MPLS

Matěj Grégr
igregr@fit.vutbr.cz

Agenda

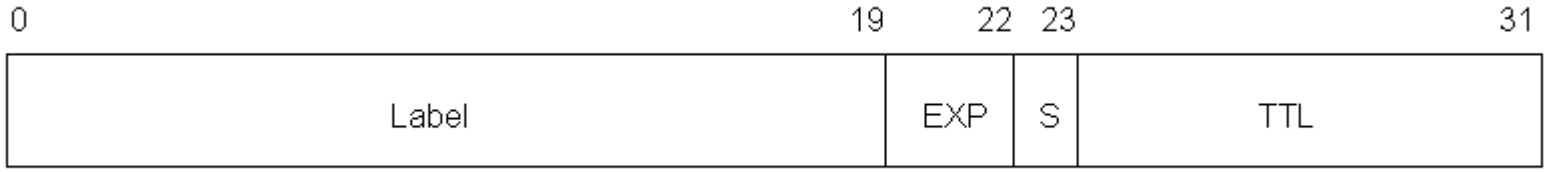
- MPLS
- LDP
- MPLS – VPN
- VPLS

MPLS – MultiProtocol Label Switching

- Why?
  - Necessity to have protocol-independent transport
  - Speed?
  - MPLS – VPN, VPLS
  - QoS
  - BGP free core

MPLS – MultiProtocol Label Switching

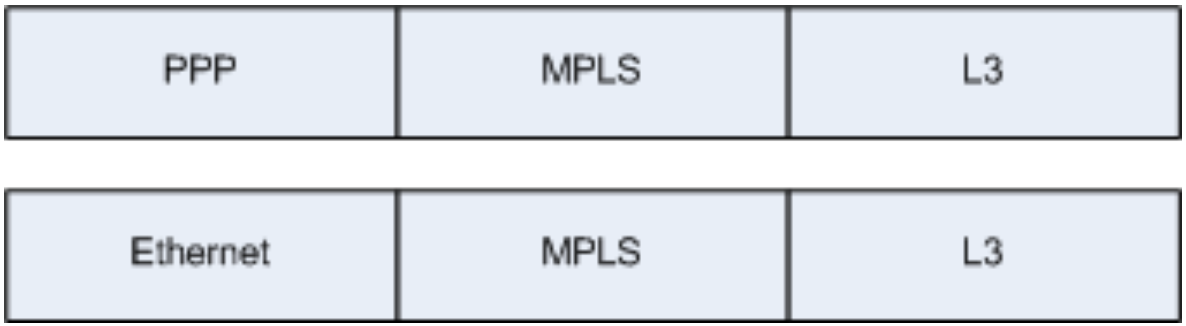
- Use labels for packet forwarding

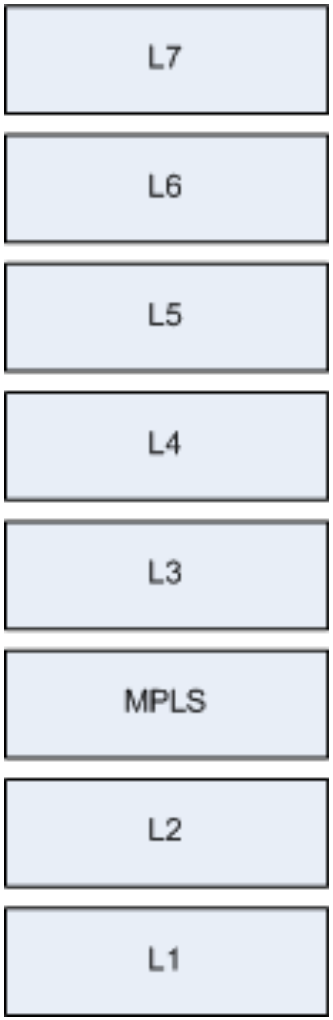


- Label
  - labels 0 – 15 reserved
- Exp – Experimental bits = QoS
- S – Bottom of stack – the last label
- TTL – the same as in IP protocol

MPLS – label

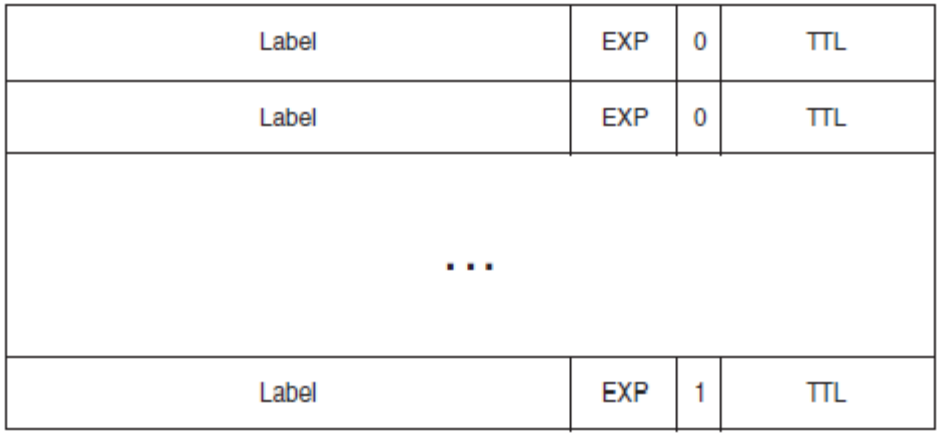
- MPLS label is inserted between L2/L3 layer
- Eliminates dependancy on L2 technology





