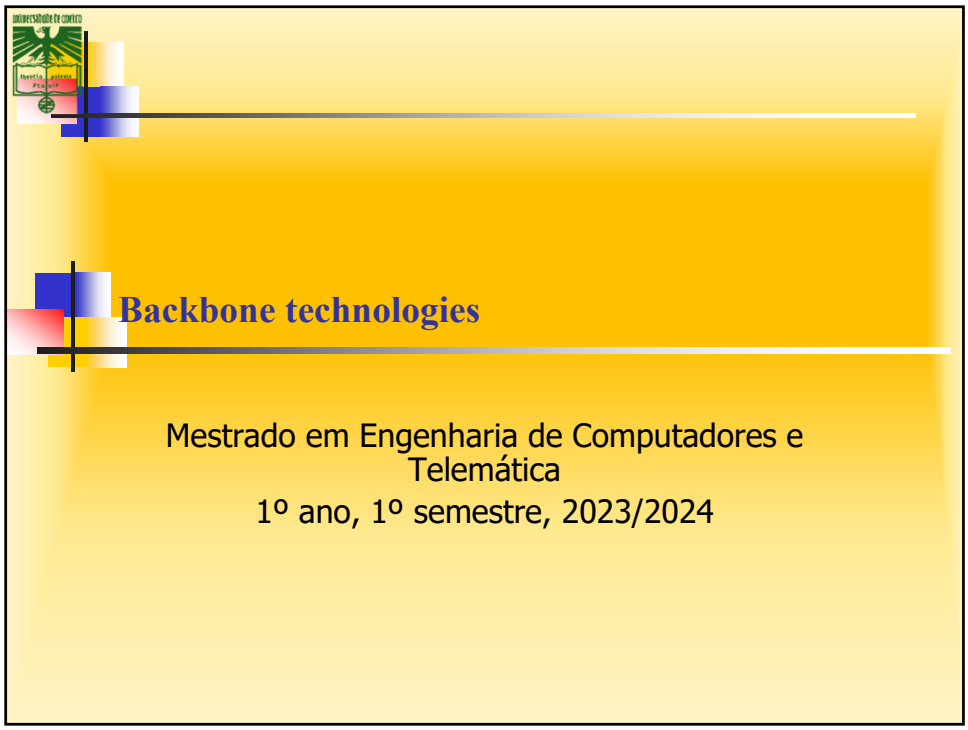


1

Traffic Engineering (TE)

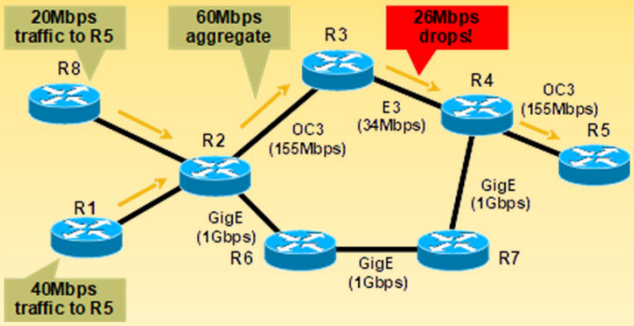
- Network Engineering
  - Build your network to carry your predicted traffic!
  - Traffic patterns are impossible to predict!
  - Routing is based on the destination and does not allow to take the maximum possible advantage of the network resources.
  - IP source routing (using options field of IP header) is not usable in practice due to security reasons.
- Traffic Engineering
  - Manipulate your traffic path to fit your network!
    - Can be done with routing protocol costs (difficult deployment), or MPLS.
    - With RIP or OSPF or ANY OTHER IGP it is not possible to condition multiple traffic flows.
  - Increase efficiency of bandwidth resources.
    - Prevent over-utilized (congested) links whilst other links are under-utilized.
  - Ensure the most desirable/appropriate path for some/all traffic.
    - Override the shortest path selected by the routing protocols.

2

2

Example – avoiding congestion

- On IP networks, *IntServ* and *DiffServ* are "routing independent architectures", retain the issues from routing
- IP network routing is based on the destination and does not allow to take the maximum possible advantage of the network resources
  - Shortest path will lead to congestion, even with available resources in the core
  - With **RIP or OSPF or ANY OTHER IGP** it is not possible to condition both flows.

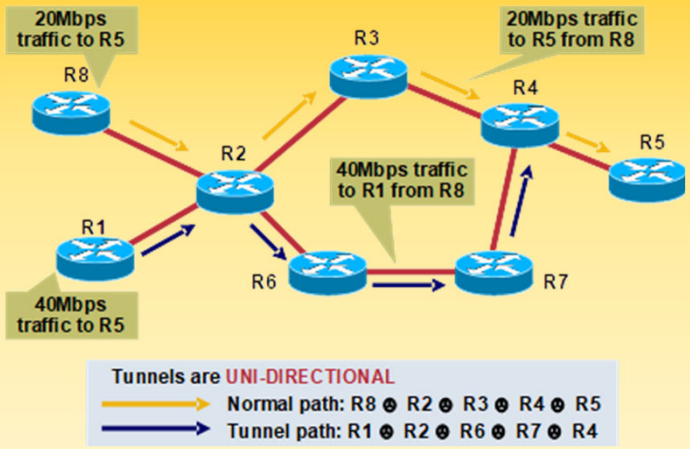


3

3

Using TE to solve… (source based routing)

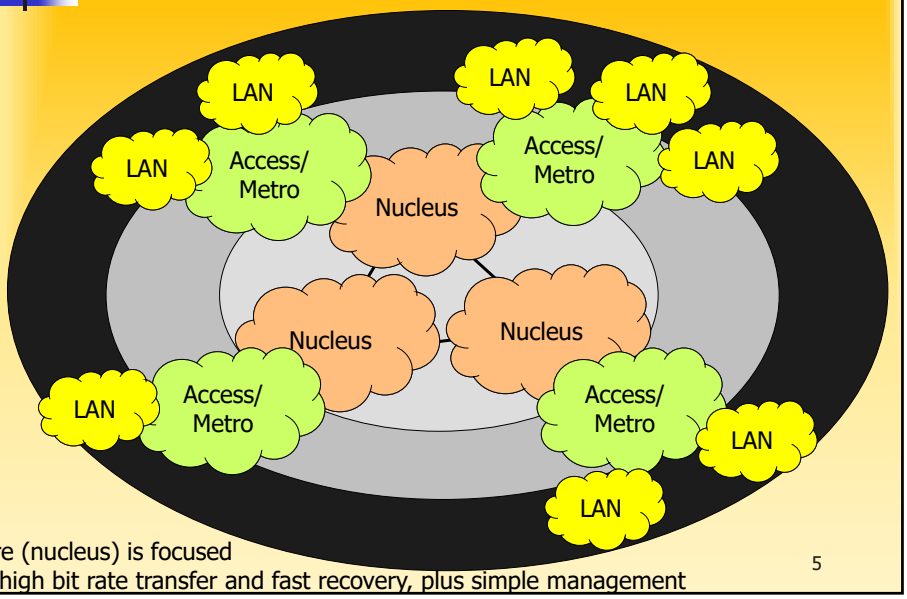
- Tunnels (virtual entities) explore all capacity
  - Packets will transport, from their source, a list of routers' addresses that define their path to the destination (*Options* field of the IP datagram header)



4

4

Network architecture of a public operator

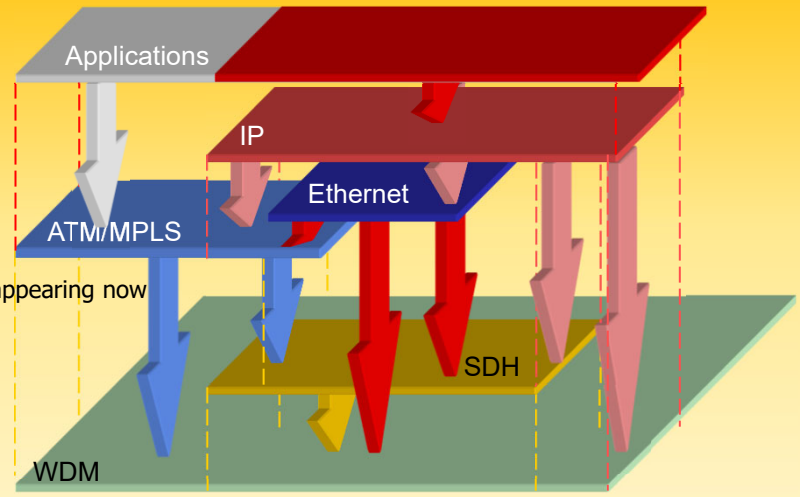


Core (nucleus) is focused
on high bit rate transfer and fast recovery, plus simple management

