ALEF NULA

DŮVĚŘUJTE SILNÝM

MPLS Applications

Karel Pouzar
CCIE#20198, CCSI#31414
karel.pouzar@alefnula.com

Agenda

ALEF NULA
DŮVĚŘUJTE SILNÝM

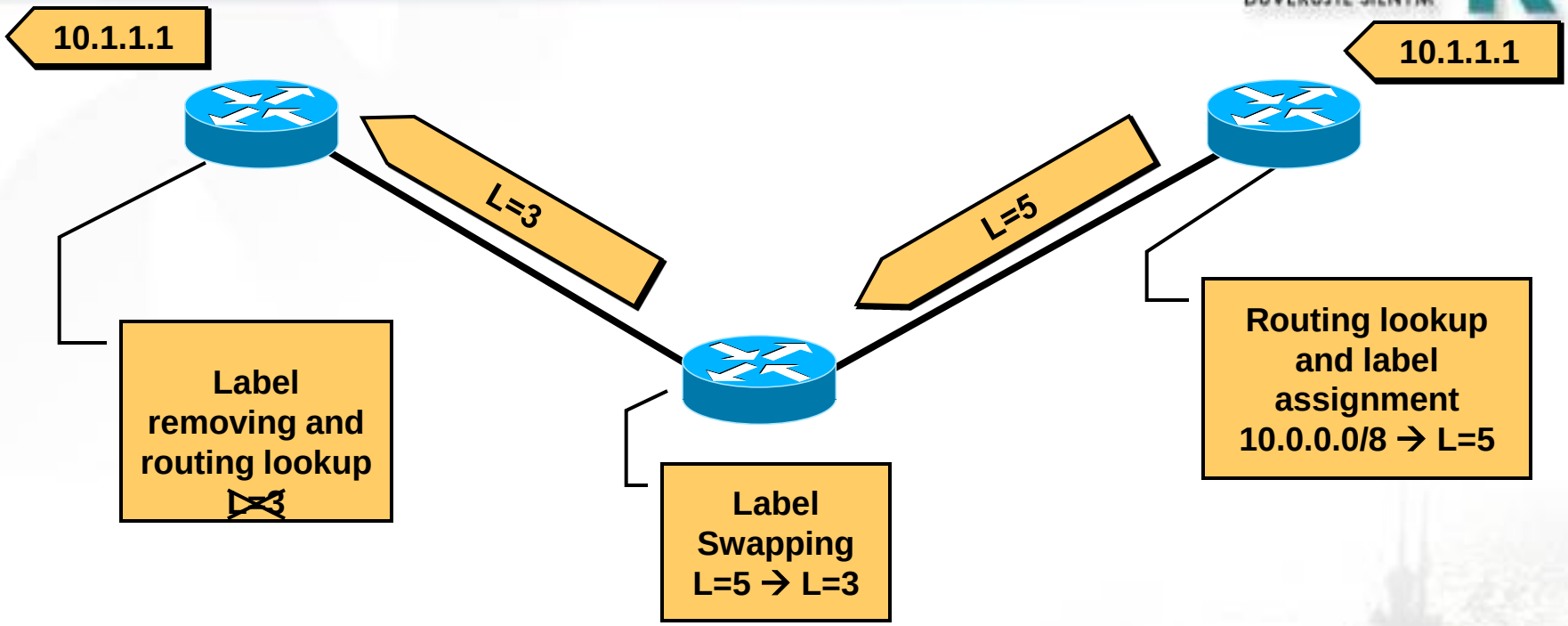
- MPLS Introduction
- MPLS VPN Architecture
- MPLS L2 Services
  - AToM, EoMPLS
  - VPLS, H-VPLS

ALEF NULA

DŮVĚŘUJTE SILNÝM

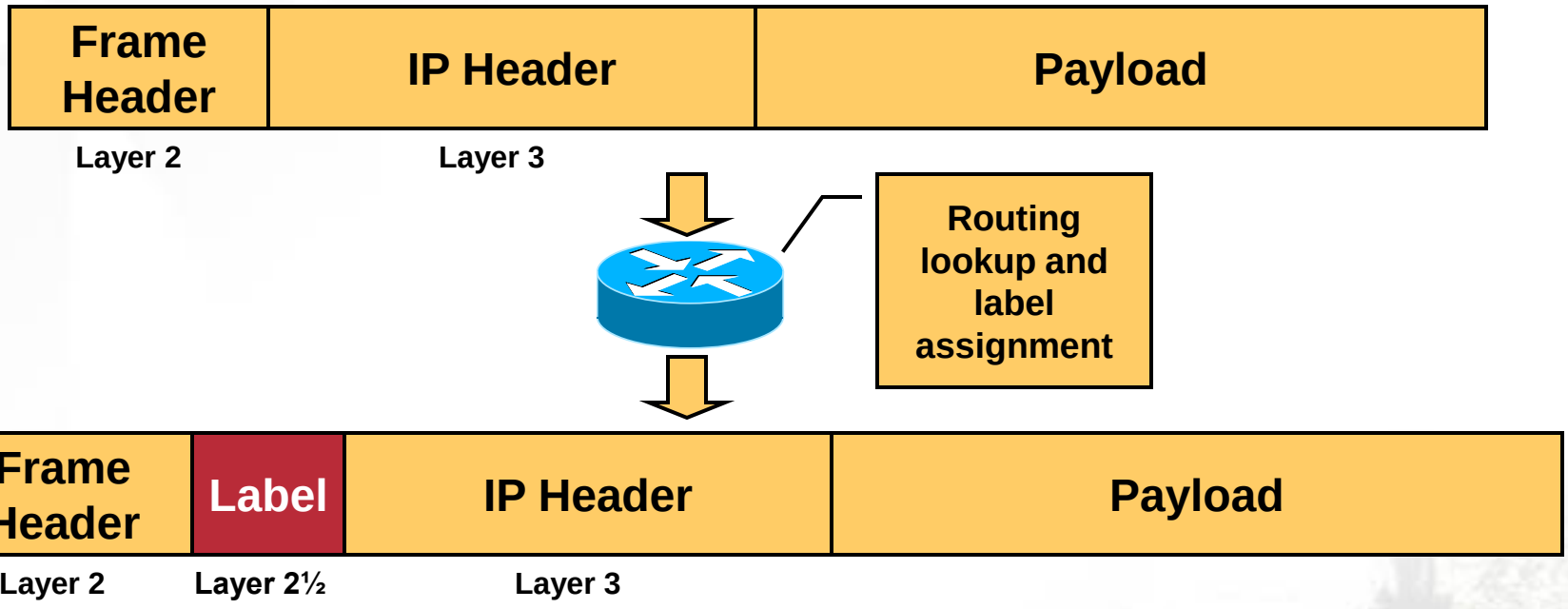
Introduction

Concepts of MPLS – example



- Routing (L3) lookup is done only by edge routers.
- Core routers switch packets according to simple label (L2) look-ups and do the label swapping as well.

Frame mode MPLS



MPLS applications

- Nowadays, there are many different applications of MPLS:
  - Unicast IP routing
  - Multicast IP routing
  - Virtual private networks (MPLS VPN – L3)
  - AToM and L2 VPN
  - Traffic Engineering (MPLS TE)
  - QoS
- Independent on application, MPLS functionality is always divided into **control plane** and **data plane**:
  - MPLS applications differ only in control plane
  - Every MPLS applications use the same label-switching data plane
  - Edge LSR Layer 3 data planes can differ
  - Generally, a label is assigned to each **forwarding equivalence class (FEC)**

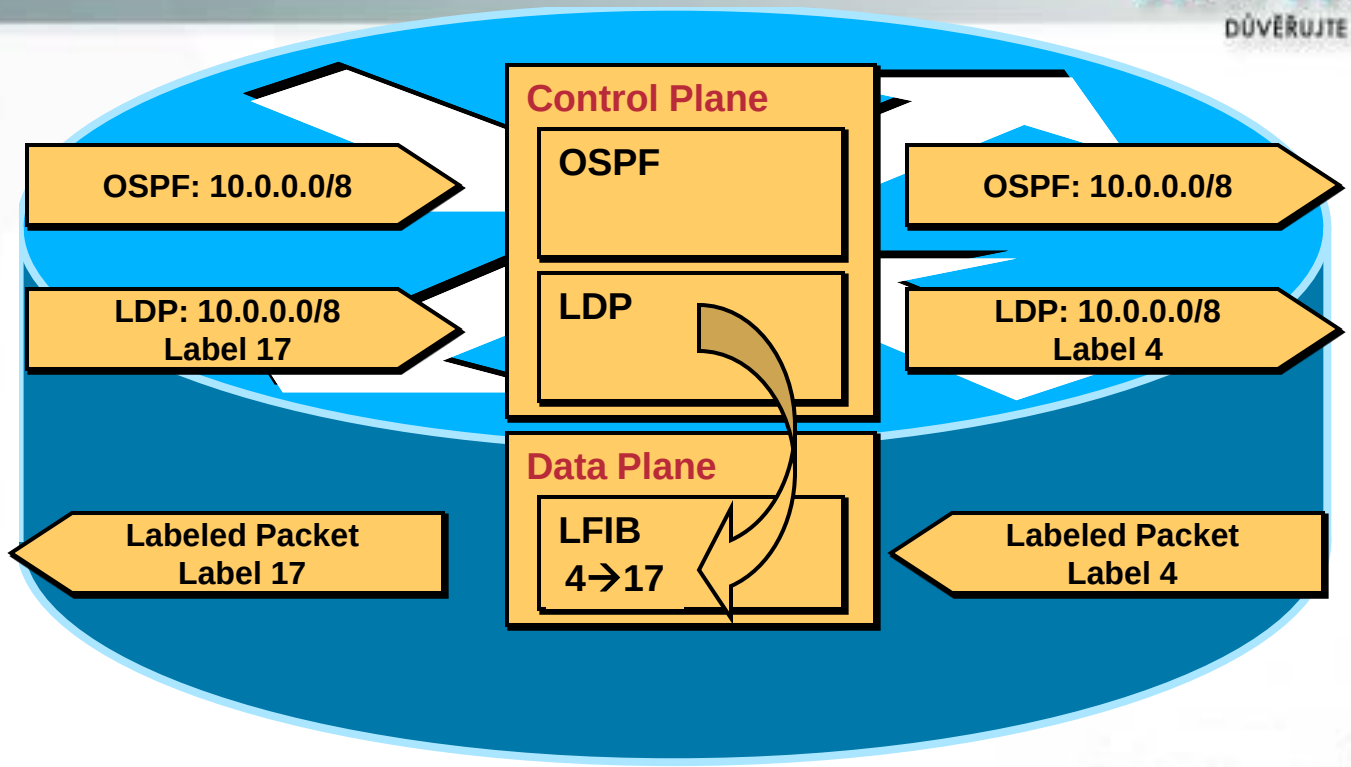
MPLS architecture

ALEF NULA
DŮVĚŘUJTE SILNÝM

- Architecture of MPLS node is divided into two major components:
  - Control component (Control plane)
    - Exchanges Layer-3 routing information and labels
  - Forwarding component (Data plane)
    - Forwards packets based on labels

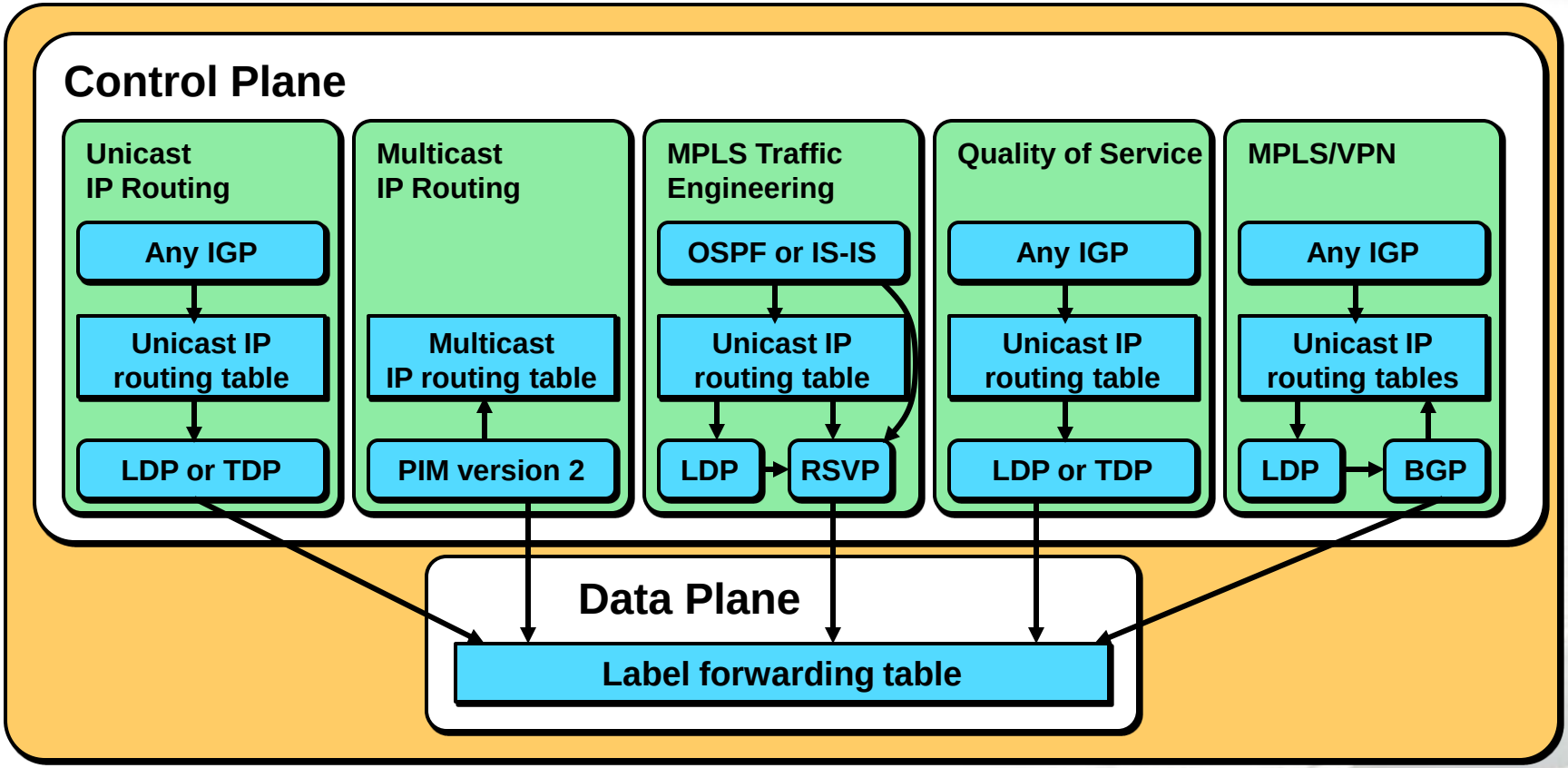
- Provides the separation of forwarding & control plane
  - Forwarding should not be interrupted by value alternation of assigned labels

MPLS architecture (cont.)



