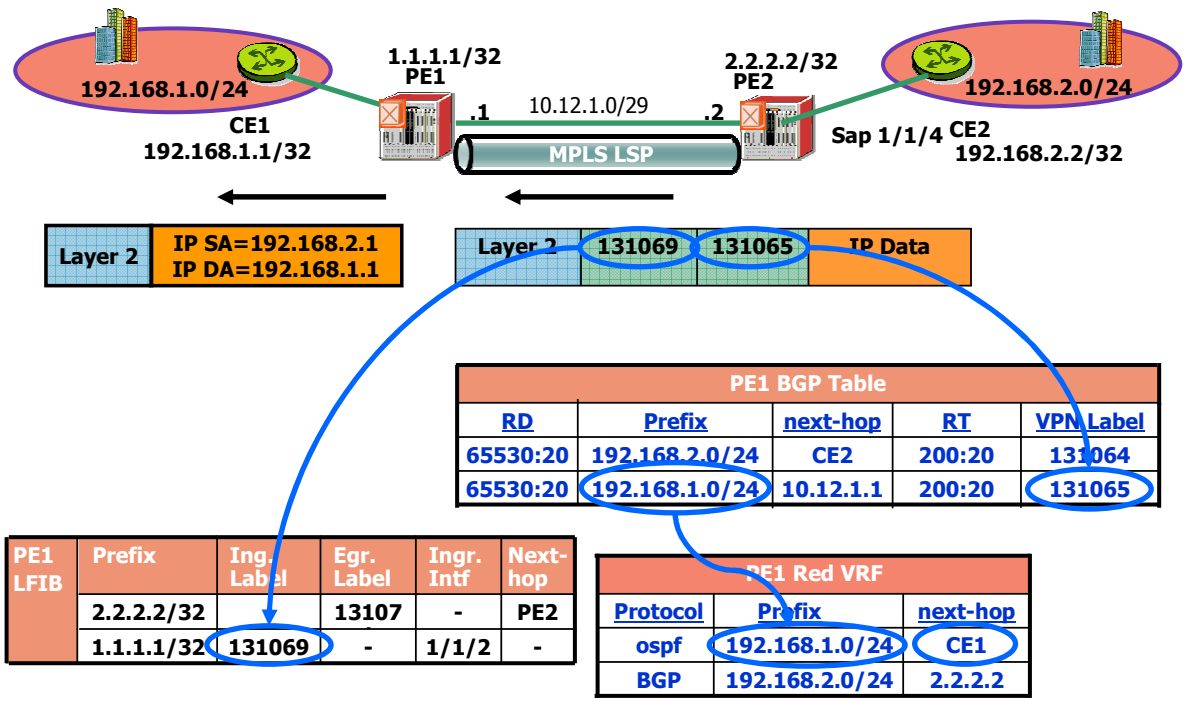
The packet will receive a second label by PE2 with a VPN label defining the service on the egress PE.

As there is a BGP route in this customers VPRN with a VPN label associated with the prefix (propagated in the MP-BGP update), the label is PUSHed onto the packet as the VPN label.

A VPRN service uses a two-level label stack. The ingress PE router pushes both a VPN label and an LSP label onto a packet.

A label will be determined for the next-hop address from the LFIB. The label associated to 1.1.1.1 is 131069.

Packet Walk: Data Plane (2)



If available, a P router along the path will SWAP the LSP label and label switch the packet towards the egress PE.

There are no changes to the IP header or to the VPN label in the core network.

The packet will be label switched across the provider core until it reaches the egress PE.

MPLS handles the forwarding between the PE routers. This means that the routers in the core of the network need not know about the routes connecting the private networks.

The egress PE will POP the LSP label as there is no egress label in the LFIB.

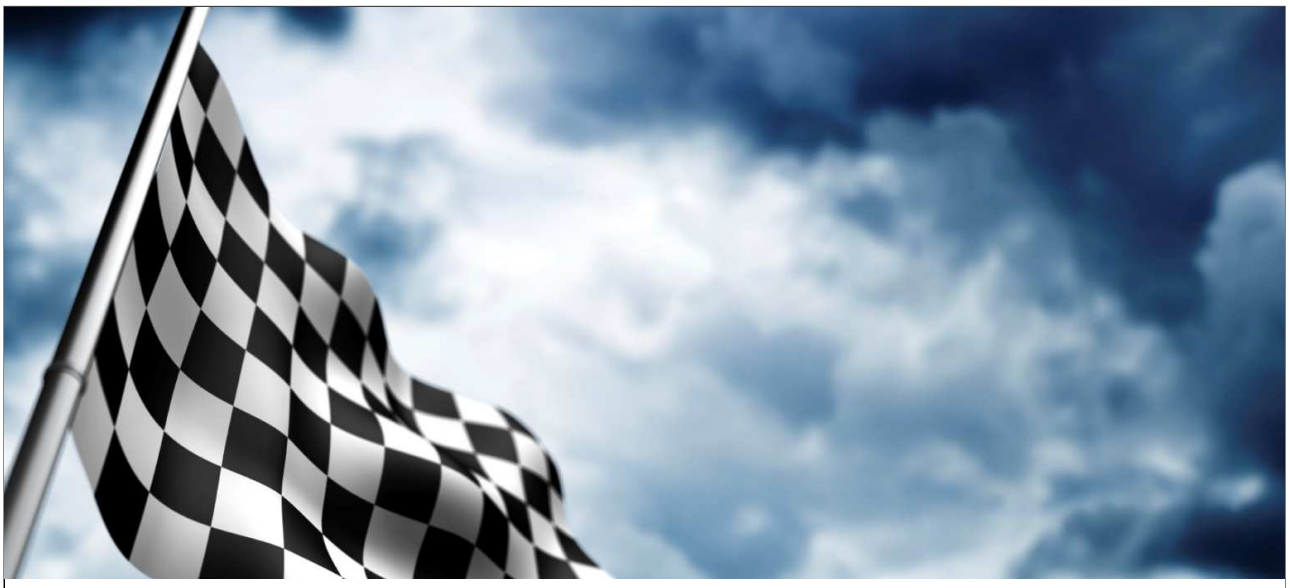
The result is an MPLS labeled packet.

The egress PE will reference the VPN label and determine the associated VRF based on the route-target.

The VPN label is POPed and the result is an unlabeled packet which will be forwarded via the VRF.

The next-hop is identified as being external to the MPLS domain, and the packet will be forwarded by PE1 to CE1 based on the IP header information.

CE1 receives an unlabeled packet for destination 192.168.1.1 and will route it via its routing table.



WRAP-UP
Module Summary

Module summary.

Module Summary

- Alcatel-Lucent's name for a L3 VPN is called a VPRN or Virtual Private Routed Network.
- MP-BGP is used as the PE to PE signaling method
- An 8-byte Route Distinguisher forms a new address called the "VPN-IPv4 address
- Route isolation between VRFs is accomplished through careful Route Target policy administration
- A VPN service label is distributed together with VPN Route information
- Static routes, RIP, eBGP and OSPF are the route distribution options available between the PE and the CE.
- OAM trace and OAM ping are the two most useful troubleshooting commands on a VPRN

Module summary:

This module was talking about the L3 VPN or VPRN.

The signaling method used to exchange the service labels is called MP-BGP.

Besides the service labels, MP-BGP also transports the route distinguisher to create the VPN-IPv4 addresses and the route target to achieve route isolation between different VRF's.

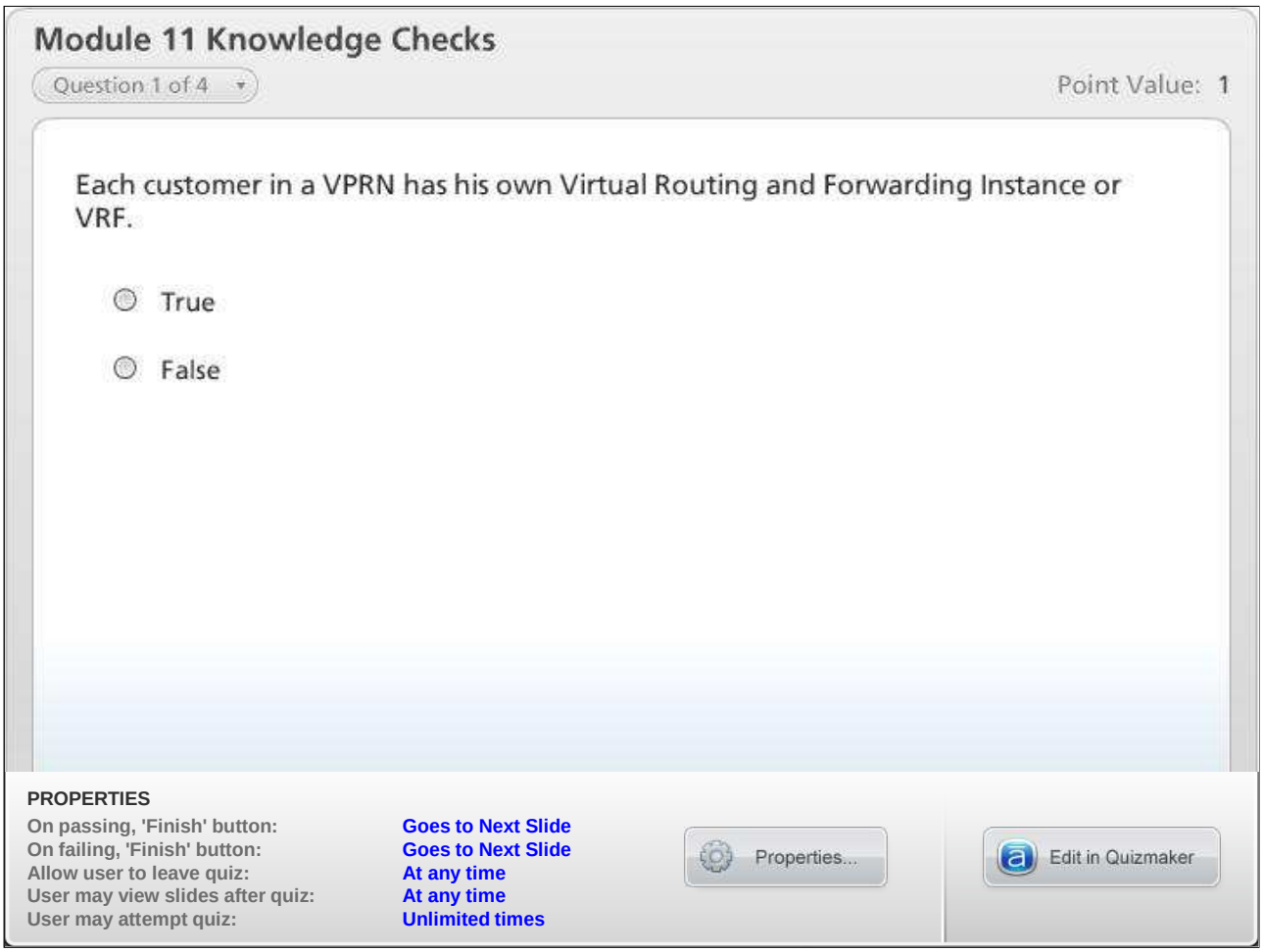
The routing protocols supported between the CE and PE are static routes, RIP, eBGP and OSPF.

There are a couple of operation and maintenance commands available for troubleshooting.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 11

Learning experience powered by Alcatel-Lucent University

Alcatel·Lucent

This completes module 11.

SR-OS Fundamentals

Module 12: Protocol Independent Multicast (PIM)

IPD Development

Available as PDF

Welcome to the 12th module of the SR-OS fundamentals course.

Table of Contents

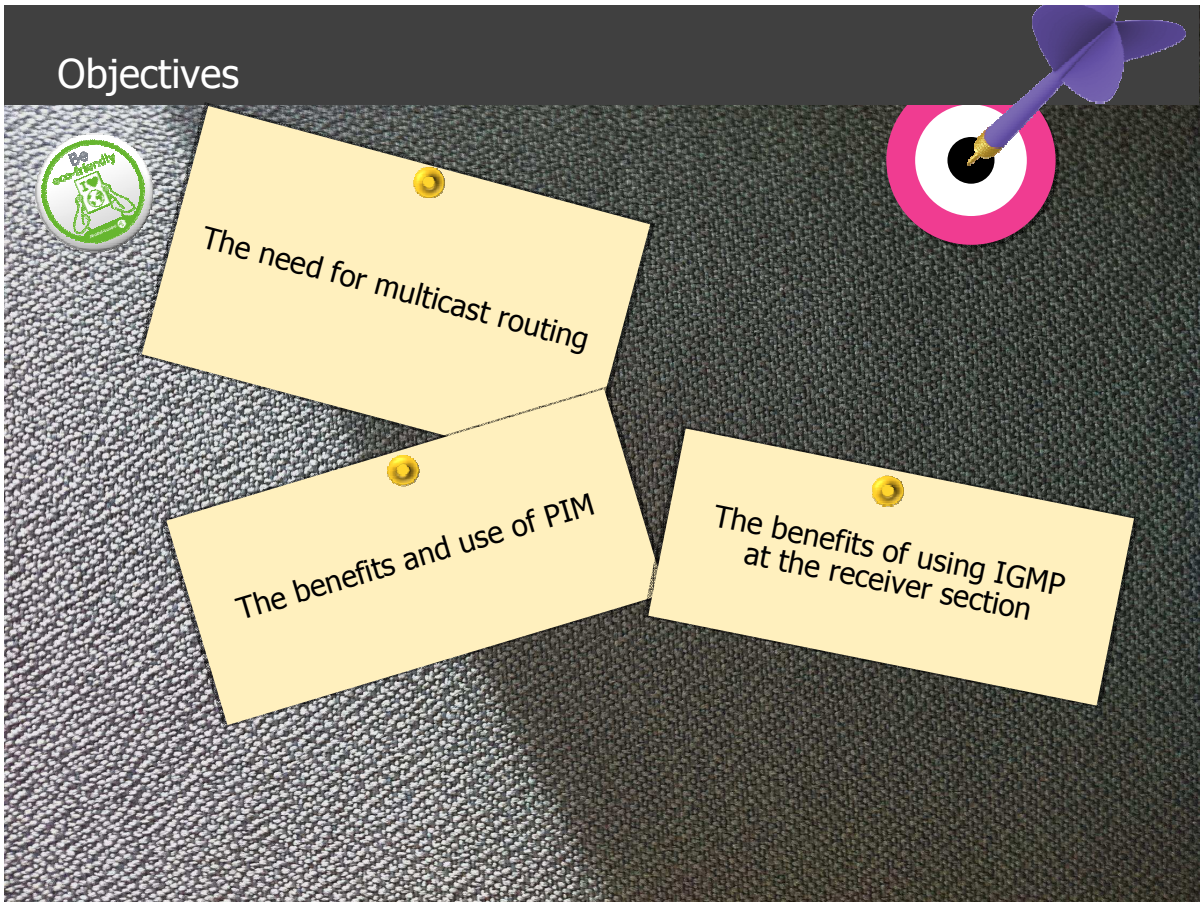
Select the chapter you are interested in or click on the Next button to follow the recommended learning path.

Module 12 is divided into three sections.

Section 1 introduces the rational behind multicast and gives an overview of the general aspects of multicast

Section 2 introduces the Internet Group Management protocol used between the receiver and the last hop router

Section 3 explains the Protocol Independent Multicast protocol used in the core network



By the end of module 12 you will be able to explain:

The need for multicast routing

The benefits of using IGMP at the receiver section

The benefits and use of PIM

Please click Next to continue.

Multicast rational

Section1: Multicast rational

In the unicast model, a source device sends a packet to a destination address that defines one other device somewhere in the network. The path taken by this packet from the source 'S' to the destination 'D' is chosen at each hop along the path by the unicast routing protocol and algorithm at work in each router, and may be independent for each packet in the sequence.

At the transport layer, TCP can provide reliable, connection-oriented services or UDP can provide connectionless best-effort services, each with their associated overhead. With TCP, an acknowledgement system is used to retransmit lost packets, while in UDP this is not done and data loss becomes the responsibility of the upper layers.

As the number of packets in the data stream or the number or receivers increase, the result is a significant burden on network resources, from sender to router to receiver.

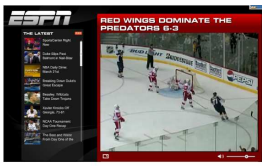
Multicast is the answer to efficient, scalable, large scale data delivery.

The source sends a single copy of a packet that is addresses to a group of receivers. This group is a logical entity that any device can choose to listen or not listen to at any time. Efficient, low overhead transport layer services are provided by UDP.

Unlike unicast traffic only one copy of any packet is required regardless of the number of receivers, as many receivers can join to the same group and therefore all can receive a copy of the same packet.

Unlike broadcast traffic, multicast will span across routers if configured to do so. Devices that don't want the packet may not receive the data at all, or if they do receive the frame, will discard it at layer 2.

Target Applications



- **One/Few to Many**
  - Live TV
  - Radio Broadcast
  - Real-Time Financial Data Updates - Stock Quotes
  - Distance Learning

- **Many to Many**
  - Audio/video conferencing,
  - Collaboration,
  - On-line Games

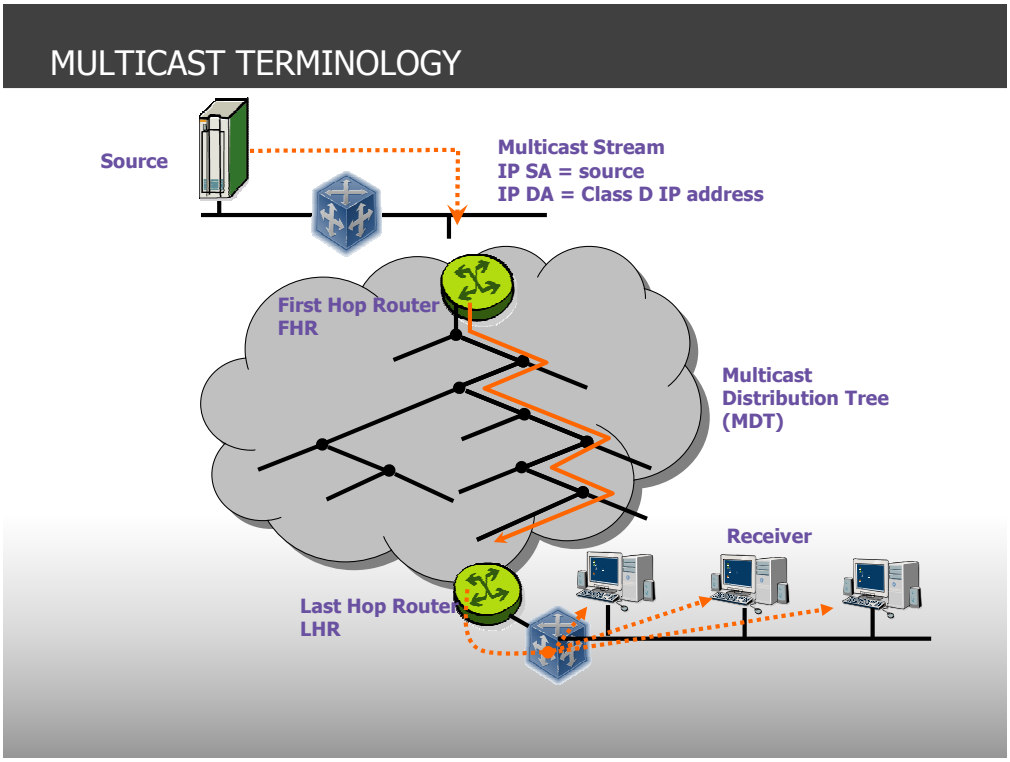
AT THE SPEED OF IDEAS

Alcatel·Lucent

Applications in general also fall into 2 categories.

The one to many application is the simplest model, and is similar to broadcast television. One source to many receivers. It is also best suited for data that is similar in nature to broadcast television, such as audio or video.

The many to many application is more complex. In this model any number of devices can be sources sending data to any number of receivers, including itself. It is best suited for multipoint applications such as a video conference, shared workspace or distance learning application.