5

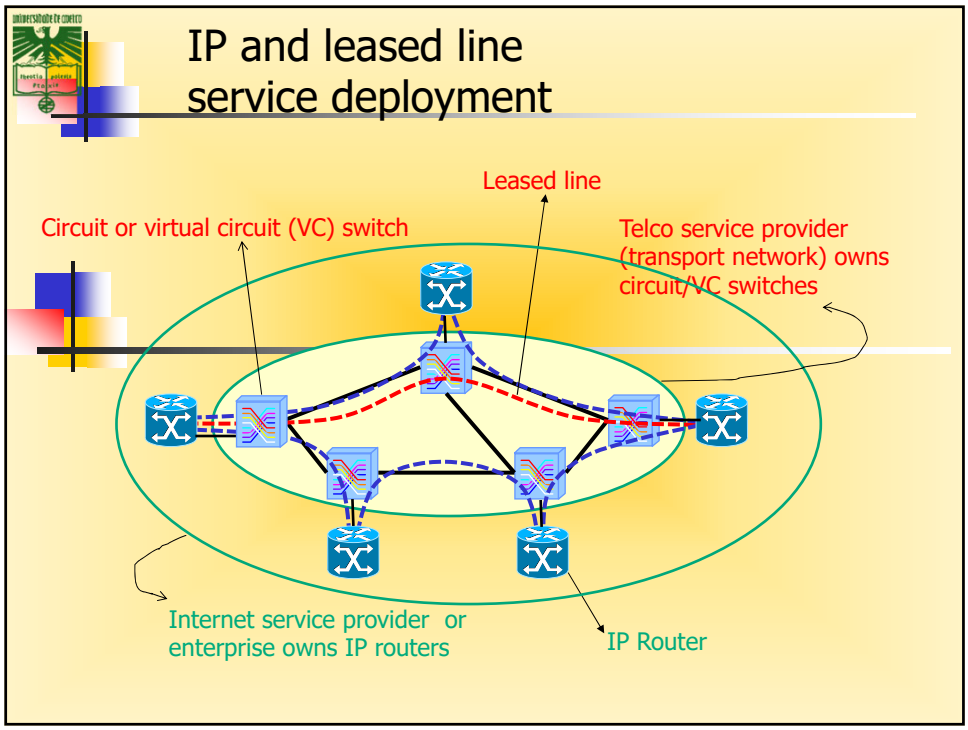
5

IP Transport Network as a Overlay

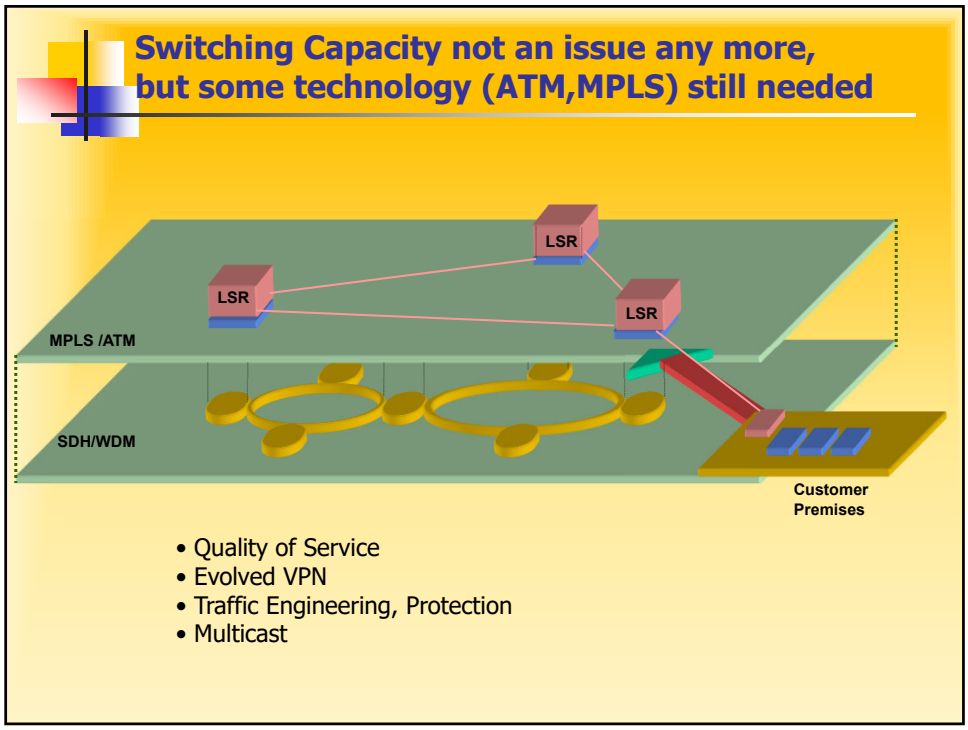


ATM is disappearing now

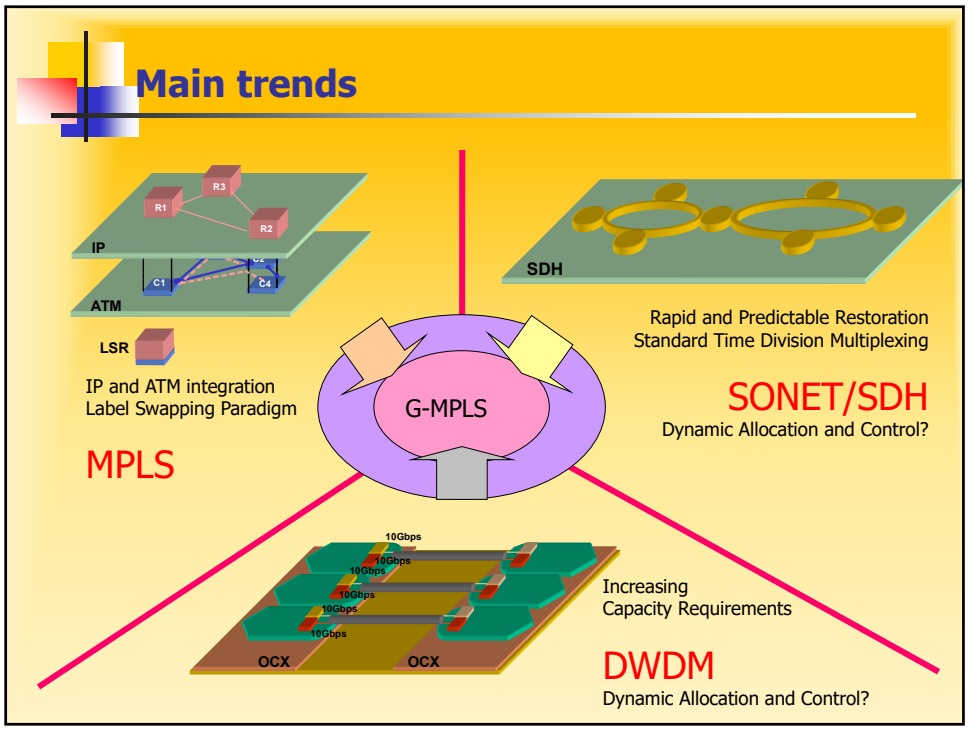
12


IP and leased line service deployment
Leased line
Circuit or virtual circuit (VC) switch
Telco service provider (transport network) owns circuit/VC switches
Internet service provider or enterprise owns IP routers
IP Router

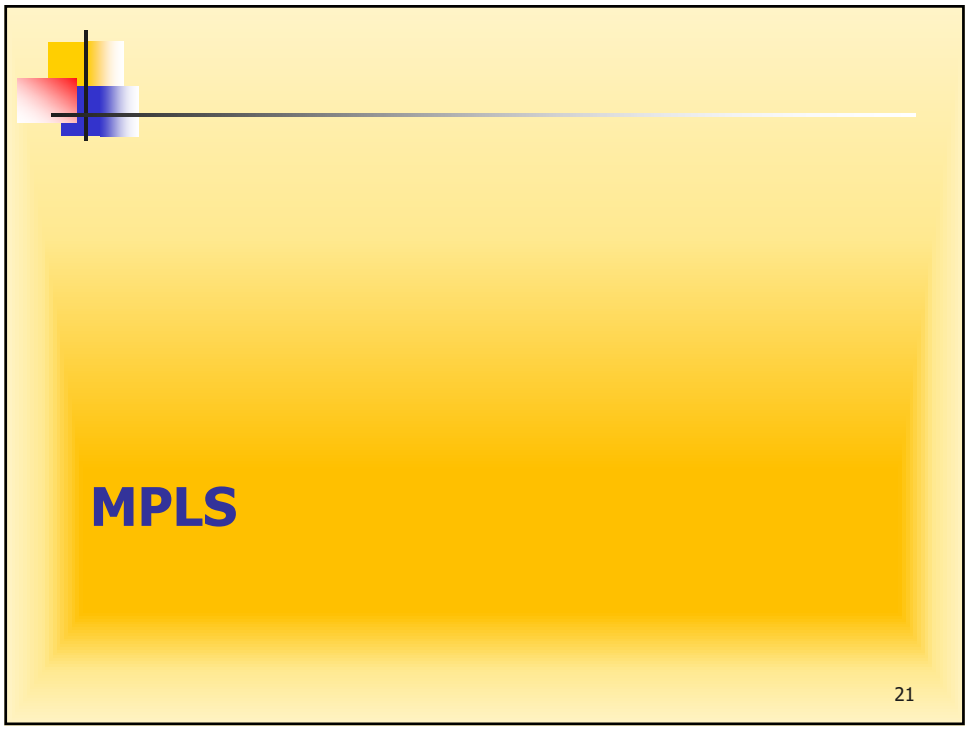
13


Switching Capacity not an issue any more, but some technology (ATM,MPLS) still needed
LSR
LSR
LSR
MPLS /ATM
SDH/WDM
Customer Premises
• Quality of Service
• Evolved VPN
• Traffic Engineering, Protection
• Multicast

14


Main trends
R1
R2
R3
IP
ATM
C1
C4
LSR
IP and ATM integration
Label Swapping Paradigm
MPLS
G-MPLS
SDH
Rapid and Predictable Restoration
Standard Time Division Multiplexing
SONET/SDH
Dynamic Allocation and Control?
10Gbps
OCX
OCX
Increasing
Capacity Requirements
DWDM
Dynamic Allocation and Control?

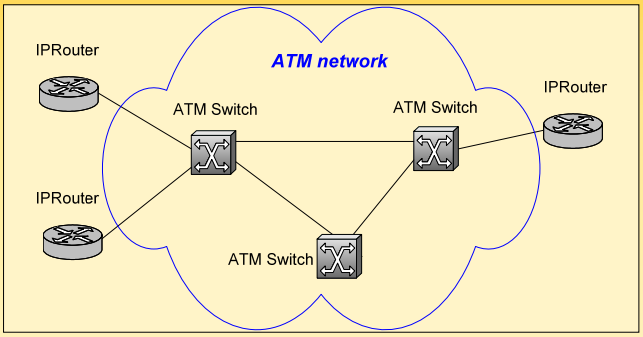
20


MPLS
21

21

IP networks over ATM

- IP routers are interconnected by an ATM network
- Connections between IP routers are implemented through virtual circuits (VCCs) or virtual paths (VPCs) on the ATM network
- It is necessary to manage two protocol layers
  (ATM is not available anymore on new networks)

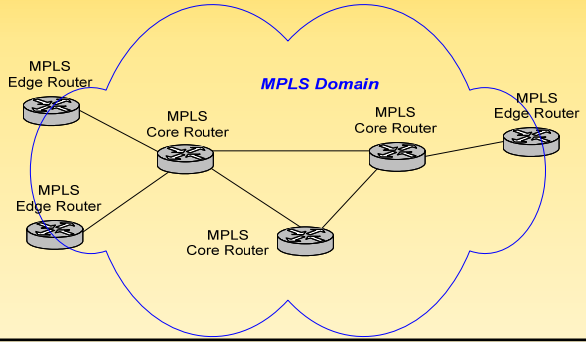


22

22

MPLS networks

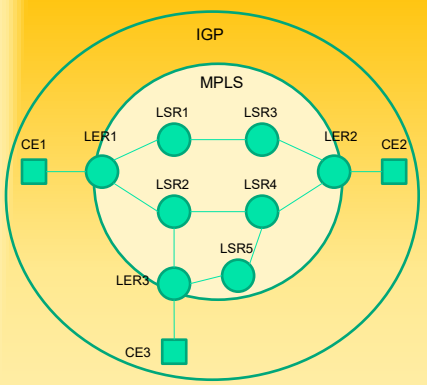
- Packets are labeled at the source with the label of the first hop
- Routers route packets based on their labels, just like ATM does with the VCI and VPI fields
- Advantages
  - Simplification of the packet routing process on routers
  - Traffic engineering capability equivalent to ATM
  - Simplification of the network management (a single protocol layer)