- Router functionality is divided into two major parts:
  - Control plane
  - Data plane

Interaction among MPLS applications



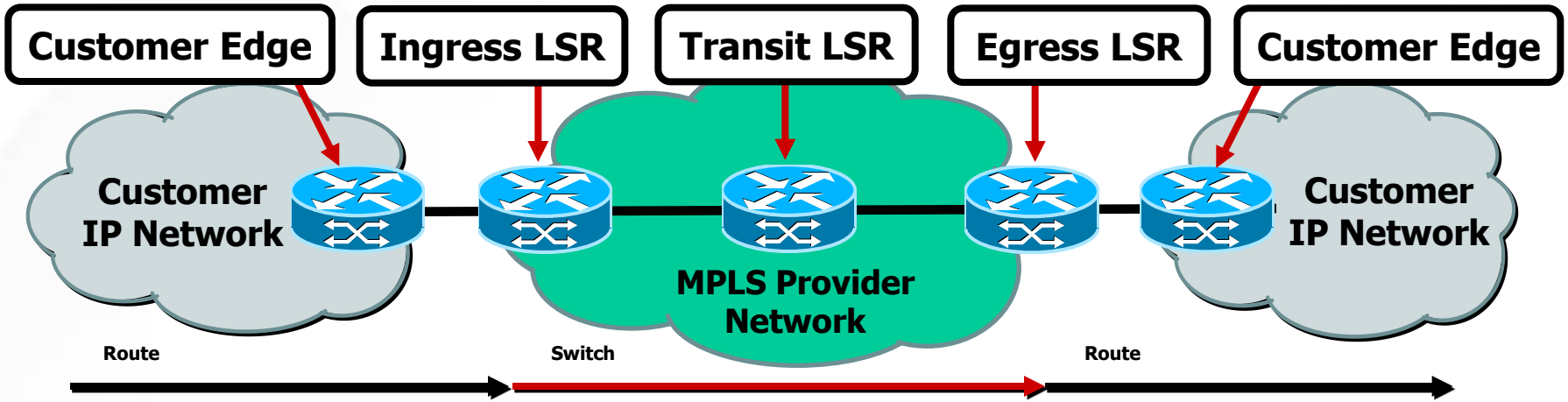
ALEF NULA

DŮVĚŘUJTE SILNÝM

MPLS VPN Architecture

MPLS Components

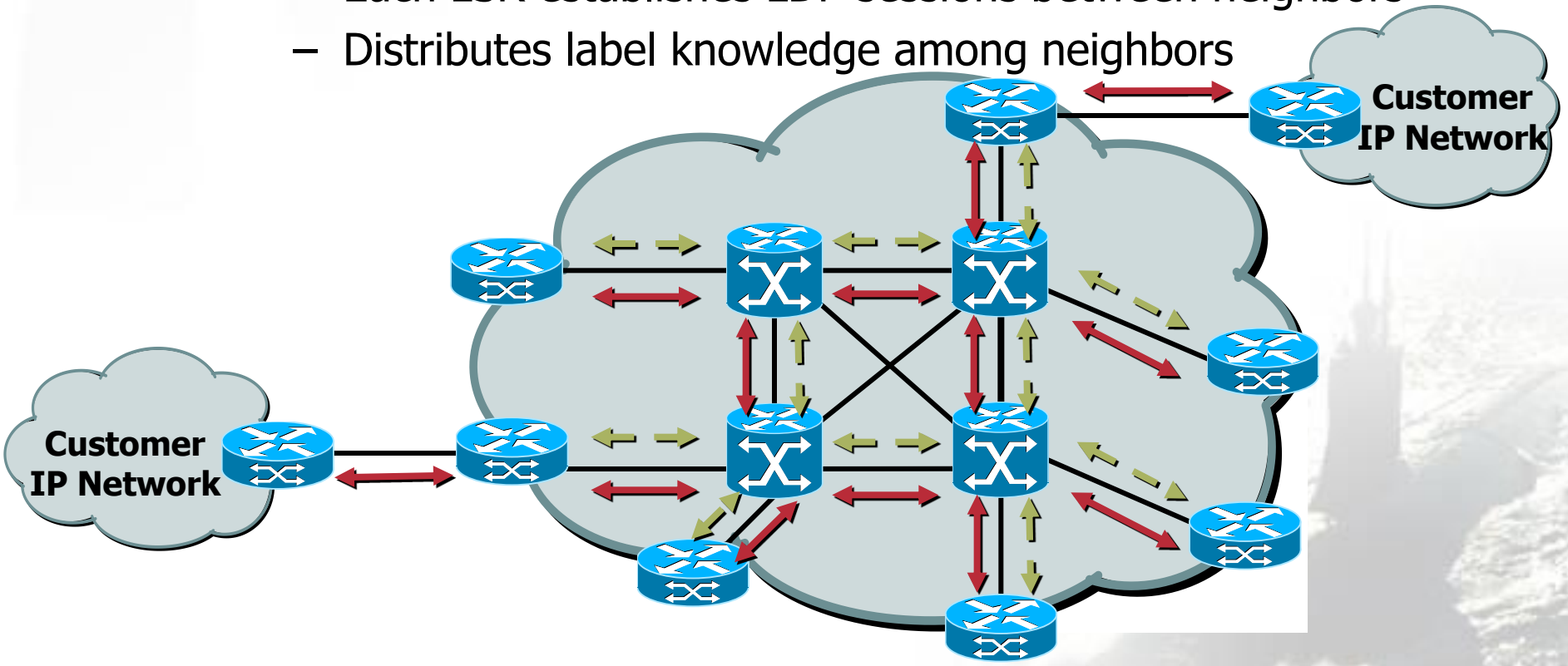
LSR = Label Switch Router



- **Customer Edge Router**
  - Connects a customer to the service provider's network, does not support MPLS
- **Ingress LSR**
  - Translates destination IP addresses into labels, edge LSR
- **Transit LSR**
  - Switches packets according to labels, core LSR
- **Egress LSR**
  - Removes labels and forwards packets to the customer, edge LSR

Label Distribution Protocol

- IGP is used to distribute IP routing information
- LDP is used for label discovery
  - Each LSR establishes LDP sessions between neighbors
  - Distributes label knowledge among neighbors

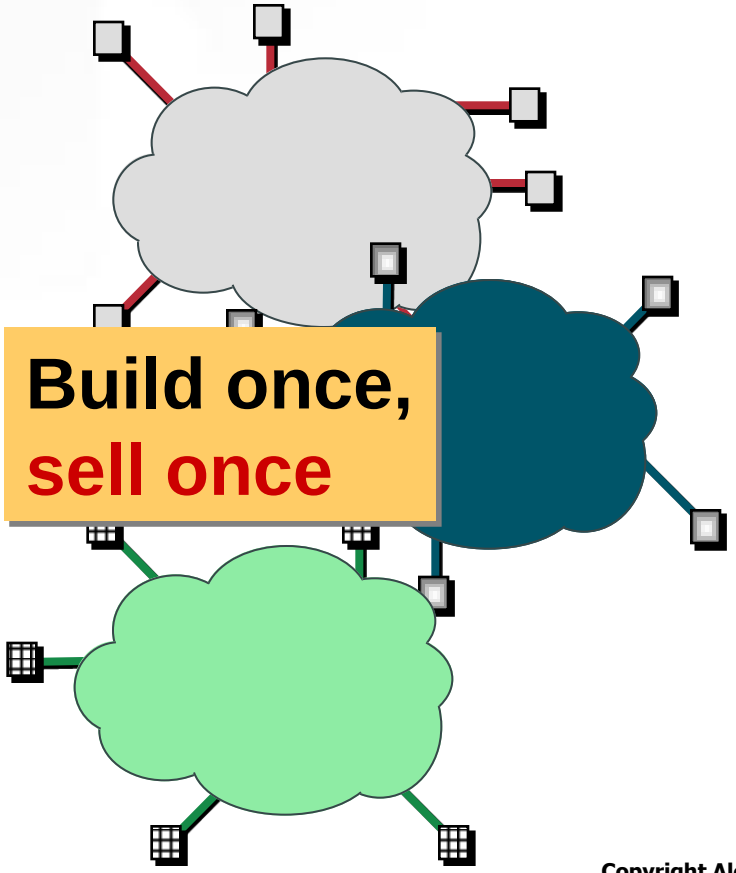


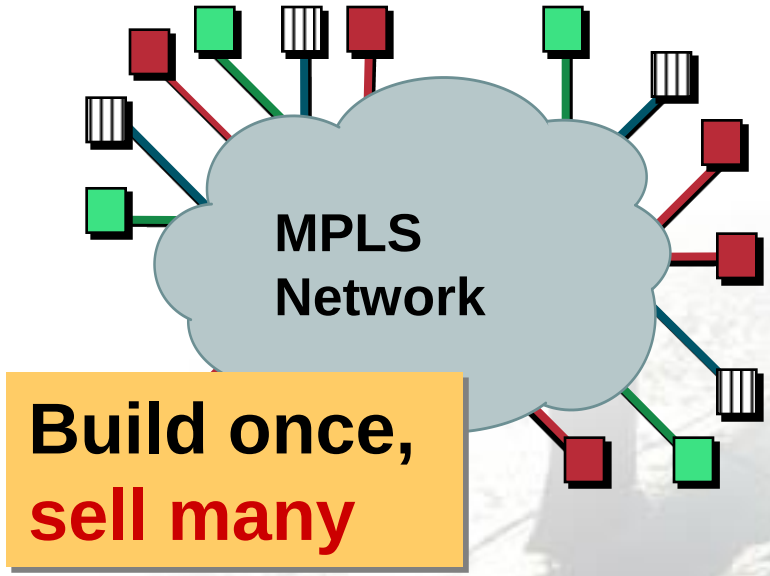
New Approach to Network Administration

Built separately
customer's private IP networks

Vs.

Single transport network for more customer's IP VPNs



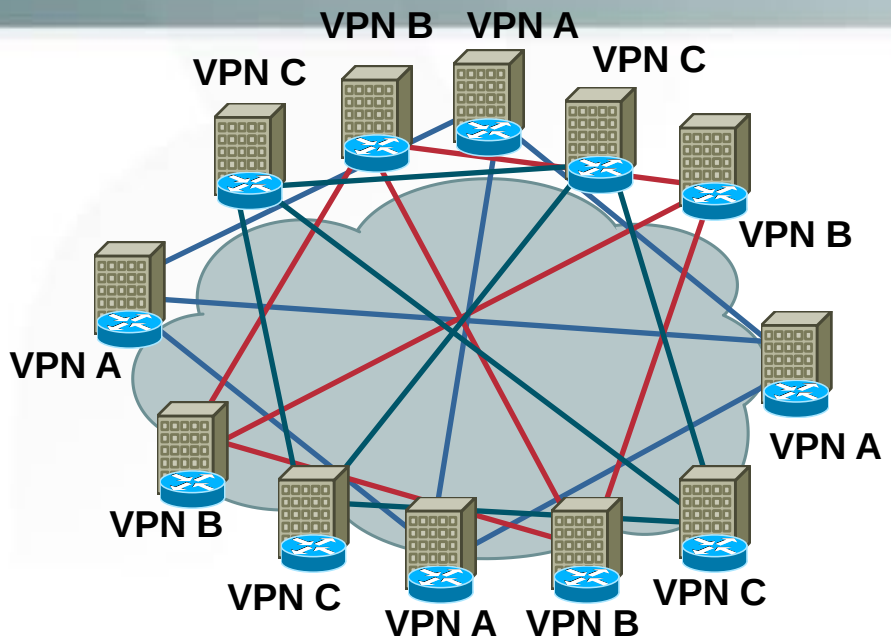


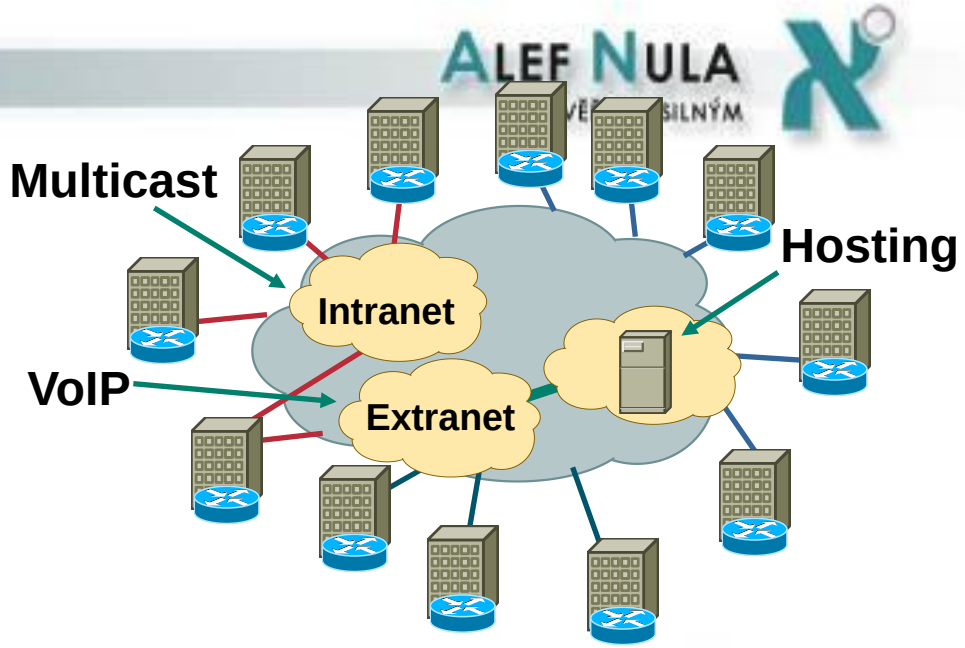
What Is a VPN?