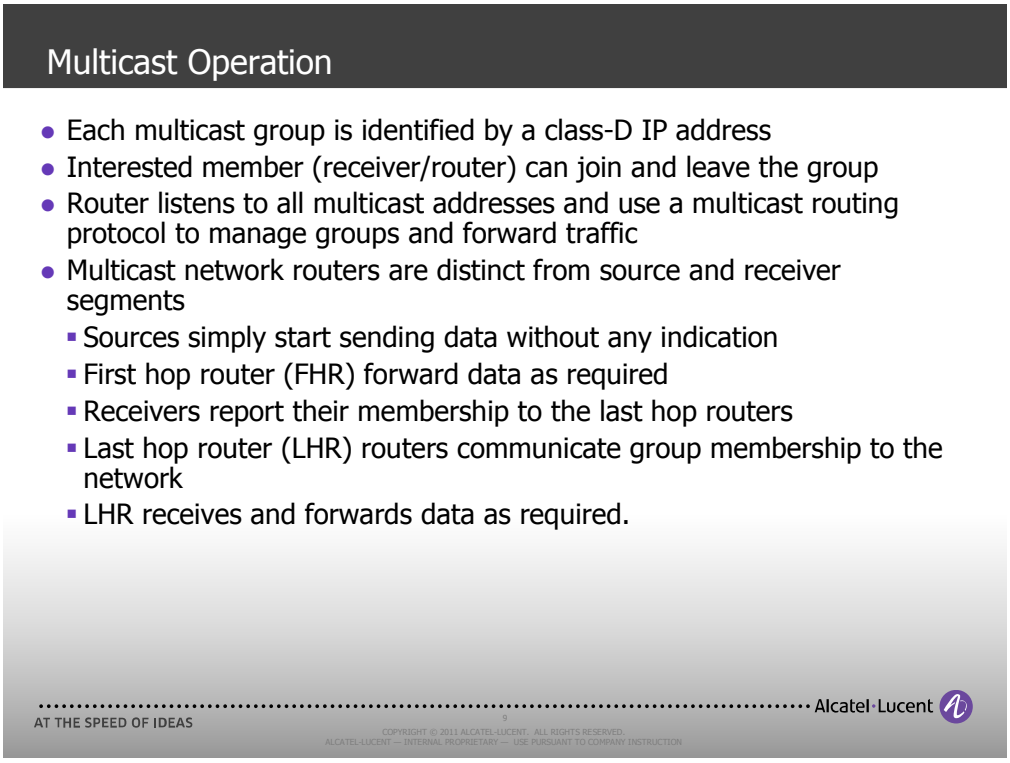
A source can be any device that generates a packet addressed to a multicast group.

A receiver can be any device that issues a request to receive multicast data, called joining a group.

A switch can be any device capable of switching multicast frames.

A router can be any device capable of forwarding and replicating multicast packets between broadcast domains.

The First Hop Router is the router closest to the source and is the router directly connected to the subnet of the source. The last hop router is the router closest to the receiver and directly connected to the receiver subnet.



Each multicast group is identified by a class-D IP address. Interested member, a receiver or router can join and leave the group. Router listens to all multicast addresses and use a multicast routing protocol to manage groups and forward traffic. Multicast network routers are distinct from source and receiver segments. Sources simply start sending data without any indication.

First hop router or FHR forwards data as required. Receivers report their membership to the last hop routers.

The last hop router or LHR routers communicate group membership to the network.

LHR receives and forwards data as required.

Multicast

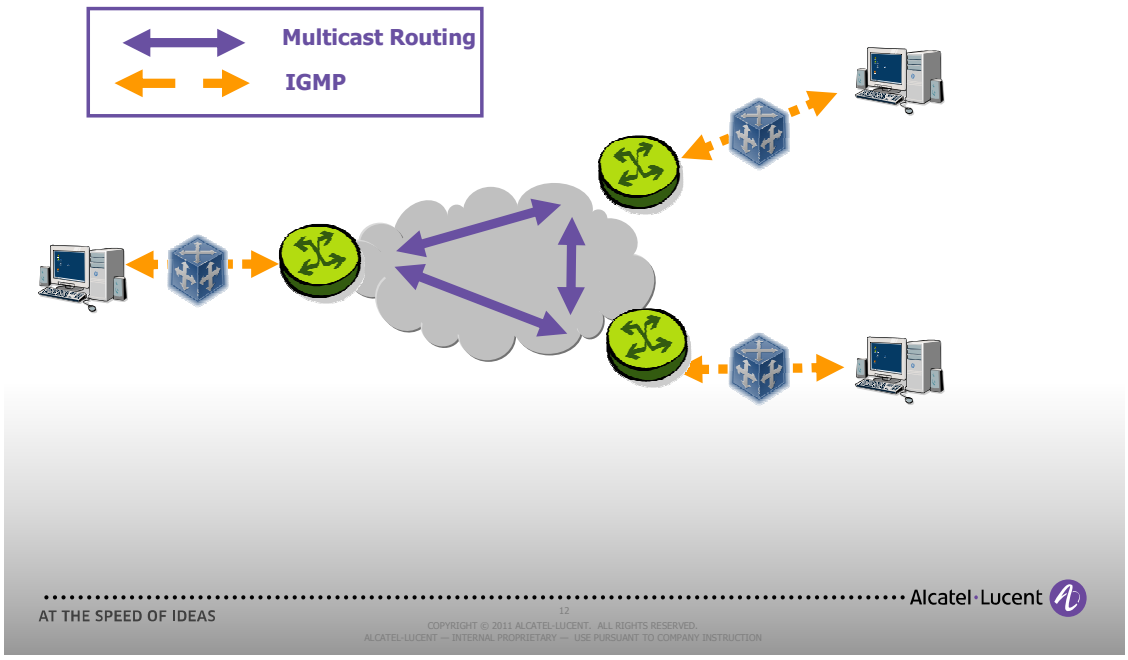
- Class D Based on Class "D" IP address
- values from 224.0.0.0 to 239.255.255.255 (224.0.0.0/4)
- 224.0.0.1 through 224.0.0.255 'reserved'
- RFC 3171 listed the reserved Class D address.

Well-Known Class D Addresses	Purpose
224.0.0.1	All hosts on a subnet
224.0.0.2	All routers on a subnet
224.0.0.4	All DVMRP routers
224.0.0.5	All MOSPF routers
224.0.0.9	RIP v2
224.0.1.1	Network Time Protocol
224.0.0.22	IGMPv3

Here you can see the IP class D range and some well known reserved Class D addresses according to the RFC 3171 standard.

IGMP

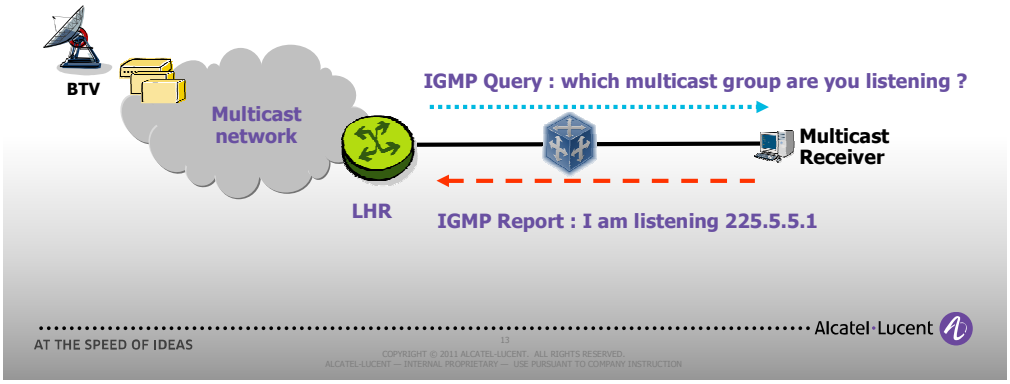
Section2: Internet Group Management Protocol or IGMP



Multicast routing protocols work in conjunction with the Internet Group Management Protocol to provide end-to-end connectivity of multicast user traffic. IGMP allows hosts to join and leave a particular multicast group, while multicast routing protocols allow multicast traffic to flow between networks. Some common multicast routing protocols are Distance Vector Multicast Routing Protocol or DVMRP and Protocol Independent Multicast or PIM.

IGMP

- IGMP: Internet Group management Protocol
  - Used by receiver hosts and Last Hop Router (LHR) to communicate with each other about the group membership
  - Allows receiver hosts to dynamically join/leave multicast groups
  - Allows LHR to query the receiver for the multicast groups it is receiving
  - IGMP version 1,2,3 are supported on 7x50

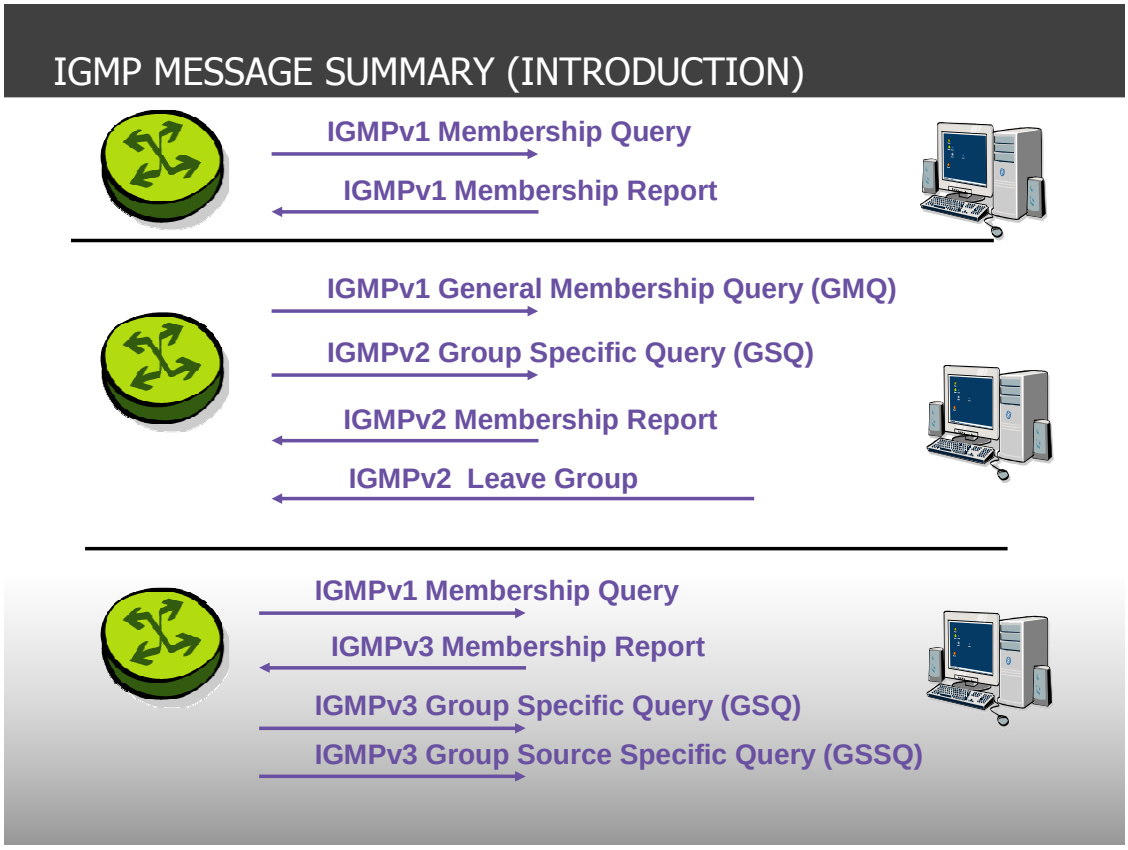


Internet Group Management Protocol is the multicast signaling protocol used between receivers and their local router. IGMP messages are encapsulated in IP packets with an IP protocol number of 2, and should always have a TTL of 1.

IGMP has multiple components, primarily a receiver component and a router component. It enables receivers to communicate group information to their local router, such as multicast groups that they wish to join or leave. It enables router to determine group membership and status, such as which groups are active on which of the routers local interfaces.

LAN switches may be IGMP aware, however as these devices historically were strictly Layer 2 devices, many LAN switches did not perform well when asked to interpret IGMP messages, and performance suffered greatly. Recent switches have overcome most of these limitations, due to faster CPU's and more efficient mechanisms for processing Layer 3 IGMP packets.

If the receiver doesn't report it has a particular multicast group for certain period of time, LHR will stop forwarding traffic



The Internet Group Management Protocol is the method used by end users for joining and leaving multicast groups. The original version 1 of IGMP specifies that the multicast router periodically transmits a Host Membership Query message onto the broadcast segment to determine which multicast groups still have interested members.

Receiving hosts respond to the Query message with a Host Membership Report for each group they currently belong to. The periodic transmission of the Query messages allows the router to maintain or update its knowledge of the group members present on each interface. IGMPv1 also allows for an end user to explicitly join a multicast group through the transmission of a Host Membership Report for the interested group address. This allows the multicast router to immediately associate the group with the broadcast interface. When the end user host decides to leave the multicast group, it does so silently. This means that the router continues forwarding traffic on the segment until the next Query message or couple of more query messages goes unanswered.

Version 2 of the IGMP specification defines both a Group- Specific Leave and Query message. This allows a departing host to inform the multicast router to stop the flow of traffic onto the segment. In addition, IGMPv2 defines the election of a querier for each LAN. The router with the lowest IP address on the LAN becomes the querier router and is responsible for sending Query messages on the segment.

IGMPv3 introduces the concept of source-specific multicast or SSM forwarding. This allows an end user to request multicast traffic for a particular group from a specific source of the traffic. With IGMPv1 and v2, a host receives multicast group traffic from all possible sources.

Enabling IGMP

- IGMP is disabled by default
- IGMP is enabled when the context is created

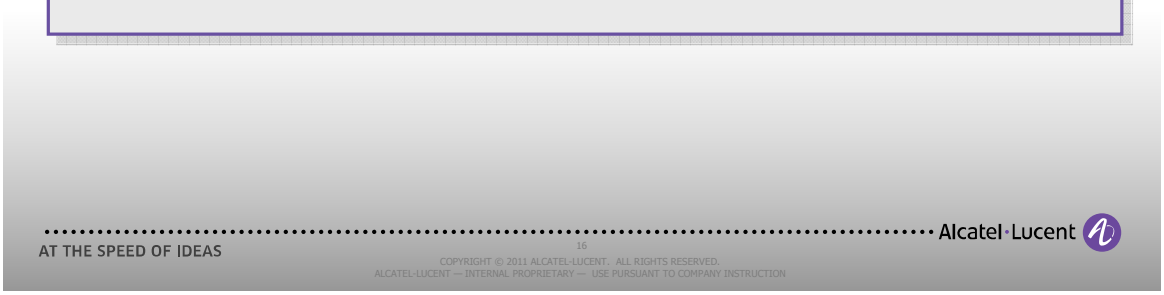
```
Context: config>router>

Syntax:  igmp

Example: config>router> igmp
         config>router>igmp$
```

IGMP is disabled by default. "configure router IGMP" brings you to the IGMP context.

Enabling an Interface

- This command enables the context to configure an IGMP interface
- IGMP should be enabled on interfaces where receivers are connected

```
Context: config>router>igmp

Syntax:  interface ip-int-name

Example: config>router>igmp$ interface toRouterB
```



The command "configure router igmp" enables the context to configure an IGMP interface. IGMP should be enabled on interfaces where receivers are connected.

IGMP Version

- Specifies the IGMP version for the interface
- IGMP version 3 is the default

```
Context: config>router>igmp>if

Syntax:  version version

Example: config>router>igmp>if$ version 3
```


Version 3 is the default version for IGMP but can be changed as desired.

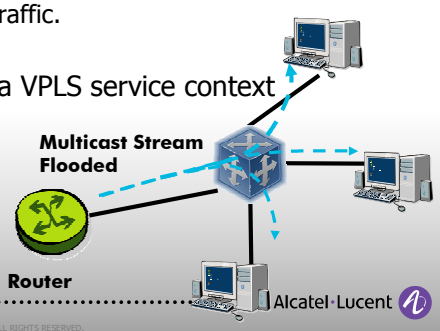
(S,G) and (*,G) notation

- Source IP address (of the source) represented by S,
- Group address represented by G.
- Also commonly used is (*,G)
  - * is a Wild card
  - Also called Starg (Pronounce Star G)
  - Any source for a particular group G
  - Use in IGMP Static and PIMSM Shared Tree

To represent a multicast group and the source from where it is coming, the (S,G) notation is used. S is the source IP address of the streaming source. G is the multicast group address. Sometimes a star is used as source. This means any source and not any specified one.

IGMP snooping

- The solution to the multicast flooding in the L2 network is Multicast Snooping.
- The L2 switch will monitor all the switch port for L3 IGMP messages.
- As the switch port receives IGMP request and forward it to the LHR, the switch will have the knowledge of the multicast join for the particular group.
- As the multicast traffic arrives from the LHR:
  - traffic will be forwarded to the switch port which IGMP membership report for the particular Group was received and forward
  - Other port will not receive the multicast traffic.
- IGMP snooping is only configurable in a VPLS service context
  - VPLS
  - SAP
  - Spoke-SDP
  - Mesh-SDP



Regular switches do flood multicast frames as they do not have the capability to look at the L3 layer. More modern switches can, once enabled by configuration, look at the L3 header and define if this is a multicast packet and which multicast group is joined or left. This is what we call multicast snooping. Snooping can be seen as monitoring the packet without changing anything.

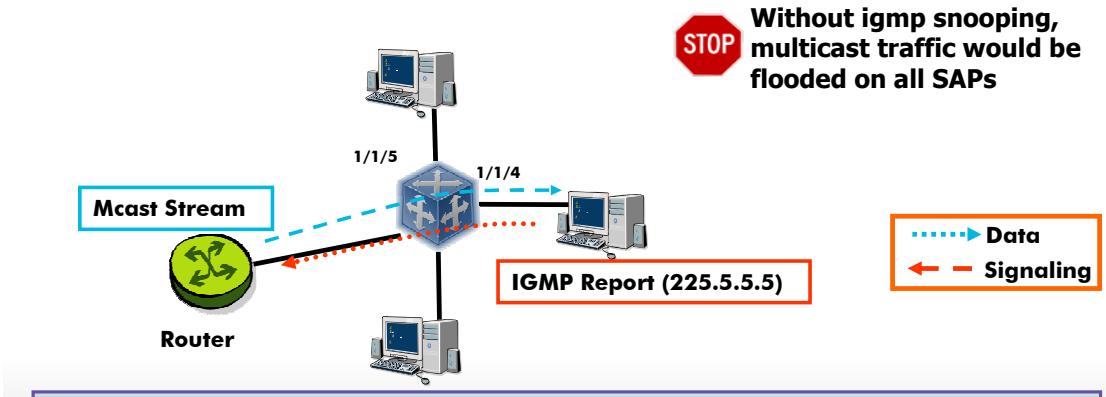
As a switch port receives an IGMP request and forwards this it to the LHR, the switch will have the knowledge of the multicast join for the particular group. As the multicast traffic arrives from the LHR, traffic will be forwarded to the switch port where the IGMP membership report for the particular Group was received.

Other ports will not receive the multicast traffic.

IGMP snooping is only configurable on the service router in a VPLS service context on the VPLS, SAP, spoke-SDP or mesh-SDP instance.