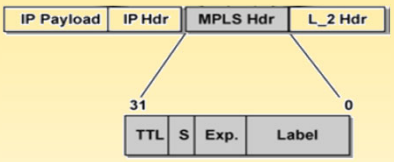


23

23

IP/MPLS





- **FEC – Forwarding Equivalence Class**
  - Identifies group of MPLS packets that are handled in the same way;
- **LABEL**
  - Local identifier based on the concept of FEC;
- **LSR – Label Switching Router**
  - LSRs in the same domain form an MPLS domain, using label forwarding;
- **LER – Label Edge Router**
  - LERs interacts with devices outside the domain, using IP technologies;

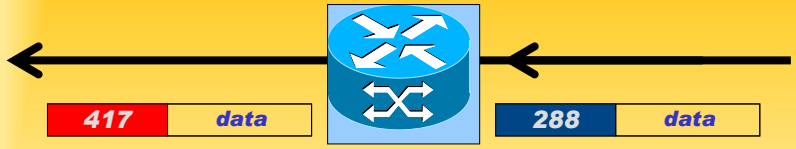
24

IP/MPLS

- **LSP – Label Switched Path**
  - Path along which FEC travels through an MPLS network;
- **LDP – Label Distribution Protocol**
  - The protocol used by MPLS for control that classifies FECs, distributes labels, and establish and maintains LSPs;
- **LFIB – Label Forwarding Information Base**
  - Forwarding table created by routing protocols and labels;
- **LSP tunneling**
  - Explicit route LSP between two LSRs that are not directly connected (indicated for TE);
- **Multi-level label stack**
  - Label stack carried by MPLS packets organized as LIFO (last-in first-out) stack;

25

Forwarding via Label Swapping

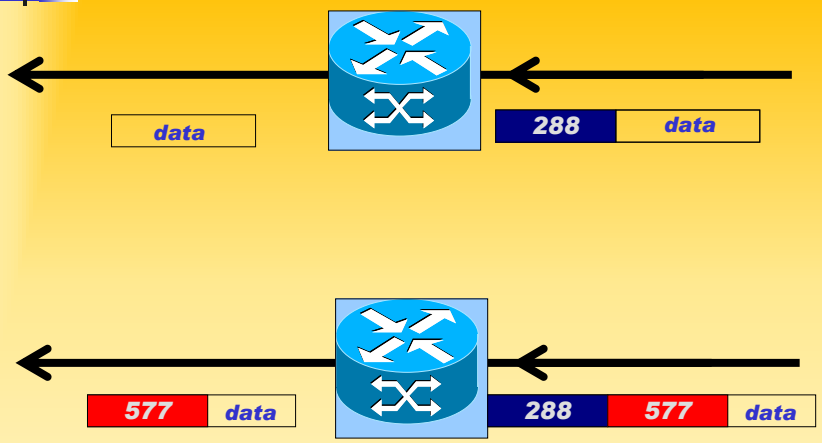


Labels are short, fixed-length values.