ALEF NULA
DŮVĚŘUJTE SILNÝM

- VPN is a set of sites which are allowed to communicate with each other
- VPN is defined by a set of administrative policies
  - –Policies determine both connectivity and QoS among sites
- Flexible inter-site connectivity
  - – Ranging from complete to partial mesh
- Sites may be either within the same or in different organizations
  - – VPN can be either intranet or extranet
- Site may be in more than one VPN
  - – VPNs may overlap
- Not all sites have to be connected to the same service provider
  - – VPN can span multiple providers

L2 (with IP) vs. MPLS L3 VPNs





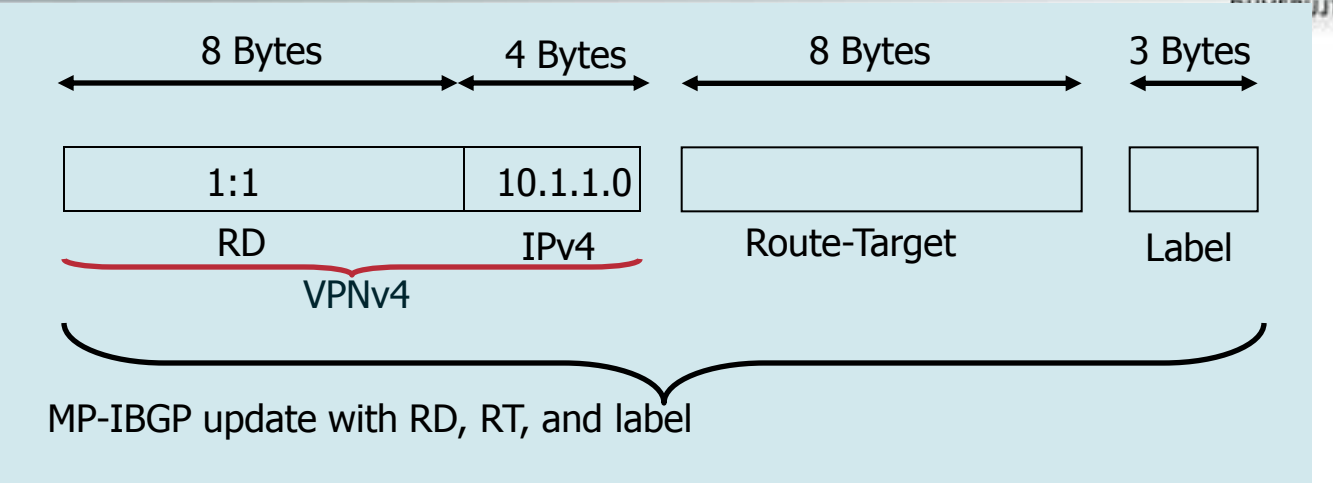
Overlay VPN

- ACLs, ATM/FR, IP tunnels, IPsec, etc. requiring n*(n-1) peering points
- Transport dependent
- Groups endpoints, not groups
- Pushes content outside the network
- Costs scale exponentially
- NAT necessary for overlapping address space
- Limited scaling
- QoS complexity

Peer-to-Peer MPLS-Based VPNs

- Point to cloud single point of connectivity
- Transport independent
- Easy grouping of users and services
- Enables content hosting inside the network
- "Flat" cost curve
- Supports private overlapping IP addresses
- Scalable to over millions of VPNs
- Per VPN QoS

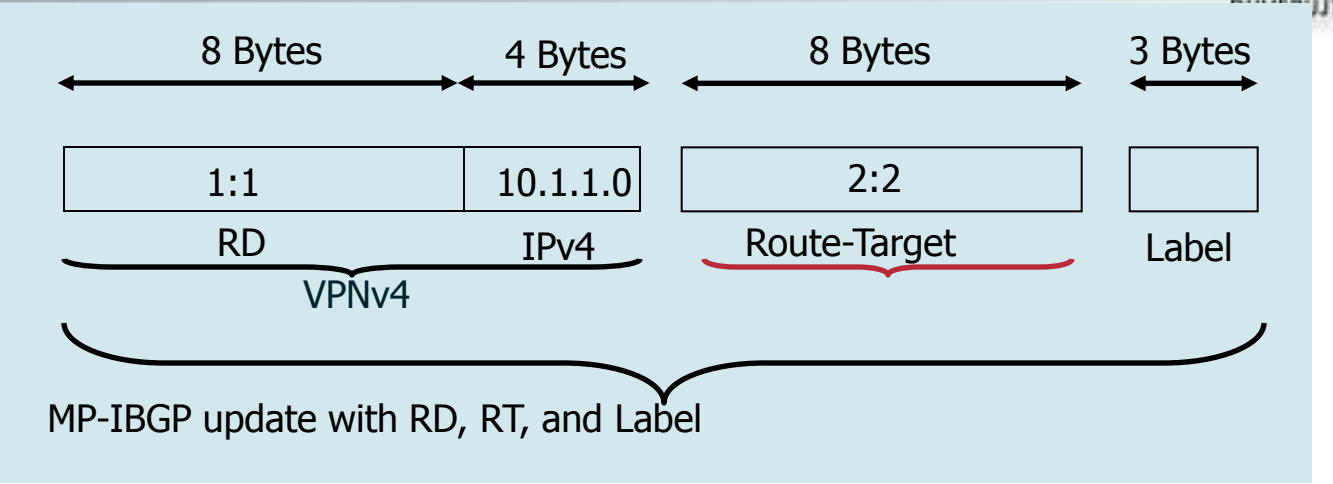
MP-BGP Update Components: VPNv4 Address



- To convert an IPv4 address into a VPNv4 address, RD (Route Distinguisher) is appended to the IPv4 address i.e. 1:1:10.1.1.0
  - Makes the customer's IPv4 route globally unique
- Each VRF must be configured with an RD at the PE
  - RD is what that defines the VRF

```
!
ip vrf v1
 rd 1:1
!
```

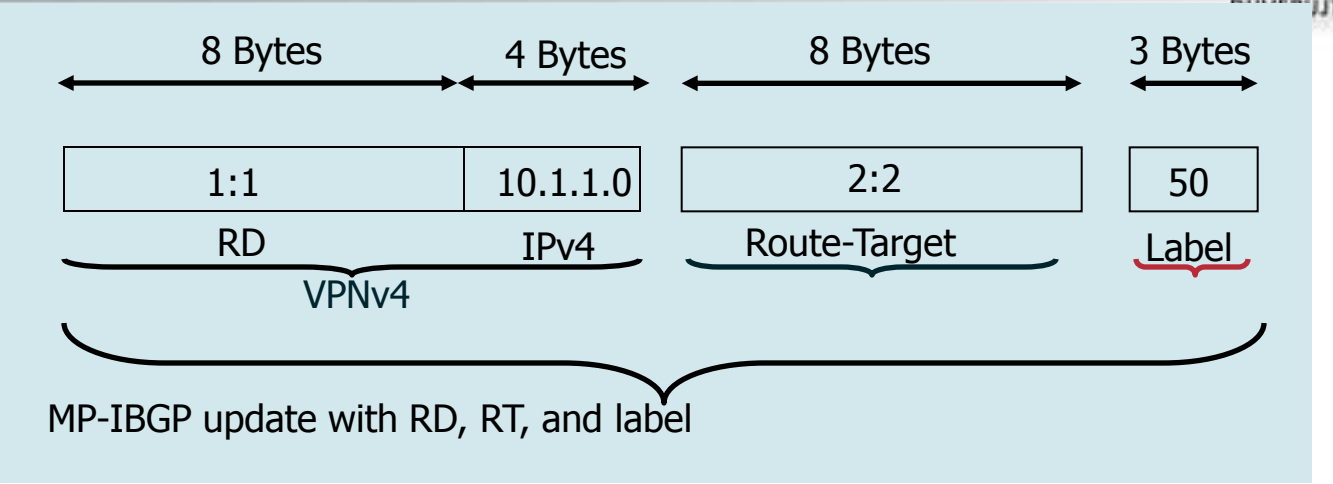
MP-BGP Update Components: Route-Target



- Route-target (RT): Identifies the VRF for the received VPNv4 prefix. It is an 8-byte extended community (a BGP attribute)
- Each VRF is configured with RT(s) at the PE
  - RT helps to color the prefix

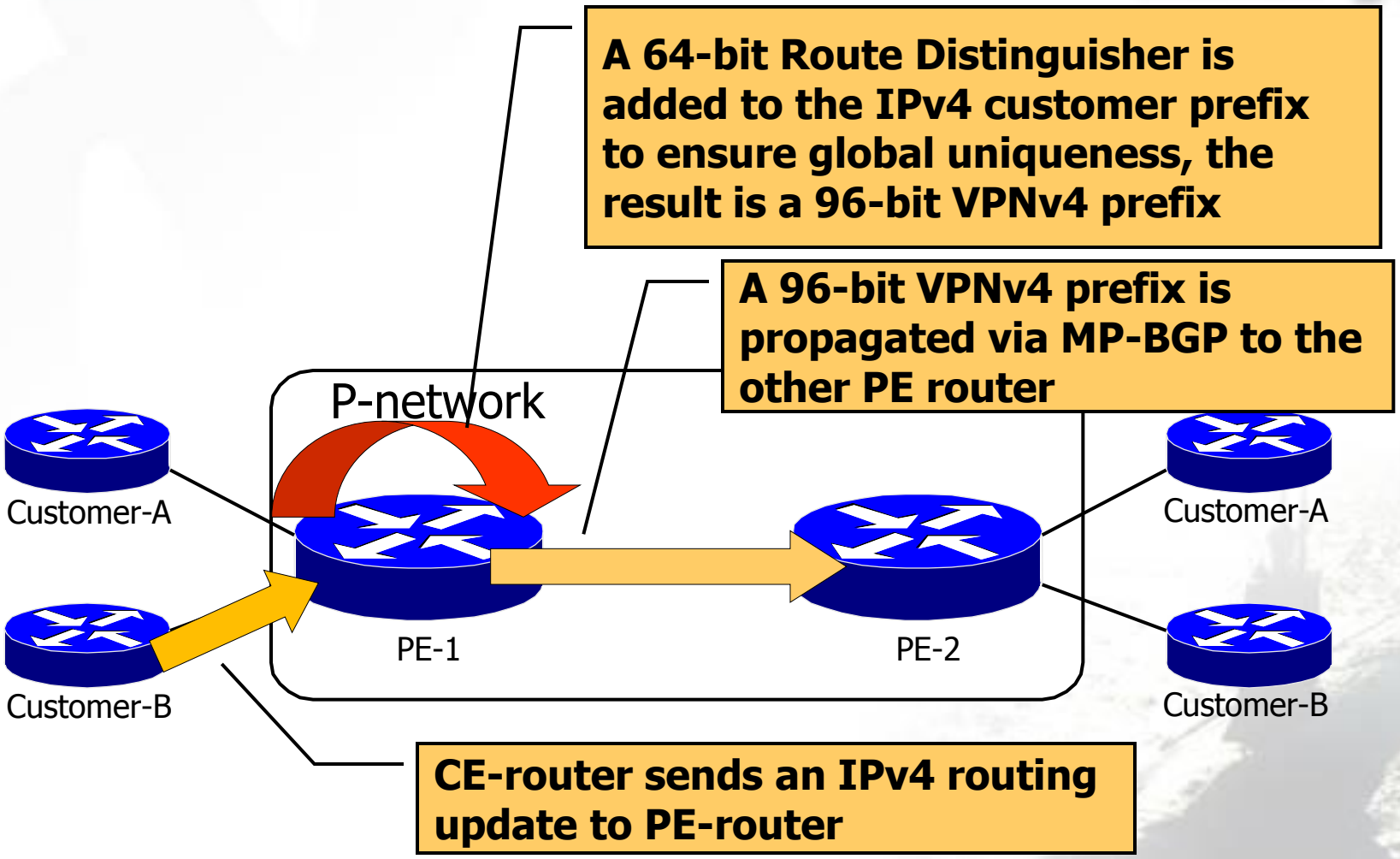
```
!
ip vrf v1
 route-target import 1:1
 route-target export 1:2
!
```

MP-BGP Update Components: Label

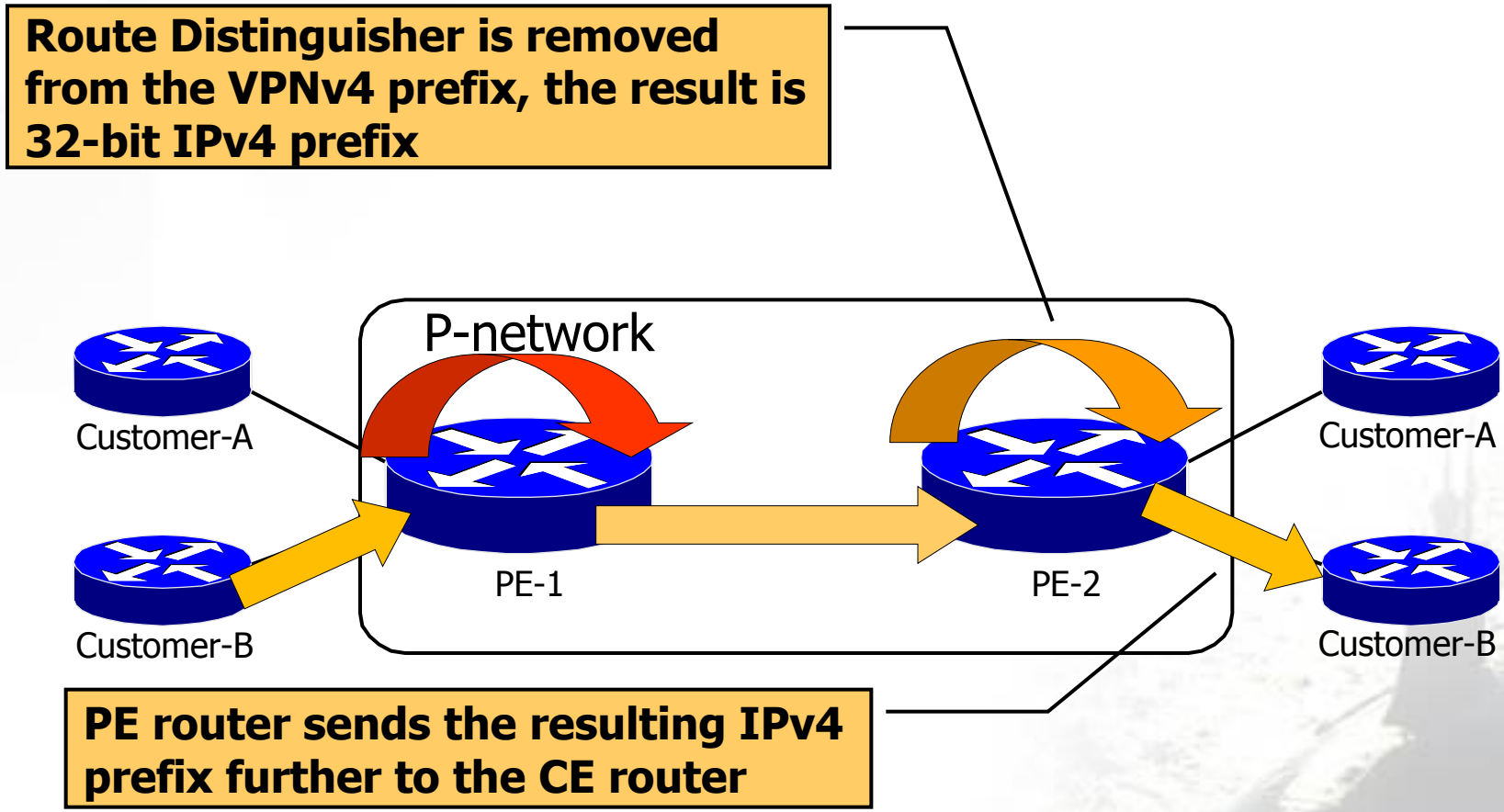


- The Label (for the VPNv4 prefix) is assigned only by the PE whose address is the next-hop attribute
  - PE routers rewrite the next-hop with their own address (loopback)
  - "Next-hop-self" towards MP-iBGP neighbors by default
- PE addresses used as BGP next-hop must be uniquely known in the backbone IGP
  - DO NOT summarize the PE loopback addresses in the core