IGMP Snooping

- Forwarding Tables created by IGMP Snooping
  - IMGP Reports and Leave are intercepted
  - Multicast forwarding table updated accordingly




```
Source Address  Group Address   Sap/Sdp Id              Svc Id   Fwd/Blk
------------------------------------------------------------------------
*               225.5.5.5       sap:1/1/4               Local    Fwd
------------------------------------------------------------------------
```

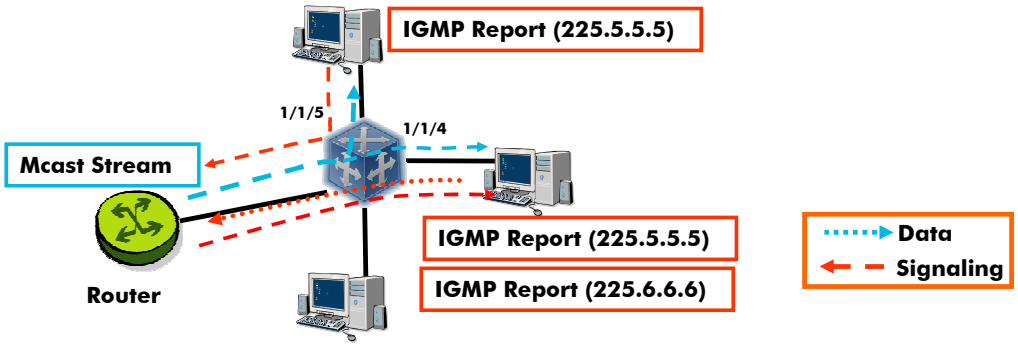

Shown here is an example case where IGMP snooping is enabled on the LAN switch to eliminate multicast flooding. A receiver sends an IGMP Host Membership Report. The switch will snoop the message and add the multicast IP address into the forwarding table associated to the port the IGMP message was received on.

Note that the switch does not stop the message on its way to its intended destination of the router, he simply looks at it.

If a multicast packet arrives on the switch with a destination address of 225.5.5.5, it will no longer be flooded, but switched only to port 1/1/4 as indicated in the forwarding table.

IGMP Snooping

- Same group can be joined by multiple receivers
- Multiple groups can be registered on same SAP




```
Source Address  Group Address   Sap/Sdp Id              Svc Id   Fwd/Blk
------------------------------------------------------------------------
*               225.5.5.5       sap:1/1/5               Local    Fwd
*               225.5.5.5       sap:1/1/4               Local    Fwd
*               225.6.6.6       sap:1/1/4               Local    Fwd
------------------------------------------------------------------------
```

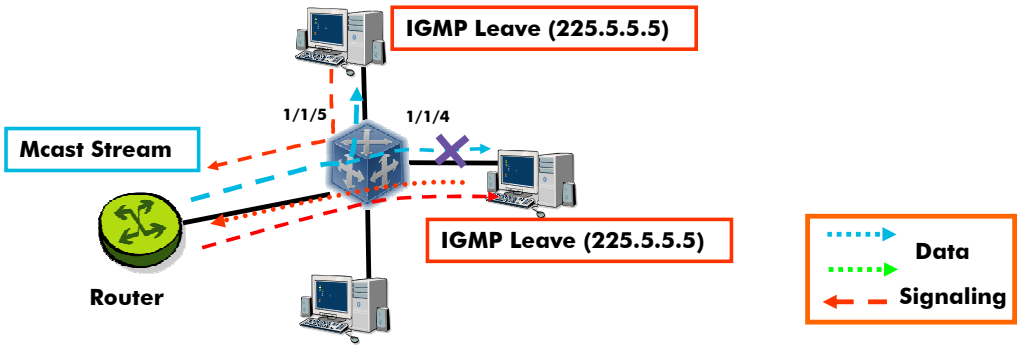

Shown here is an example case where IGMP snooping is enabled on the LAN switch to eliminate multicast flooding. A receiver has already been receiving stream 225.5.5.5. Another receiver sends an IGMP Host Membership Report. The switch will snoop the message and add the multicast IP address into the forwarding table associated to the port the IGMP message was received on.

Note that the switch does not stop the message on its way to its intended destination of the router, he simply looks at it.

If a multicast packet arrives on the switch with a destination address of 225.5.5.5, it will no longer be flooded, but switched only to port 1/1/4 and 1/1/5 as indicated in the forwarding table.

IGMP Snooping - Leave

- IGMP Leave intercepted
  - Sap is removed from the forwarding table




```
Source Address  Group Address   Sap/Sdp Id                Svc Id   Fwd/Blk
-------------------------------------------------------------------------
*               225.5.5.5       sap:1/1/5                 Local    Fwd
*               225.6.6.6       sap:1/1/4                 Local    Fwd
-------------------------------------------------------------------------
```

AT THE SPEED OF IDEAS

22

Shown here is an example case where IGMP snooping is enabled on the LAN switch to eliminate multicast flooding. A receiver has already been receiving stream 225.5.5.5. Another receiver sends an IGMP Group Leave. The switch will snoop the message and remove the multicast IP address in the forwarding table associated to the port the IGMP message was received on.

If a multicast packet arrives on the switch with a destination address of 225.5.5.5, it will no longer be sent to 1/1/4, but switched only to port 1/1/5 as indicated in the forwarding table.

Forwarding Table

- Displays the multicast forwarding table for the specified service

```
A:RouterA# show service id 499 mfib
===============================================================================
IGMP Snooping MFIB for service 499
===============================================================================
Source Address    Group Address     Sap/Sdp Id        Fwd/Blk
-------------------------------------------------------------------------------
*                 235.6.6.6         sap:1/1/4         Fwd
*                 235.4.4.4         sap:1/1/1         Fwd
-------------------------------------------------------------------------------
Number of entries: 2
===============================================================================
A:RouterA#
```


This command displays the multicast forwarding table for service 499.

Label	Description
Source Address	IPv4 unicast source address
Group Address	IPv4 multicast group address
SAP/SDP ID	Indicates the SAP/SDP to which the corresponding multicast stream will be forwarded/blocked
Forwarding/Blocking	Indicates whether the corresponding multicast stream will be blocked/forwarded
Number of Entries	Number of entries in the MFIB
Forwarded Packets	Indicates the number of multicast packets forwarded for the corresponding source/group
Forwarded Octets	Indicates the number of octets forwarded for the corresponding source/group

Verifying IGMP Snooping

- Verifies IGMP snooping is enabled on a per service basis

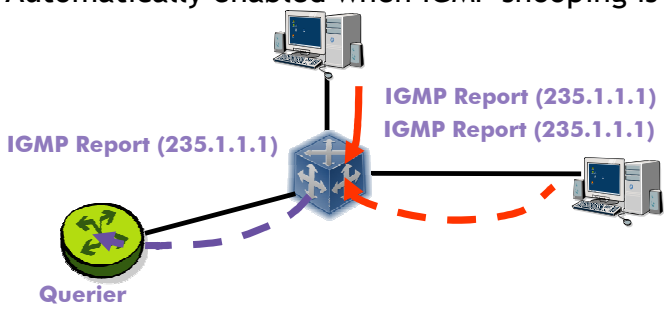
```
A:Access1# show service id 100 igmp-snooping base
===============================================================================
IGMP Snooping Base info for service 100
===============================================================================
Admin State : Up
Querier     : 192.16.1.1 on SDP 200:100
-------------------------------------------------------------------------------
Sap/Sdp                  Oper    MRtr  Send      Max Num   MVR        Num
Id                       State   Port  Queries   Groups    From-VPLS  Groups
-------------------------------------------------------------------------------
sap:1/1/1                Up      No    Disabled  No Limit  Local      0
sdp:200:100              Up      No    Disabled  No Limit  N/A        0
===============================================================================
```


The **'show service id 10 igmp-snooping base'** command displays the IGMP snooping information for the specified service instance. Shown here are the service interfaces associated to the service and enabled for IGMP snooping.

Label	Description
Source Address	IPv4 unicast source address
Group Address	IPv4 multicast group address
SAP/SDP ID	Indicates the SAP/SDP to which the corresponding multicast stream will be forwarded/blocked
Forwarding/Blocking	Indicates whether the corresponding multicast stream will be blocked/forwarded
Number of Entries	Number of entries in the MFIB
Forwarded Packets	Indicates the number of multicast packets forwarded for the corresponding source/group
Forwarded Octets	Indicates the number of octets forwarded for the corresponding source/group

IGMP Proxy

- Allow a switch to send IGMP messages on behalf of connected receivers
- Automatically enabled when IGMP snooping is enabled



- Single IGMP leave transmitted to querier for last receiver

If both receivers sends an IGMP reports responding to an IGMP Query, the switch will send a single IGMP report in response to the two IGMP reports messages that it receives from the receivers.

If one of the receivers sends a leave but the other receiver is still 'listening' to a given multicast group, the leave is not transmitted to the querier. The leave will be transmitted to the querier only when the last receiver sends a leave.

This is called IGMP proxy. IGMP proxy is automatically enabled when IGMP snooping is enabled.

PIM

Section 3: Protocol Independent Multicast.

PIM

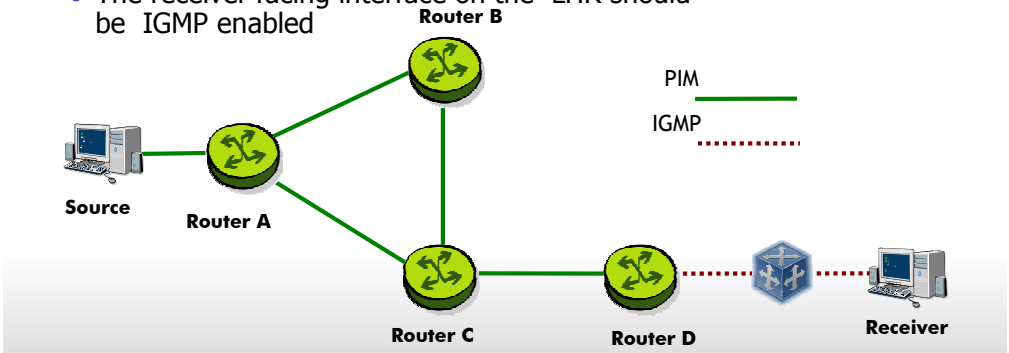
- PIM: Protocol Independent Multicast
  - Used to propagate multicast forwarding state between routers
  - It establishes and maintains Multicast Distribution Tree (MDT)
  - Enables routers to build a delivery tree between the sender(s) and the receivers
  - It is "protocol Independent" and it simply uses the unicast routing table for its operations

The name PIM means Protocol Independent Multicast. This implies that the multicast is independent of the underlying unicast routing. Meaning that any unicast routing protocol may be used.

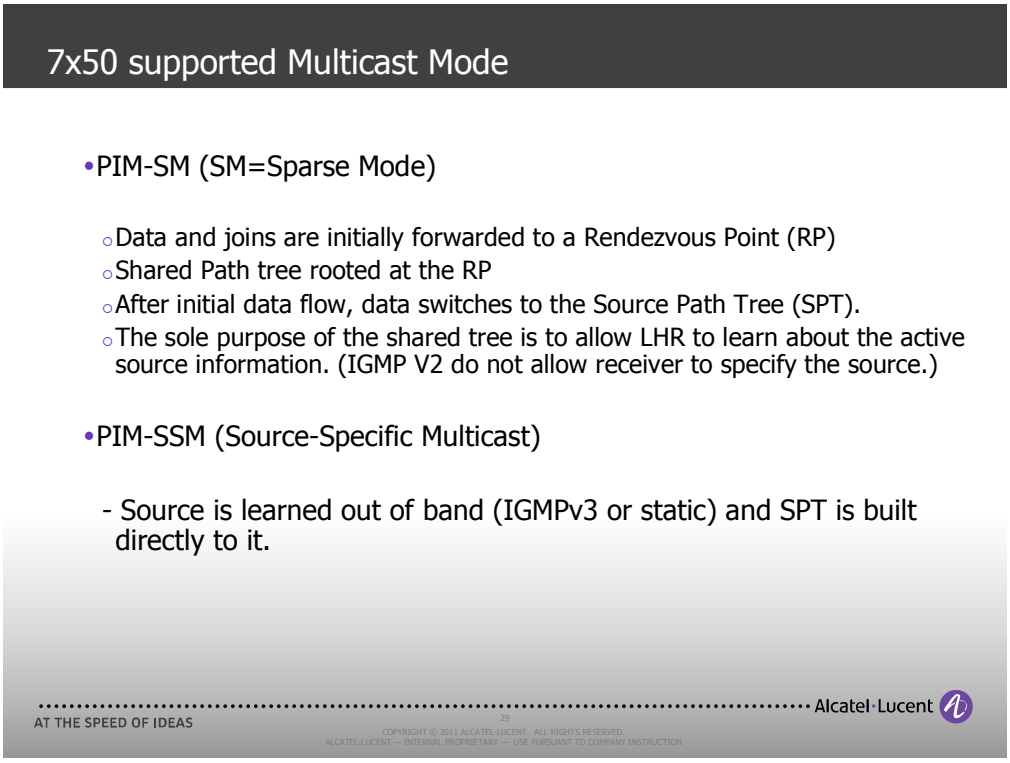
PIM is used to propagate multicast forwarding states between the first and last hop router. This propagation of states result in a tree that is called the Multicast Distribution Tree or MDT. Once this tree is built, multicast data can be delivered between sender and receiver.

Where to configure PIM and IGMP

- PIM must be enabled on
  - all core facing interfaces
  - All source facing interfaces
  - All PIM Router system addresses
- The receiver facing interface on the LHR should be IGMP enabled



Between the first and last hop router, PIM must be enabled on all core facing interfaces, source facing interfaces and all PIM router system addresses. The router interface facing the receiver segment should not be put into PIM, but should be IGMP enabled.



Multicast data forwarding is based on the behavior of trees. As stated earlier, the logical path taken through the network is called the Multicast Distribution Tree or MDT. The type of MDT will vary depending on the choice of Multicast routing protocol configured. The role of the multicast routing protocol is to create the MDT by signaling, so that packets may then follow this logical path.

Sparse Mode protocols may use a Source Path Tree, but it must first start with the Shared Path Tree. The Shared Tree is a tree rooted at the RP or Rendez-vous point, which is the meeting point in the network where the source and receivers must first meet to establish the multicast flow. It is only after this initial meeting has been completed that Sparse Mode protocols will utilize the Source Path Tree.

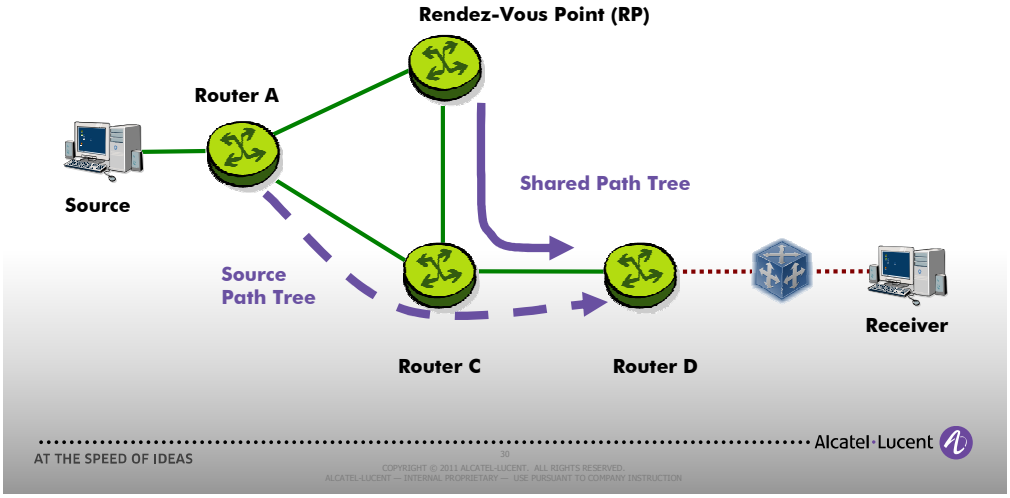
The concept of the Rendezvous Point is also a Sparse Mode characteristic, something not seen in the Dense Mode model. The RP allows multicast data flows between sources and receivers to 'meet' at a predefined network location. It is primarily required since Sparse Mode operation is based on the principle that the sources and receivers do NOT know of each others existence, and therefore require an intermediate device to establish the flow.

This is method of operation is referred to as Any Source Multicast or ASM.

The RP is seen by many as a limitation itself, and other Sparse Mode protocols have eliminated this requirement. Still, the Sparse Mode model is the basis for most multicast implementations today.

Shared vs Source Path Tree

- Shared Path tree is rooted at the RP and is represented by (*,G)
- Source Path Tree is rooted at the source and is represented by (S,G)

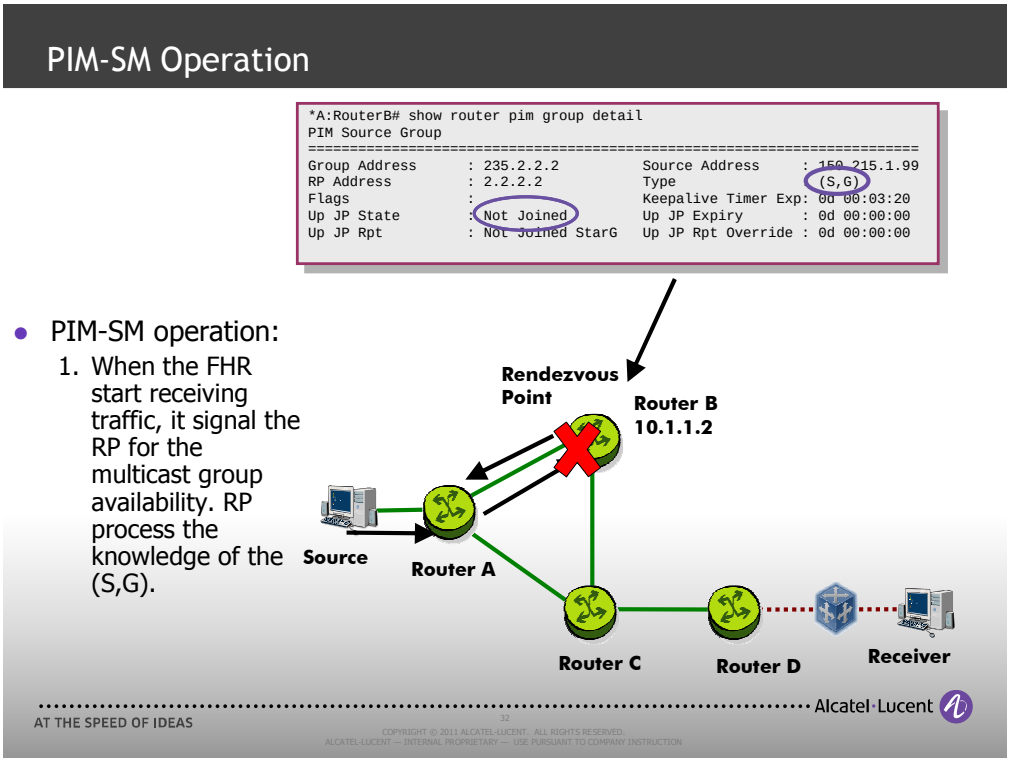


The Shared Path Tree is rooted at the RP while the Source Path Tree is rooted at the source. The Source PT diverges from the Shared PT because the RP is not on the shortest path from the LHR to the source. The Source PT always follow the Shortest Path from FHR to LHR.

PIM-SM: RP

- RP= Rendezvous Point
- At least one router in the PIM network needs to be act as a RP.
- RP allows FHR to register their receiving multicast group.
- RP allows LHR to perform the initial (*,G) multicast join.
- RP provide source information to LHR to build a Source Path Tree.

For Sparse Mode to work and Shared Path Trees to exist, the Rendezvous Point must exist in the network and be defined in the configuration. If the RP is not known to even one of the routers, the Sparse Mode network may not function correctly.



This is an illustration of a PIM SM network in operation, focusing on the creation of the Source Path Tree between the first-hop router from the source and the RP. Note that the Shared Path Tree is not present on the RP in this case, indicating there are no current receivers for this data at this moment.