26

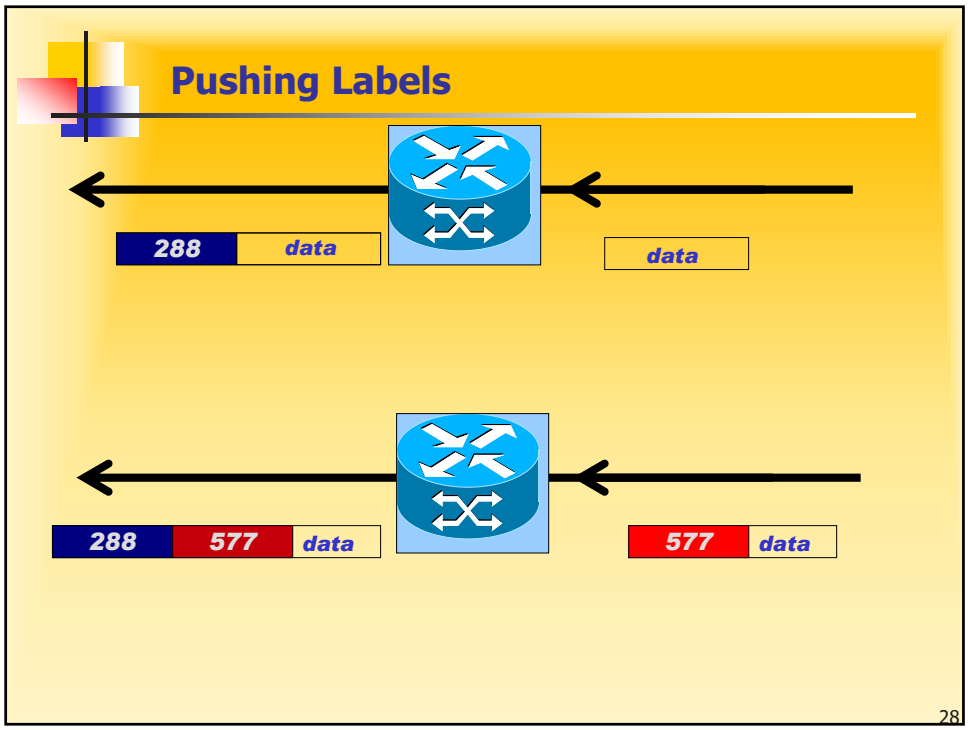
26

Popping Labels

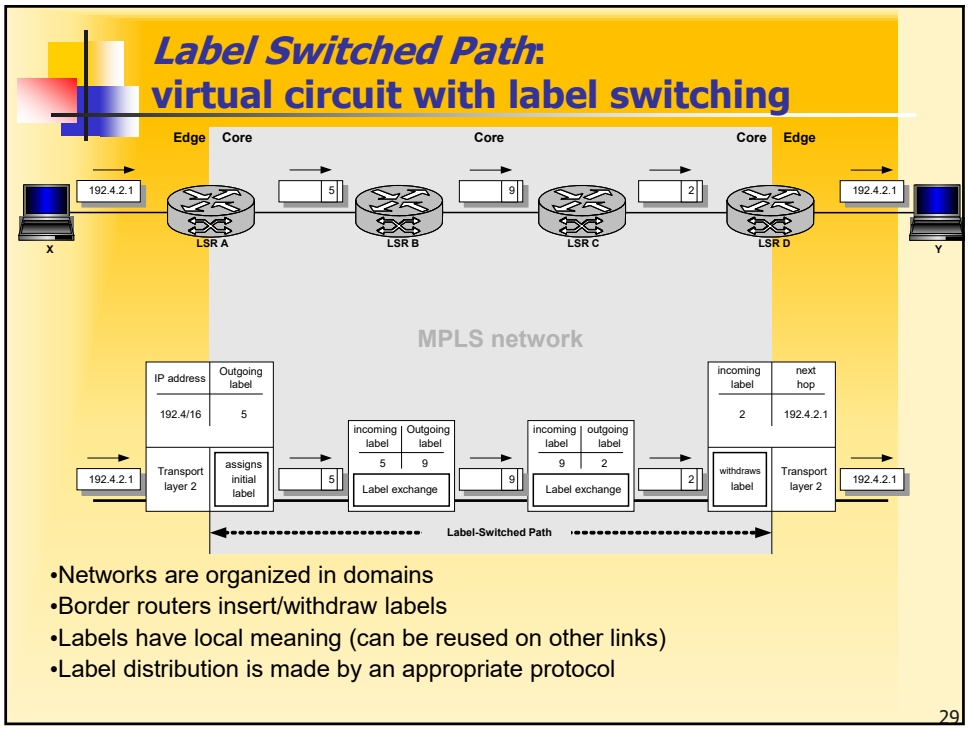


27

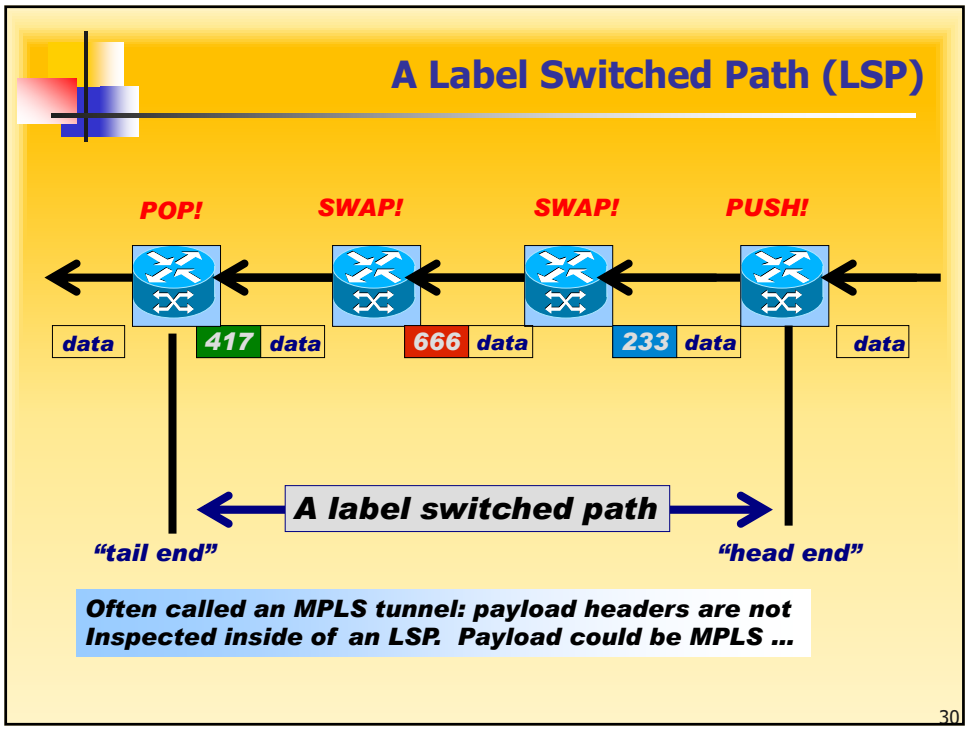
27


Pushing Labels
288
data
data
288
577
data
577
data
28

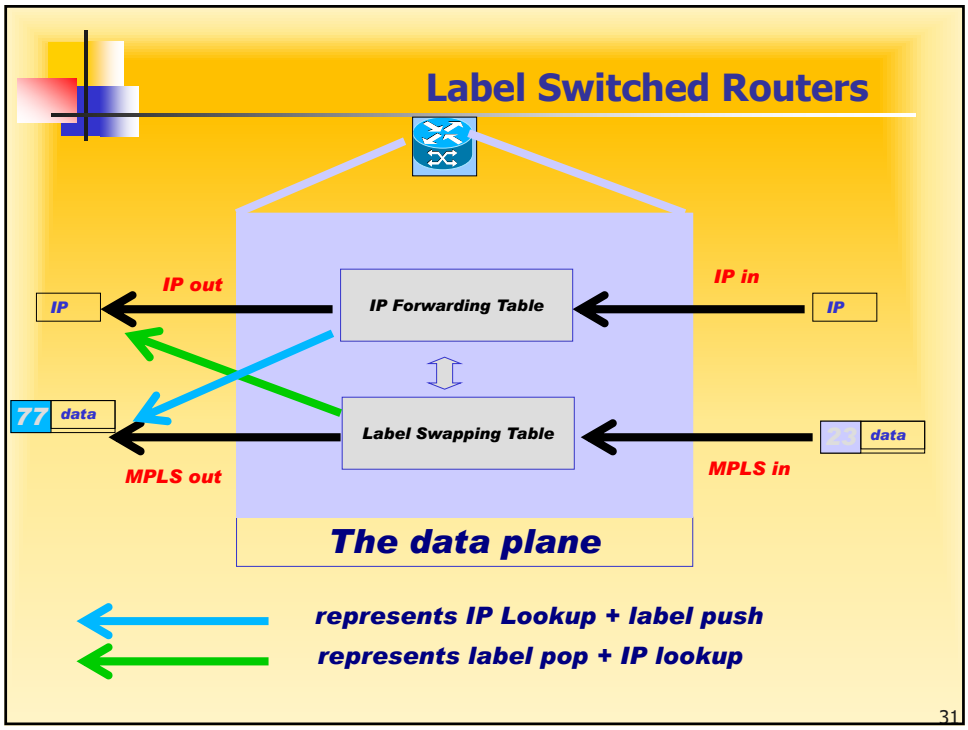
28


Label Switched Path:
virtual circuit with label switching
Edge
Core
Core
Core
Edge
192.4.2.1
5
9
2
192.4.2.1
X
LSR A
LSR B
LSR C
LSR D
Y
MPLS network
IP address
Outgoing label
192.4/16
5
incoming label
Outgoing label
5
9
incoming label
outgoing label
9
2
incoming label
next hop
2
192.4.2.1
192.4.2.1
Transport layer 2
assigns initial label
5
Label exchange
9
Label exchange
2
withdraws label
Transport layer 2
192.4.2.1
Label-Switched Path
•Networks are organized in domains
•Border routers insert/withdraw labels
•Labels have local meaning (can be reused on other links)
•Label distribution is made by an appropriate protocol
29

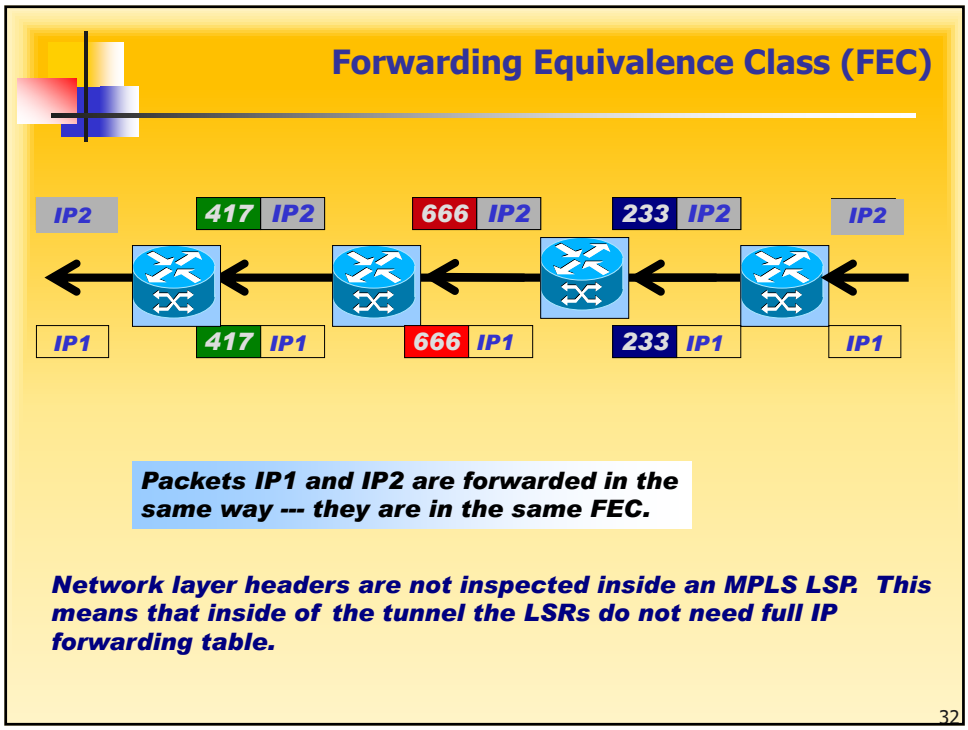
29


A Label Switched Path (LSP)
POP!
SWAP!
SWAP!
PUSH!
data
417 data
666 data
233 data
data
A label switched path
"tail end"
"head end"
Often called an MPLS tunnel: payload headers are not Inspected inside of an LSP. Payload could be MPLS ...
30

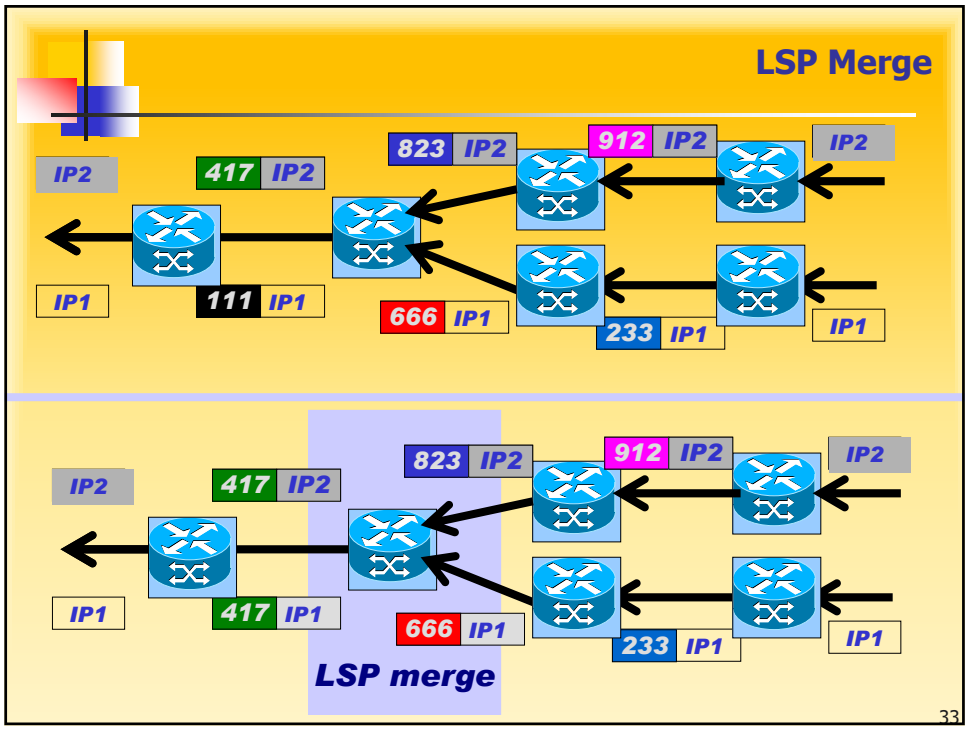
30


Label Switched Routers
IP out
IP in
IP
IP Forwarding Table
IP
77 data
Label Swapping Table
23 data
MPLS out
MPLS in
The data plane
represents IP Lookup + label push
represents label pop + IP lookup
31

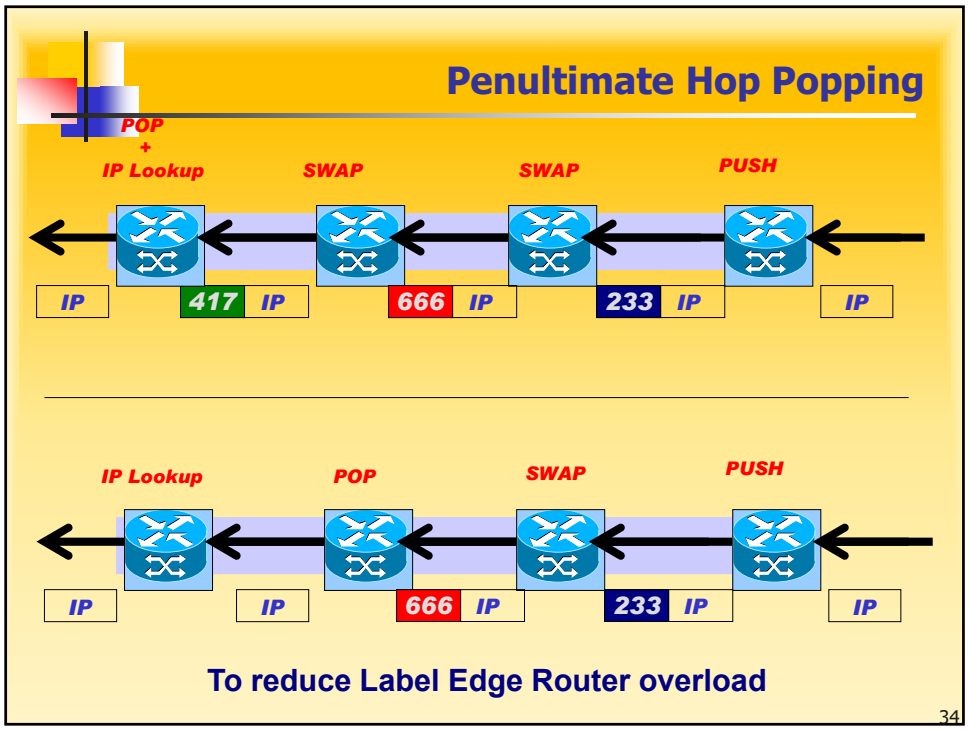
31


Forwarding Equivalence Class (FEC)
IP2
417 IP2
666 IP2
233 IP2
IP2
IP1
417 IP1
666 IP1
233 IP1
IP1
Packets IP1 and IP2 are forwarded in the same way --- they are in the same FEC.
Network layer headers are not inspected inside an MPLS LSP. This means that inside of the tunnel the LSRs do not need full IP forwarding table.
32