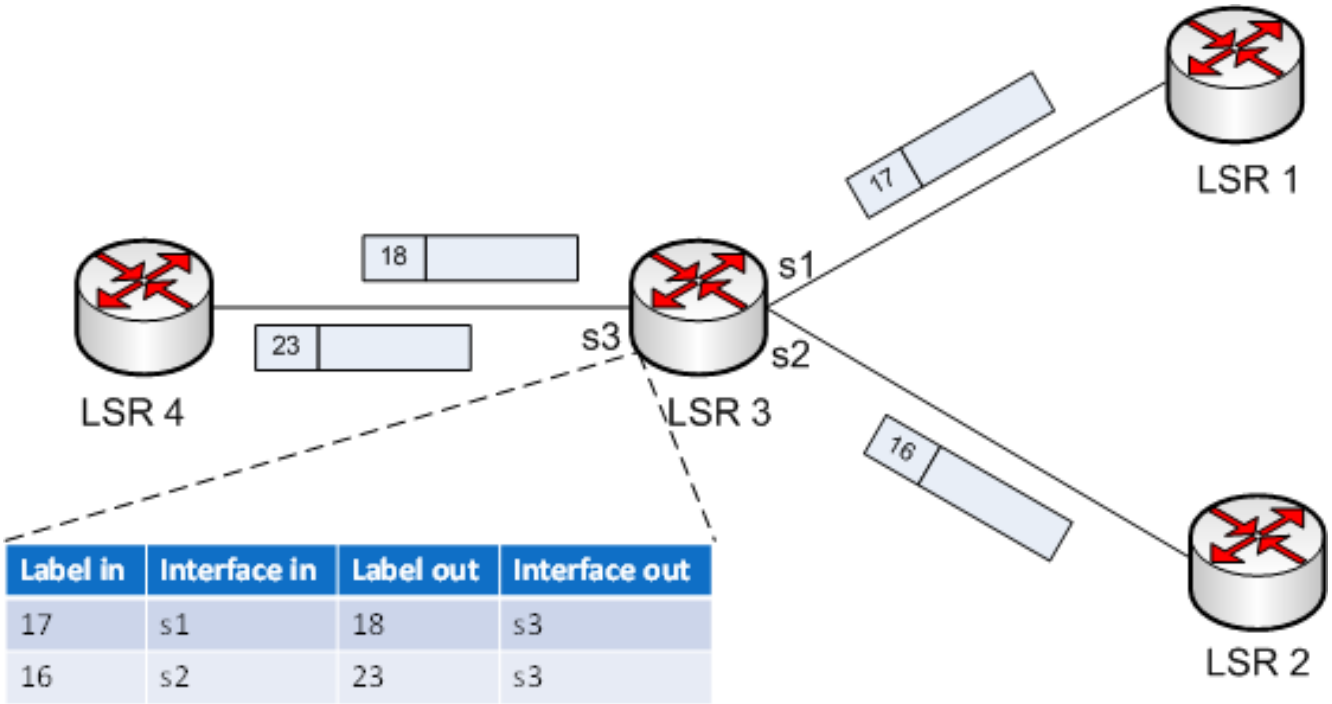
Label stack

- Several labels can be stacked
- MPLS switches operate with the topmost label only
- Last label in the stack is marked with the S bit



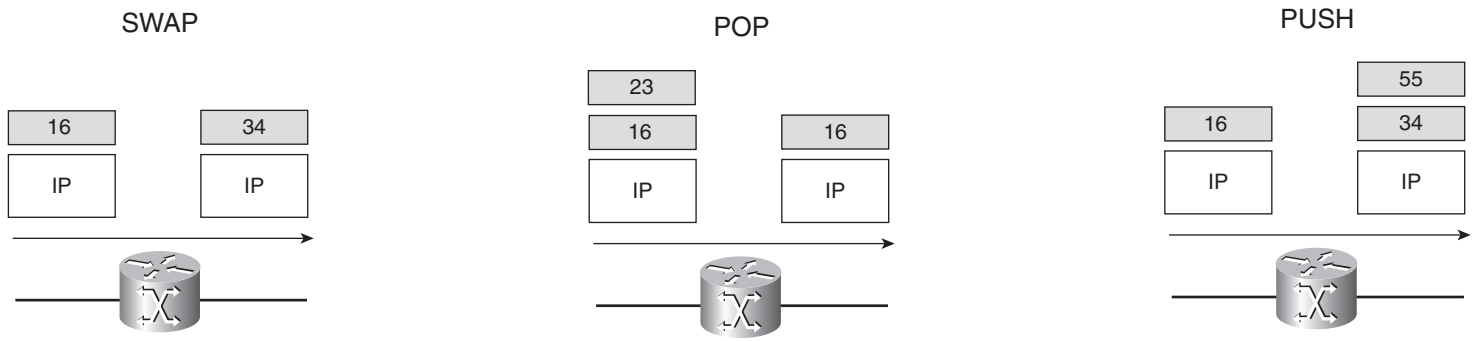
Label switching

- Labels are **locally** significant



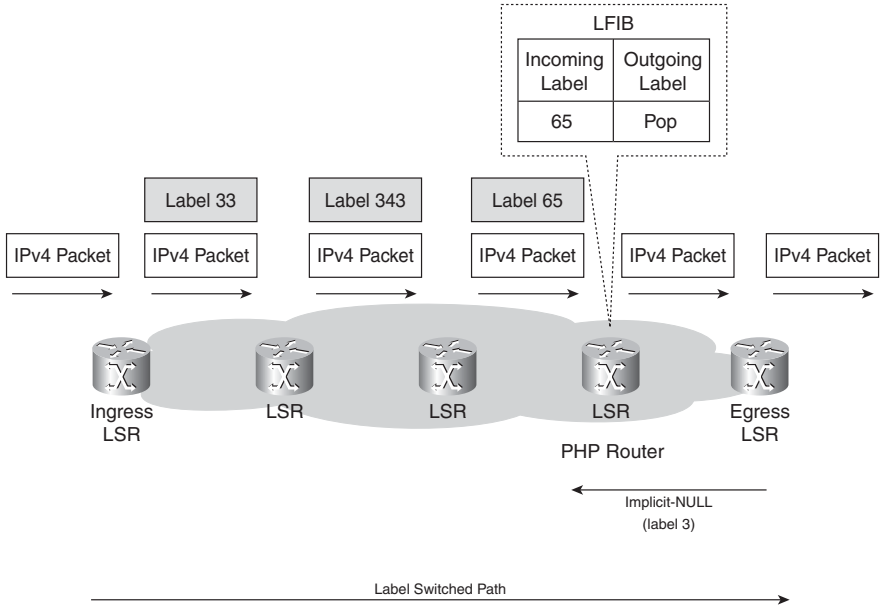
Label in	Interface in	Label out	Interface out
17	s1	18	s3
16	s2	23	s3