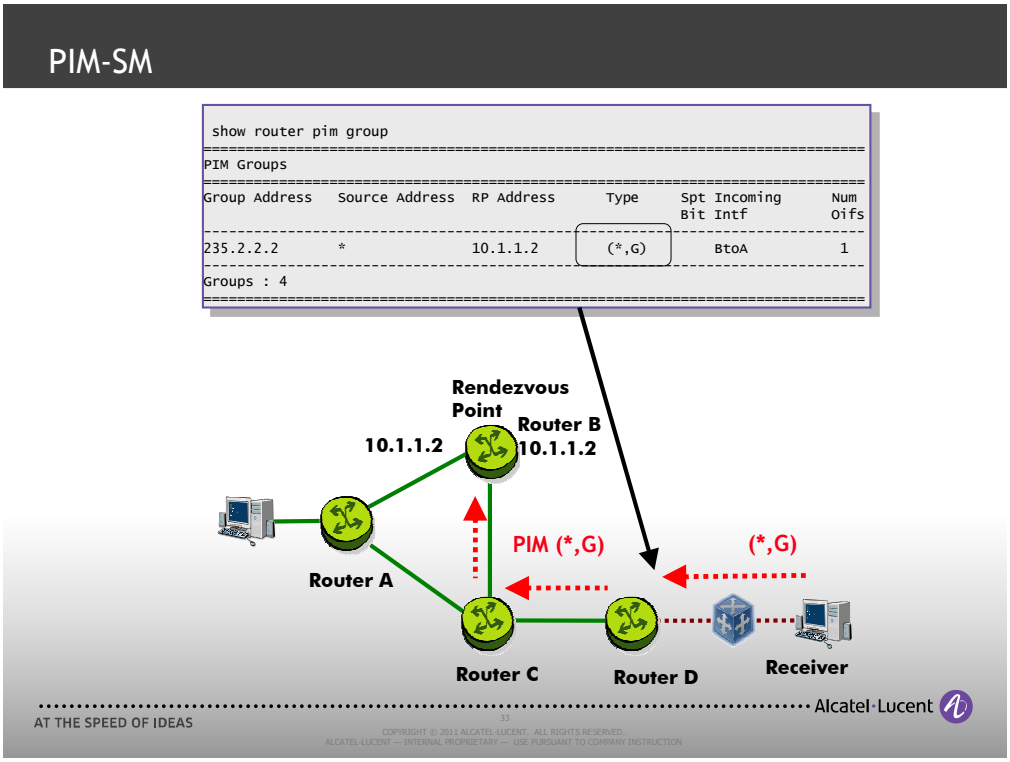
The source sends data for different groups. No signaling is required. This creates a Source Path Tree between the source device and Router A.

Router A notes the presence of a connected source sending to groups, determines the RP for the groups by checking the local RP-set. It sends a unicast Register message containing an encapsulated multicast packet from source to the RP. Router A will continue to do this at whatever rate the packets are arriving from the source until it is requested by the RP to stop.

Since a Shared Path Tree does not exist on Router B, the RP, for the groups, the data from the Register is de-encapsulated and discarded.

Since there are no receivers for this group, the RP immediately sends a Register Stop to the first hop router to terminate the registration .

The forwarding tables will be maintained for the lifetime of the source. They are maintained on the first hop router from the source by the receipt of data from the source, and on the RP by the periodic Registration that will occur to notify the RP that the source is still active.



The receiver connected to Router D issues an IGMP (*, G) unsolicited host membership report. Router D notes the presence of the connected receiver, determines the RP for group G and sends a PIM (*, G) Join towards the.

Router C receives the PIM (*, G) Join, determines the RP for group G and sends a PIM (*, G) Join towards the RP.

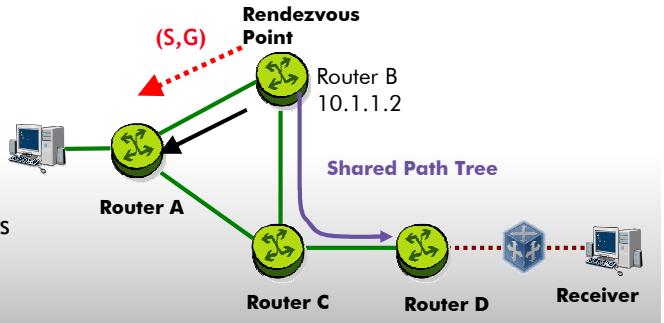
Router B receives the PIM (*, G) Join, determines the RP for group G is itself and does not propagate the Join any further.

The Shared Path Tree has been created for group G, all routers on the Shared Path Tree will have a (*, G) entry.

PIM-SM

- PIM-SM operation:
  - When the RP receives the (*,G) join, the RP initiates an (S,G) join to the source and RP receives Multicast traffic
  - LHR which is D now forms a shared tree with B and C or RPT for the multicast traffic.
  - Switchover to Source Path Tree (SPT) occurs immediately to find better path to the source

```
*A:RouterB# show router pim group detail
===============================================================================
PIM Source Group
===============================================================================
Group Address      : 235.2.2.2          Source Address     : 150.215.1.99
RP Address         : 2.2.2.2            Type               : (S,G)
Flags              : spt, rpt-prn-des   Keepalive Timer Exp: 0d 00:03:22
MRIB Next Hop      : 10.12.1.1          MRIB Src Flags     : remote
Up Time            : 0d 00:02:18        Resolved By        : rtable-u

Up JP State        : Joined             Up JP Expiry       : 0d 00:00:52
Up JP Rpt          : Pruned             Up JP Rpt Override : 0d 00:00:00
```



AT THE SPEED OF IDEAS

Alcatel·Lucent

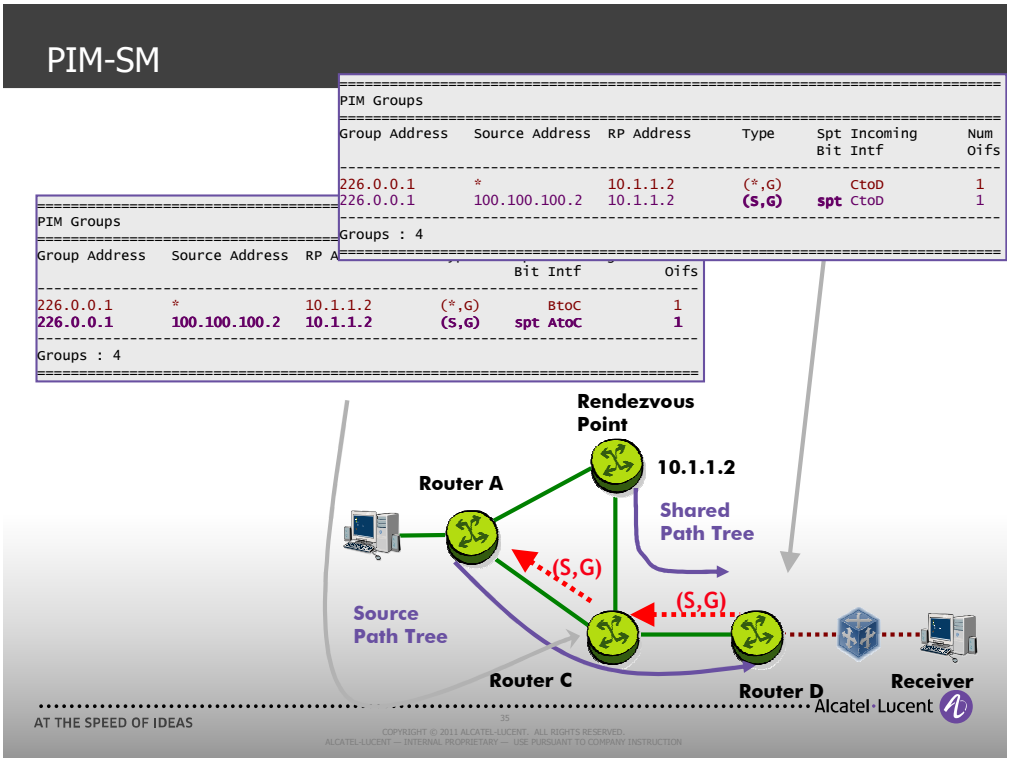
Since a Shared Path Tree exists on Router B, the RP, for group G, the data from the Register is forwarded on the existing branch of the Shared Path Tree.

Router B sends a PIM (S, G) Join towards the source to create the Source Path Tree

Once the Source Path Tree is completed, which could be several hops away, the source data will flow over the Source Path Tree to the RP, and then be forwarded to the receiver on the existing branch of the Shared Path Tree. Since the Registration is still occurring, there are duplicate packets forwarded on the Shared Path Tree.

After receiving the packet on the Source Path Tree, the RP sends a Register Stop to the first hop DR to terminate the registration.

The Source Path Tree is now created between the source S and the RP. The data is flowing from the source S via the Source Path Tree to the RP and over the branch of the Shared Path Tree from the RP to group G.



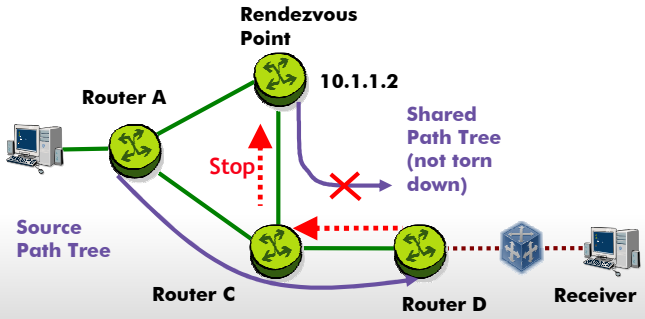
Router D, connected to the receiver, signals PIM messages for a switchover to the Source Path Tree to occur by sending a PIM (S, G) Join to the source .

Router C receives the PIM (S, G) Join and sends a PIM (S, G) Join towards the source S source.

Router A receives the PIM (S, G) Join and determines it is connected to the source. The Source Path Tree is completed and data will flow towards Receiver on the shortest path (IGP).

PIM-SM

```
*A:RouterD# show router pim group
===============================================================================
PIM Groups
===============================================================================
Group Address   Source Address  RP Address      Type     Spt Incoming      Num
                                                         Bit Intf          Oifs
-------------------------------------------------------------------------------
235.2.2.2       *              2.2.2.2          (*,G)       tope2         1
235.2.2.2       150.215.1.99   2.2.2.2          (S,G)       tope1         1
```



The switchover to the Source Path Tree has been completed for Router D, but duplicate packets are once again being received at Router D, one from the newly created Source Path Tree, and a second from the preexisting Shared Path Tree. The Shared Path Tree packets must stop arriving at Router D.

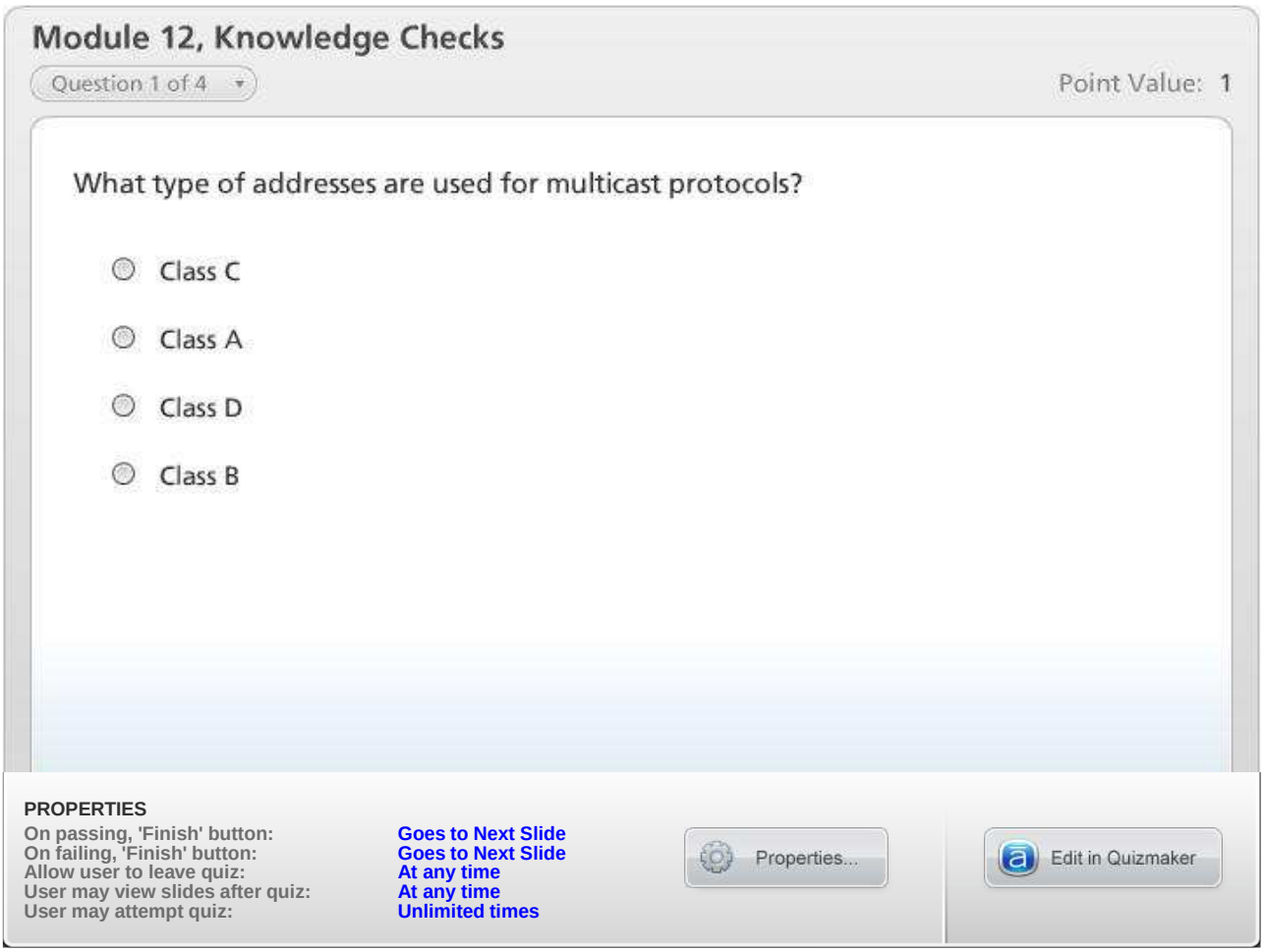
Router C is where the Shared and Source Tree diverge. To eliminate the packets from source S for group G from arriving via the RP, it must send an (S, G) prune message to Router B. Since the source interface is indicated on the interface leading to Router B, the (S,G) prune would be discarded. In order to avoid it , a flag must be set in the prune message indicating that this message is to be treated as an exception.

The duplicate packets have been eliminated and the data is flowing to Receiver via the Source Path Tree. The Shared PT will be maintained by the periodic PIM Join messages originated by Router D and propagated towards the source router for as long as the local receiver is a member of the group.

Knowledge Checks

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempts at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 12

This completes module 12. This module introduced the need for multicast and explained the basic operation of IGMP and PIM.

SR-OS Fundamentals

Module 13: QoS Overview

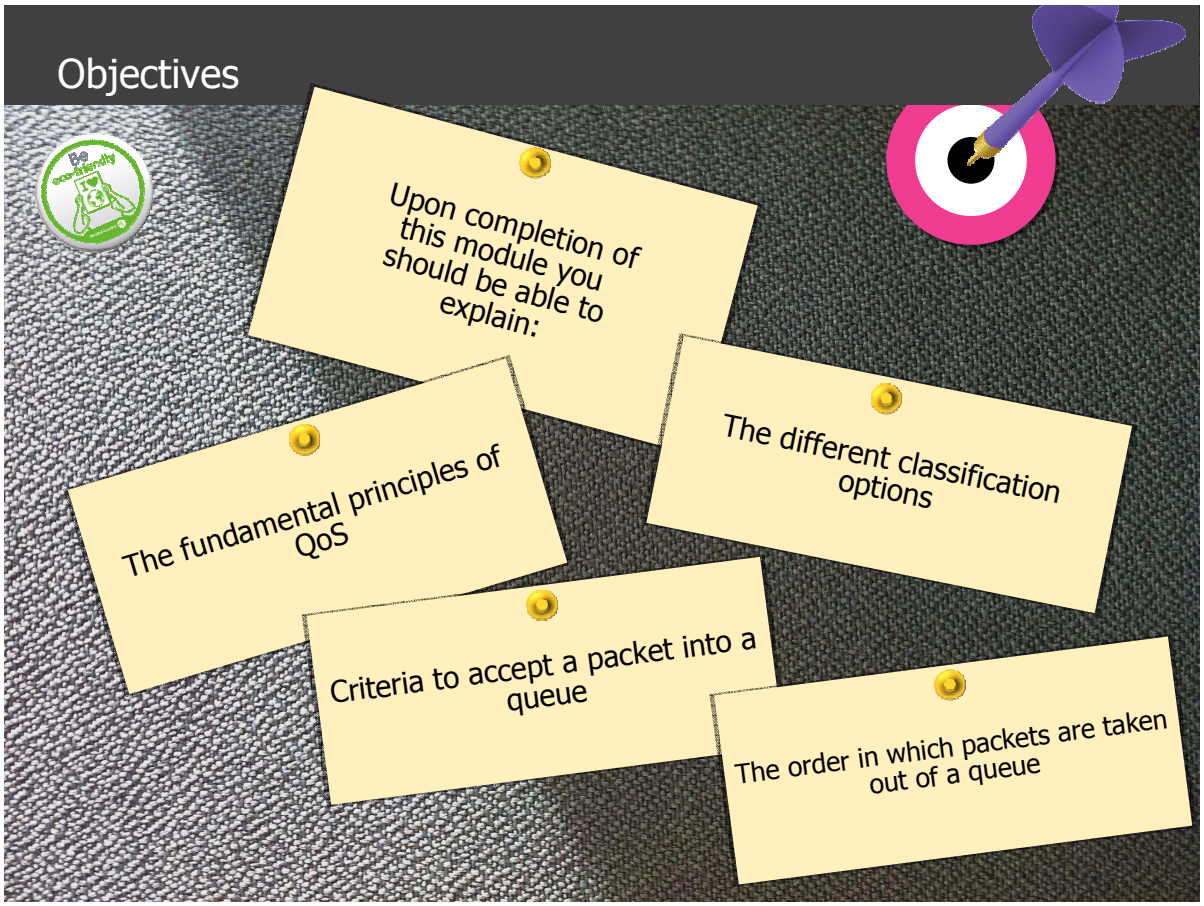
IPD Development

Welcome to the 13th module of the SR-OS fundamentals course.
This module is a first introduction to Quality of Service or QoS.

Agenda

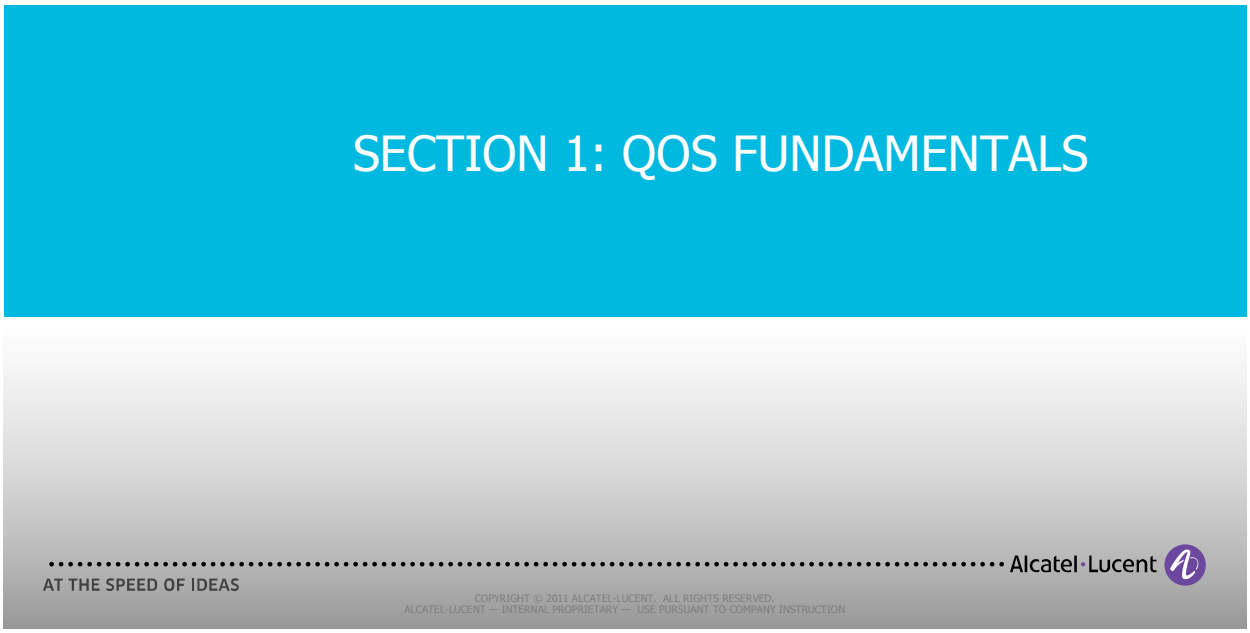
- Module 13:
  - Section 1:
    - QoS fundamentals
  - Section 2:
    - Classification
  - Section 3:
    - Buffers and queues
  - Section 4:
    - Scheduling

Module 13 is divided into four sections.
Section 1 gives an overview of the fundamental principles of QoS.
Section 2 explains the options to do classification.
Section 3 talk about buffers and queues.
And section 4 explains how traffic is scheduled out of the queues



By the end of module 13 you will be able to explain:

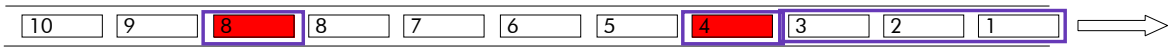
The fundamental principles of QoS, the different classification options, criteria to accept a packet into a queue and the order in which packets are taken out of a queue.