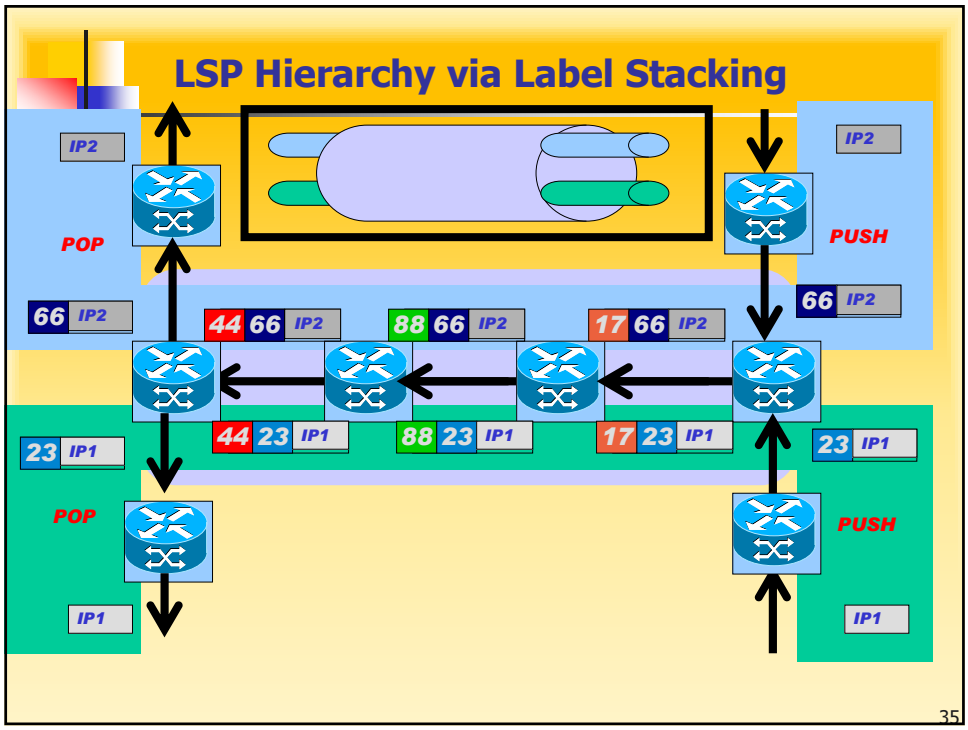
32


LSP Merge
IP2
417 IP2
823 IP2
912 IP2
IP2
IP1
111 IP1
666 IP1
233 IP1
IP1
IP2
417 IP2
823 IP2
912 IP2
IP2
IP1
417 IP1
666 IP1
233 IP1
IP1
LSP merge
33

33

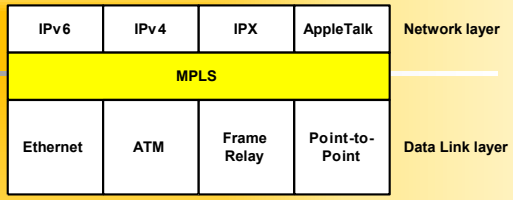

Penultimate Hop Popping
POP
+
IP Lookup
SWAP
SWAP
PUSH
IP
417 IP
666 IP
233 IP
IP
IP Lookup
POP
SWAP
PUSH
IP
IP
666 IP
233 IP
IP
To reduce Label Edge Router overload
34

34

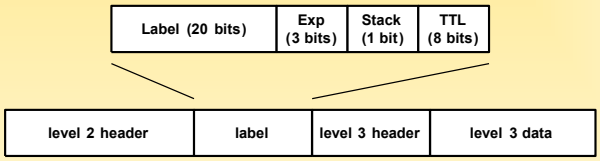

LSP Hierarchy via Label Stacking
IP2
POP
66 IP2
44 66 IP2
88 66 IP2
17 66 IP2
IP2
PUSH
66 IP2
23 IP1
44 23 IP1
88 23 IP1
17 23 IP1
23 IP1
POP
IP1
PUSH
IP1
35

35

Labels



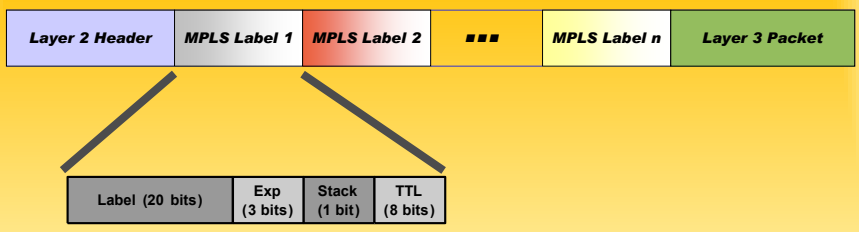
- On some Data Link (level 2) technologies, label is given by the appropriate fields of their header
  - ATM technology : VPI (Virtual Path ID) and VCI (Virtual Channel ID) fields
  - Frame Relay technology: DLCI (Data Link Connection Identifier) field
- On other Data Link technologies (Point-to-Point, Ethernet), the label is inserted between layer 2 and layer 3 headers



36

36

Generic MPLS Encapsulation



RFC 3032. MPLS Label Stack Encoding

- ***Label:*** ***Label Value, 20 bits***
- ***Exp:*** ***Experimental, 3 bits***
- ***S:*** ***Bottom of Stack, 1 bit***
- ***TTL:*** ***Time to Live, 8 bits***

37

37

IP/MPLS Network Establishment

- **Discovery link and topology**
  - IP routing table is built
    - LSRs and LERs use routing protocols to discover network topology eg. OSPF, ISIS, (BGP);
    - CEs advertise their addresses using routing protocols into MPLS cloud;
  - Forwarding Information Base (FIB) is built, initially without label information
- **Label Assignment**
  - FECs creation
    - LSRs classify with the same FEC all packets handled on the same way;
  - Allocate Labels
    - Every LSR allocates locally labels for every destination in the IP routing table (LIB and LFIB setup);

38

Label Distribution Protocols

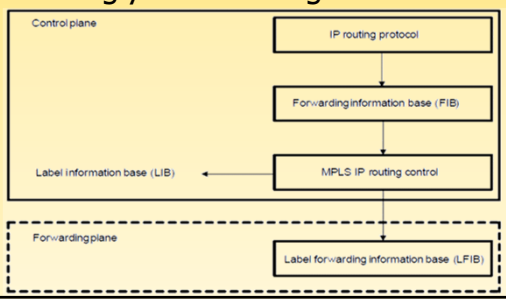
- Unconstrained routing
  - Label Distribution Protocol (LDP).
  - Path is chosen based on IGP shortest path.
- Constrained routing
  - Constrained by explicit path definition and/or performance requirements (e.g., available bandwidth).
  - Resource Reservation Protocol with Traffic Engineering (RSVP-TE).
    - Evolution of RSVP to support traffic engineering and label distribution.
  - Constrained based Routing LDP (CR-LDP).
    - Evolution of LDP to support constrained routing.
    - Deprecated!
- MPLS VPN scope
  - MP-BGP using address family VPN IPv4 and family specific MP_REACH_NLRI attribute.

39

39

IP/MPLS Network Establishment

- **Label distribution operation and LSP Establishment**
  - Discovery
    - Basic Discovery – LSRs send LDP link Hellos UDP (multicast) for directly connected peers.
    - Extended Discovery – LSRs send LDP targeted Hellos UDP for a specific (remote) IP peer.
  - Session Establish and Maintenance
    - TCP session is established and it is maintained through periodically Keep-Alive messages
  - LFIBs are established accordingly with routing and Label tables.



40

Label Distribution Protocol (LDP)

- Dynamic distribution of label binding information.
- LSR discovery.
- Reliable transport with TCP.
- Incremental maintenance of label swapping tables (only deltas are exchanged).
- Designed to be extensible with Type-Length-Value (TLV) coding of messages.
- Modes of behavior that are negotiated during session initialization
  - Label distribution control (ordered or independent).
  - Label retention (liberal or conservative).
  - Label advertisement (unsolicited or on-demand).

41

41

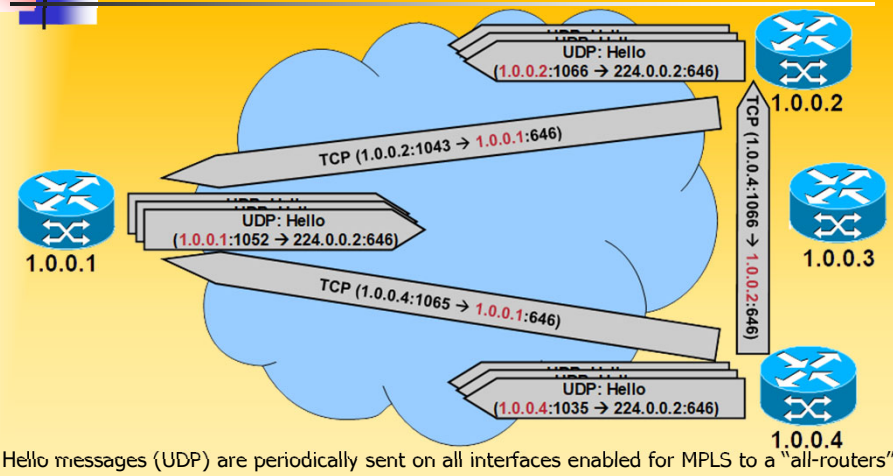
LDP Messages

- Discovery messages
  - Announce and maintain the presence of an LSR in a network.
  - **Hello Messages** (UDP) sent to "all-routers" multicast address.
  - Once neighbor is discovered, a LDP session is established over TCP.
- Session messages
  - Establish (**Initialization Message**) and maintain (**KeepAlive Message**) sessions between LDP peers.
- Advertisement messages
  - When a new LDP session is initialized and before sending label information an LSR advertises its interface addresses with one or more **Address Messages**.
  - An LSR withdraw previously advertised interface addresses with **Address Withdraw Messages**.
  - Create, change, and delete label mappings for FECs.
    - **Label Mapping, Label Request, Label Abort Request, Label Withdraw,** and **Label Release Messages.**
- Notification messages
  - Provide advisory information and to signal error information.