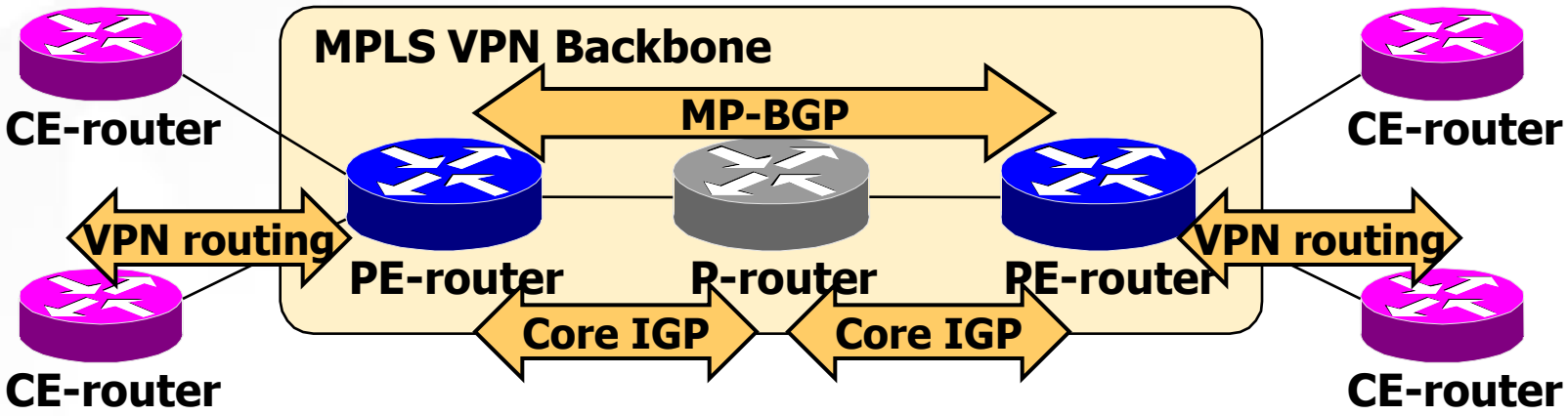
Usage of Route Distinguisher in MPLS VPN



Usage of Route Distinguisher in MPLS VPN (cont.)

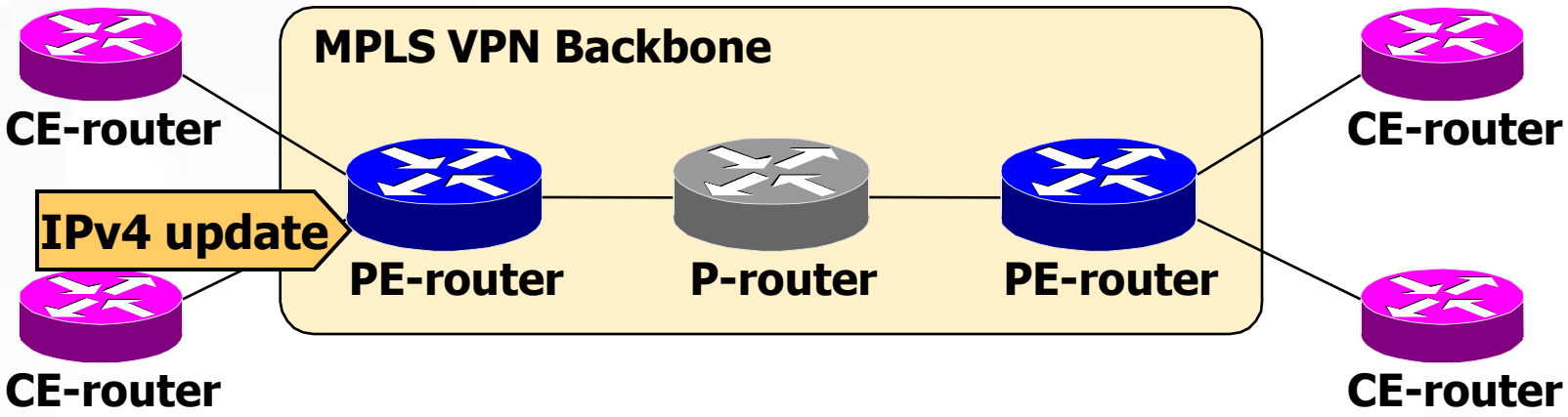


MPLS VPN Routing: PE-Router view



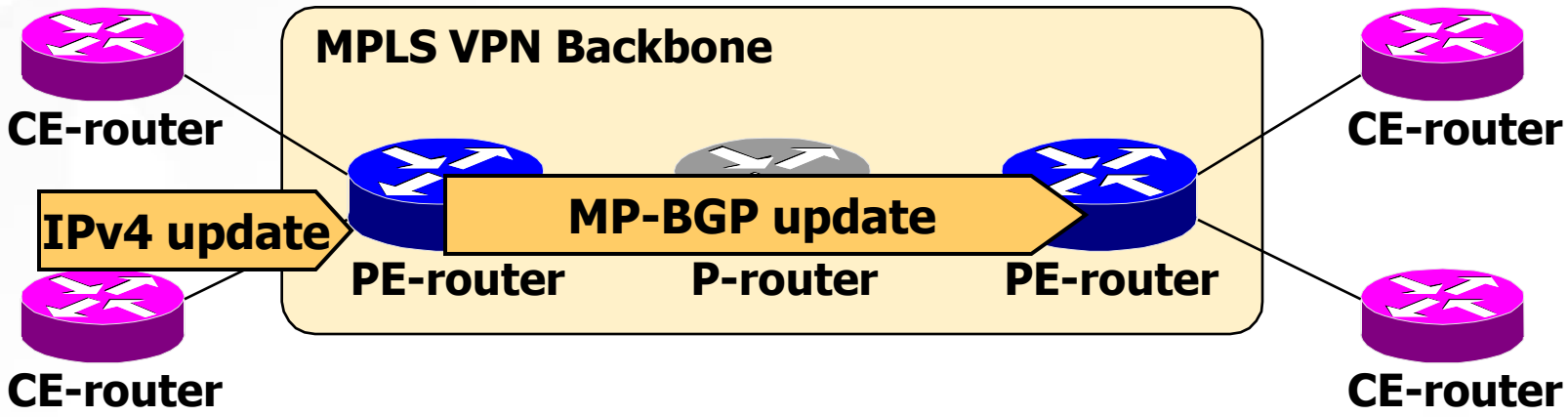
- PE-routers:
  - Exchange VPN routes with CE-routers through per-VPN routing protocols
  - Exchange core routes with P- and PE-routers over the core IGP
  - Exchange VPNv4 routes with other PE-routers over multi-protocol iBGP sessions

MPLS VPN End-to-End Routing Information Flow (1/3)



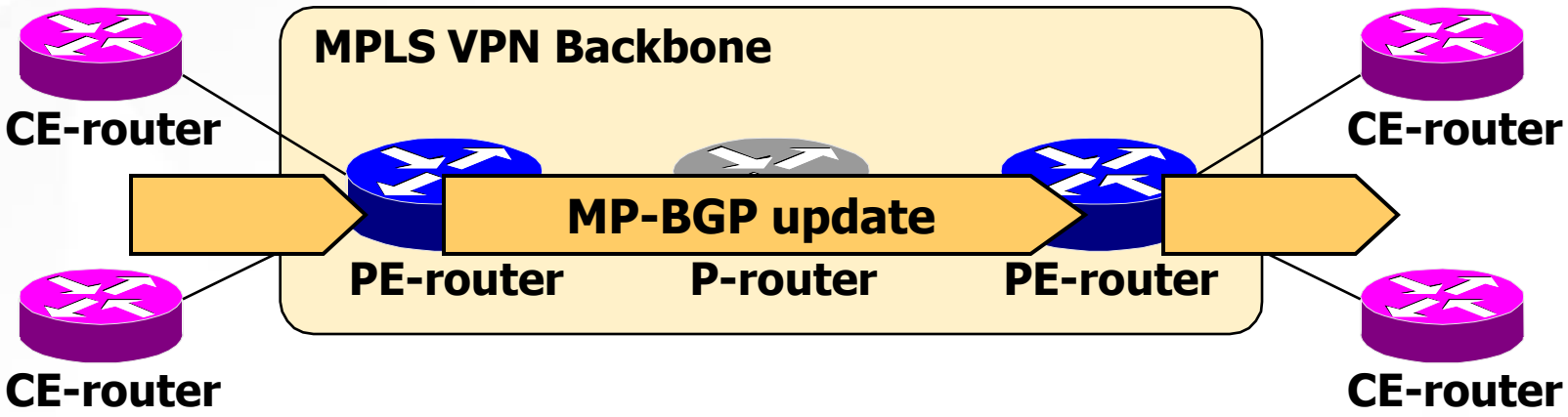
– PE-routers gets IPv4 routing updates from CE-routers and install them into appropriate Virtual Routing and Forwarding (VRF) tables

MPLS VPN End-to-End Routing Information Flow (2/3)



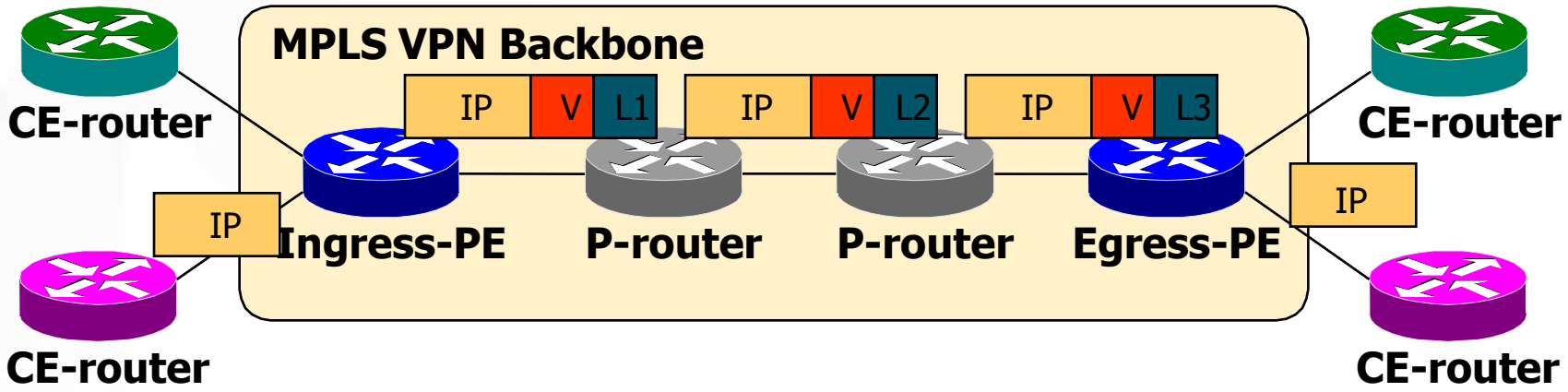
- PE-routers export VPN routes from VRF to MP-IBGP and propagate them as VPNv4 routes to other PE-routers
- A full mesh of IBGP sessions is required among PE-routers

MPLS VPN End-to-End Routing Information Flow (3/3)



- The receiving PE-router imports the incoming VPNv4 routes to the appropriate VRF based on route targets that are attached to the routes
- Routes installed into VRF are propagated to CE-routers

VPN Packet Forwarding over MPLS VPN Backbone



Q: How will PE routers forward VPN packets over the MPLS VPN backbone?

A: Mark VPN packets with label stack. Use LDP label of egress PE-router as a top label, VPN label assigned by PE-router as a second label in the stack.