Label stack – possible actions



Reserved labels

- Implicit NULL
  - Value 3
  - Egress LSR requests the upstream LSR to perform a pop operation
    - Called PHP - Penultimate Hop Popping
  - The result - egress LSR receives an IP packet and only needs to perform an IP lookup to be able to forward the packet



Reserved labels

- Implicit NULL label removes the whole label – the EXP (QoS) bits are thus removed as well
- Explicit NULL label
  - Value 0 (2 for IPv6)
  - Egress LSR signals the explicit NULL label to the penultimate hop router
    - receives labeled packets with a label of value 0 (2) as the top label
  - Can remove the label without FIB search
  - QoS bits are preserved
  - However, label must be removed (POP action)

MPLS – basic terms

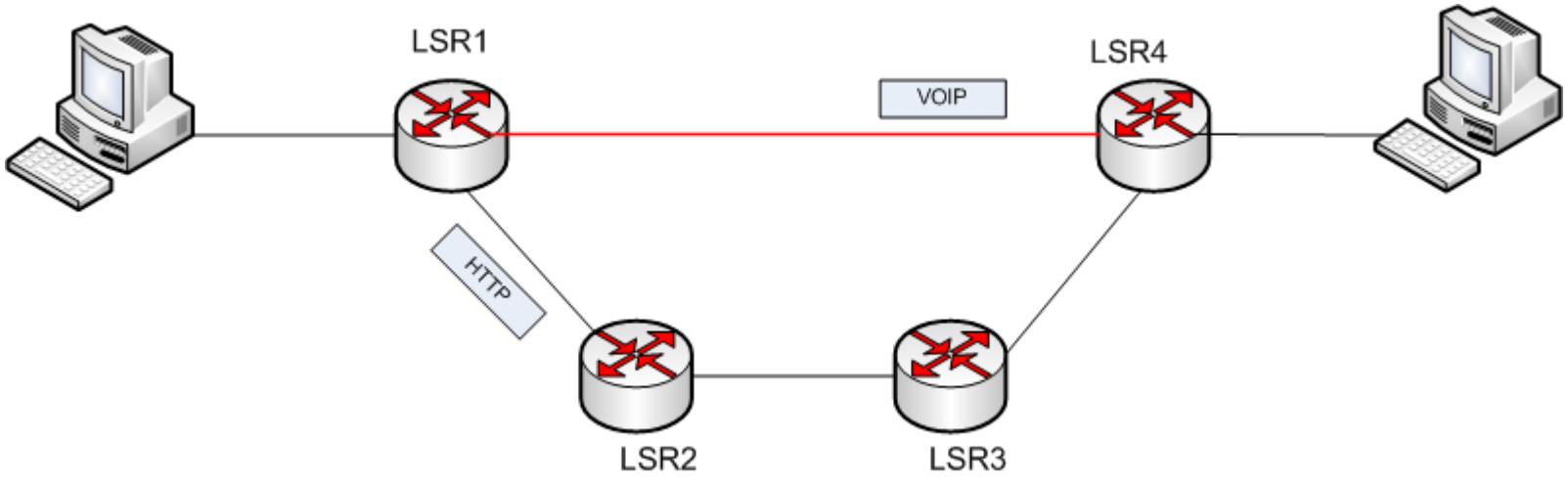
- Label
- Forwarding Equivalence Class (FEC)
- Label stack
- Label switching router (LSR)
- Label-switched path (LSP)
- Label Forwarding Information Base (LFIB)
- Label Distribution Protocol (LDP)
- Ingress/Egress MPLS node
- Edge MPLS node

FEC

- flow of packets that are forwarded along the same path and are treated the same
- All packets belonging to the same FEC have the same label
- Usually:
  - Packets with Layer 3 destination IP addresses matching a certain prefix
  - Multicast packets belonging to a certain group
  - the same BGP next hop
- Other examples:
  - VoIP traffic