42

42

LDP Neighbour Discovery



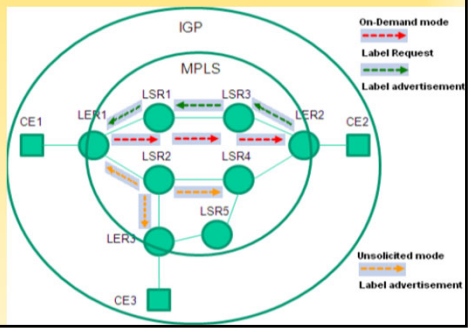
- Hello messages (UDP) are periodically sent on all interfaces enabled for MPLS to a "all-routers" multicast address (224.0.0.2).
- If there is another router on that interface it will respond by trying to establish a LDP/TCP session with the source of the hello messages.
- Both TCP and UDP messages use well-known LDP port number 646
- LDP Session is started by the router with higher IP address.

43

43

IP/MPLS Network Establishment

- LSP Establishment and Maintenance
  - Downstream On-demand mode– Upstream LSR sends a label request message (with FEC description) to its Downstream:
    - Ordered mode – a LSR only sends label (response) to its Upstream when it receives the label from its Downstream;
    - Independent mode – a LSR sends the label (response) when receives any request label;
  - Downstream Unsolicited mode – a Downstream LSR advertises label binding information to its Upstream LSR unsolicited after session to be established, without request ;
  - Label retention mode
    - Conservative mode
      - LSRs keeps only the labels from next hops
      - Indicated for limited label space
    - Liberal mode
      - LSRs keeps any labels, even if those are not from next hops
      - Indicated for quick adaptation of route changes

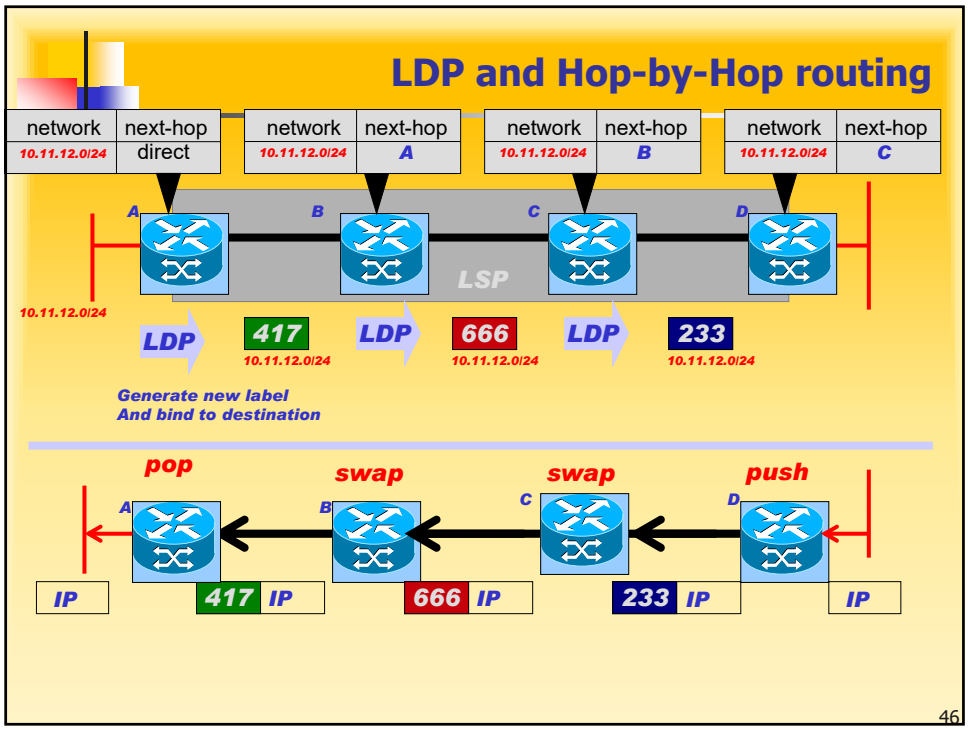


44

Multiprotocol Label Switching (MPLS) – TE usage

MSC Engenharia de Computadores e Telemática
1º ano, 1º semestre, 2023/2024

45



46

46

Constrained based Routing

- A cost is associated to each link
- Each link has a further set of attributes that represent performance metrics

The routing objective is to determine the lowest cost path that does not violate the restrictions that were assigned

- Restrictions can be associated to a set of performance characteristics, like for example, bandwidth, delay, priority, etc.
  - For the bandwidth case, the restriction that is imposed to the routing algorithm is that the path must have, on each connection it traverses, a bandwidth higher than a certain threshold.
  - In this case, the connection attribute used is the available bandwidth.

47

47

Constraint Based Routing?

Basic components

1. Specify path constraints
2. Extend topology database to include resource and constraint information
3. Find paths that do not violate constraints and optimize some metric
4. Signal to reserve resources along path
5. Set up LSP along path (with explicit route)
6. Map ingress traffic to the appropriate LSPs

Problem here: OSPF areas hide information for scalability. So these extensions work best only within an area...

Extend Link State Protocols (IS-IS, OSPF)

Extend RSVP or LDP or both!

Note: (3) could be offline (e.g. network operational center – NOC - systems) or online (perhaps an extension to OSPF)

Problem here: what is the "correct" resource model for IP services?

48

48

Resource Reservation + Label Distribution

Two emerging/competing/dueling approaches:

Add label distribution and explicit routes to a resource reservation protocol

RSVP-TE = **RSVP**

CR-LDP = **LDP**

Add explicit routes and resource reservation to a label distribution protocol

RSVP-TE: Extensions to RSVP for LSP Tunnels
draft-ietf-mpls-rsvp-lsp-tunnel-08.txt

Constraint-Based LSP Setup using LDP
draft-ietf-mpls-cr-ldp-05.txt

As of February 2003, the IETF MPLS working group deprecated CR-LDP and decided to focus purely on RSVP-TE.

RFC 3468: The Multiprotocol Label Switching (MPLS) Working Group decision on MPLS signalling protocols

49

49

LSPs establishing protocols

- RSVP-TE (*Resource Reservation Protocol – Traffic Engineering*)
  - Extension of the RSVP protocol
- CR-LDP (*Constrained based Routing – Label Distribution Protocol*)
  - Extension of the LDP protocol, deprecated

- Both protocols enable:
  - The specification of a route to a LSP
  - To chose the labels on each link of the route
  - To make resources reservation for the LSP
- Routes are previously determined:
  - By management (Traffic engineering), in a NOC
  - By a *Constrained based Routing* type protocol

50

50

RSVP-TE vs. CR-LPD

RSVP-TE	***CR-LDP***
• ***Soft state periodically refreshed***	• ***State maintained incrementally***
• ***IntServ QoS model***	• ***New QoS model derived from ATM models***

And the QoS model determines the additional information attached to links and nodes and distributed with extended link state protocols…

And what about that other Internet QoS model, diffserv?

51

51

Recall....
ReSerVation Protocol (RSVP)

- ReSerVation Protocol (RSVP) was developed to communicate resource needs between hosts and network devices
  - Associated to the Intserv QoS model
- RSVP allows:
  - The source to describe the characteristics of the IP packets flow.
  - Destinations to describe the reservation they want.
  - Routers to know how to process the packets flow in order to fulfil the requested reservation.
- Encapsulated on IP (protocol type = 46 (0x2E))
- Signalling is based on PATH and RESV messages.
  - PATH announces the traffic characteristics at the sender.
  - RESV achieves reservations that were initiated by the receivers.
  - If the reservation is not possible, a RESV ERR message is sent.
- The routers reservation states have to be periodically refreshed (soft states).

54

54

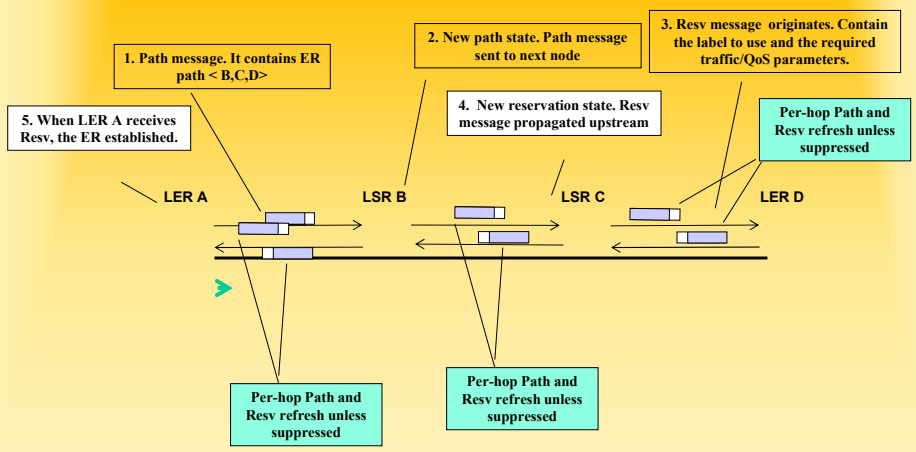
Resource Reservation Protocol with Traffic Engineering (RSVP-TE)

- Evolution of RSVP
  - RFC 3209: RSVP-TE: Extensions to RSVP for LSP Tunnels. (12/2001)
  - RFC 5151: Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions. (2/2008)
- To map traffic flows onto the physical network topology through label switched paths, resource and constraint network information are required
  - Provided by Extend Link State Protocols (IS-IS or OSPF with TE extensions).
    - RFC 3630: Traffic Engineering (TE) Extensions to OSPF Version 2. (9/2003)
    - RFC 5305: IS-IS Extensions for Traffic Engineering. (10/2008)