ALEF NULA

DŮVĚŘUJTE SILNÝM

L2 MPLS VPNs

Layer 3 and Layer 2 VPN Characteristics

Layer 3 VPNs

- SP devices forward customer packets based on Layer 3 information (e.g. IP addresses)
- SP is involved in customer IP routing
- Support for any access or backbone technology
- IP specific
- Foundation for L4-7 Services!
- Example: RFC 2547bis VPNs (L3 MPLS-VPN)

Layer 2 VPNs

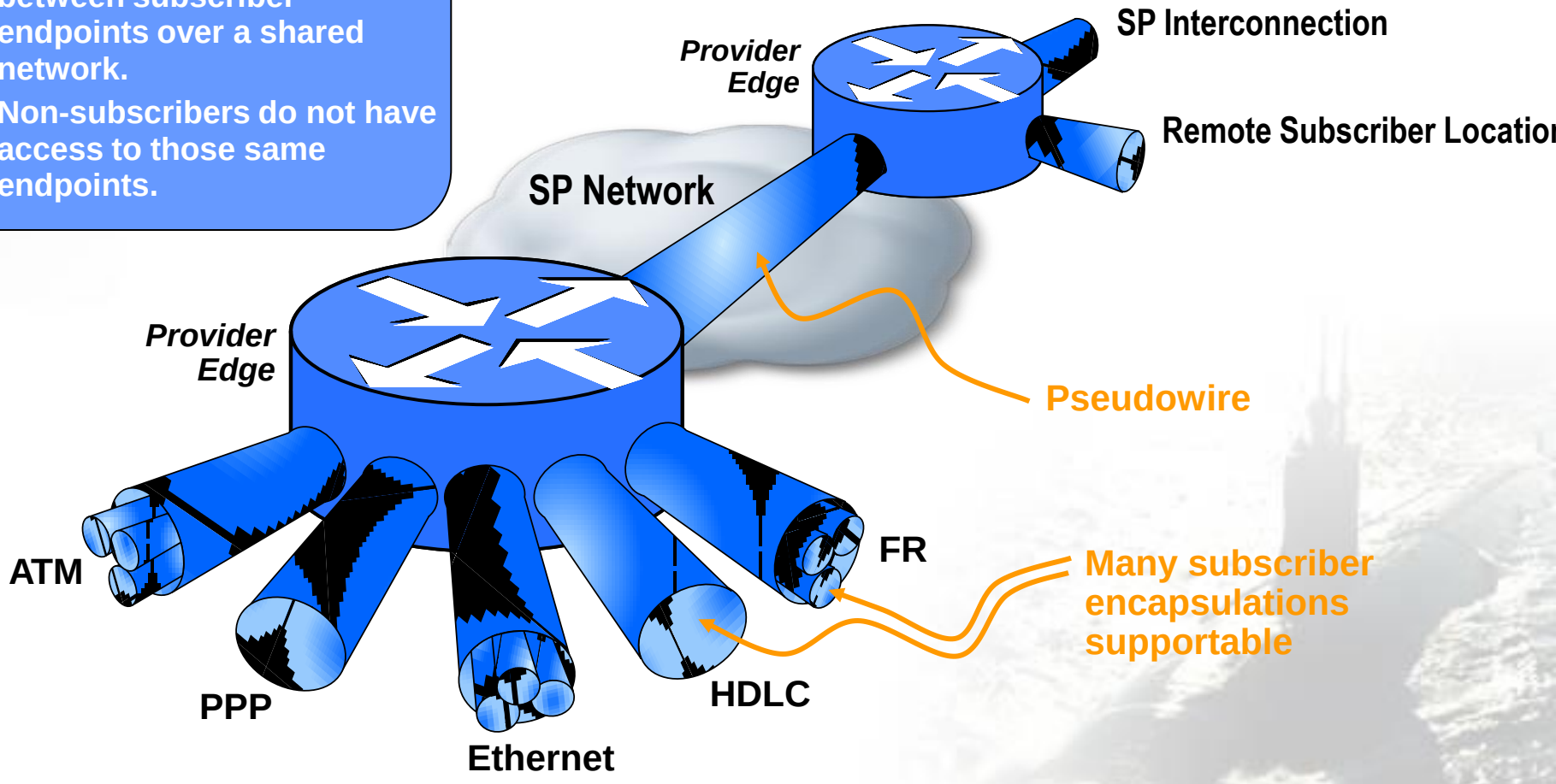
- SP devices forward customer frames based on Layer 2 information (e.g. DLCI, VPI/VCI, MAC)
- Enterprise stays in control of L3 policies (Routing, QoS)
- Access technology is determined by the VPN type
- Multiprotocol support
- Example: FR—ATM—Ethernet

The Choice of L2VPN over L3VPN Will Depend on How Much Control the Enterprise Wants to Retain
L2 VPN Services Are Complementary to L3 VPN Services

IETF L2VPN reference model

- An L2VPN is comprised of switched connections between subscriber endpoints over a shared network.

  Non-subscribers do not have access to those same endpoints.

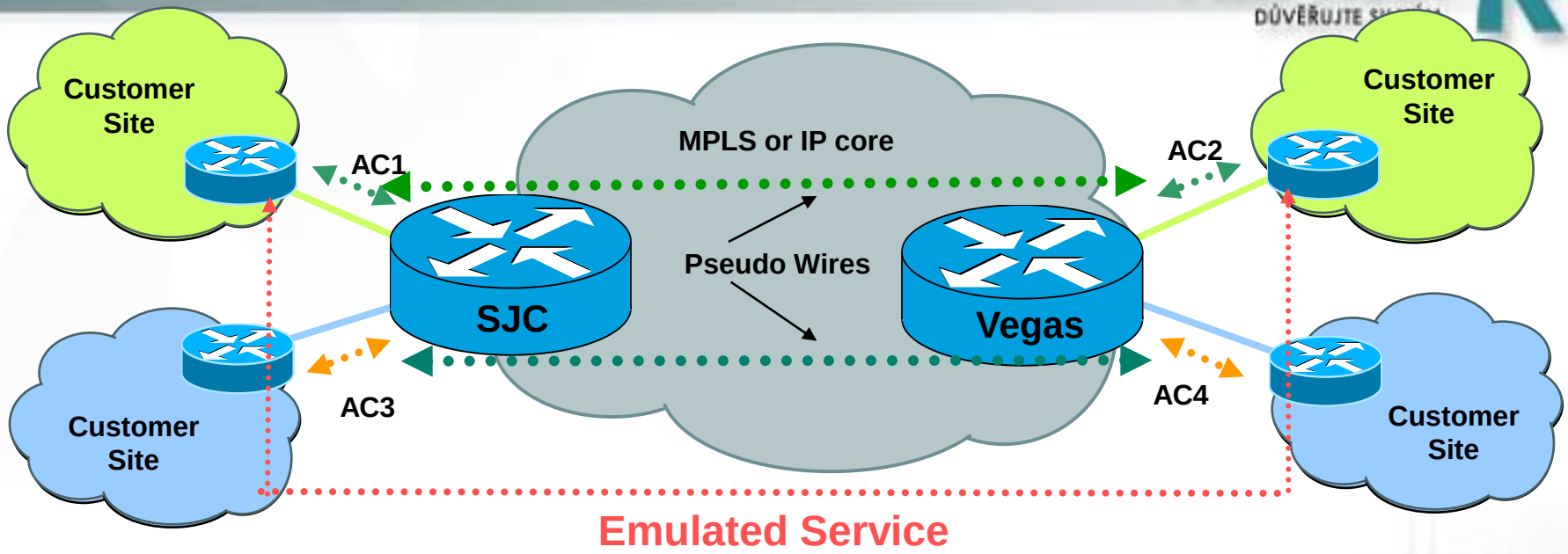


ALEF NULA

DŮVĚŘUJTE SILNÝM

Any Transport Over MPLS (AToM)

Transport Types



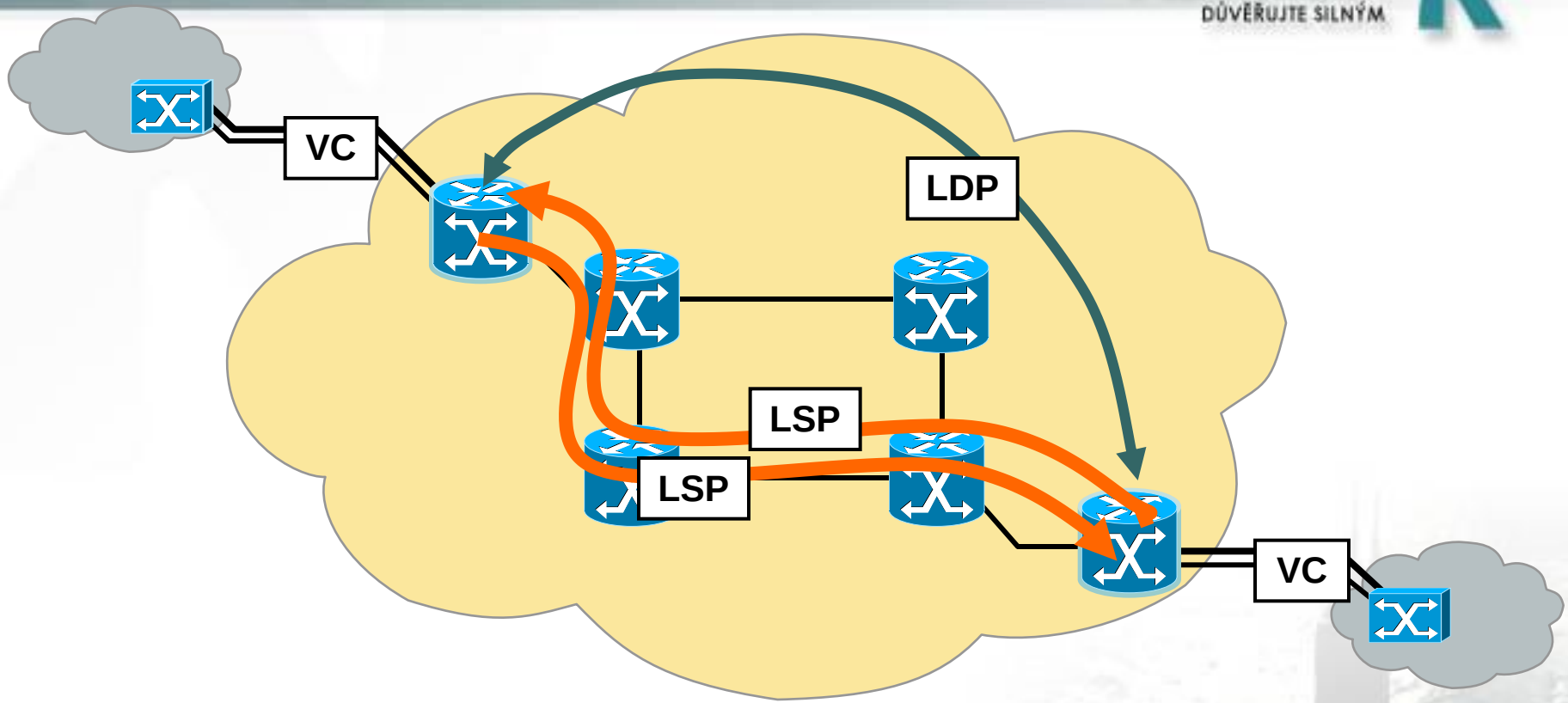
- AToM enables the following types of Layer 2 frames and cells to be directed across an MPLS backbone:
  - Ethernet, Ethernet VLAN
  - ATM adaptation layer 5 (AAL5), ATM cell relay
  - Frame Relay
  - Point-to-Point Protocol (PPP)
  - High-Level Data Link Control (HDLC)
  - SDH/PDH – Circuit emulation over packet - CeOP

How AToM Works

AToM frames are carried across an MPLS backbone in the following manner:

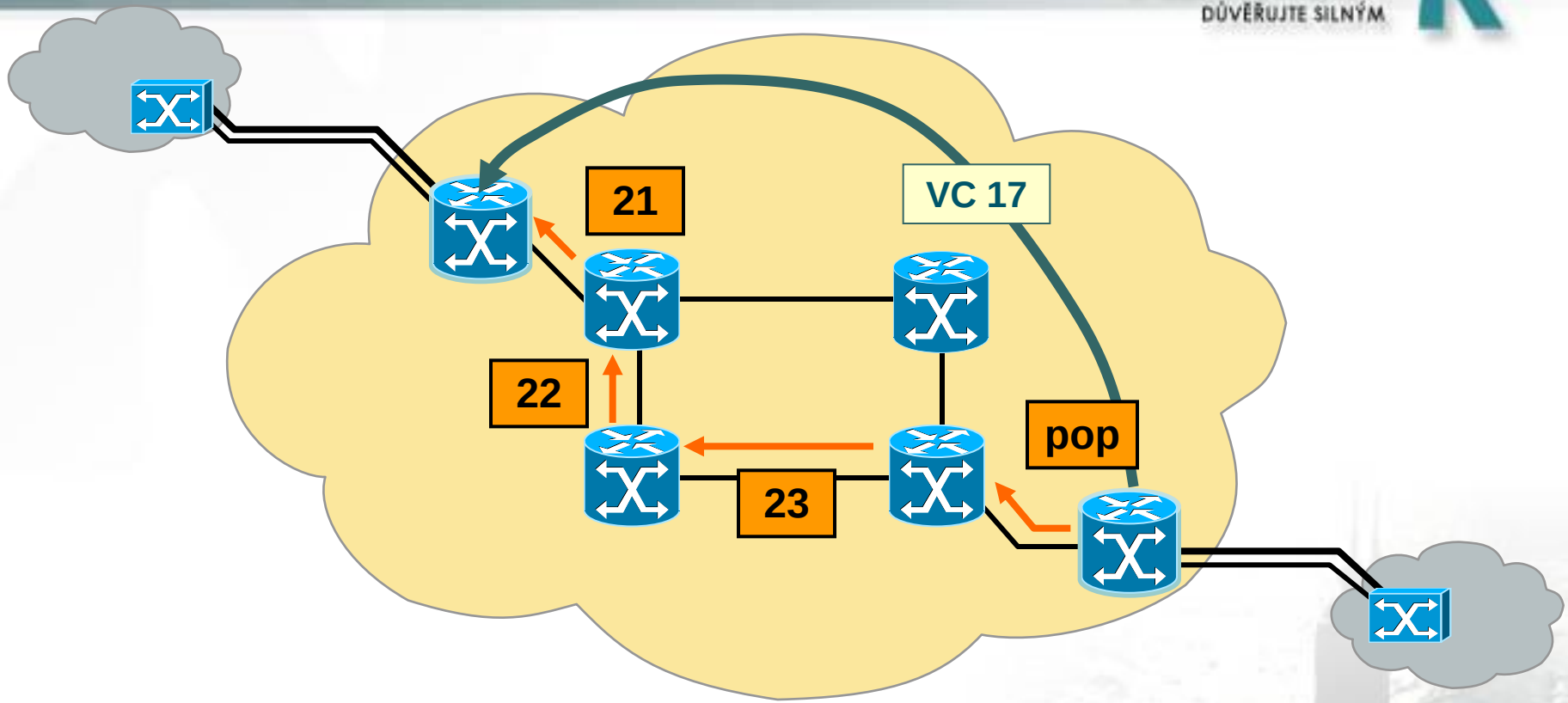
- Ingress and egress interfaces are non-MPLS interfaces.
- Ingress PE encapsulates frame into MPLS; egress PE decapsulates.
- Label stack of two labels is used (similar to an MPLS VPN).
  - Topmost label ("tunnel label") used for LSP PE to PE.
  - Second label ("VC label") identifies outgoing interface in the egress PE.
- LDP has been extended to carry virtual circuit forwarding equivalence class (VC FEC).
- A directed (multihop) LDP session is used from PE to PE.