Section 1: QoS fundamentals

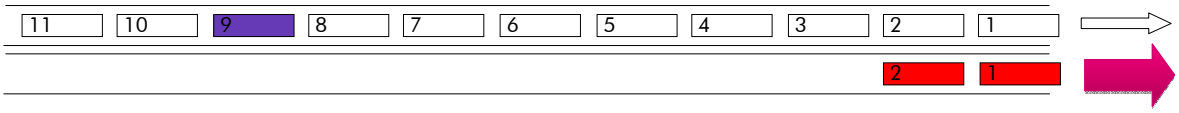
Introduction: Quality of Service Principle: De-serialization

FIFO: First In First Out



Congestion in a FIFO network → High Priority traffic suffers:

- Delay
- Jitter (Variable Delay)
- Packet Loss



QoS Basic tools:

1. Identify the High Priority packets and Classify into correct Forwarding Classes.
2. Queue the classified packets into dedicated hardware buffers.
3. Release or Schedule the packets from respective queues at assigned speeds.

When there is no additional QoS mechanism on a service router configured, the FIFO or First In First Out mechanism would be used. This is what is used by default on a service router and could be sufficient for some types of traffic, but in most cases this leads to packet delay, variable delay or jitter and packet loss.

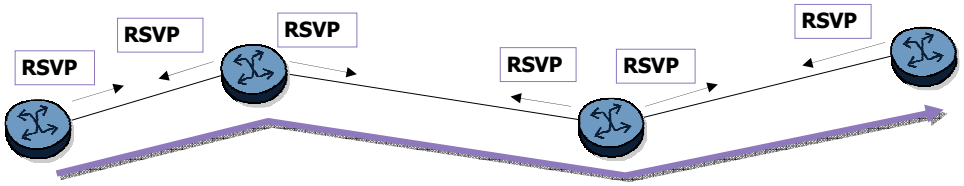
Traffic that comes first in a router or a queue will be sent out first. This sounds like a fair mechanism, but if there is congestion and packets have to wait before getting transmitted, higher priority packets are stuck inside this long frame of packets in a queue. From the moment the congestion gets solved, the packets 1, 2 and 3 in this example are than sent out first while it would be better to serve packet 4 and 8 first, as these are packets with a higher priority.

The solution to this is the de-serialisation of packets like different lanes on a highway. If there is an accident on the road and there is a limited amount of cars that could pass the location of the accident, only one lane could pass through. Which lane and, hence, which cars can pass, depend on the importance of the cars at that moment. This importance is given by the police. For example, an ambulance would get a much higher priority to pass than a normal car.

When talking about packets, this means that there is first an identification process of which packets come into the high priority lane or here we talk about a queue and forwarding class. Which forwarding class gets eventually first priority depends on the mechanism to take out packets out of a queue. In a service router, this task is done by the scheduler.

So the three main basic tasks of QoS are identification of packets, queuing and scheduling.

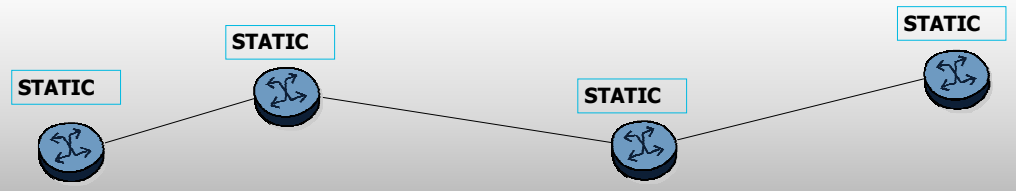
Integrated Service versus Differentiated Services



Integrated Service: A Protocol reserves resources through network dynamically per microflow, such as RSVP.

	Integrated Services (Legacy)	Differentiated Services (7750)
Disadvantage	High Overhead when many microflows	Heavy task for administrator
Advantage	Automatically created QoS by protocol	No protocol overhead, full control

Differentiated Services: An administrator configures statically resources per macroflow, containing multiple microflows.



Reserving resources has to be done along the entire path of routers the packet is sent through.

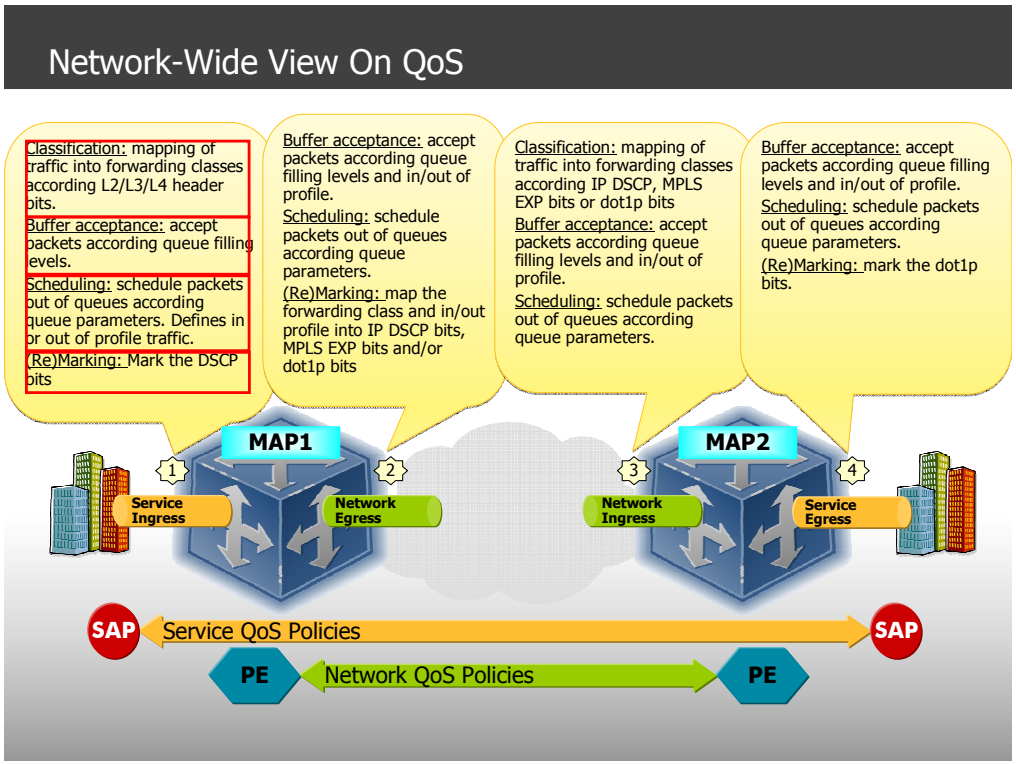
In the highway example, the high priority lane has to be available on the entire highway.

There are two main principles to reserve resources. Integrated services and differentiated services.

Integrated services uses a protocol that signals the resource requirements along the path. This is a fully automatic process that does not require an administrator to configure each router individually. But, the big disadvantage is the high signalling overhead. For each flow or customer data flow, this protocol needs to signal the path upfront. This is absolutely not scalable and therefore not used as the preferred mechanism to reserve resources.

The industry standard today is the differentiated services approach. An administrator configures statically on each service router a way on how the majority of flows or microflows will be map to a macroflow. This macro flow is also called a behaviour aggregate. It can be compared to the lane on the highway. The cars on one lane might have different needs but are mapped to only one lane with a certain throughput. Only three or four lanes are available per highway per direction. On a service router, up to eight behaviour aggregates or macro flows can be configured. There in no protocol and hence no protocol overhead and there is no need to set up resources for each packet flow. However, the set up of this macro flow is a heavy task for the administrator that has to go on each router one by one to set the right configuration. This sounds like a disadvantage but there are methods to optimize this. Like using policies that are controlled by a central management system like the 5620 SAM.

The diff serve or differentiated service approach is the mechanism use on the Alcatel-Lucent service routers.



The four main operations of a QoS mechanism are classification, buffer acceptance, scheduling, and marking or remarking.

This slide explains the four operations on an example of a point to point service from SAP to SAP over an IP/MPLS network.

When a packet is coming from the CE router or switch, it will be classified, accepted into the queue, scheduled out of the queue and maybe remarked at the SAP ingress of the service.

Classification is analyzing the traffic and doing the mapping into classes, called forwarding classes. The criteria of which packet belongs to which forwarding class can be based on the received L2, L3 or L4 header information.

The mapping of packets to forwarding classes is like adding a stamp to a packet. Based on this internal stamping, it will be assigned to a queue or buffer space.

Depending on the utilisation of the buffer space and parameters of the queue, the packet is or is not accepted. This process is called buffer acceptance.

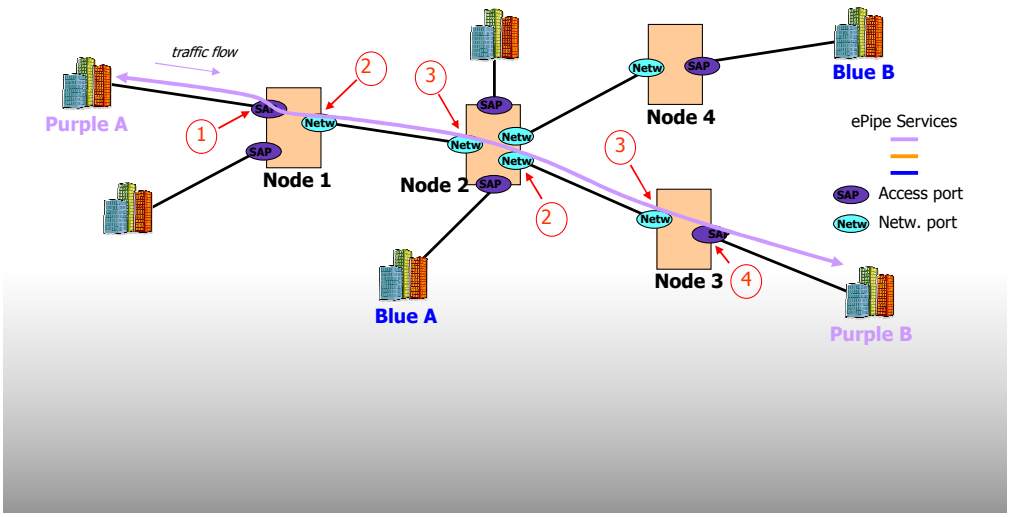
Once each packets belongs to a certain forwarding class and is accepted into a queue, the packet is taken out of the queue by the scheduler. This scheduler is an important function, as it decides which queue to serve first.

Optionally the packet can get a marking based on the forwarding class it was mapped to. If a packet already got a marking, it can be remarked or left untouched.

The process of classification on multiple header fields happens only on the SAP ingress side of a router. Once it is classified, it will gets its forwarding class for the entire data path. Classification on subsequent routers is not done on multiple field but on the LSP exp, bits, dscp or dot1q bits.

Buffer acceptance, scheduling and marking process is also performed at the egress side of each router after passing the switch fabric.

Overview of Different QoS Policies



In a real life example, there are more service routers and services in place.

Let's have a look to the blue packet flow.

Traffic is classified into one or more forwarding classes and placed into queues. Classification and queue characteristics are set by the SAP-ingress policies [*number 1*]. Queuing is done on a per-SAP basis.

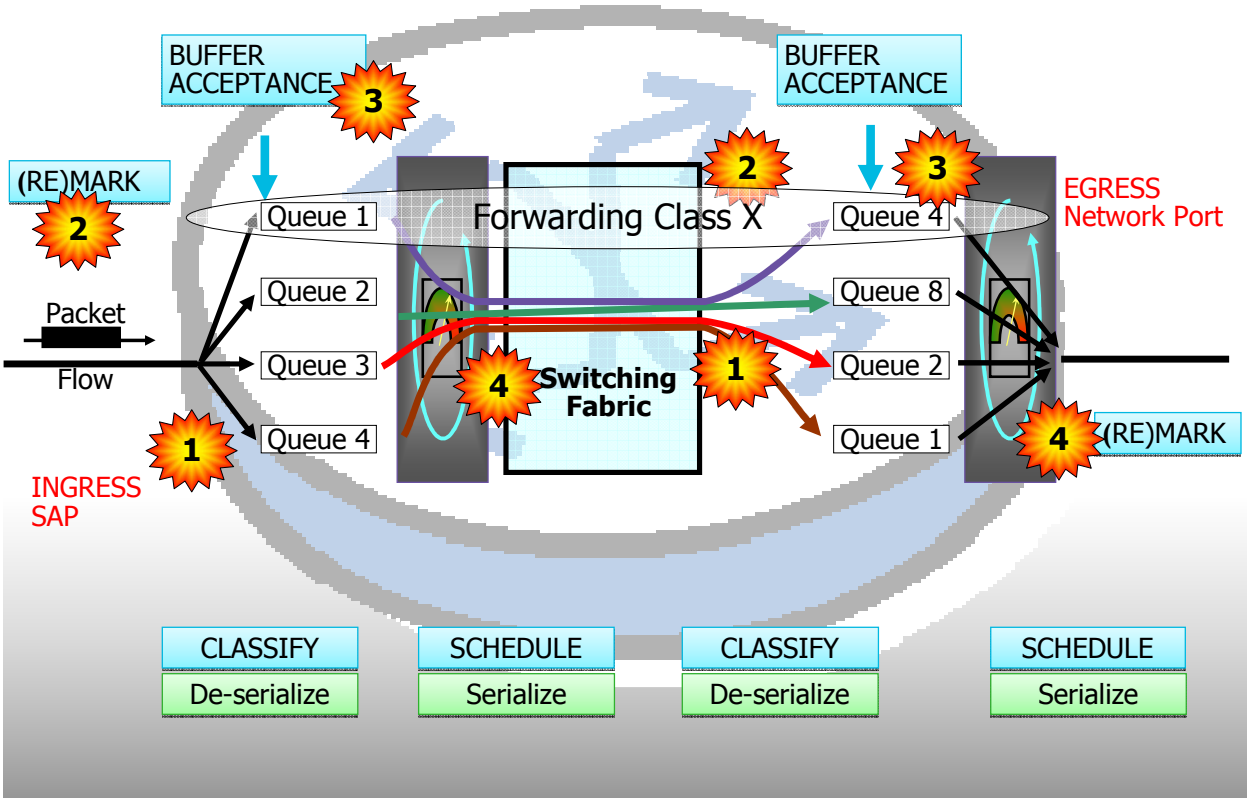
The Traffic forwarding class is mapped to the tunnel header QoS markings, EXP, DSCP and dot1p bits, as defined in a network QoS policy [*number 2*]. All traffic belonging to the same FC on a single port is queued together. Network queue parameters are defined in a network-queue QoS policy.

On node 2 [*number 3*], traffic is mapped to FCs based on tunnel header markings. Tunnel header to FC mapping is defined in a network QoS policy. All traffic belonging to the same FC on a single forwarding complex is queued together.

The same process is happening at the egress of node 2 and ingress of node 3 [*number two and three emphasis AGAIN*].

SAP-Egress QoS policies [*number 4*], define queuing and dot1p packet marking based on the traffic FC. Queuing here is also done per SAP.

QoS 7750 Architecture



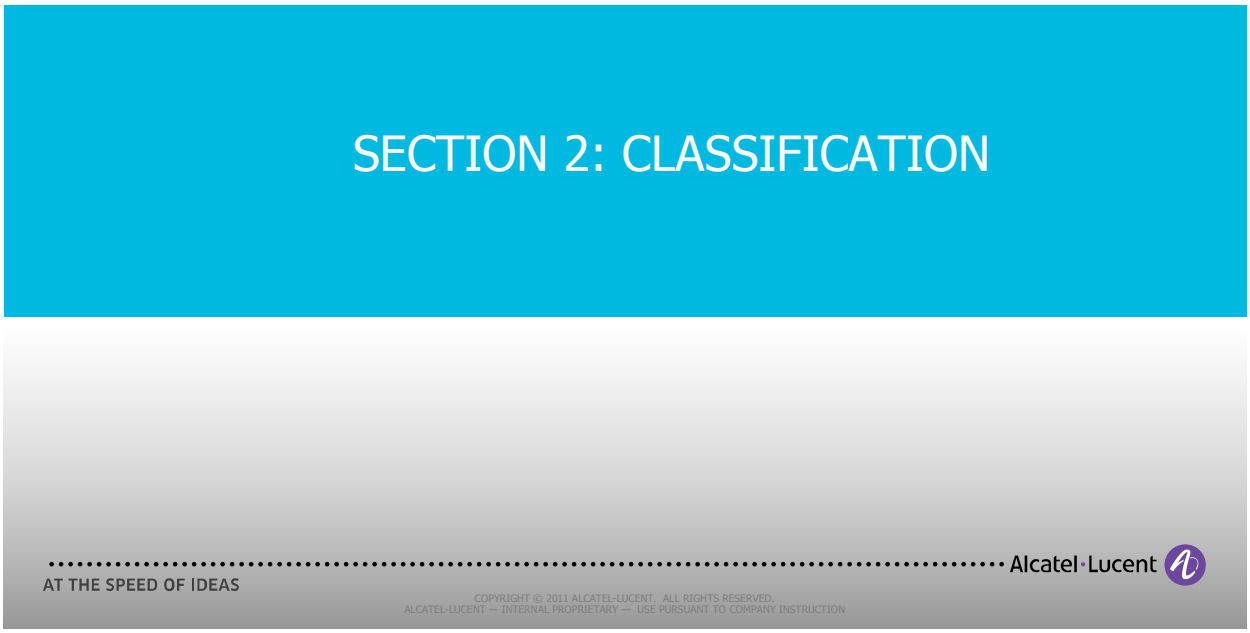
This slide shows the whole process of QoS on one router. Two times the four operational processes of classifying, marking, buffer acceptance and scheduling.

The first two operations are the classification and/or marking or remarking of the packets.