FEC example

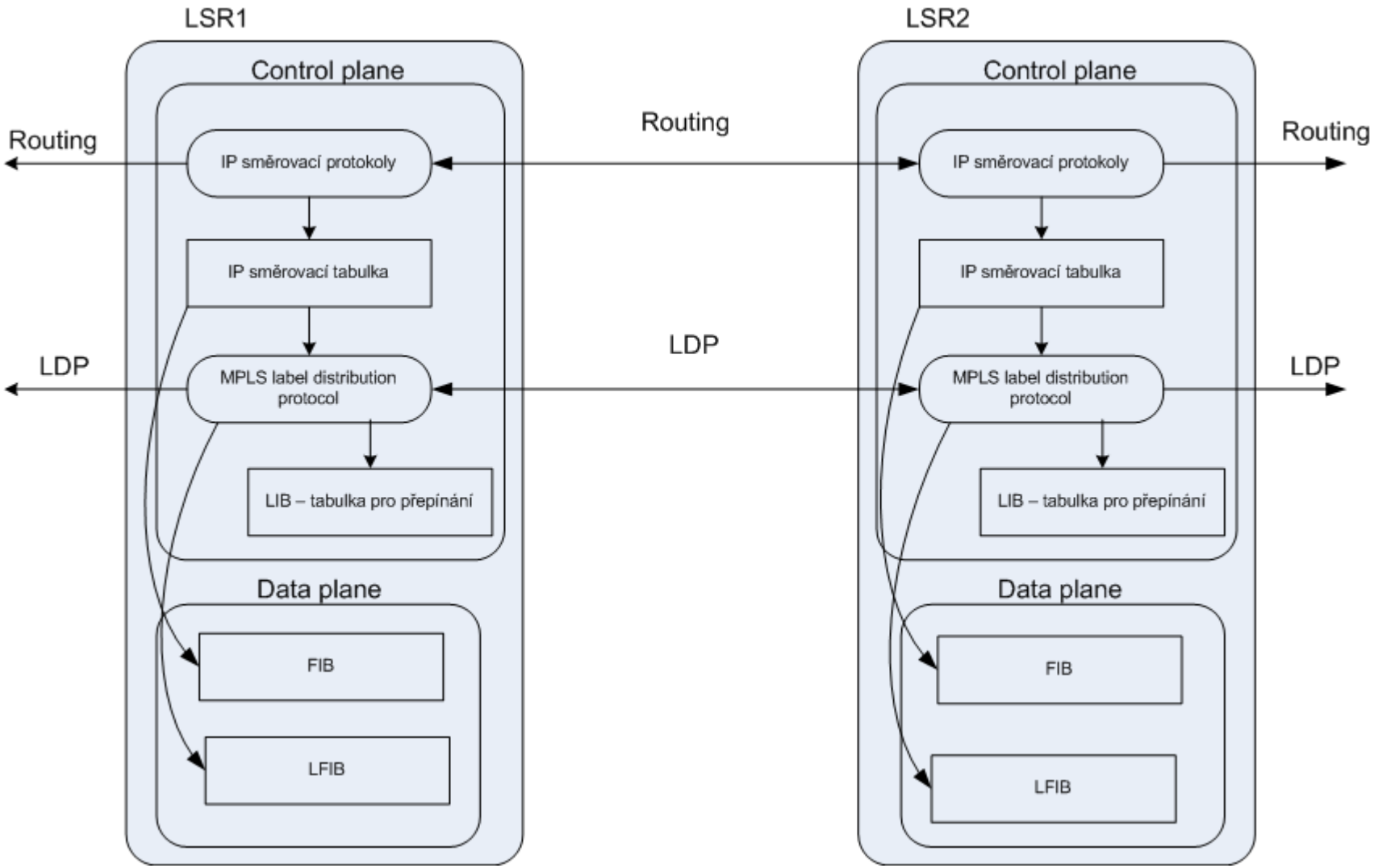
- LSP Paths
  - LSP1 = (LSR1, LSR4), LSP2 = (LSR1, LSR2, LSR3, LSR4)
  - LSP1 is faster than LSP2
  - FEC1 – all VoIP data, LSR1 adds label so all FEC1 packets go through LSP1
  - FEC2 – rest of the traffic use LSP2



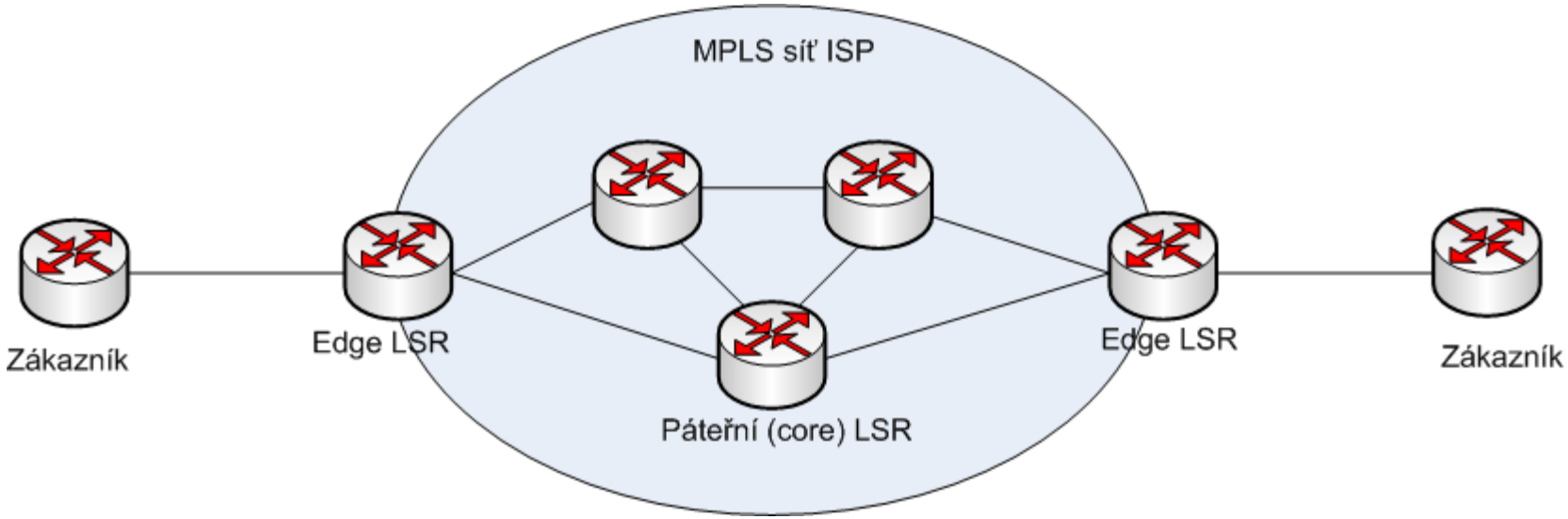
LIB - Label Information Base

- Routers use IGP protocol internally
- After network convergence – LIB table is created
- Can be tied with
  - CEF (FIB) – cache
  - LFIB – FIB, but for MPLS