How AToM Works (Cont.)



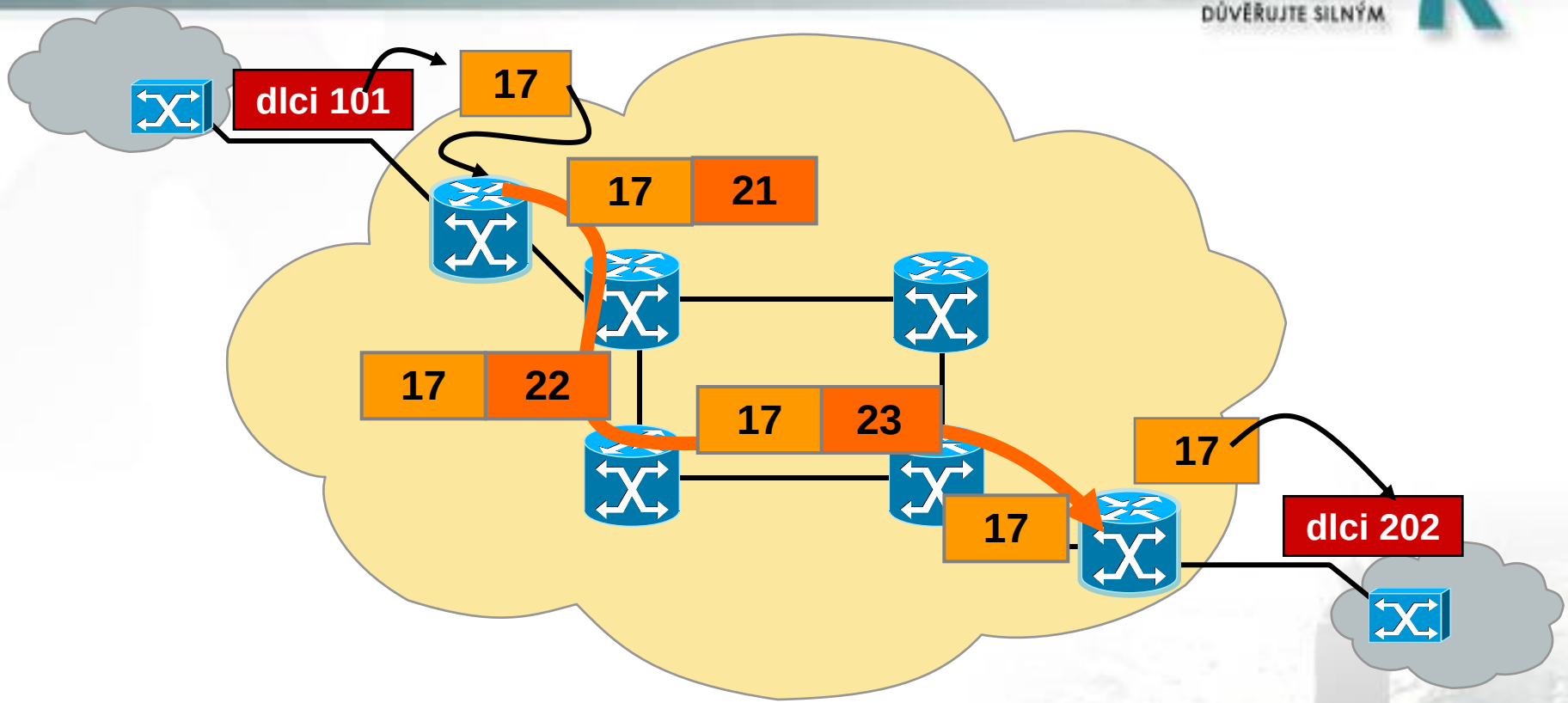
- The IGP and the LDP between directly connected LSRs establish one LSP in each direction.
- **A directed LDP session between PE routers is established.**

How AToM Works (Cont.)



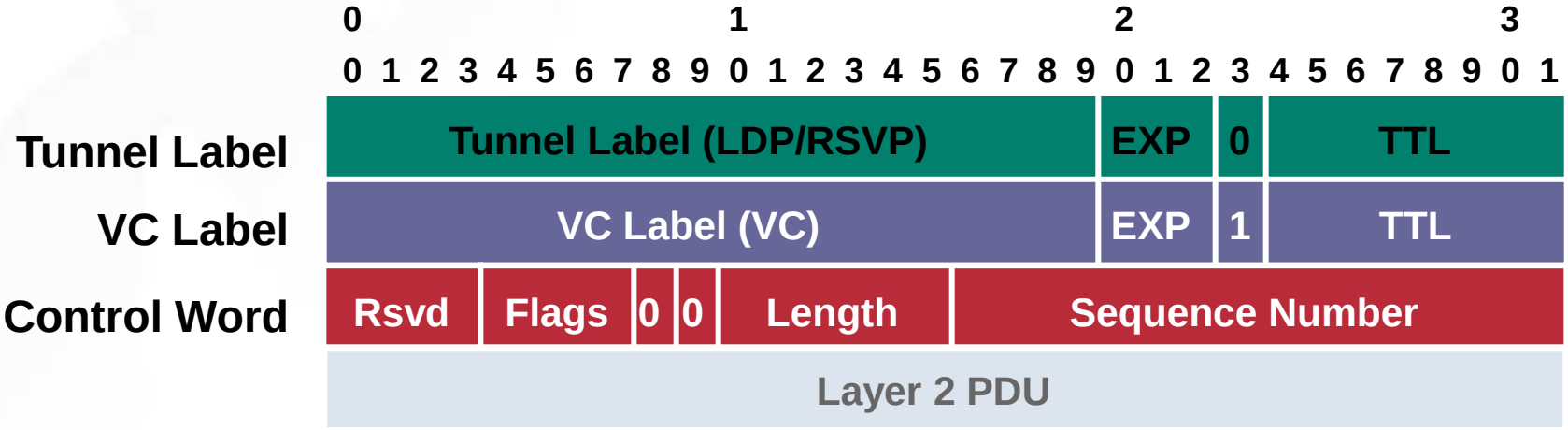
- **LDP between directly connected LSRs generates an LSP.**
- The egress PE allocates VC label 17.
- The directed LDP session between PE routers propagates the VC label.

How AToM Works (Cont.)



- The ingress PE receives a frame on DLCI 101.
- The frame is encapsulated and forwarded along the LSP.
- The egress PE reconstructs the frame.

AToM Control Word



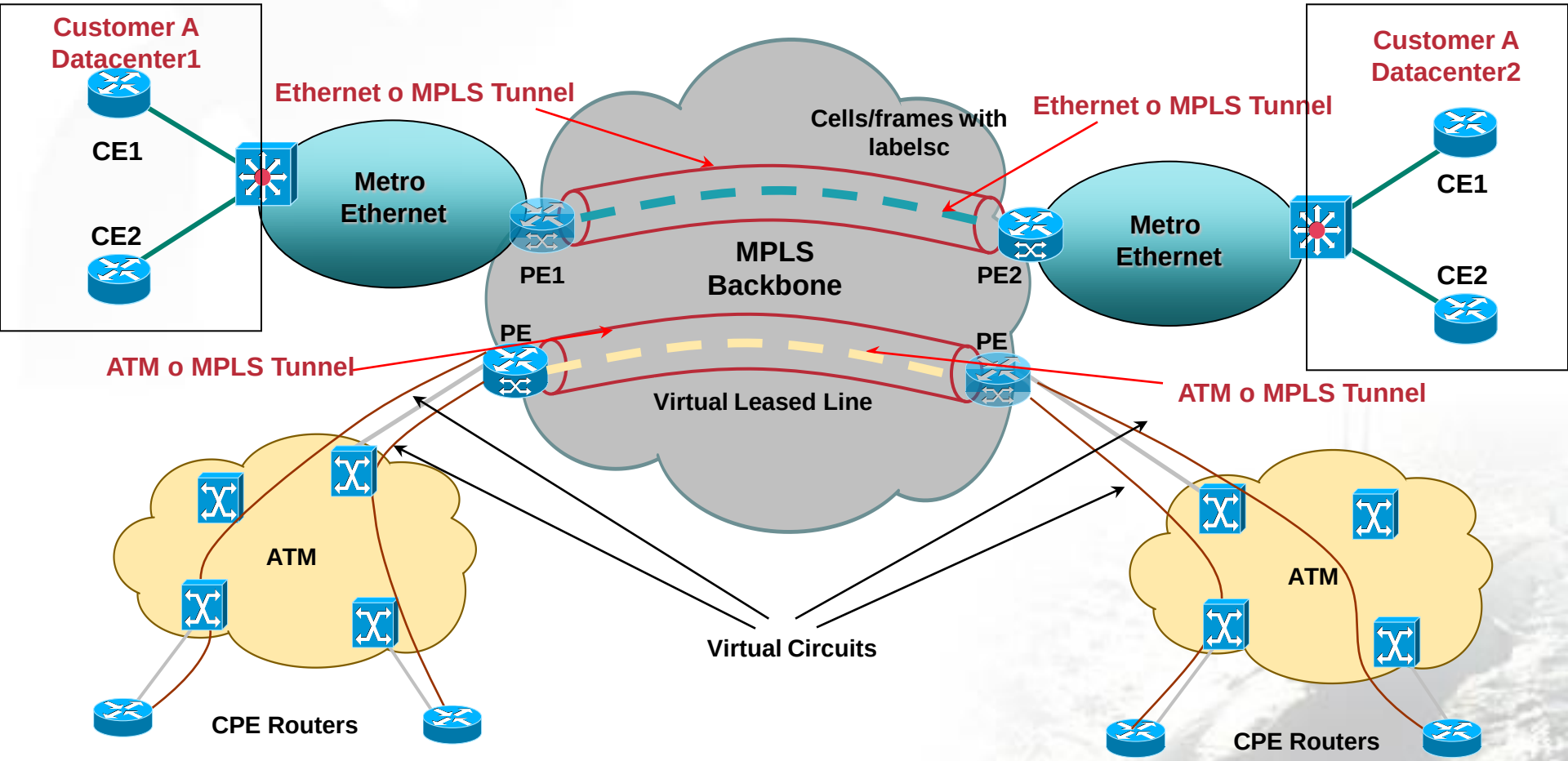
- The control word is optional.
- It is transmitted after the label(s) and before the Layer 2 PDU.
- Flag field carries different bits for different Layer 2 protocols:
  - Frame Relay: FECN, BECN, DE, C/R
  - ATM: AAL5 or cell, EFCI, CLP, C/R
- Sequence number 0 indicates that no sequencing is done.

ALEF NULA

DŮVĚŘUJTE SILNÝM

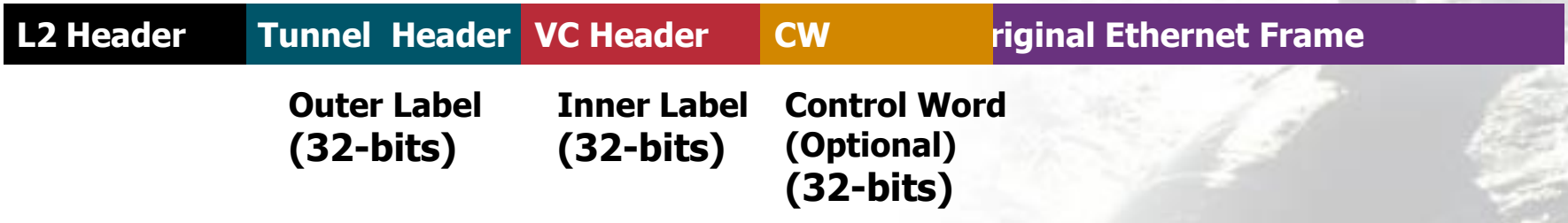
Ethernet Over MPLS (EoMPLS)

AToM Deployment Example



EoMPLS: Data Plane Overhead

- At imposition, PE encapsulates CE's Ethernet or VLAN packet to route across MPLS cloud
- These are the associated overheads:
  - Transport header is 6 Bytes DA + 6 Bytes SA + 2 Bytes etype + optional 4 Bytes of VLAN Tag
  - There's (at least) 2 levels of MPLS header (Tunnel + VC) each contributing with 4 Bytes
  - There is an optional 4-Byte control word



MEF

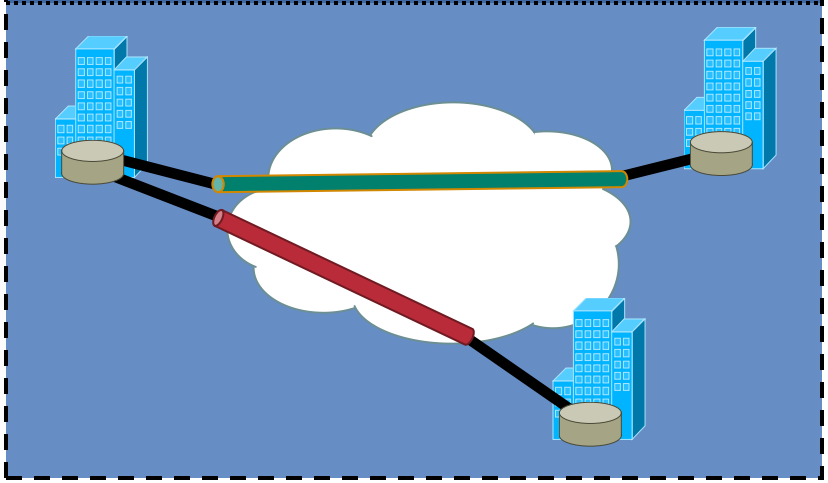
Metro Ethernet Forum

- New Standardization Subject
- http://metroethernetforum.org
- Carrier Ethernet (formerly Metro Ethernet)
- Focused on services based on L2 – mainly Ethernet