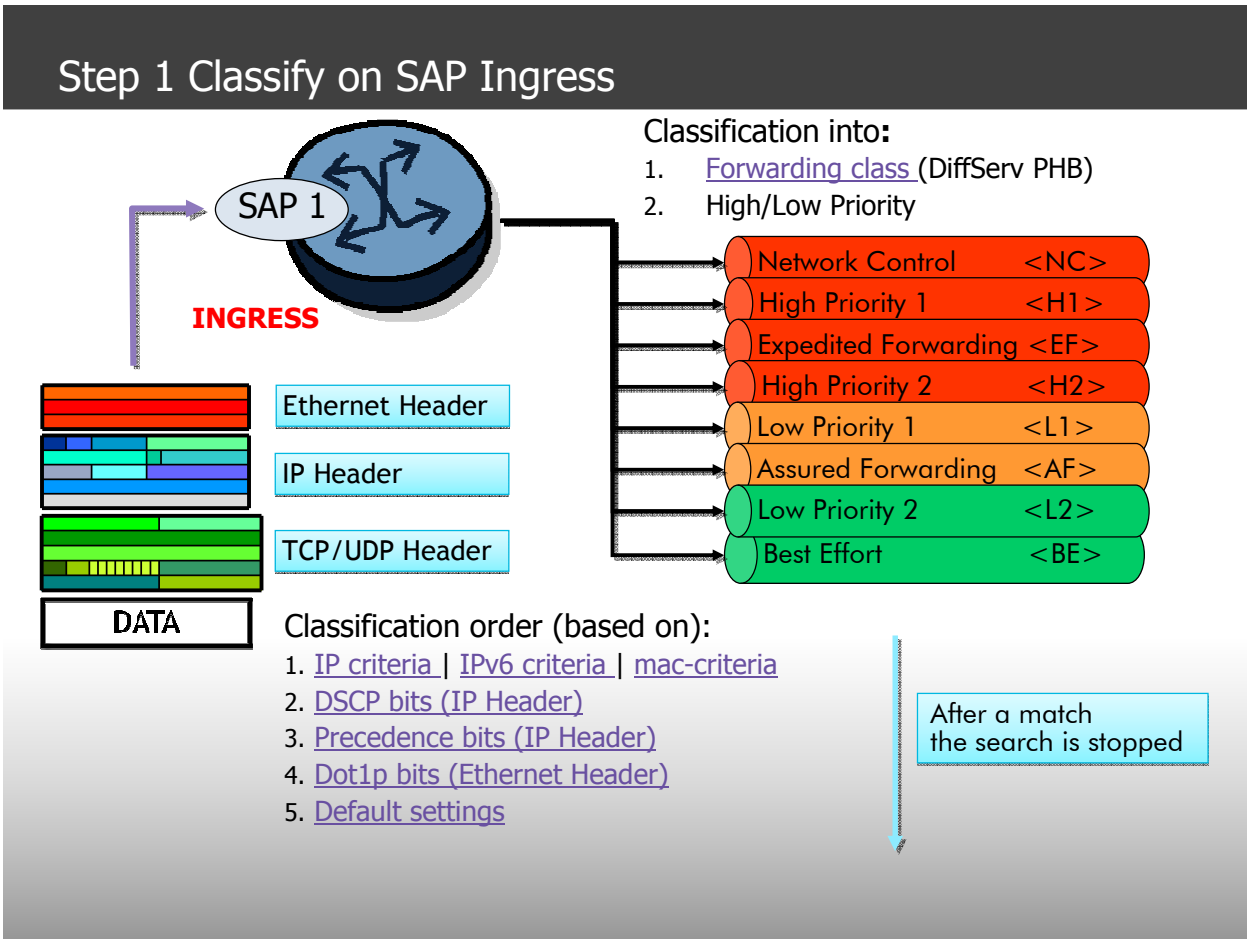
Thirdly, packets are accepted into the appropriate queues or not.

As a last operation, packets are taken out of the queues.

Once the packets have been sent over the switch fabric, this process starts again, but classification is very limited. It is just a matter of mapping forwarding class to a queue.



Section 2: Classification



Traffic arriving on a SAP will be classified based on the SAP ingress policy. This policy can contain a couple of criteria.

First the IP or MAC criteria are analyzed, then the DSCP bits, precedence bits, dot1p bits and when no match is found the default forwarding class is taken. There is always a default forwarding class specified. So when no criteria are configured in the policy by the administrator or no match is found for a certain packet, the default forwarding class setting is taken. The order of classification is a fixed rule defined on the service router. The order is important because whenever there is a match found, the traffic is mapped to this specified forwarding class and the process of analyzing all the criteria one by one is stopped.

Eventually, all traffic will be mapped to any of the 8 forwarding classes. There is also a high/low priority tag added to the packet which has an effect on the queue acceptance later on. So, the forwarding class is not the only important aspect of defining the importance of a packet, there is an additional level inside the forwarding class which is called the priority.

Forwarding Classes (FC)

FC	FC Name	Class type	Notes
NC	Network control	Real time	For network control traffic.
H1	High-1	Real time	For delay/jitter sensitive data
EF	Expedited	Real time	For delay/jitter sensitive data
H2	High-2	Real time	For delay/jitter sensitive data
L1	Low-1	NRT – Assured	For assured traffic
AF	Assured	NRT – Assured	For assured traffic
L2	Low-2	NRT – Best effort	For BE traffic
BE	Best Effort	NRT – Best effort	For BE traffic

Up to 56 forwarding sub-classes available

Eight forwarding classes are defined on a service router divided into two categories: the expedited forwarding classes, NC, H1, EF and H2; and the best effort forwarding classes, L1, AF, L2 and BE. The two categories are important later on when talking about the scheduling order.

The class type and notes field give an idea which type of traffic is mapped to each of the forwarding classes. But this is best practices and only a suggestion. The QoS designer has the flexibility to define his own rules of classification and scheduling.

There are options to firstly classify the traffic into sub-classes. But these will eventually be mapped in to one of the eight main classes. Up to 56 sub-classes are available. The main purpose for this is the pre-marking of a packet or to make a difference in priority for a different packet belonging to the same forwarding class.

Multi-Field (MF) Classification on SAP Ingress: IP-Criteria

```
A:7750# configure qos sap-ingress 10 ip-criteria entry 10 match ?

  - match [protocol <protocol-id>]
  - no match

 <protocol-id>        : [0..255] - protocol numbers accepted in DHB
                        keywords - none|crtp|crudp|egp|eigrp|encap|ether-ip|
                        gre|icmp|idrp|igmp|igp|ip|ipv6|ipv6-frag|ipv6-icmp|
                        ipv6-no-nxt|ipv6-opts|ipv6-route|isis|iso-ip|l2tp|
                        ospf-igp|pim|pnni|ptp|rdp|rsvp|stp|tcp|udp|vrrp


 [no] dscp            - Specify DSCP match
 [no] dst-ip          - Specify destination IP and mask match
 [no] dst-port        - Specify destination TCP/UDP port match
 [no] fragment        - Specify match criteria applies to IP fragments
 [no] src-ip          - Specify source IP and
 [no] src-port        - Specify source TCP/UD
```

```
sap-ingress 500 create
    queue 1 create
    exit
    fc "ef" create
        queue 1
    exit
    ip-criteria
        entry 10 create
            match
                dst-ip 10.0.0.0/8
            exit
            action fc "ef"
        exit
    exit
```

IPv4 header

Let us have a look to one of the criteria options in the classification process, the IP criteria.

Inside a SAP ingress policy, different entries are possible with different IP criteria. These entries have a number and the lowest number is analyzed first. Inside an entry, one or several match criteria are found. If more than one is set, this works as an AND function and there is only a match if the packet does match all the criteria.

IP-criteria can be a protocol criteria like "match protocol IGMP" or dscp criteria like "match dscp value of 34" or a destination or source IP addresses like "match IP address range 10.10.10.0/24".

Destination or source port numbers are also options in the classification. The same goes for the fragment flag in the header of an IP packet.

The example of SAP ingress policy 500 matches all packets with the destination IP address range of 10.0.0.0/8 to the forwarding class "ef".

This packet is than put into queue number one. Queue one and the relation between the queue and the forwarding class has been configured upfront in the SAP ingress policy. Other traffic not matching entry 10 in the IP criteria are mapped to the "be" forwarding class as this is the default setting.

Multi-Field (MF) Classification on SAP Ingress: IPv6-Criteria

```
A:7750# configure qos sap-ingress 10 ipv6-criteria entry 10 match ?

  - match [next-header <next-header>]
  - no match

 <next-header>        : [1..42|45..49|52..59|61..255] - protocol numbers
                        accepted in DHB
                        keywords - none|crtp|crudp|egp|eigrp|encap|ether-ip|
                        gre|icmp|idrp|igmp|igp|ip|ipv6|ipv6-icmp|
                        ipv6-no-nxt|isis|iso-ip|l2tp|
                        ospf-igp|pim|pnni|ptp|rdp|rsvp|stp|tcp|udp|vrrp


 [no] dscp            - Specify DSCP match
 [no] dst-ip          - Specify destination IPv6 address and mask match
 [no] dst-port        - Specify destination TCP/UDP port match
 [no] src-ip          - Specify source IPv6 add
 [no] src-port        - Specify source TCP/UDP
```

IPv6 headers

```
sap-ingress 500 create
    queue 1 create
    exit
    fc "ef" create
        queue 1
    exit
    ipv6-criteria
        entry 10 create
            match protocol udp
            src-port 1234
            exit
            action fc "ef"
        exit
    exit
```

Also IPv6 criteria can be used as a matching criteria for classification.

In the example of SAP ingress policy 500, all traffic matching protocol UDP AND use source port 1234 will be mapped to forwarding class "ef" and gets queued at queue number 1.

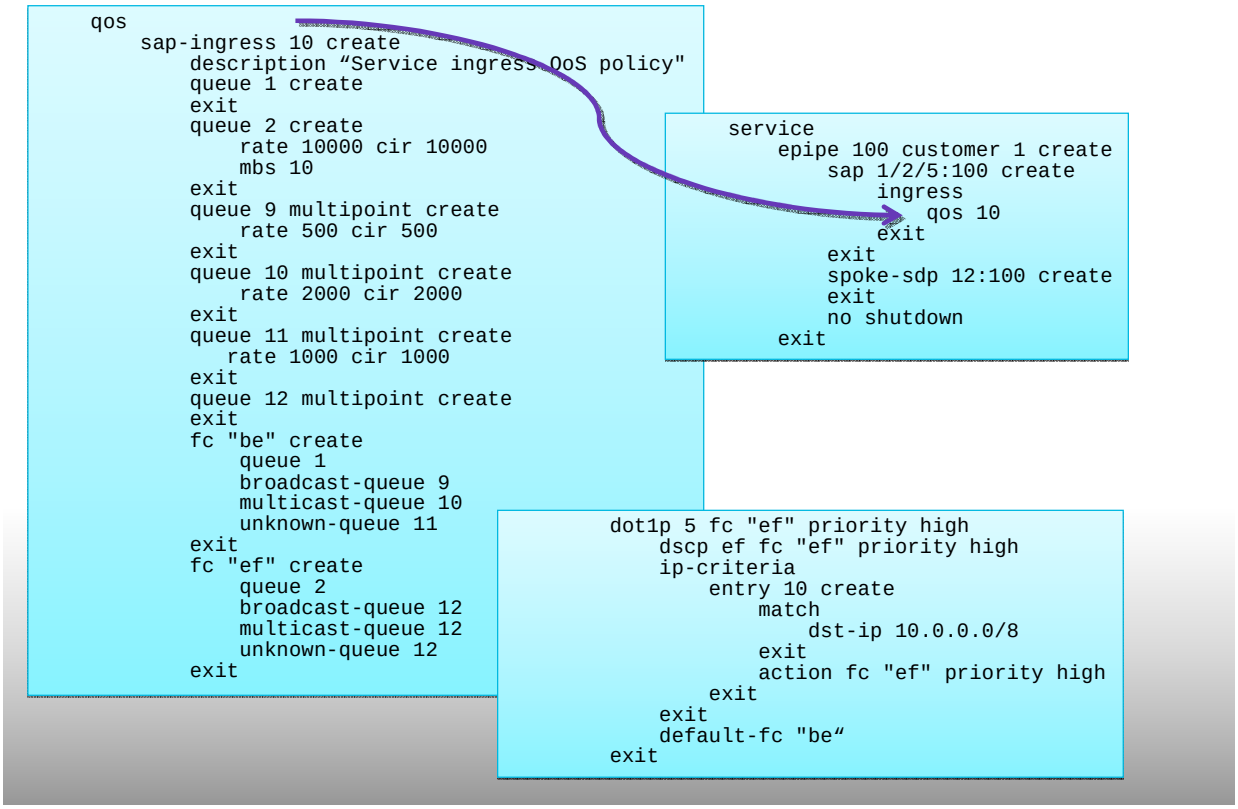
SAP ingress example

```
Configure qos  sap-ingress 88 create
            queue 1 create
            exit
            queue 2 profile-mode create
                rate 15000 cir 5000
            exit
            queue 3 create
            exit
            queue 11 multipoint create
            exit
            fc "h2" create
                queue 2
            exit
            fc "l1" create
                queue 3
           exit
            ip-criteria
                entry 10 create
                    match protocol udp
                        dst-port eq 5015
                    exit
                    action fc "h2" priority low
                exit
                entry 20 create
                    match protocol udp
                        dst-port eq 5020
                    exit
                    action fc "l1" priority low
                exit
            exit
        exit
```

Another example of a SAP ingress policy where up to four queues and two forwarding classes are configured. Two entries are set in the IP criteria. First entry ten is analyzed. All traffic matching protocol UDP and destination port equals 5015 will be mapped to forwarding class "h2" and gets its queue acceptance priority of "low". If no match is found match, entry 20 is analyzed. This entry has a match criteria of protocol UDP and destination port equals to 5020. The action is set to forwarding class "l1". Traffic that was mapped to forwarding class h2 is queued into queue 2 and traffic that is mapped to forwarding class l1 is queued to queue 3. Queue 1 and 3 need to be created before the IP criteria maps to any of the two queues.

Other traffic not matching the IP criteria will be mapped to "be" and get queued to the default queue number one.

Applying the Qos SAP ingress policy



Once created, the SAP ingress policy needs to be applied to take effect. This is done under the service, under the SAP and in the correct direction. Here the direction is ingress.

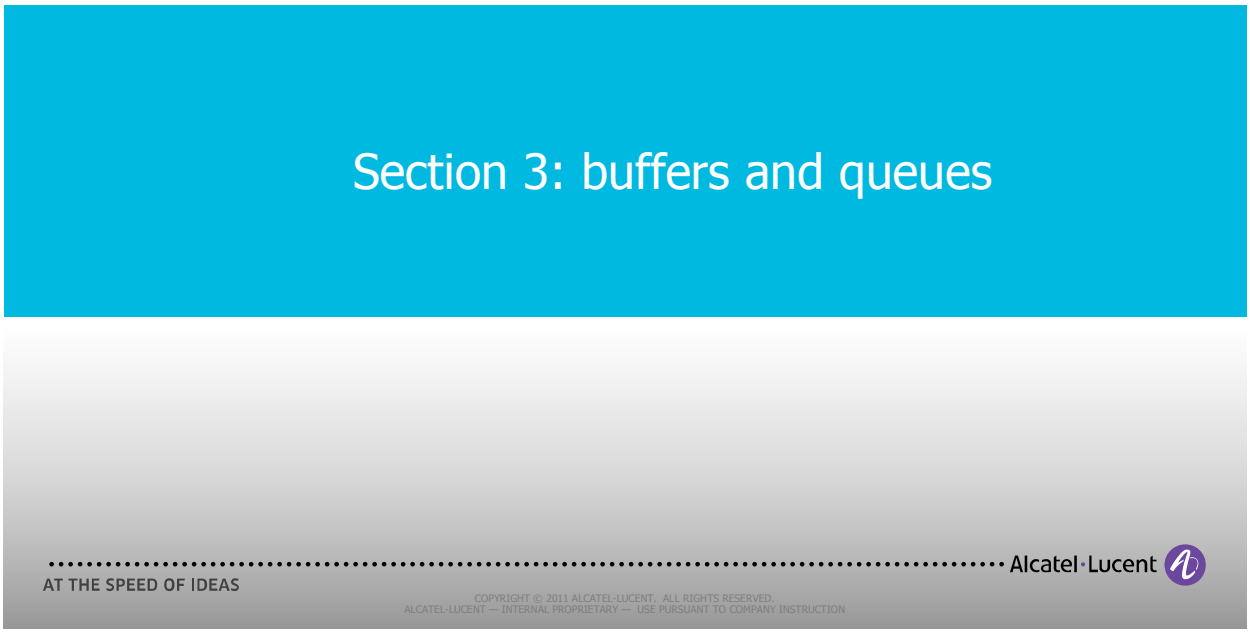
Default Classification on SAP Ingress

- default-fc
  - If no other rule applies, the packet is classified as belonging to the default-fc forwarding class (default = be)
- default-priority:
  - If no other rule applies, the packet is classified as having queuing priority = default-priority (default = low)
  - Notes:
    - The queuing priority of a packet determined at service ingress classification is only used at the service ingress buffer acceptance phase (see further)
    - Do not confuse with scheduling priorities

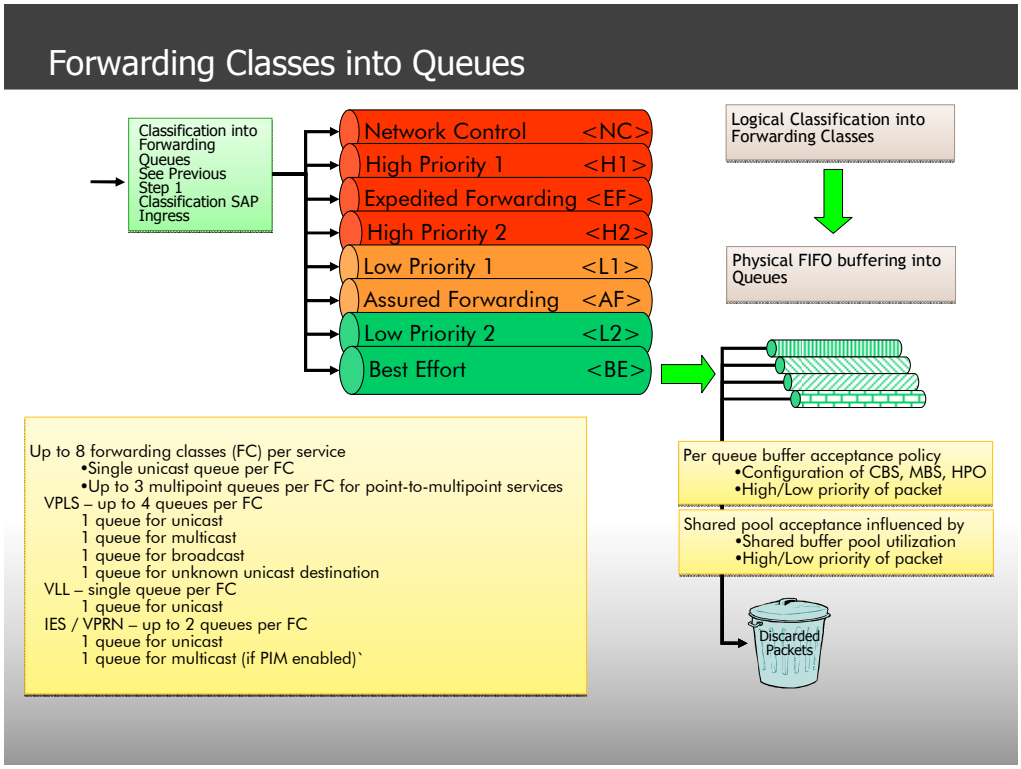
```
qos
    sap-ingress 12 create
        …
        default-fc "ef"
        default-priority high
    exit
```

The default setting of the default forwarding class is set to "be". However, this default setting can be set to any of the eight forwarding classes.

The default queue acceptance priority is set to "low". This can be changed to "high" by configuration. The queuing priority of a packet determined at the service ingress classification is only used at the service ingress buffer acceptance phase, which will be explained in more detail in the next section.