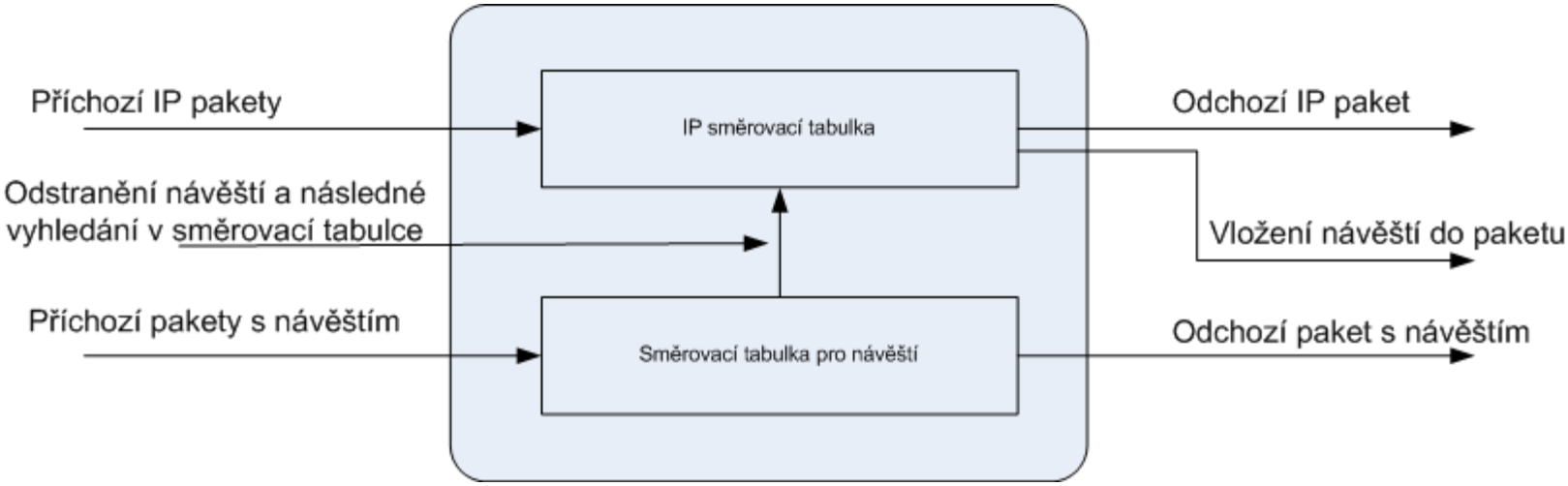
Router scheme



MPLS - network



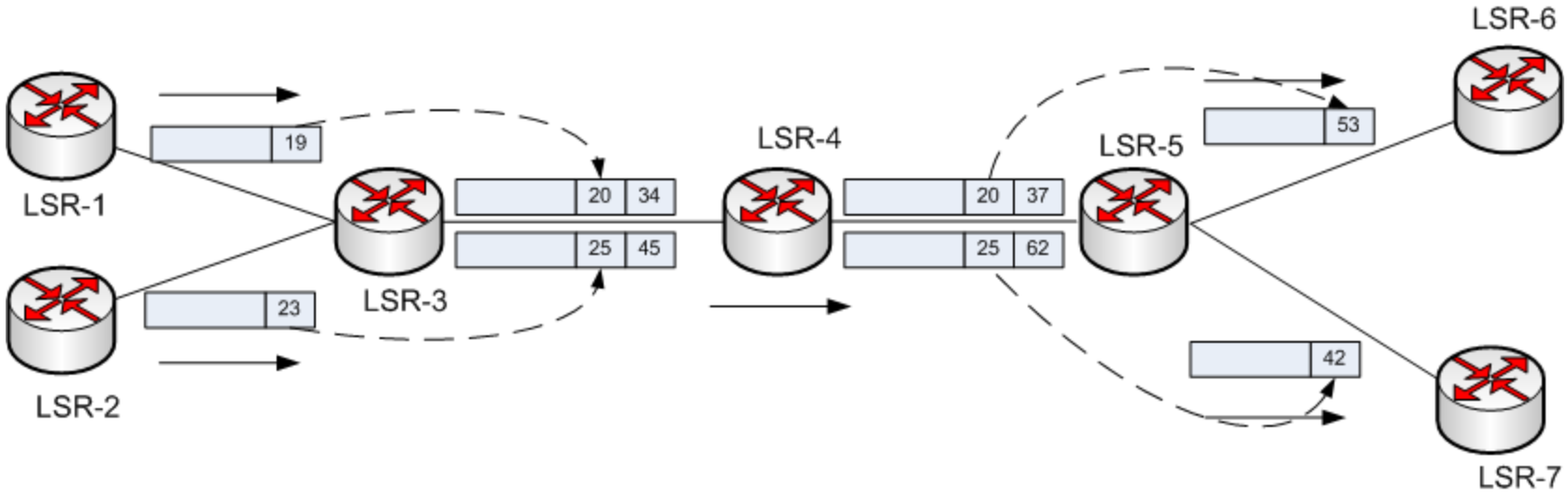
Edge LSR


Příchozí IP pakety
Odstranění návěští a následné vyhledání v směrovací tabulce
Příchozí pakety s návěštím
IP směrovací tabulka
Směrovací tabulka pro návěští
Odchozí IP paket
Vložení návěští do paketu
Odchozí paket s návěštím

LSP – label switched path

- LSP – path of LSRs in MPLS network
- First LSR in LSP path – ingress LSR, last – egress LSR
- LSP is unidirectional
- Ingress LSR does not have to be the first router that adds label – labels can be stacked

LSP – tunnel


LSR-1
LSR-2
LSR-3
LSR-4
LSR-5
LSR-6
LSR-7
19
23
20
34
25
45
20
37
25
62
53
42

LDP – Label Distribution protocol

- Labels are locally significant
- LDP protocol is used to distribute labels between routers

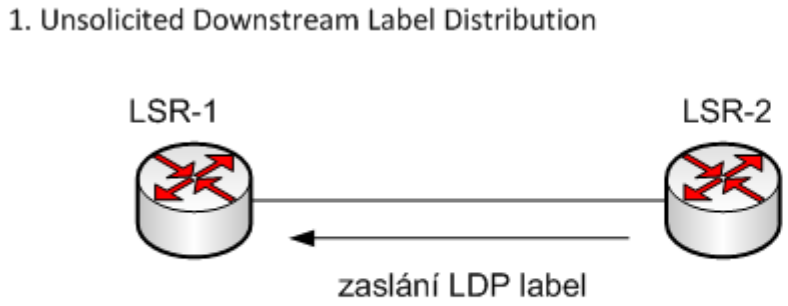
- Discovery messages – discover other LSR in network, UDP multicast
- Session messages – connection between LDP nodes
- Advertisement messages – changes, creation, deletion labels for FEC
- Notification messages – information messages and management signalization