55

55

Establishment of a LSP with the RSVP-TE protocol



56

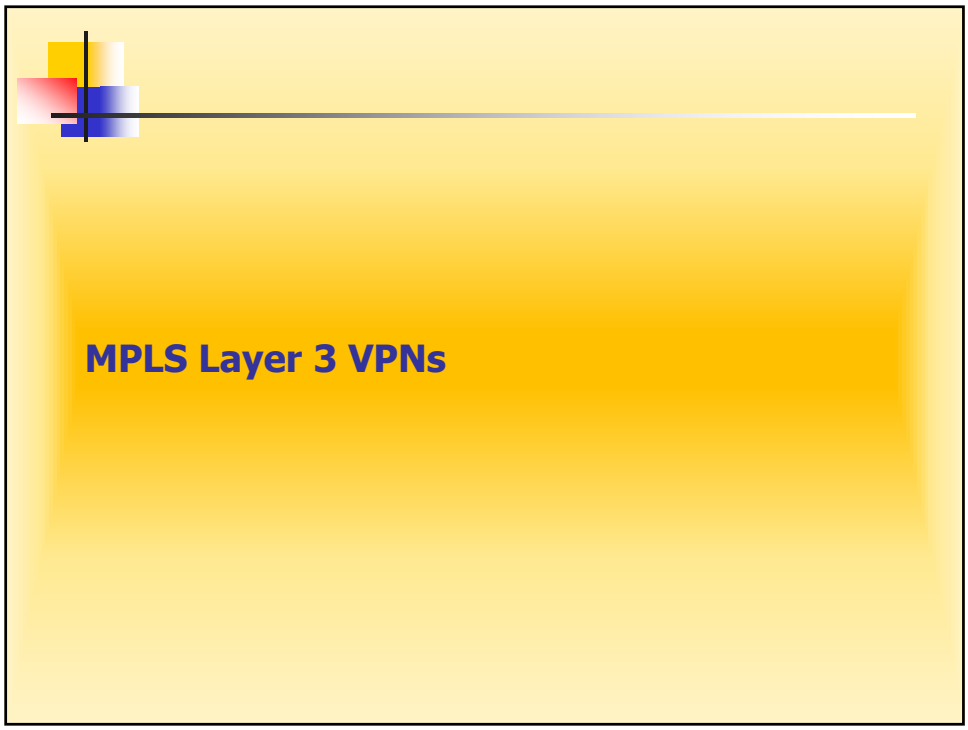
56

REVIEW: MPLS - Major Drivers

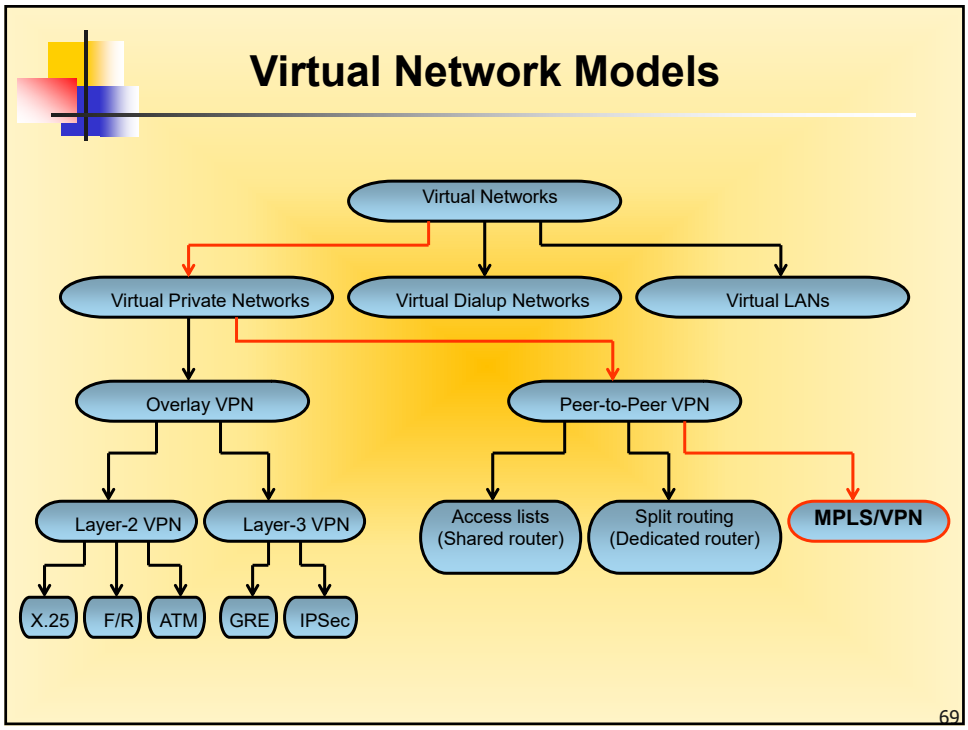
- Provide IP VPN Services *"the leased line"*
  - Scalable IP VPN service – Build once and sell many
  - Managed Central Services – Building value added services and offering them across VPNs
- Managing traffic on the network using MPLS Traffic Engineering
  - Providing tighter SLA/QoS (Guaranteed BW Services)
  - Protecting bandwidth - Bandwidth Protection Services
- Integrating Layer 2 & Layer 3 Infrastructure
  - Layer 2 services such as ATM (or Frame Relay) over MPLS
  - Mimic layer 2 services over a highly scalable layer 3 infrastructure

57

57


MPLS Layer 3 VPNs

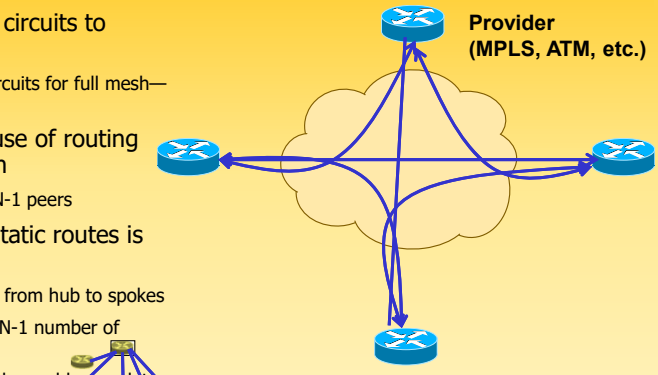
68

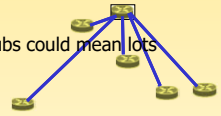

Virtual Network Models
Virtual Networks
Virtual Private Networks
Virtual Dialup Networks
Virtual LANs
Overlay VPN
Peer-to-Peer VPN
Layer-2 VPN
Layer-3 VPN
Access lists (Shared router)
Split routing (Dedicated router)
MPLS/VPN
X.25
F/R
ATM
GRE
IPSec
69

69

Overlay Network

- Provider sells a circuit service
  - customer wants to deploy its own network over those services
- Customers purchases circuits to connect sites, runs IP
  - N sites, (N*(N-1))/2 circuits for full mesh—expensive
- scalability issue because of routing peers in mesh approach
  - N sites, each site has N-1 peers
- Hub and spoke with static routes is simpler,
  - still buying N-1 circuits from hub to spokes
  - suffers from the same N-1 number of routing peers
  - Spokes distant from hubs could mean lots of long-haul circuits





70

70

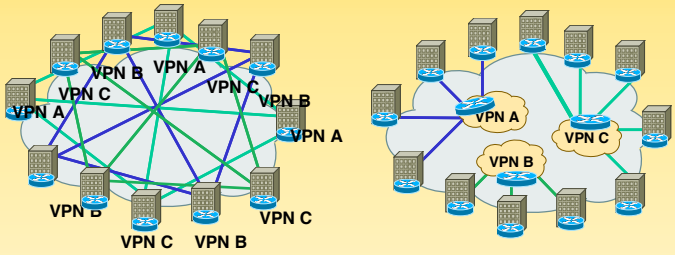
IP/MPLS Applications

- **MPLS-based VPNs**
  - MPLS L3VPN – VPN-IP over MPLS
    - VPN is a secure IP-based network between geographically dispersed sites that can communicate securely over a shared backbone;
    - MPLS VPNs provides the capability to deploy and administer scalable Layer 3 VPN backbone services to business customers
  - MPLS L2VPN – Any Technology (AT) or Transport over MPLS (e.g.: EoMPLS)
    - AT over MPLS transport Layer2 packets over MPLS network;
    - Allow the use of MPLS network to provide connectivity between customer sites with existing Layer2 networks;
  - MPLS-TE
    - Extends existing IP protocols and makes use of MPLS forwarding capabilities to provide TE
    - Brings explicit routing capabilities to MPLS networks

72

MPLS L3 VPNs using BGP (RFC2547)

- End user perspective
  - Virtual Private IP service
  - **Simple routing – just point default to provider**
  - Full site-site connectivity without the usual drawbacks (routing complexity, scaling, configuration, cost)
- Major benefit for provider – scalability

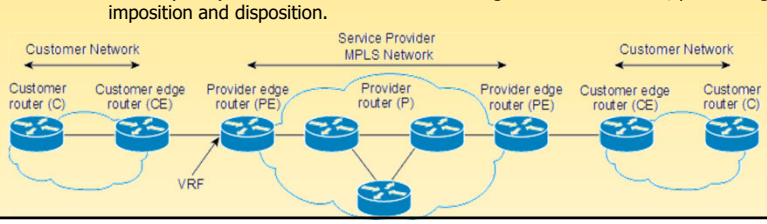


73

73

MPLS VPN Terminology

- Customer router (C) is connected only to other customer devices.
- Customer Edge (CE) router peers at Layer 3 to the Provider Edge (PE).
  - The PE-CE Interface runs either a dynamic routing protocol (eBGP, RIPv2, EIGRP, or OSPF) or has static routing (Static, Connected).
- Provider (P) router, resides in the core of the provider network.
  - Participates in the control plane for customer prefixes. The P router is also referred to as a Label Switch Router (LSR), in reference to its primary role in the core of the network, performing label switching/swapping of MPLS traffic.
- Provider Edge (PE) router, sits at the edge of the MPLS SP network.
  - In an MPLS VPN context, separate VRF routing tables are allocated for each user group.
  - Contains a global routing table for routes in the core SP infrastructure.
  - The PE is sometimes referred to as a Label Edge Router (LER) or Edge Label Switch Router (ELSR) in reference to its role at the edge of the MPLS cloud, performing label imposition and disposition.



74

74

VPN Routing and Forwarding Instance (VRF)

- PE routers maintain separate routing tables.
  - Virtual Routing and Forwarding (VRF) instance is separate from the global routing table that exists on PE routers

– Global routing table
  - Contains all PE and P routes (perhaps BGP)
  - Populated by the VPN backbone IGP

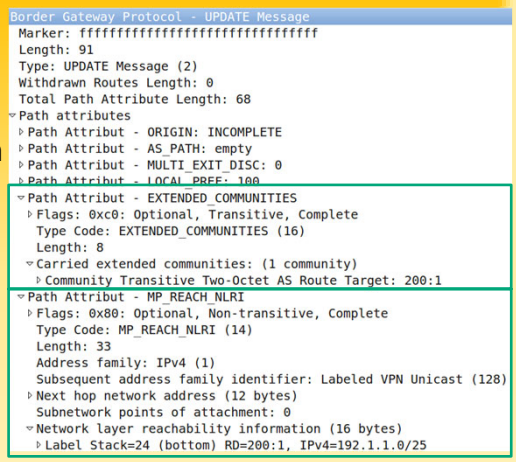
– VRF (VPN routing and forwarding)
  - Routing and forwarding table associated with one or more directly connected sites (CE routers)
  - VRF is associated with any type of interface, whether logical or physical (e.g. sub/virtual/tunnel)
  - Interfaces may share the same VRF if the connected sites share the same routing information
  - Routes are injected into the VRF from the CE-PE routing protocols for that VRF and any MP-BGP announcements that match the defined VRF.