Section 3: Buffers and queues



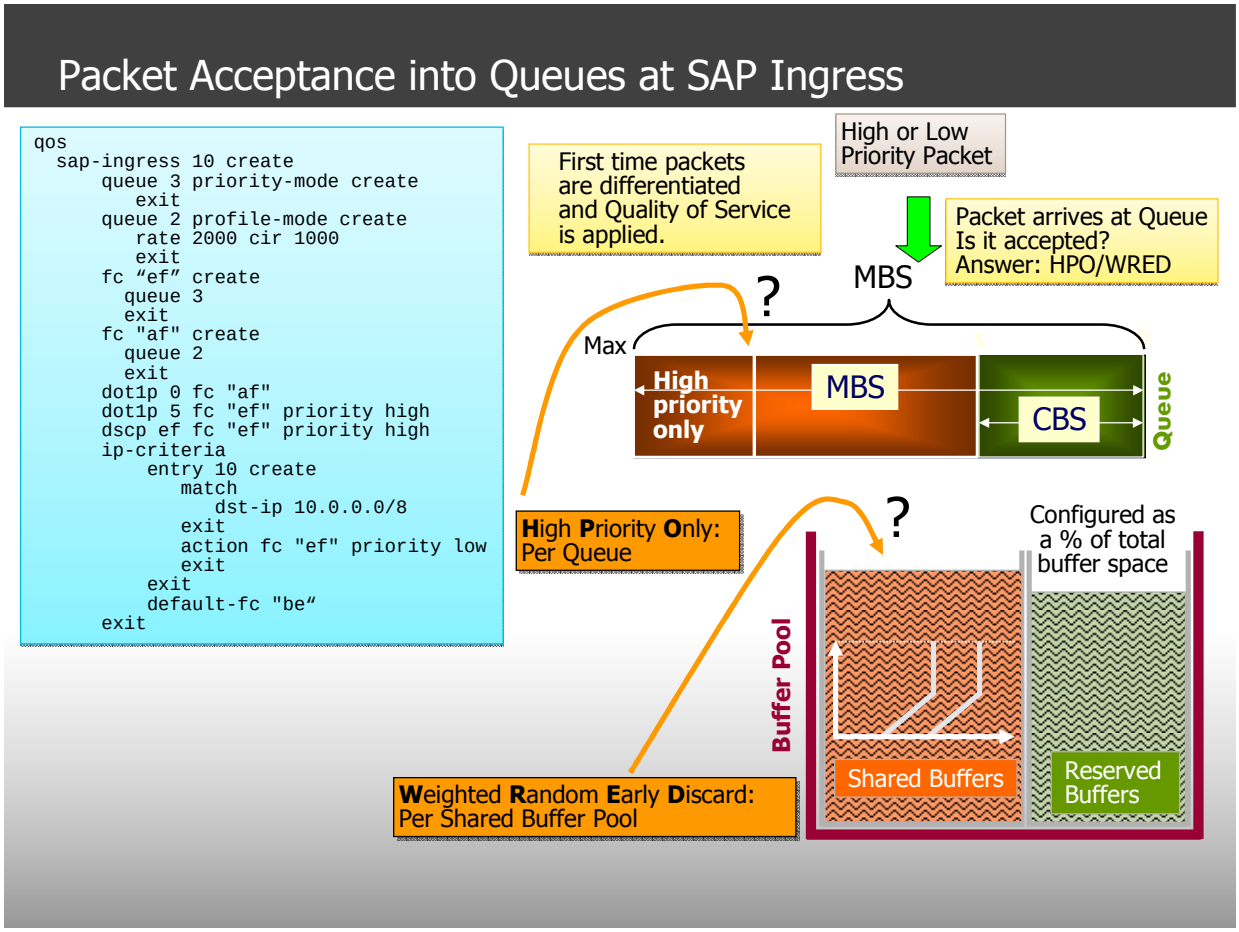
The former section explains the logical classification of traffic into eight forwarding classes. Let's now have a closer look at how the packets are accepted and buffered into the queues. In order to understand this, new parameters that describe a queue are introduced like CBS, MBS and HPO. The parameter CBS or committed burst size defines the committed size or length of a queue. The MBS or Maximum Burst Size is the maximum size of a queue but MBS – CBS is a portion of a queue that is not committed to be available at a certain time. It is taken from a shared portion of the available pool of memory, called the shared buffer space.

The acceptation of low and high priority packets into a queue depends on the overall utilization of the shared buffer space, queue depth and the HPO value of the queue. The shared buffer space is always shared between queues.

The usage of the shared buffer space and queue parameters are explained in more detail later on.

Let us first look to how many queues a forwarding class can be mapped to. The answer is one. However, based on the nature of the traffic, up to four different queues can be used for the same forwarding class. The nature of the traffic means if the traffic is unicast, multicast, broadcast or unknown unicast traffic. Within a VPLS service all of these types of traffic could be available, meaning that four queues per forwarding class are possible. This means for eight forwarding classes, 32 queues per SAP can be configured.

A VLL or point to point service does not care about the different types of traffic and has 1 queue per forwarding class. This is because of the point to point type of service. IES and VPRN do have an additional queue per forwarding class available when using PIM multicast.



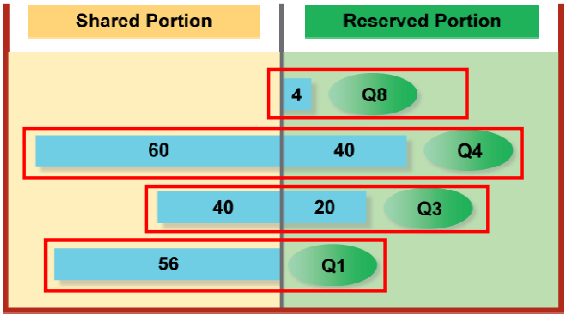
Traffic in the SAP-ingress policy with destination IP address 10.10.10.1 will be mapped to forwarding class "ef" and get a priority tag of "low". This traffic is than put into queue 3. But is the traffic accepted in to the queue at all times? Well to answer this question we need to further explain the parameters of the queue and the priority setting.

The queue space is taken out of a general pool of memory. This memory is assigned per MDA or per IOM. The memory is also divided into a shared and into a reserved portion. How much space the reserved and shared portion takes can be configured.

Now each queue takes a portion out of the reserved part based on the CBS value of the queue and a part is taken out of the shared portion based on the MBS minus the CBS value. If not oversubscribed, the CBS is committed and available at all times for that particular queue. The MBS minus the CBS space is shared and therefore not guaranteed to be available. It depends on other queues that might have taken some portion of the shared buffer space at a certain point in time.

The HPO or High Priority Only parameter expressed in percent is the portion of the queue that is only reserved for the high priority traffic. So if set to 10%, 10% of the queue can only be used by traffic that was tagged as high priority traffic in the SAP ingress policy.

Queues and Buffer Pool



- On SAP-ingress — expressed in KB
- On network port/MDA — expressed as percentage of total port/MDA buffer pool

```
7x50>config>qos>network-queue#
info
    queue 4 create
        mbs 50
        cbs 15
    exit
    fc "af" create
        queue 4
    exit
```

Percent

```
7x50>config>qos>sap-ingress# info
    queue 1 create
        mbs 56
        cbs 0
    exit
    queue 3 create
        mbs 60
        cbs 20
    exit
    queue 4 create
        mbs 100
        cbs 40
    exit
    queue 8 create
        mbs 4
        cbs 4
    exit
    fc "af" create
        queue 4
    exit
    fc “h2" create
        queue 4
    exit
    fc “l1" create
        queue 3
    exit
    fc “l2" create
        queue 1
    exit
    fc “nc" create
        queue 8
    exit
```

Kilo Byte

A couple of examples of queue configurations. The MBS and CBS are expressed in kilobyte. On a network port or MDA this is expressed as a percent of the total port or MDA buffer pool.

Queue 8 has an MBS and CBS of 4 which takes four kilobyte out of the reserved portion and zero out of the shared portion because MBS minus CBS equals zero.

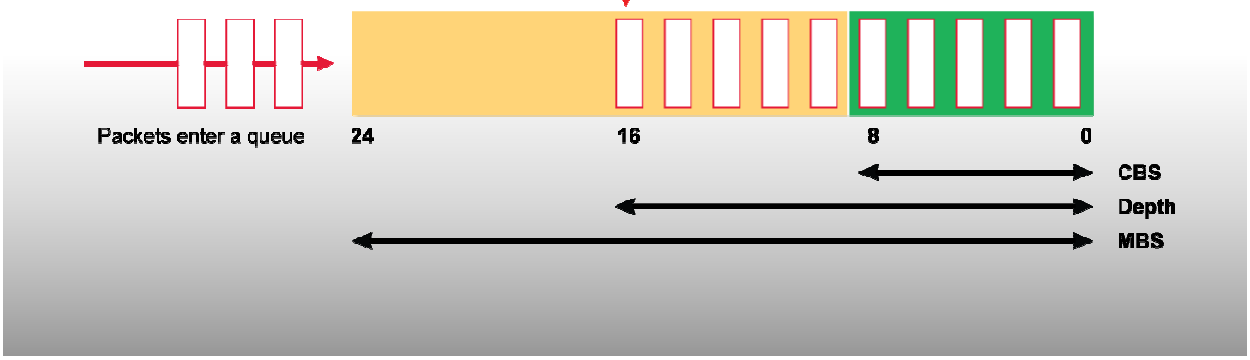
Queue 4 has an MBS of 100 and CBS of 40 which takes 100 kilobyte out of the reserved portion and 60 out of the shared portion because MBS minus CBS equals 60.

Queue 3 has an MBS of 60 and CBS of 20 which takes 60 kilobyte out of the reserved portion and 40 out of the shared portion because MBS minus CBS equals 40.

Queue 1 will take its space out of the shared portion which means that there is no guarantee that traffic can be queued.

Queue Depth or Utilization by Other Packets

```
7x50# show pools 1/1/6 access-ingress

***** output omitted *****

-------------------------------------------------------------------------
Name                  FC-Maps       MBS       HP-Only A.PIR    A.CIR
                                    CBS       Depth   O.PIR    O.CIR
-------------------------------------------------------------------------
88->1/1/6->4          af            24        8       1000000  0
                                    8         16       Max      0
=========================================================================
```

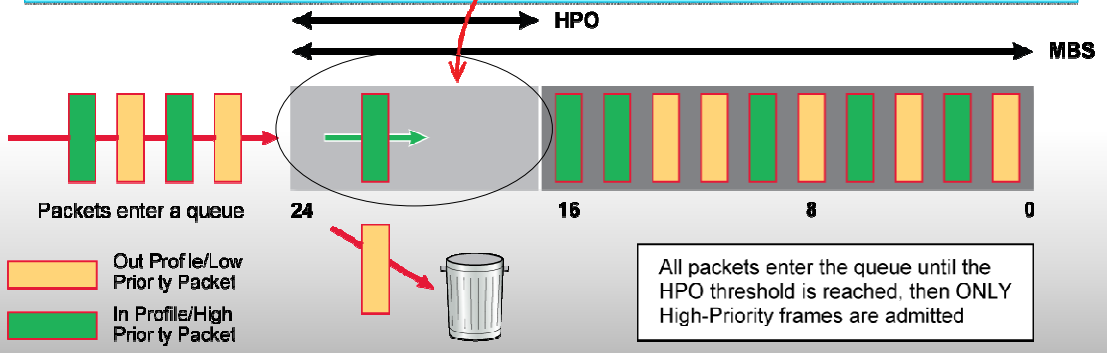


The above slide illustrates how packets arrive and are stored one by one in a queue. First the Reserved buffer space is used by the packets and, when more buffer space is needed, the queue borrows buffer space from the shared buffer pool, if available, up to the MBS value. Once a packet is accepted in a queue, it will eventually be serviced, possibly with added delay.

The **`show pools 1/1/6 access-ingress`** command gives the actual depth of the queue. This is what is consumed at time of entering the command, - in our case 16Kbytes. These will be expressed in blocks of 4KB on an IOM 2, and 3KB on an IOM 3.

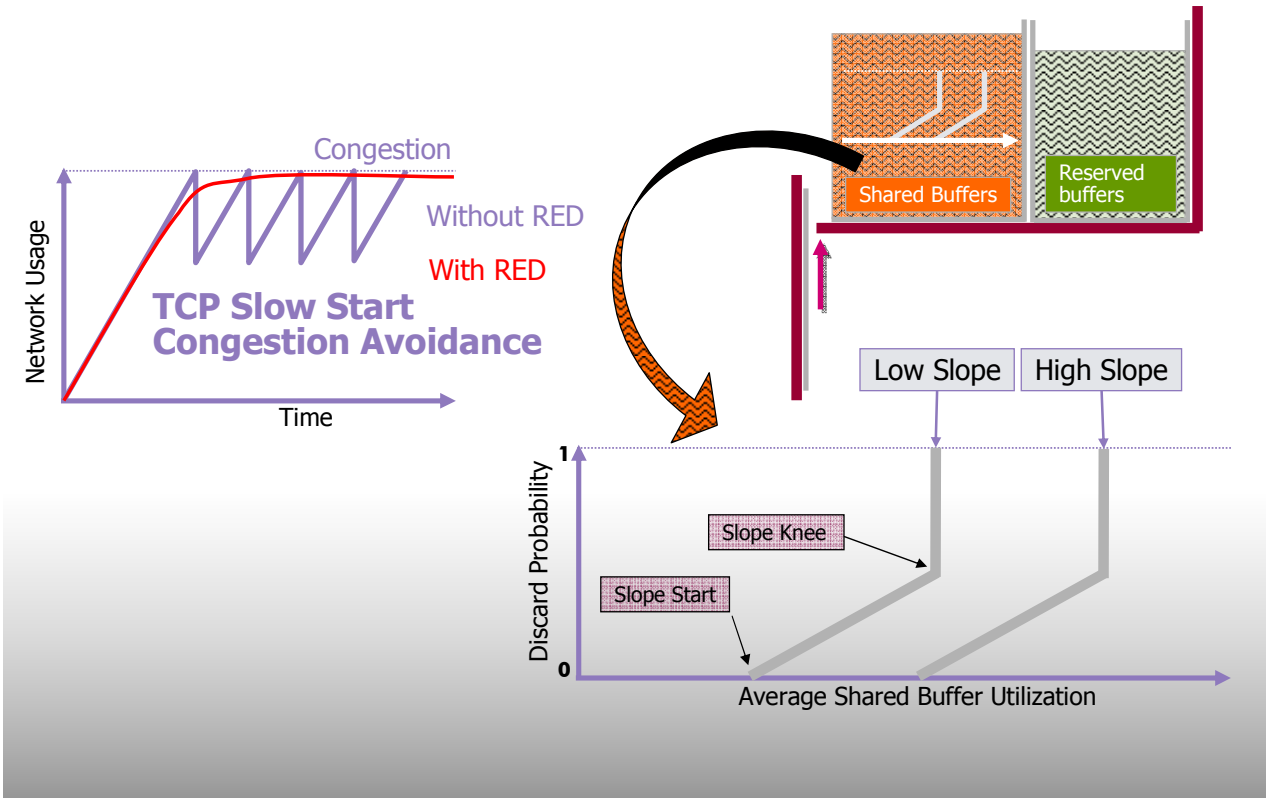
Buffer Admission Control — High Priority Only

```
7x50# show pools 1/1/6 access-ingress

***** output omitted *****

-------------------------------------------------------------------------
Name                  FC-Maps       MBS        HP-Only A.PIR    A.CIR
                                    CBS        Depth   O.PIR    O.CIR
-------------------------------------------------------------------------
88->1/1/6->4          af            24         8       1000000  0
                                    8          16       Max      0
=========================================================================
```



The High Priority Only parameter specifies the amount of buffer space that is reserved for queuing high priority only traffic. HPO is specified as a percentage of the MBS value.

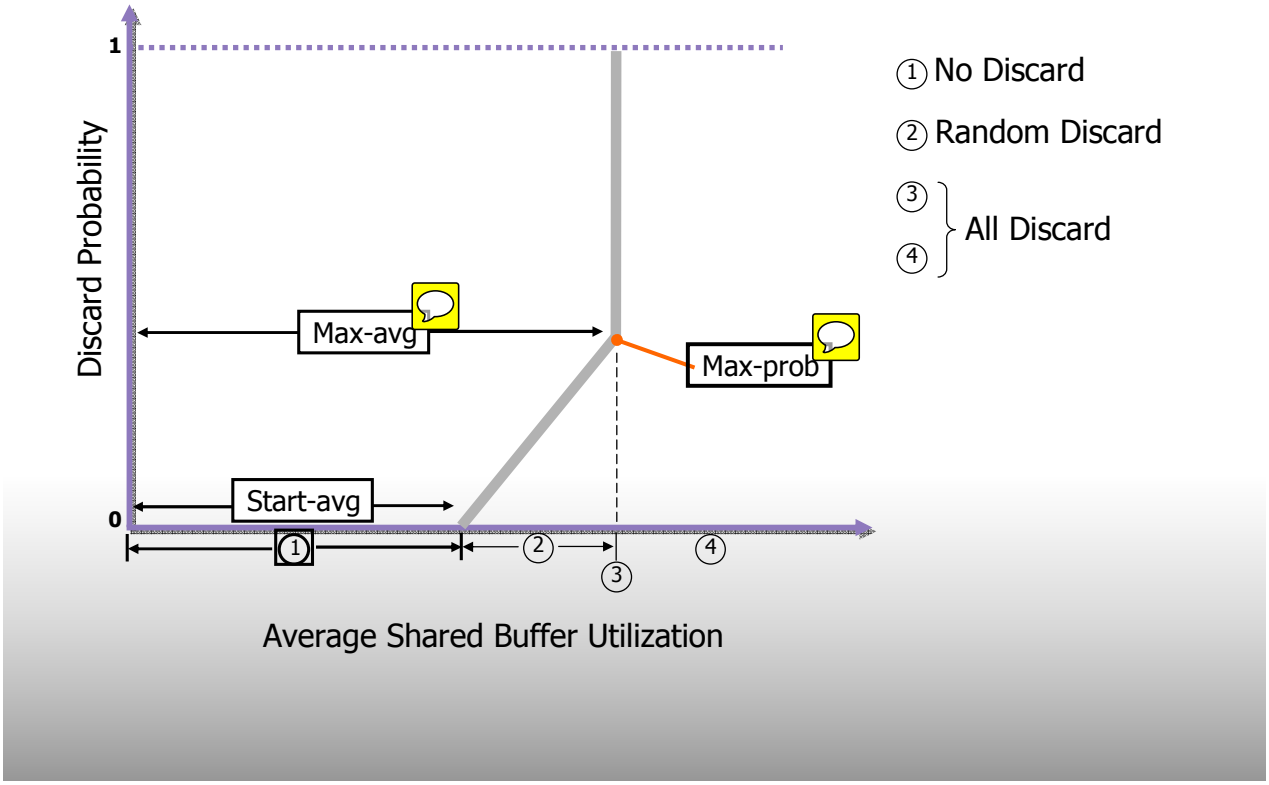
WRED Shared Buffer Management



Queue acceptance also depends on the availability of a slope policy on the shared buffer space. A slope policy, by default disabled, solves the problem of the slow start congestion avoidance mechanism which is typical for TCP traffic. It implies that traffic will be discarded upfront even if there is space available for queuing. Two slopes are available within one slope policy, one for the low priority traffic and one for the high priority traffic.

The slope policy defines the probability in relation with average utilization of the shared buffer when to drop a packet or not.

WRED Slope Configuration Parameters



Looking at one slope, there is the area of absolutely no discard, an area where randomly packets are discarded and an area where all packets are discarded.