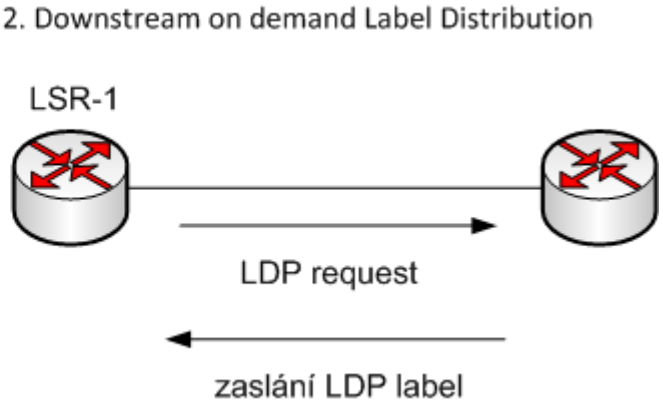
LDP – discovery

- UDP packet to all-routers multicast (224.0.0.2), or UDP unicast port (646), to a specific IP address
- After detection – TCP LDP relation

LDP – labels distribution ①

1. Unsolicited Downstream Label Distribution
2. Downstream on Demand Label Distribution





LDP – label distribution ②

1. Unsolicited Downstream Label Distribution
   - LSR-2 finds a new next hop for a FEC
   - Create label and send to LSR-1

2. Downstream on Demand Label Distribution
   - LSR-1 discover that next hop for FEC is LSR-2
   - LSR-1 asks LSR-2 for a label
   - LSR-2 sends the label LSR-1

Label distribution ③

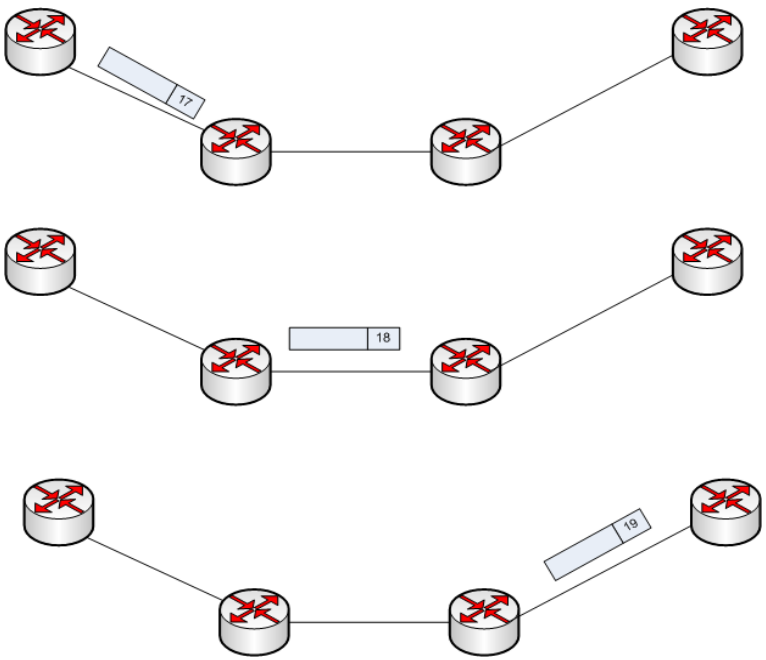
- LDP is used to create a LSP
- Independent:
  - LSR creates a local binding for a FEC independently from the other LSRs
  - each LSR creates a local binding for a particular FEC as soon as it recognizes the FEC
    - Usually, this means that the prefix for the FEC is in its routing table.
  - some LSRs begin to label switch packets before the complete LSP is set up end to end
- Ordered:
  - LSR only creates a local binding for a FEC
    - if it recognizes that it is the egress LSR for the FEC
    - if the LSR has received a label binding from the next hop for this FEC

Label distribution using IGP protocol

- Extensions for IGP protocols exist, allowing label distribution
- Supported in e.g. IS-IS, OSPF
  - Often too complicated, usually LDP is preffered
- BGP – often use for labels in MPLS – VPN