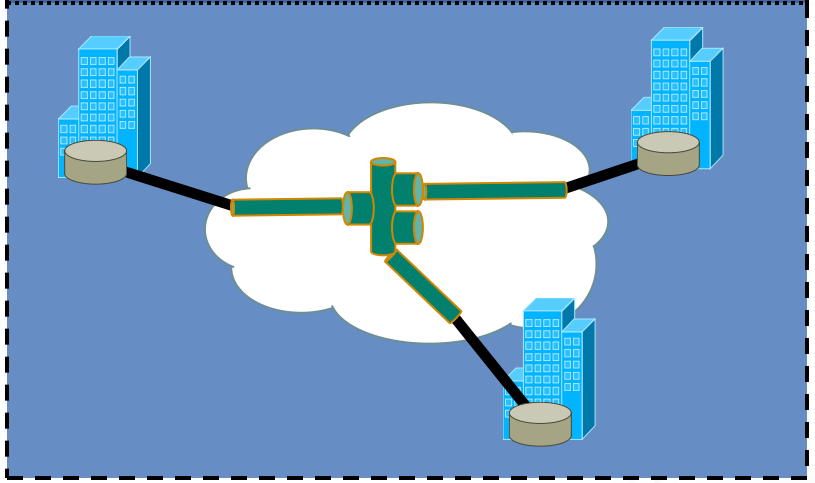
Carrier Ethernet Services

ALEF NULA
DŮVĚŘUJTE SILNÝM

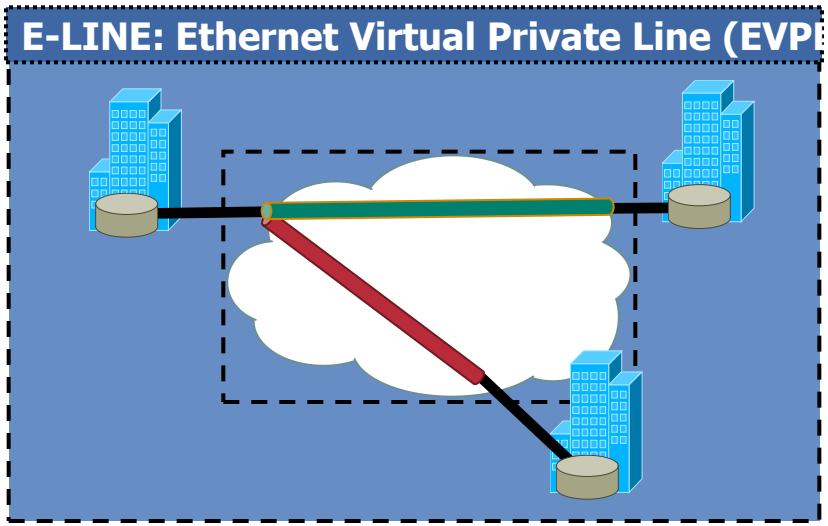
E-LINE: Ethernet Private Line (EPL)



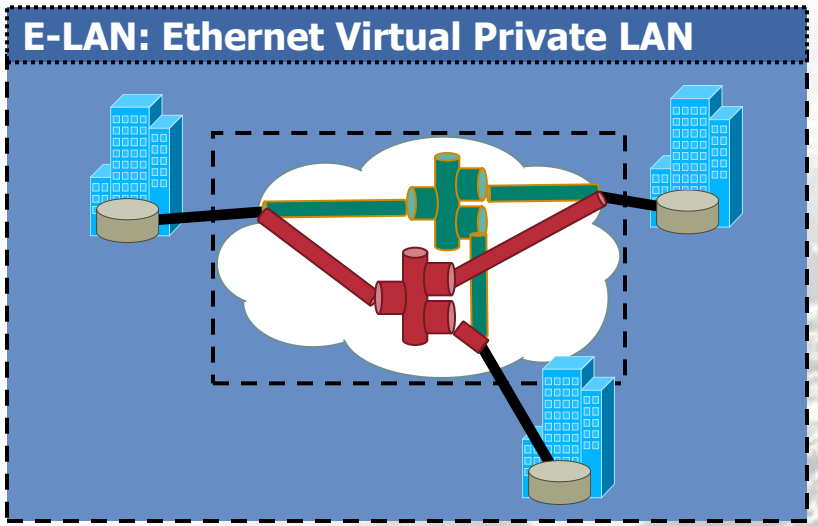
E-LAN: Ethernet Private LAN



E-LINE: Ethernet Virtual Private Line (EVPL



E-LAN: Ethernet Virtual Private LAN



MEF Services Framework

UNI Attributes

UNI ID	Arbitrary String		
Speed	10Mb	100Mb	1000Mb
Mode	Full	Auto	
MAC Layer	802.3		
Service Multiplexing	Yes	No	
Bundling	Yes	No	
All to One Bundling	Yes	No	
Ingress BW Profile Per UNI	No	CIR, CBS EIR, EBS	
Ingress BW Profile Per EVC	No	CIR, CBS EIR, EBS	
Ingress BW Profile Per Class of Svc ID	No	CIR, CBS EIR, EBS	
CE-VLAN ID/EVC Map	$\{\text{CE-VLAN ID} \rightarrow \text{EVC(i)}\}_i$		

EVC Attributes

EVC Type	PTP (E-Line)	MP (E-LAN)
EVC ID	Arbitrary string	
CE-VLAN ID Preservation	Yes	No
CE-VLAN COS Preservation	Yes	No
Unicast Frame Delivery	Conditional	Un-conditional
MCAST Frame Delivery	Conditional	Un-conditional
BCAST Frame Delivery	Conditional	Un-conditional
Class of Service ID	EVC	
	EVC, L2 CoS	
	EVC, L3 DSCP	
EVC Performance	Delay, Delay-Variation, Loss, ...*	

*see MEF10, section 6.7

L2 Control Processing

	Peer	Discard	Pass
802.3x	☐	☐	☐
LACP	☐	☐	☐
802.1x	☐	☐	☐
GARP	☐	☐	☐
STP	☐	☐	☐
Bridge MCAST	☐	☐	☐
CDP	☐	☐	☐
VTP	☐	☐	☐
PAgP	☐	☐	☐
UDLD	☐	☐	☐

ALEF NULA

DŮVĚŘUJTE SILNÝM

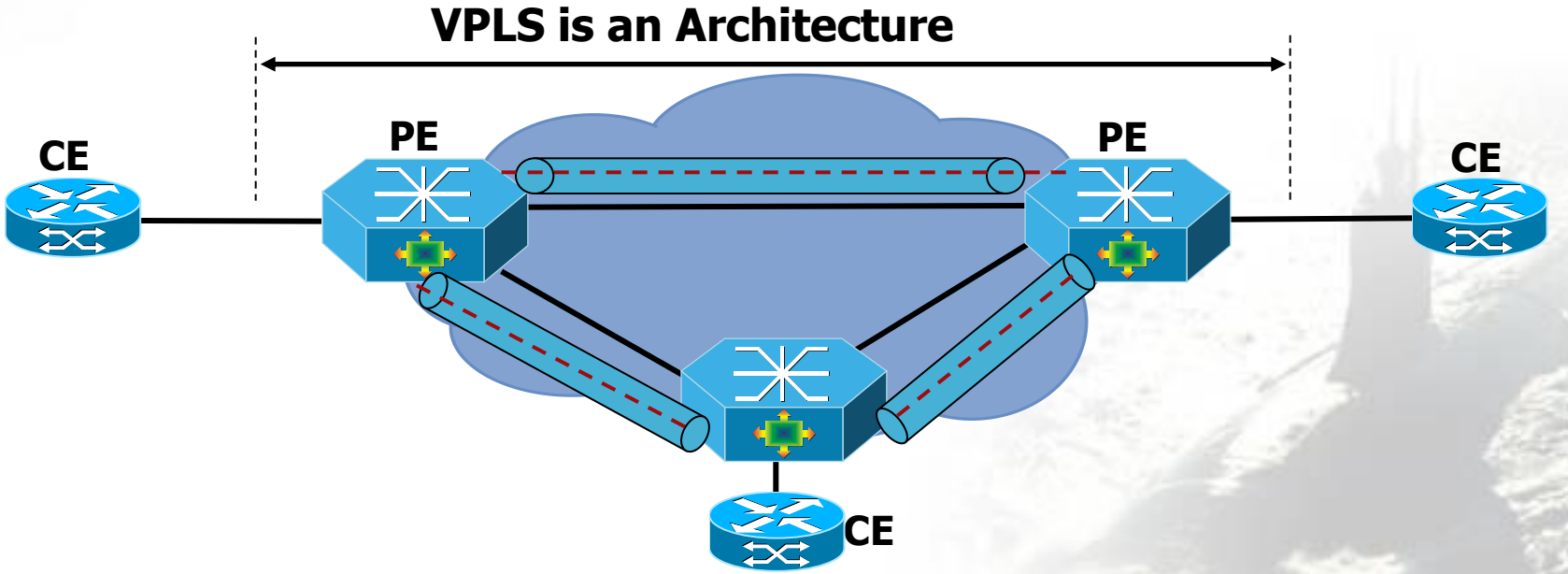
VPLS

Virtual Private LAN Service (VPLS)

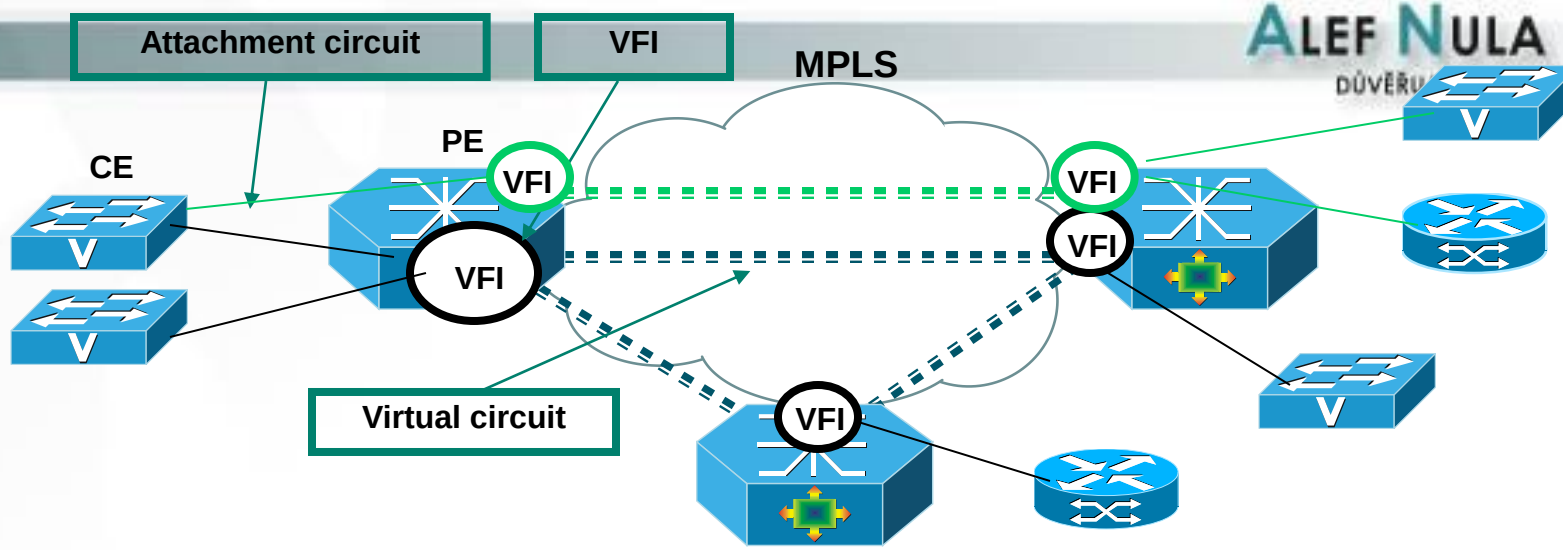
ALEF NULA
DŮVĚŘUJTE SILNÝM

- VPLS defines an architecture allows MPLS networks offer Layer 2 multipoint Ethernet Services
- SP emulates an IEEE Ethernet bridge network (virtual)
- Virtual Bridges linked with MPLS Pseudo Wires

Data Plane used is same as EoMPLS (point-to-point)

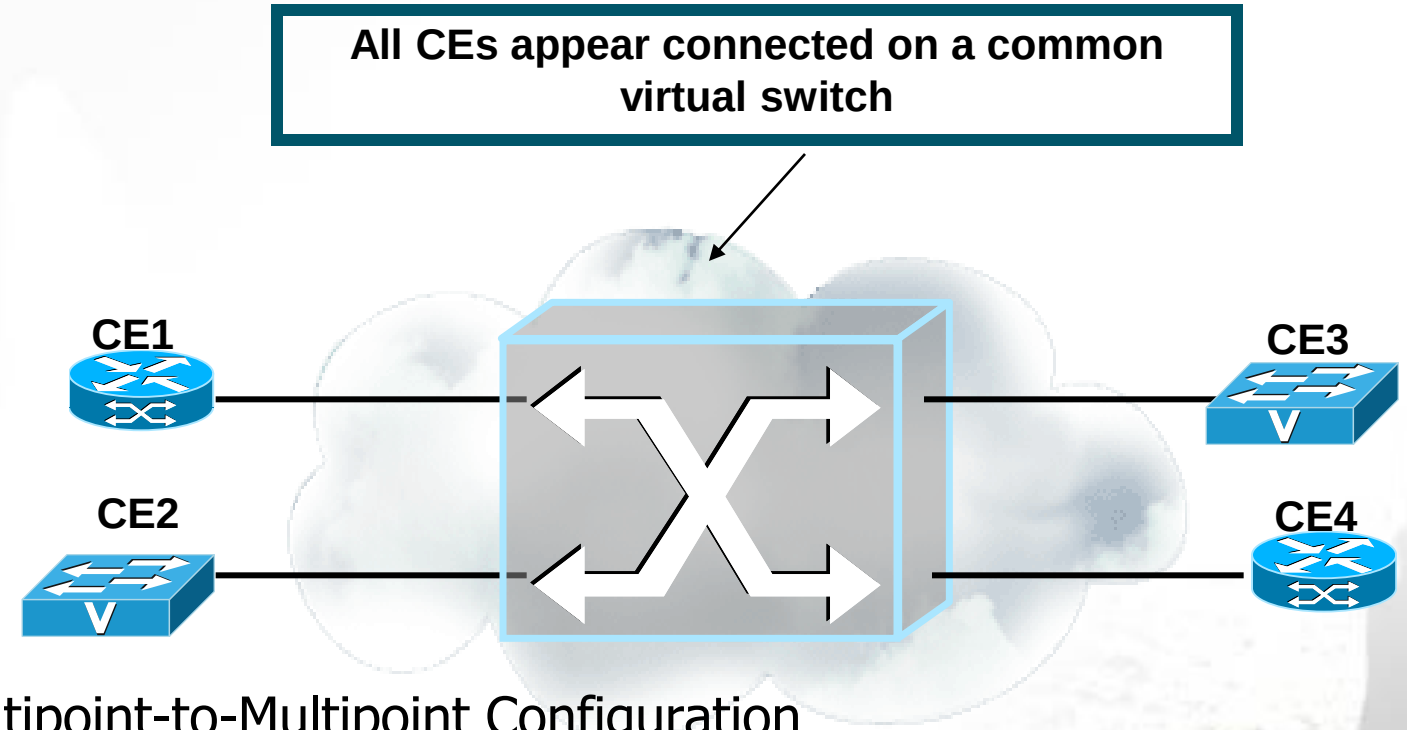


VPLS Components



- AC (Attachment Circuit)
  - – Connect to CE device, it could be Ethernet physical or logical port, ATM bridging (RFC1483), FR bridging (RFC1490), even AToM pseudo wire. One or multiple ACs can belong to same VFI
- VC (Virtual Circuit)
  - –EoMPLS data encapsulation, tunnel label is used to reach remote PE, VC label is used to identify VFI. One or multiple VCs can belong to same VFI
- VFI (Virtual Forwarding Instance)
  - – Also called VSI (Virtual Switching Instance). VFI create L2 multipoint bridging among all ACs and VCs. It's L2 broadcast domain like VLAN
  - – Multiple VFI can exist on the same PE box to separate user traffic like VLAN