75

Carrying VPN Routes in BGP

- VRFs by themselves aren't all that useful
  - Need some way to get the VRF routing information off the PE and to other PEs
  - This is done with BGP

- Additions to MP-BGP to Carry MPLS-VPN Info
  - RD: Route Distinguisher
  - RT: Route Target
  - VPNv4 address family
  - MPLS Label



76

76

Terminology, 1/2

- RR—Route Reflector
  - A router (usually not involved in packet forwarding) that distributes BGP routes within a provider's network
- PE—Provider Edge router
  - The interface between the customer and the MPLS-VPN network; only PEs (and maybe RRs) know anything about MPLS-VPN routes
- P—Provider router
  - A router in the core of the MPLS-VPN network, speaks LDP/RSVP but not necessarily VPNv4
- CE—Customer Edge router
  - The customer router which connects to the PE; does not know anything about labels, only IP (most of the time)
- LDP—Label Distribution Protocol
  - Distributes labels with a provider's network that mirror the IGP, one way to get from one PE to another
- LSP—Label Switched Path
  - The chain of labels that are swapped at each hop to get from one PE to another

77

77

Terminology, 2/2

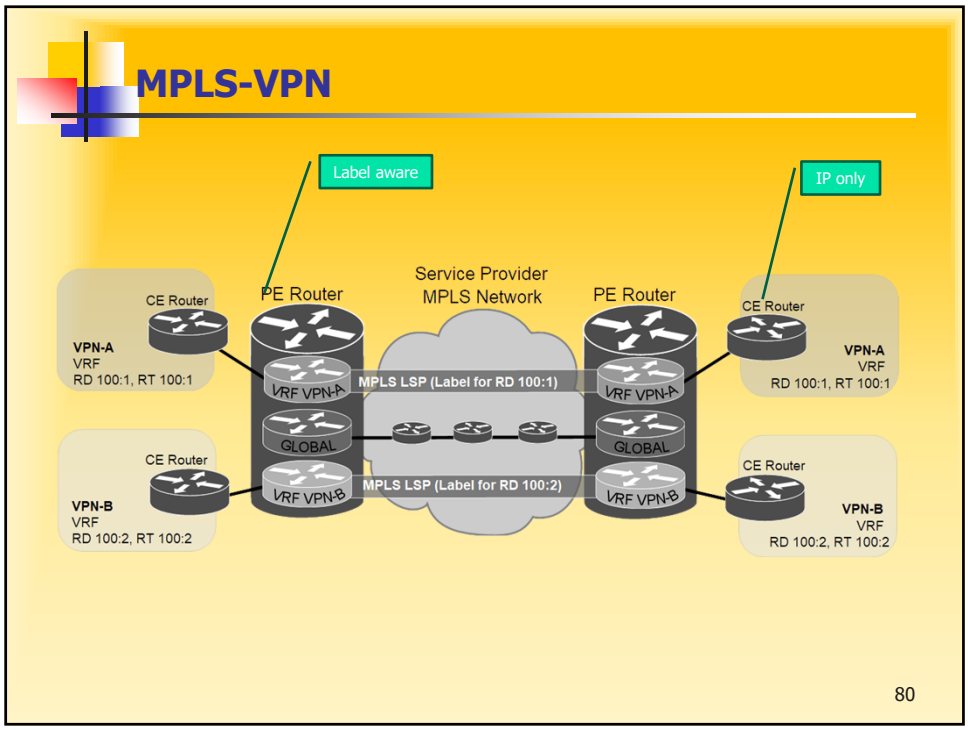
- VPN—Virtual Private Network
  - A network deployed on top of another network, where the two networks are separate and never communicate
- VRF—Virtual Routing and Forwarding instance
  - Mechanism in IOS used to build per-interface route and forwarding information bases (RIB and FIB)
  - VRF exports and imports one or more RT (route targets).
- VPNv4
  - Address family used in BGP to carry MPLS-VPN routes
- RD
  - Route Distinguisher, used to uniquely identify the same network/mask from different VRFs (i.e., 10.0.0.0/8 from VPN A and 10.0.0.0/8 from VPN B)
  - objective: make routes unique, hide routes from different customers
- RT
  - Route Target, used to control import and export policies, to build arbitrary VPN topologies for customers
  - exported RTs can be carried in BGP

Example:

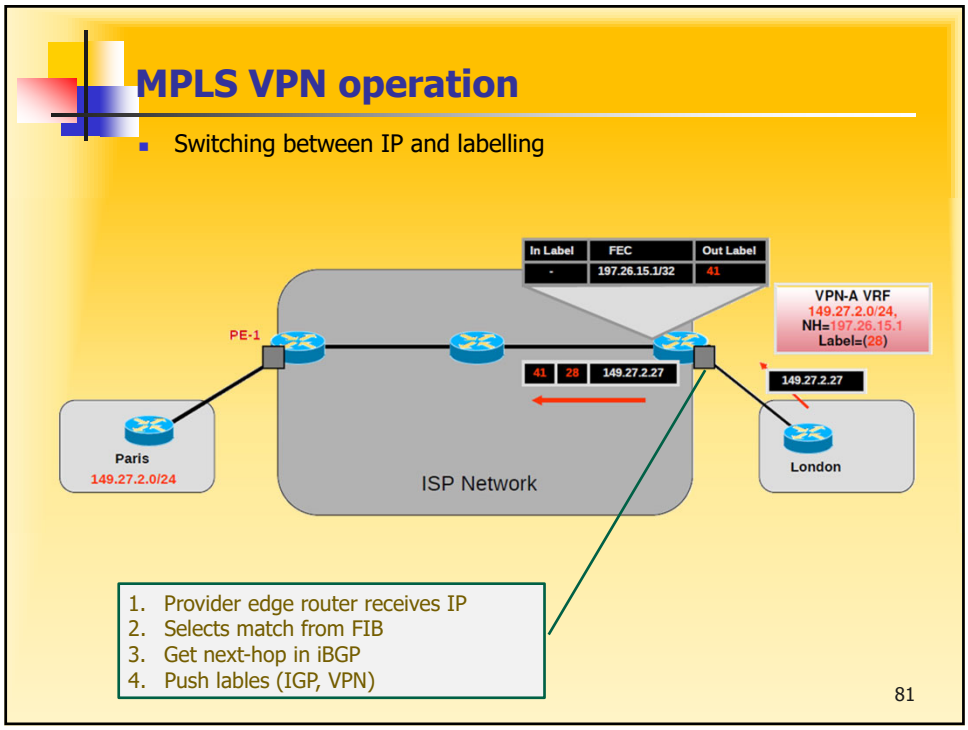
```
ip vrf VPN-A
rd 100:1
route-target export 100:1
route-target import 100:1
```

78

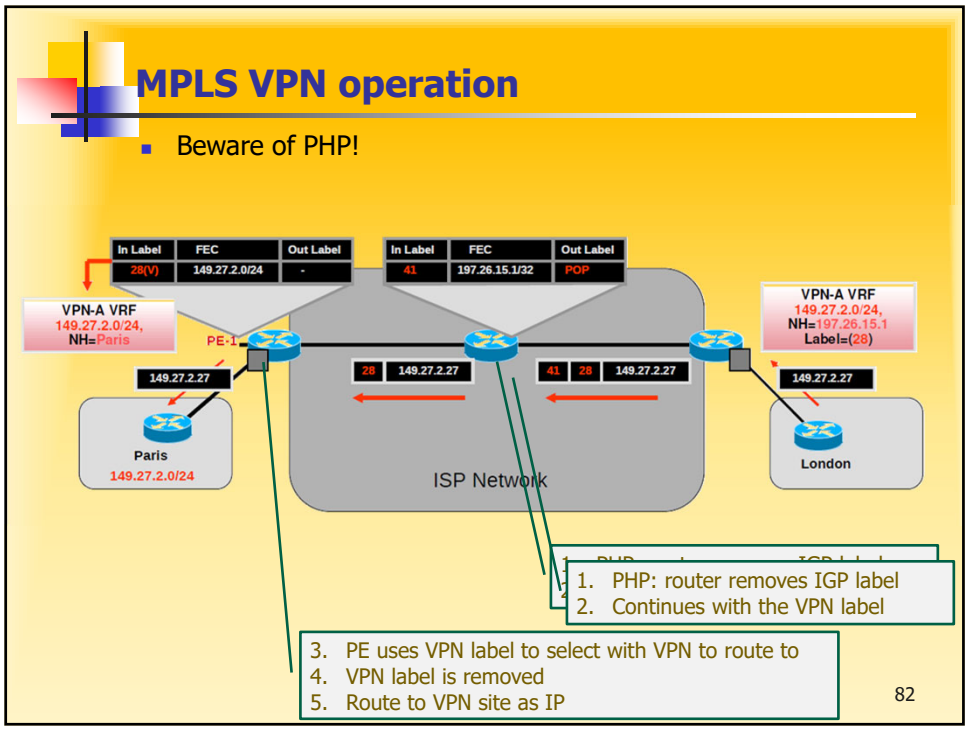
78


MPLS-VPN
Label aware
IP only
Service Provider
MPLS Network
CE Router
PE Router
PE Router
CE Router
VPN-A
VRF
RD 100:1, RT 100:1
VRF VPN-A
MPLS LSP (Label for RD 100:1)
GLOBAL
MPLS LSP (Label for RD 100:2)
VRF VPN-B
CE Router
VPN-B
VRF
RD 100:2, RT 100:2
80

80


MPLS VPN operation
Switching between IP and labelling
In Label
FEC
Out Label
-
197.26.15.1/32
41
VPN-A VRF
149.27.2.0/24,
NH=197.26.15.1
Label=(28)
PE-1
41
28
149.27.2.27
149.27.2.27
Paris
149.27.2.0/24
ISP Network
London
1. Provider edge router receives IP
2. Selects match from FIB
3. Get next-hop in iBGP
4. Push lables (IGP, VPN)
81

81


MPLS VPN operation
Beware of PHP!
In Label
FEC
Out Label
28(V)
149.27.2.0/24
-
In Label
FEC
Out Label
41
197.26.15.1/32
POP
VPN-A VRF
149.27.2.0/24,
NH=Paris
PE-1
VPN-A VRF
149.27.2.0/24,
NH=197.26.15.1
Label=(28)
149.27.2.27
28
149.27.2.27
41
28
149.27.2.27
149.27.2.27
Paris
149.27.2.0/24
ISP Network
London
1. PHP: router removes IGP label
2. Continues with the VPN label
3. PE uses VPN label to select with VPN to route to
4. VPN label is removed
5. Route to VPN site as IP
82

82