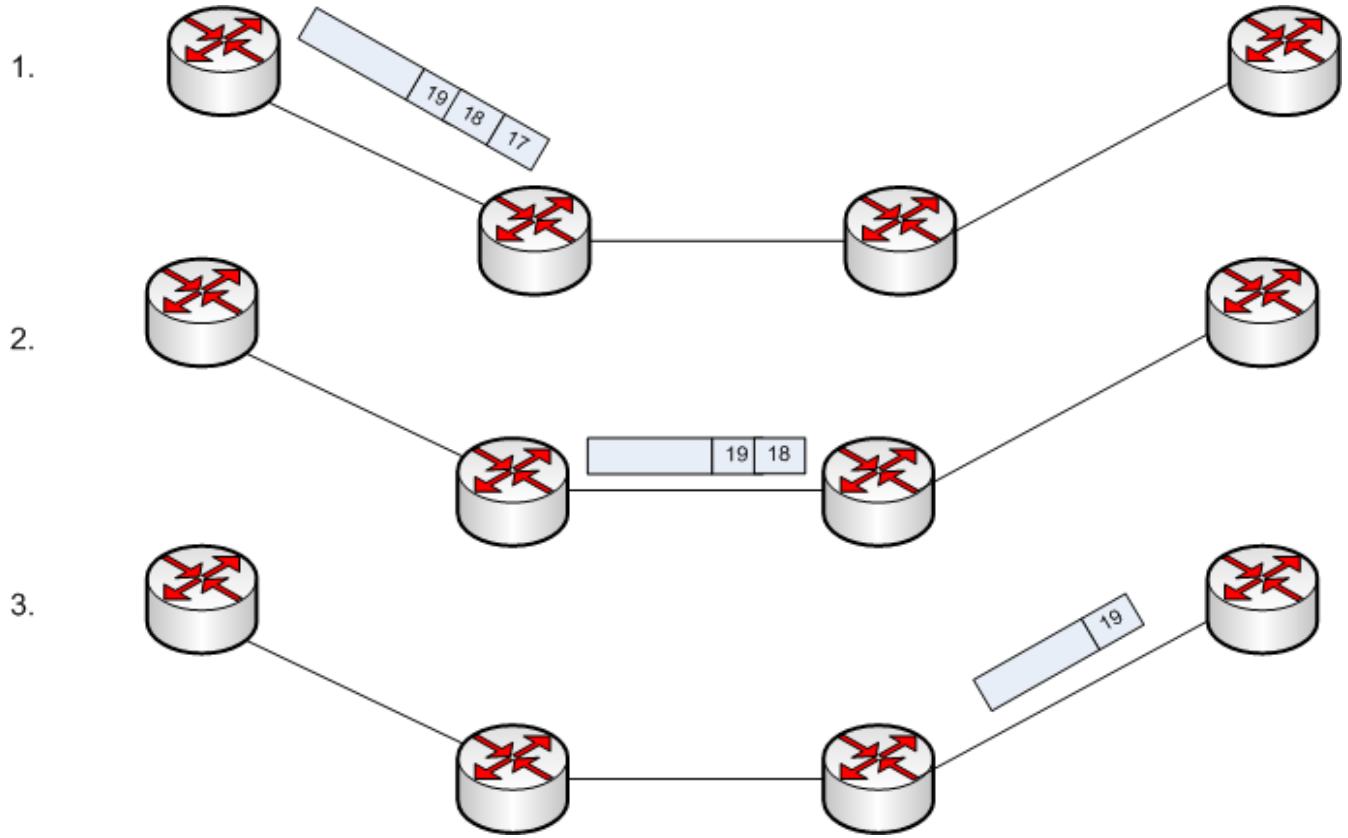
Hop-by-hop routing

- Every LSR use SWAP operation
- LSP – the same paths as classic routing

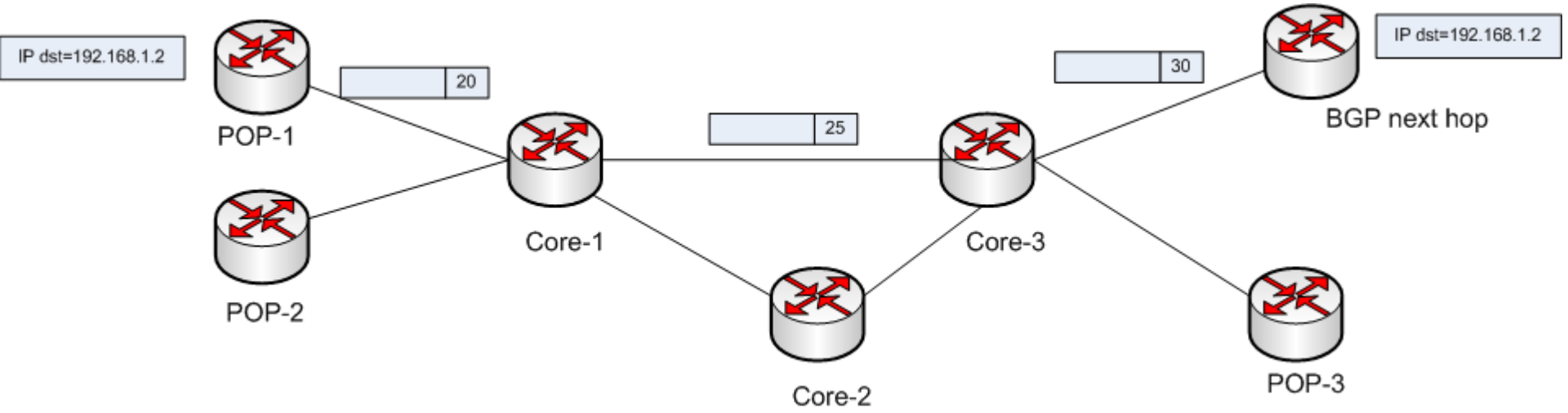


Explicit routing

- Pre-defined path in the network



BGP free core



- POP-1 – accepts packet with destination IP 192.168.1.2
- POP-1 – finds network in routing table, push MPLS label, MPLS label = BGP – Next hop
- Core-1 – Core-3 do not have to run BGP – only MPLS switching
- BGP-Prague removes MPLS label, standard IP lookup

MPLS VPN

MPLS – VPN

- The mot popular use case for MPLS
- Full-meshed connectivity for customers
- ISP infrastructure is hidden for customer
- Provider Edge routers can be shared between several customers (cheaper)