WRED CLI example

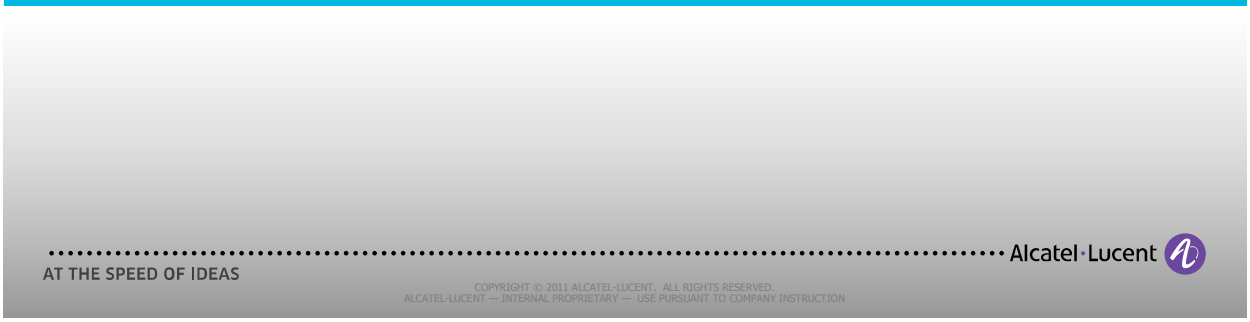
```
port 9/2/1
    network
        egress
            pool
                slope-policy "myred"
            exit
        exit
    access
        egress
            pool
                slope-policy “myred”
            exit
        ingress
            pool
                slope-policy “myred”
            exit
    no shutdown
    exit
exit
card 1 mda 1
    network
        ingress
            pool
                slope-policy “myred”
            exit
    no shutdown
exit
```

```
qos
    slope-policy "myred" create
        high-slope
            start-avg 55
            max-avg 100
            max-prob 100
            no shutdown
        exit
        low-slope
            start-avg 40
            max-avg 50
            max-prob 100
            no shutdown
        exit
```

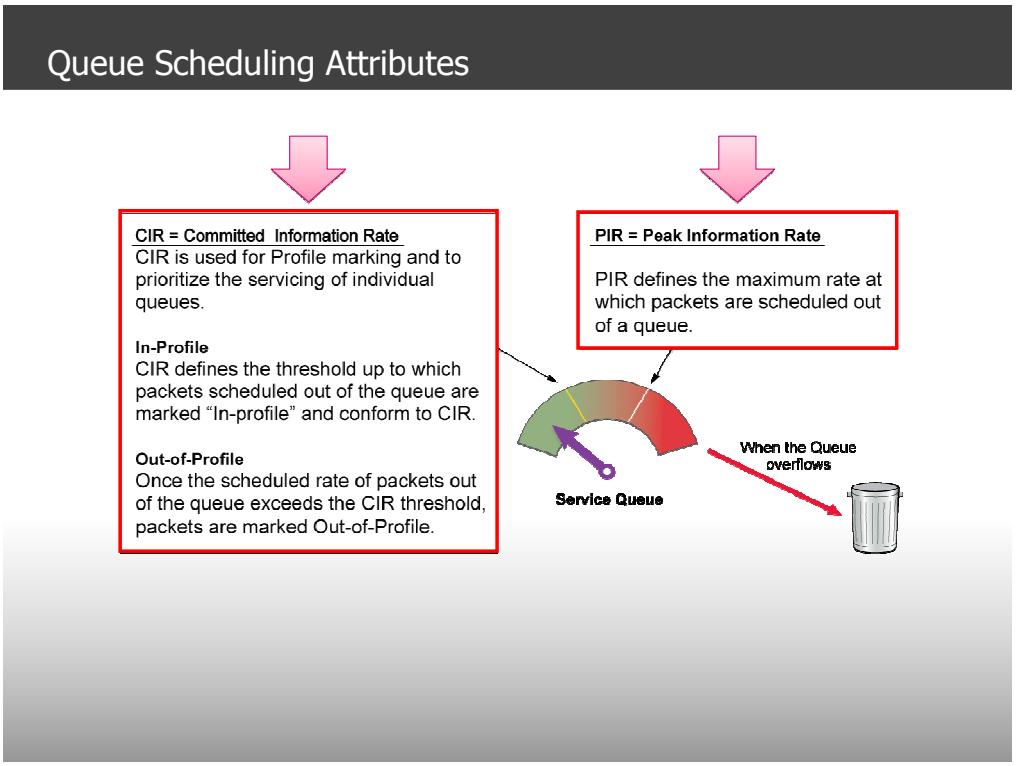
```
A:7750# show pools network-egress 9/2/1

===============================================================================
Pool Information
===============================================================================
Port                : 9/2/1
Application         : Net-Egr           Pool Name          : default
Resv CBS            : Sum
-------------------------------------------------------------------------------
Utilization                 State       Start-Avg    Max-Avg    Max-Prob
-------------------------------------------------------------------------------
High-Slope                  Up              55%       100%        100%
Low-Slope                   Up              40%        50%        100%

Time Avg Factor     : 7
Pool Total          : 20480 KB
Pool Shared         : 12288 KB          Pool Resv          : 8192 KB

High Slope Start Avg : 8192 KB          High slope Max Avg : 12288 KB
Low Slope Start Avg  : 5120 KB          Low slope Max Avg  : 6144 KB

Pool Total In Use    : 10240 KB
Pool Shared In Use   : 8448 KB          Pool Resv In Use   : 1792 KB
WA Shared In Use     : 8446 KB

Hi-Slope Drop Prob   : 7                Lo-Slope Drop Prob : 100
```

This example shows the configuration of the slope policy “myred”, the context where to apply it and the verification command.

SECTION 4: SCHEDULING



Section 4: Scheduling



Scheduling is taking packets out of a queue and deciding amongst different queues which one to serve first.

Before the two main schedulers are explained it is important to understand the Committed Information Rate or CIR parameter configured at each queue and the Peak Information Rate or PIR.

The CIR for a queue influences the scheduling priority .The scheduler prioritizes individual queues based on their current CIR and PIR states. Queues operating at or below their CIR are always serviced before queues operating above their CIR.

The CIRs for ingress and egress service queues are provisioned within the SAP-ingress and SAP-egress QoS policies respectively and defined in kbps (kilobits/sec). The CIR for a network queue is defined within the network queue QoS policy as a percentage of the network interface bandwidth.

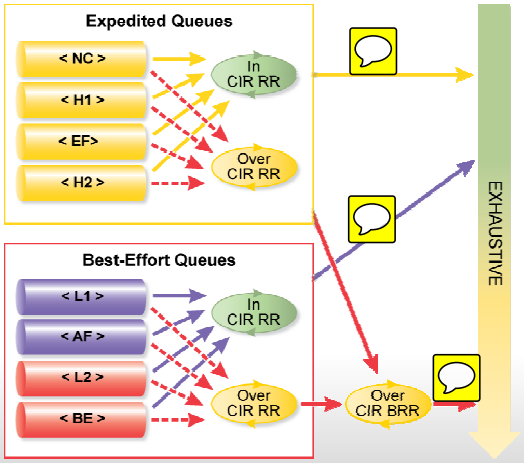
The SR/ESS allows the overbooking of the CIR. However, care should be taken to not overbook the CIR of high priority class queues because they might not be possible to meet their service commitments.

The PIR defines the maximum rate at which packets are allowed to exit a queue. It does not specify the maximum rate at which packets may enter a queue, however; that is dictated by the queue's ability to absorb bursts, and is defined by the maximum burst size.

The PIR is provisioned on ingress and egress service queues within SAP-ingress and SAP-egress QoS policies respectively. The PIR of service queues is defined in kilobits/sec. The PIR for a network queue is defined within the network-queue QoS policy as a percentage of the network interface bandwidth.

The PIR and CIR for a queue specified by the service provider are the administrative values. The operational PIR and CIR values depend on the administrative value, specified adaptation rule, and hardware scheduling rates supported in the SR/ESS.

IOM 1 and 2 Default Scheduler



- Default method for scheduling queues in the SR/ESS
- Packets are scheduled out of a queue according to the following criteria:
  1. Expedited <= CIR (In-profile)
  2. Best Effort <= CIR (In-profile)
  3. Expedited and Best Effort <= PIR (Out-of-profile)
- Parameters to configure:
  - CIR/PIR of the queue
  - Type of the queue (EX/BE)

Default schedulers are implemented in the hardware and are the default method of scheduling queues in the SR/ESS. Therefore, if no explicit hierarchical or egress port scheduler policy is defined or applied, queues are scheduled with single-tier scheduling. There are no explicit configurable parameters for basic scheduling other than a queue's CIR, PIR, and type. Single-tier scheduling is robust and provides a scalable solution to share bandwidth fairly among competing services.

Basic schedulers schedule queues based on the forwarding class of the queue and the operational status relative to the queue's CIR and PIR. Queues operating within their CIR values are serviced before queues operating above their CIR values, with expedited forwarding class queues given preference over best effort or non-expedited forwarding class queues.

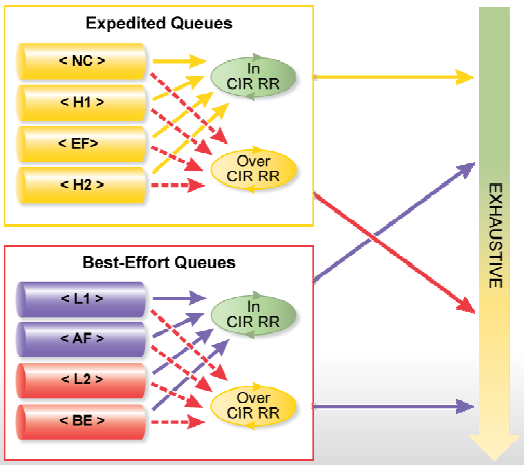
Pairs of schedulers send traffic to a switch fabric, service access port, or network port.

Queues are serviced in the following order:

- First Pass — Round Robin between expedite queues operating within their CIR.
- Second Pass — Round Robin between best effort queues operating within their CIR.
- Third Pass — Biased Round Robin between all queues operating within their PIR, but above their CIR.

In the third pass, the queues are serviced with the Biased Round Robin because the expedited queues obtain at least 50% of third pass bandwidth if there is enough traffic in those queues. Biased Round Robin is basically a round-robin scheduler. However, each time it is interrupted by the two higher priority passes, it resumes servicing the expedite queues.

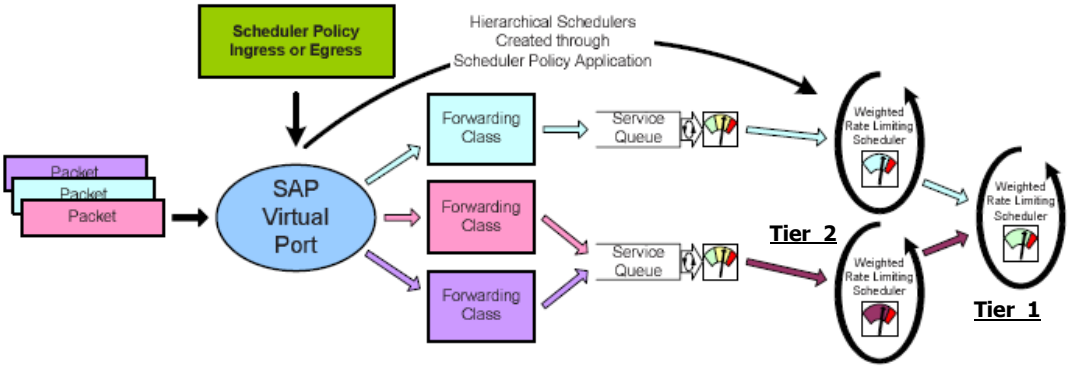
IOM 3 Default Scheduler

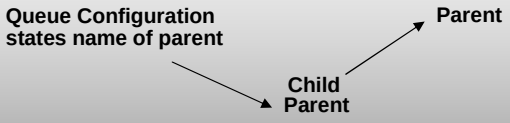


- Packets are scheduled out of a queue according to the following criteria:
  1. Expedited <= CIR (In-profile)
  2. Best Effort <= CIR (In-profile)
  3. Expedited <= PIR (Out-of-profile)
  4. Best Effort <= PIR (Out-of-profile)
- Parameters to configure:
  - CIR/PIR of the queue
  - Type of the queue (EX/BE)

IOM 3 default scheduling is improved in the sense that there are up to 4 scheduling loops instead of the 3 scheduling loops in the IOM 1 and 2. This means that expedited out-of-profile packets can now be serviced exhaustively before the best effort out-of-profile packets.

H-QoS Scheduling Policy





Another, more sophisticated scheduler is the hierarchical scheduler. Once configured and applied, it overrides the default one tier scheduler.

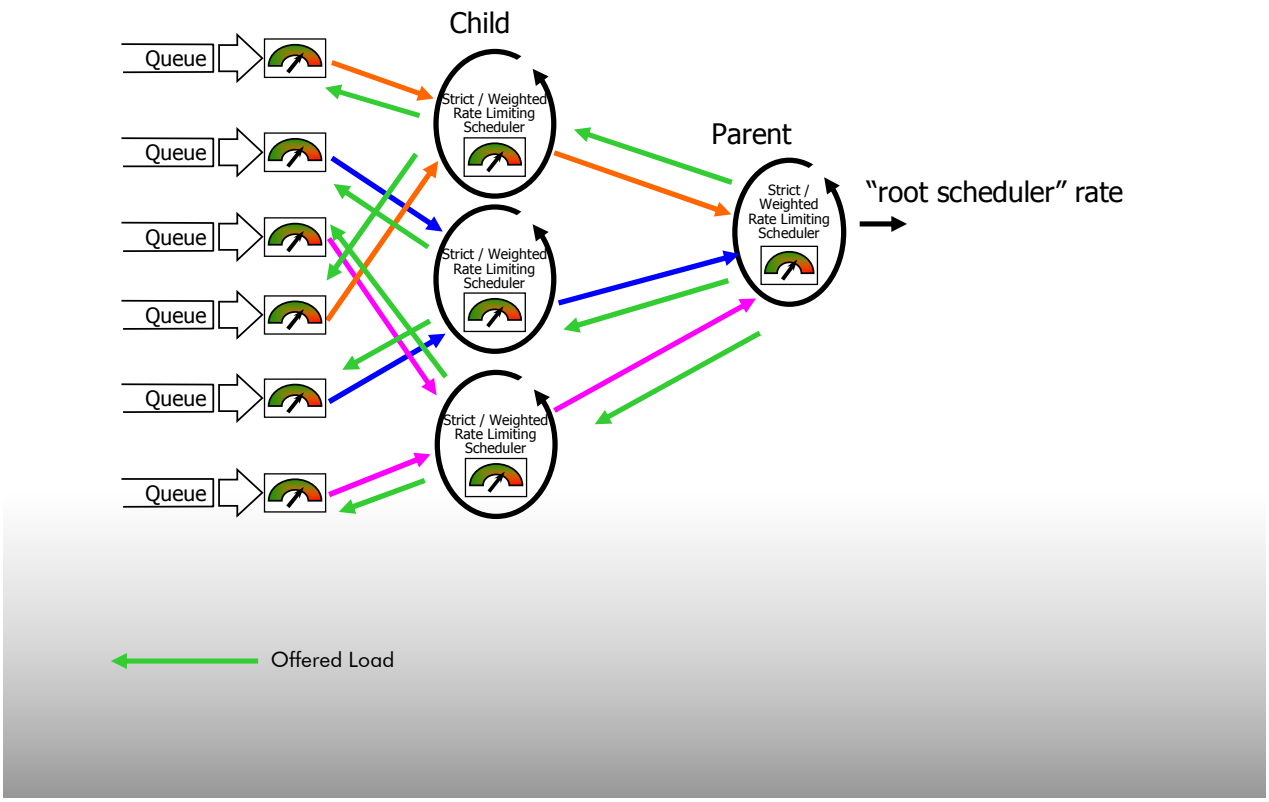
Hierarchical Scheduling involves creating hierarchical scheduling policies which are then applied to services ingress and service egress queues.

Three tiers of virtual schedulers are supported; tier 1, 2 and 3. The tier level determines the scheduler's position within the hierarchy.

A scheduler policy can be applied to a SAP or a multi-service customer site. A multi-service customer site is a group of SAPs with a common originating/termination port, MDA or card.

H-QOS schedulers only work on SAP ingress and egress policies.

Hierarchical Virtual Schedulers - Bandwidth distribution



The parent scheduler works out which child gets priority for its bandwidth.

This is worked out in 2 phases:

•Within CIR for traffic from that scheduler that is within CIR

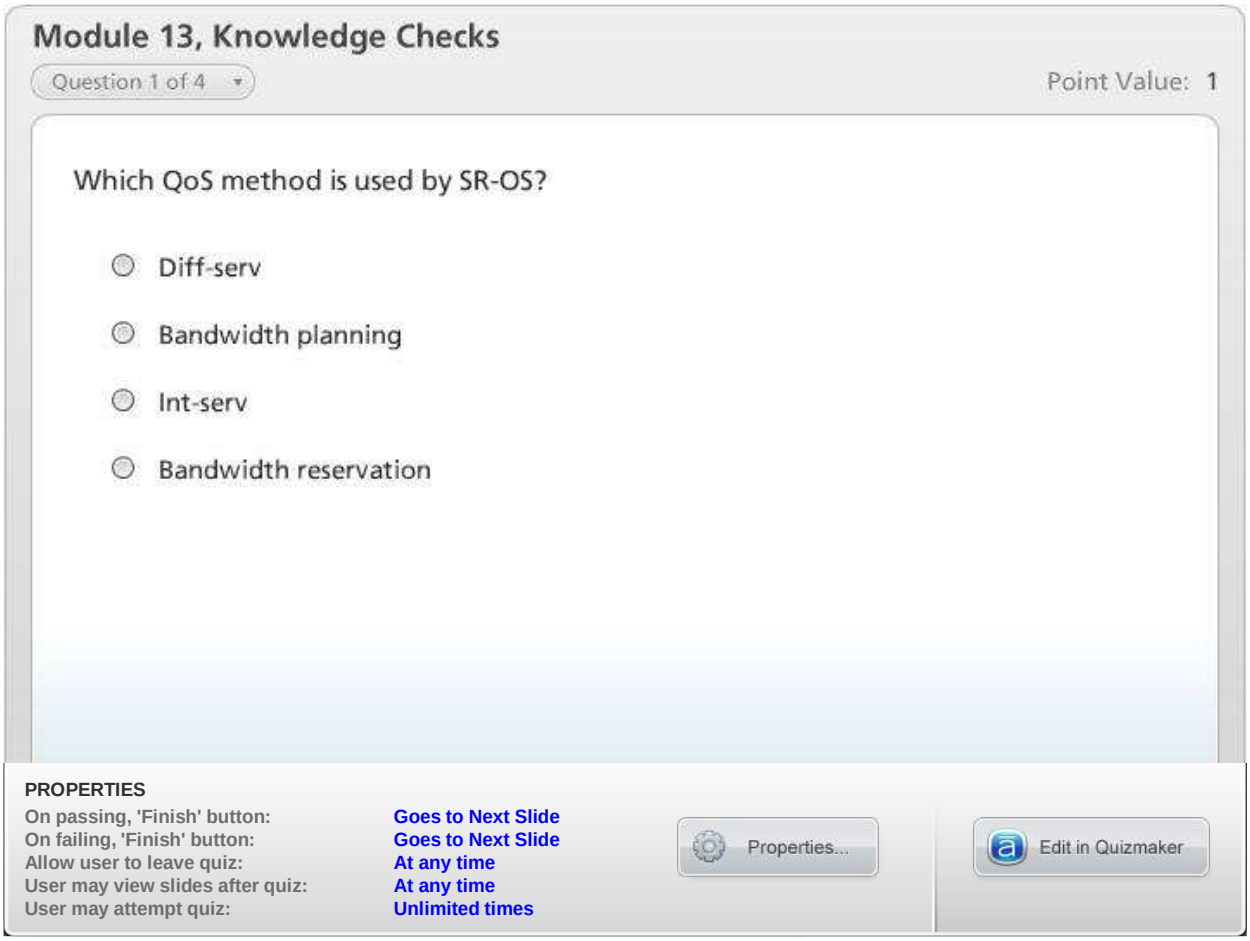
•Above CIR for traffic from that scheduler that is above CIR

No bandwidth on the parent is allocated for Above CIR traffic until all Within CIR traffic from all children is exhaustively serviced.

KNOWLEDGE CHECKS

On the next slide are 4 questions that quiz you on the key points of this mod.

You will have 1 attempt at each question and have the option to view the results at the end. You will be given feedback for each answer, but knowledge checks are not scored.



End of Module 13

This completes module 13. This module introduced some fundamentals QoS concepts from classification, queueing, buffer management and scheduling. QoS is a complex, but important topic. This is why Alcatel-Lucent has developed more detailed courses on QoS.