VPLS Customer Perspective



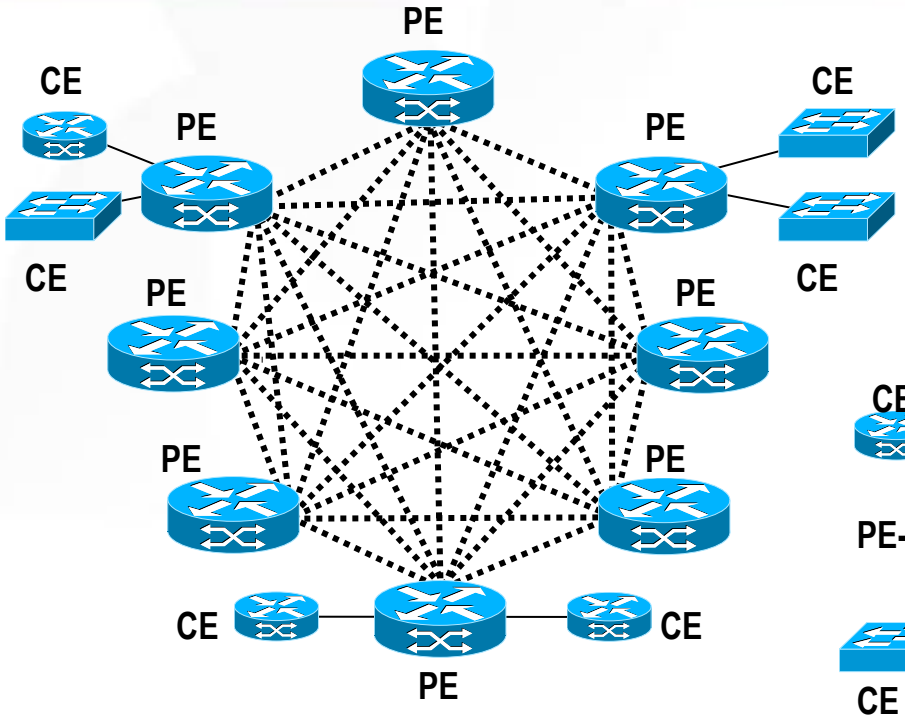
- Multipoint-to-Multipoint Configuration
- Forwarding of Frames based on Learned MAC addresses
- Uses a Virtual Forwarding Instances (VFI, like VLAN) for customer separation

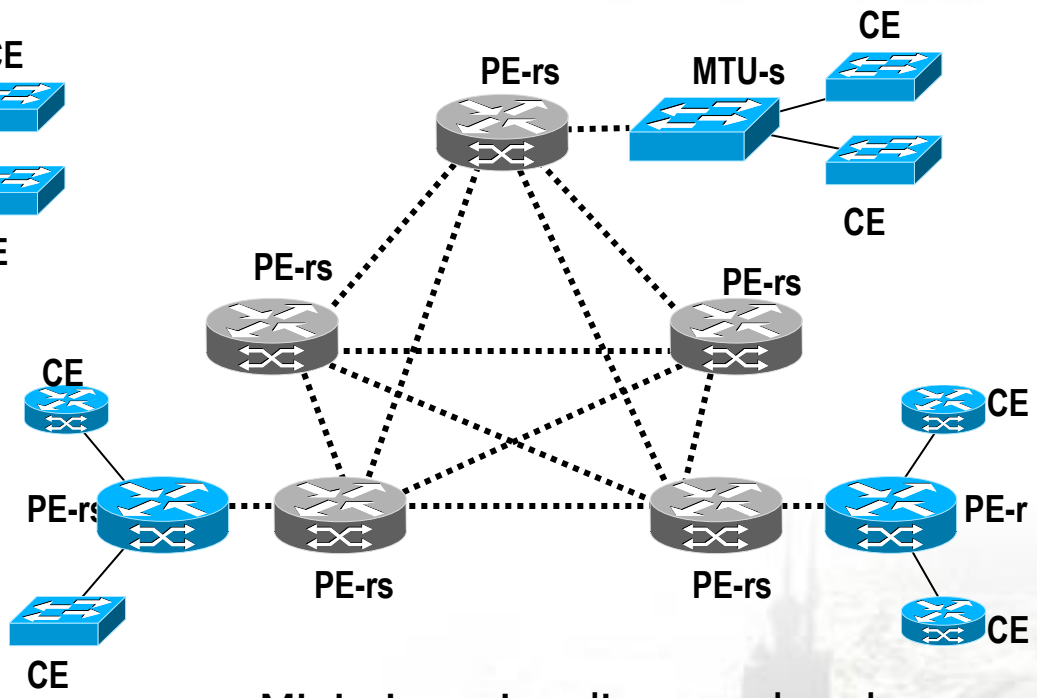
Multipoint Bridging Requirements

VPLS simulate a virtual LAN service, it MUST operate like a traditional L2 LAN switch as well

- Flooding/Forwarding
  - Forwarding based on [VLAN, Destination MAC Address]
  - Unknwon Ucast/Mcast/Broadcast – Flood to all ports (IGMP snooping can be used to constrict multicast flooding)
- MAC Learning/Aging/Withdrawal
  - Dynamic learning based on Source MAC and VLAN
  - Refresh aging timers with incoming packet
  - MAC withdrawal upon topology changes
- Loop Prevention
  - Split Horizon to avoid loop
  - Spanning Tree (possible but not desirable)

Why H-VPLS?

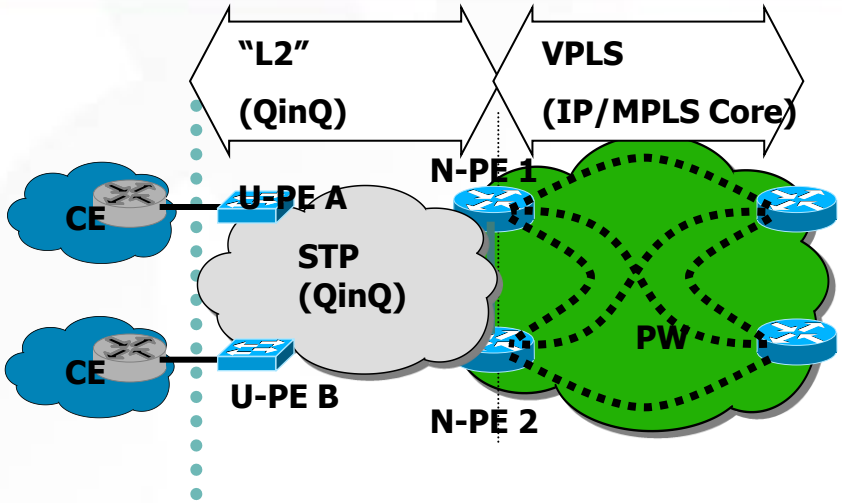




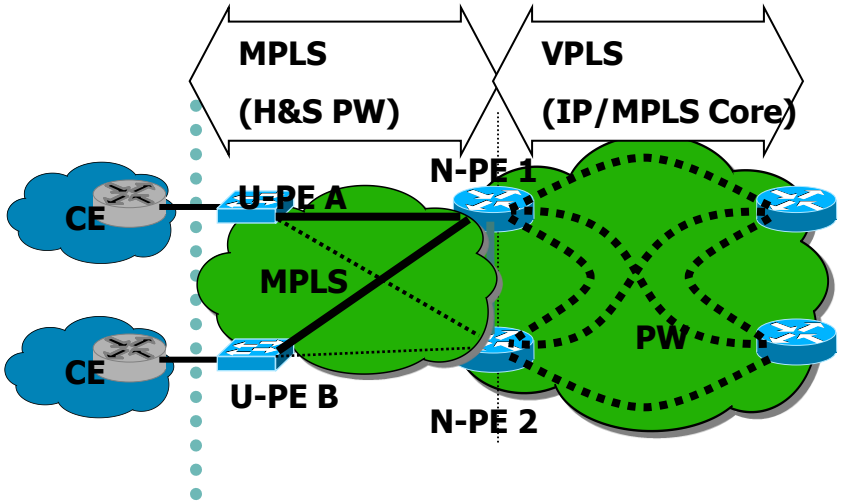
- Potential signaling overhead
- Full PW mesh from the Edge
- Packet replication done at the Edge
- Node Discovery and Provisioning extends end to end

- Minimizes signaling overhead
- Full PW mesh among Core devices
- Packet replication done the Core
- Partitions Node Discovery process

H-VPLS: 2 approaches

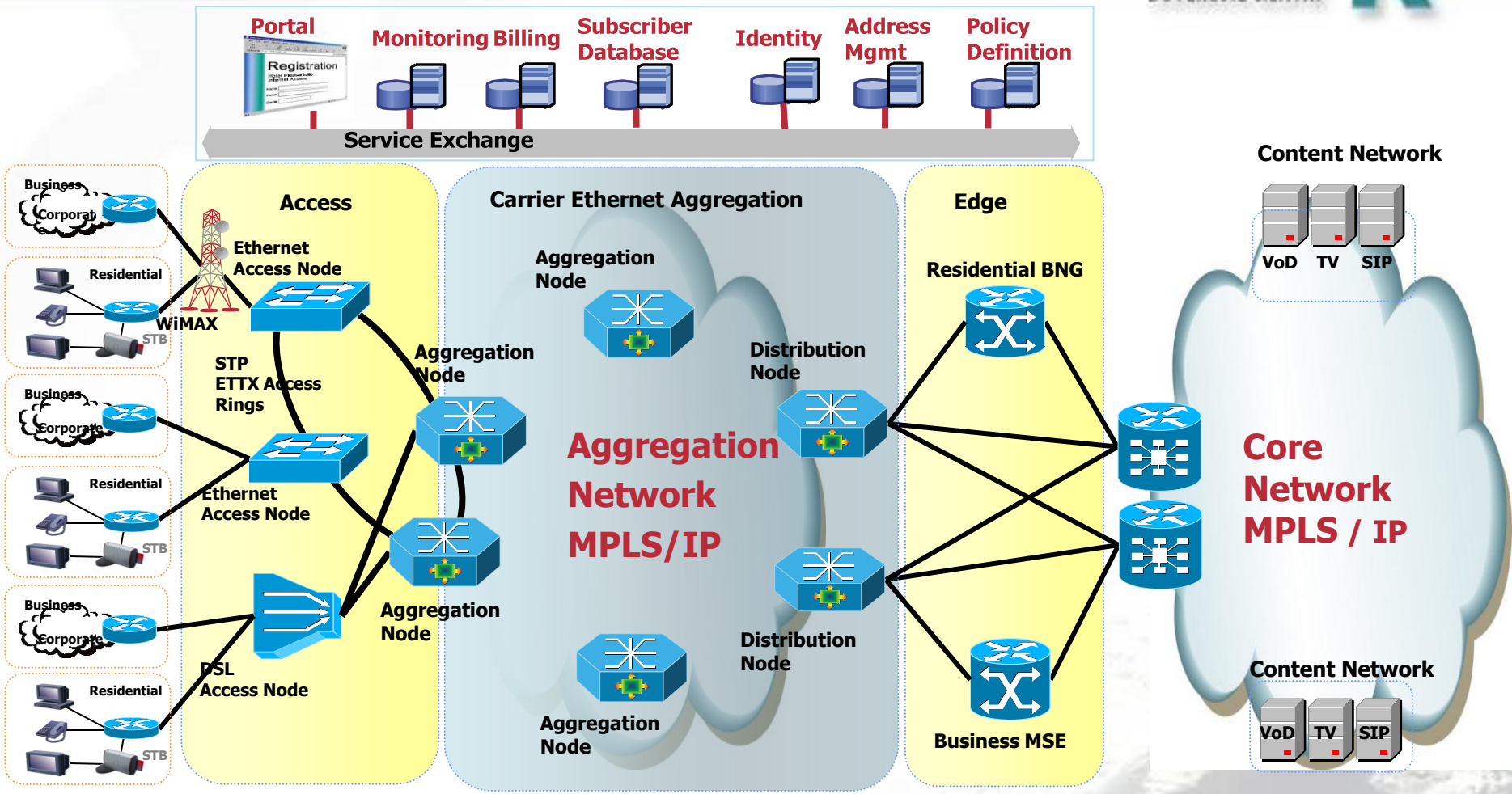


- **H-VPLS with bridge-group domain at access**
- **Access domain defined by IEEE 802.1ad**



- **H-VPLS with MPLS the edge, using PW EoMPLS circuit to backhall traffics from U-PE to N-PE**

Service Providers IP-NGN Architecture



ALEF NULA

DŮVĚŘUJTE SILNÝM

Q&A