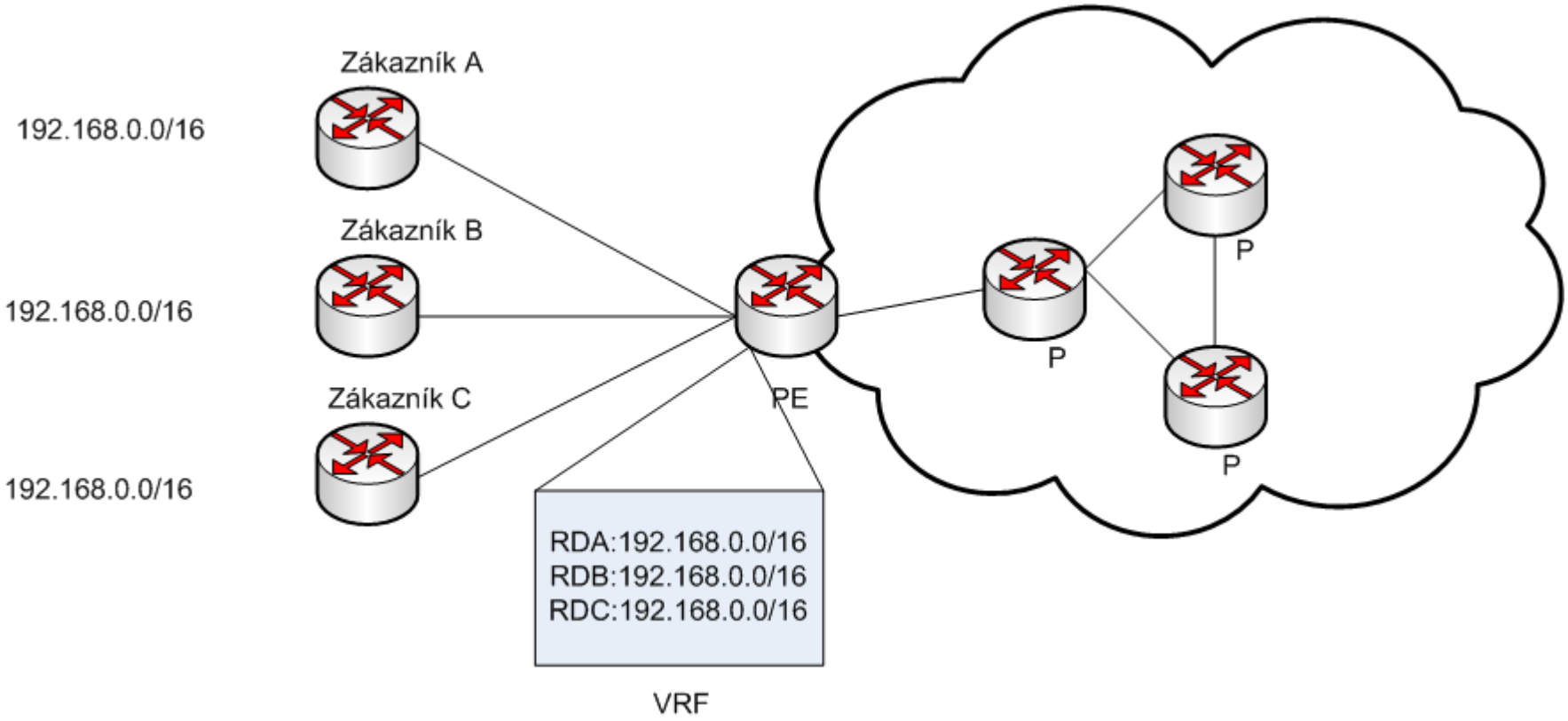
MPLS – VPN basic terms

- CE – Customer edge router
- PE – Provider edge router
- P router – Provider‘s MPLS switch/router
- PHP – Penultimate hop pop
- PoP – Point of presence
- RD – Route distinguisher
- VRF – Virtual routing and forwarding table

CE a PE

- CE
  - Router in customer‘s network
  - Typically runs IGP (OSPF, RIP, static routing)
  - Not aware of MPLS network
- PE
  - ISP‘s PoP – can connect several customers
  - Customers are distinguished based on RD
  - Own VRF for every customer (private addresses)

PE architecture



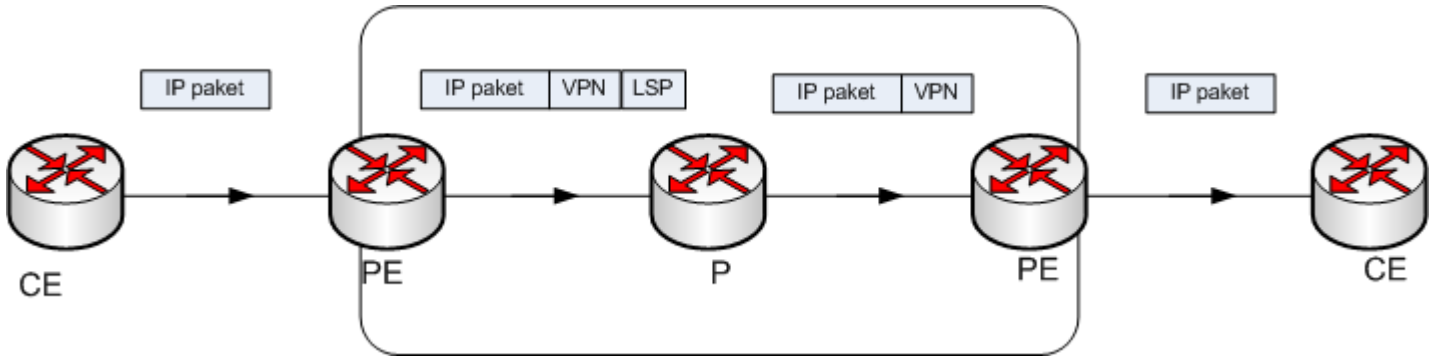
RD – Router Distinguisher

- Customers can use same private address space (RFC 1918)
- IPv4 address + RD = unique identifier
- RD = 64 bits = 96 bits VPNv4 address
- VPNv4 addresses are distributed between PE routers using BGP

MPLS VPN – labels

- VPN routing use two labels
- First label – LSP path through MPLS ISP network
- Second label – VPN
- P routers use only LSP path label

1. CE sends IP packet
2. PE push label for VPN (S bit = 1) and label for LSP
3. Penultimate router removes LSP label
4. PE router use VPN label to distinguis between customers



MPLS VPLS

VPLS

- L2 connectivity achieved accross L3 network
- Whole L2 frame can be encapsulated using MPLS
- L2 interconnectivity between geographical different buildings

VPLS ①

- CE routers see the connection as L2 switch
- VPLS needs to achieve the same properties
  - Ethernet forwarding
  - Broadcasting ethernet frames with unknown ethernet MAC address
  - Broadcast/multicast replication
  - Loop detection
  - Dynamic MAC learning

VPLS ②

- Connection between PE routers – **pseudowire**
- CE router sends Ethernet broadcast frame to PE router
  - PE sends the frame to all VPLS physical ports and VPLS pseudowires
- PE routers must form full-meshed network
  - Loop prevention, STP elimination

VPLS ③

- Two labels are used for packet forwading
- First – LSP path label for packet
- Second – pseudowire label

- full-meshed requires configuration on all PE routers
  - LDP relation does not use multicast but directed unicast