



**Chair for Network Architectures and Services – Prof. Carle**  
Department of Computer Science  
TU München

# **Master Course Computer Networks IN2097**

**Prof. Dr.-Ing. Georg Carle**

**Chair for Network Architectures and Services  
Department of Computer Science  
Technische Universität München  
<http://www.net.in.tum.de>**



Technische Universität München



## Announcements for Upcoming Lectures

- Tuesday, 19 November 2013: Exercise
  
- Monday, 25 November 2013
  - Special event: Network of Excellence in Internet Science talks
  - Time: **9:15-10:45**
  - Location: lecture hall in LRZ (HE 009)  
(ground floor, entrance to the right)
  
- Tuesday, 26 November 2013: Lecture



## Special event: Network of Excellence in Internet Science

- Advisory Committee Talks
- 09:15 – 09:45
  - “Triple revolution: the intersection of social network analysis, the far-flung personalized internet, and the mobile revolution”
  - Barry Wellman (University of Toronto)  
<http://www.internet-science.eu/users/barrywellman>
- 09:45 – 10:15
  - “Evaluating Network Architectures”
  - David Clark (MIT)  
<http://www.internet-science.eu/users/ddc>
- 10:15 – 10:45
  - “The (Moral) Responsibility of Internet Intermediaries”
  - M. Thompson (Hong Kong Univesity)  
<http://www.internet-science.eu/users/barrywellman>

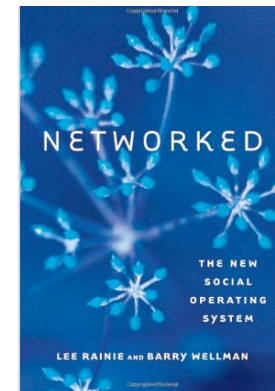




## Barry Wellman

- Professor Barry Wellman
  - University of Toronto
  - directs NetLab at the Faculty of Information
  - is the former S.D. Clark Professor at the Department of Sociology
  - is a member of the Cities Centre and the Knowledge Media Design Institute
  - co-author of the prize-winning book

Networked: The New Social Operating System (with Lee Rainie, Director of the Pew Internet and American Life Project) published by MIT Press in Spring 2012. The book analyzes the social nature of networked individualism, growing out of the Social Network Revolution, the Internet Revolution, and the Mobile Revolution.





## David Clark

- Dr. David Clark
  - MIT Computer Science and Artificial Intelligence Laboratory
  - has worked since receiving his Ph.D. there in 1973
  - since the mid 70s, leading the development of the Internet
  - from 1981-1989 acted as Chief Protocol Architect, and chaired the Internet Activities Board
  - His current research looks at re-definition of the architectural underpinnings of the Internet, and the relation of technology and architecture to economic, societal and policy considerations.
  - U.S. NSF: Future Internet Design program.
  - past chairman of the Computer Science and Telecommunications Board of the National Academies
  - co-director of the MIT Communications Futures Program, a project for industry collaboration



## Marcelo Thompson

- Dr. Marcelo Thompson
  - Assistant Professor of Law
  - Deputy Director of the LLM in Information Technology and Intellectual Property Law at the Faculty of Law, The University of Hong Kong
  - Courses: "Law and Society", "Legal Theory", "Privacy and Data Protection" and "Regulation of Cyberspace"
  - Research: intersection between law, political theory and the study of technological change
  - Doctorate at the University of Oxford, Oxford Internet Institute, on neutrality in technology law and politics.
  - LLM (Law and Technology) from University of Ottawa, on copyright reform and the human right of access to knowledge



# Virtual Private Networks

Acknowledgements:

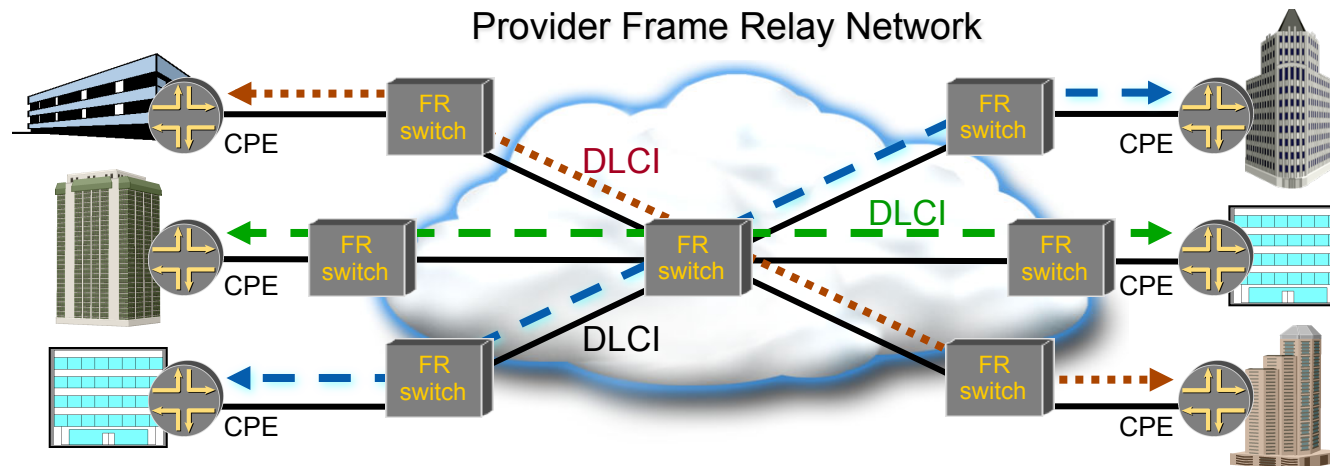
John Jamison,  
University of Illinois at Chicago

Philip Matthews  
Nortel Networks





# Deploying VPNs using Overlay Networks

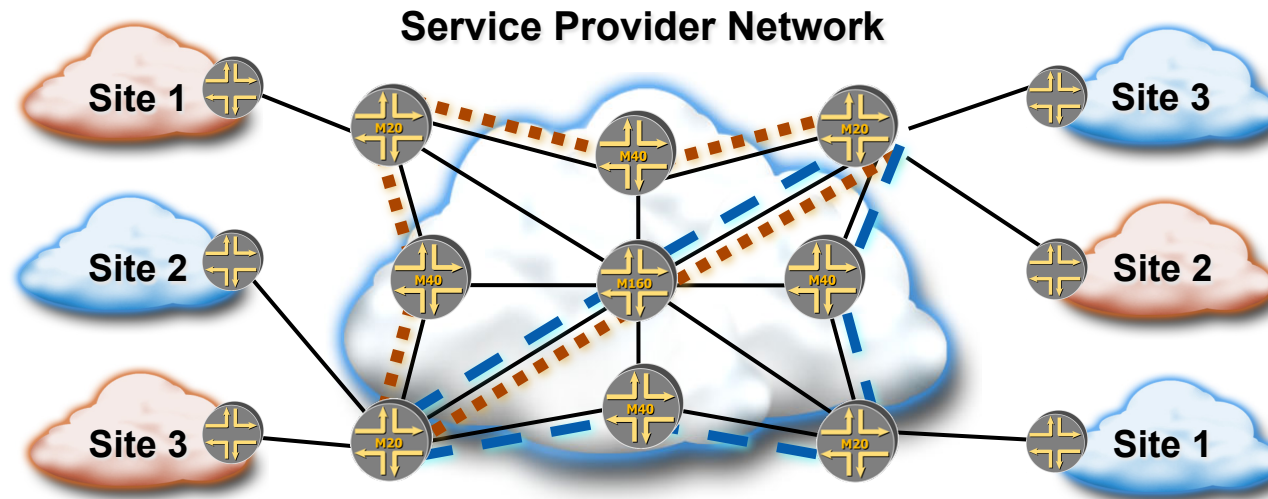


- Operational model
  - PVCs overlay the shared infrastructure (ATM/Frame Relay)
  - Routing occurs at CPE
- Benefits
  - Mature technologies
  - Inherently 'secure'
  - Service commitments (bandwidth, availability, etc.)
- Limitations
  - Scalability and management of the overlay model
  - Not a fully integrated IP solution





# MPLS: A VPN Enabling Technology



## □ Benefits

- Seamlessly integrates multiple “networks”
- Permits a single connection to the service provider
- Supports rapid delivery of new services
- Minimizes operational expenses
- Provides higher network reliability and availability

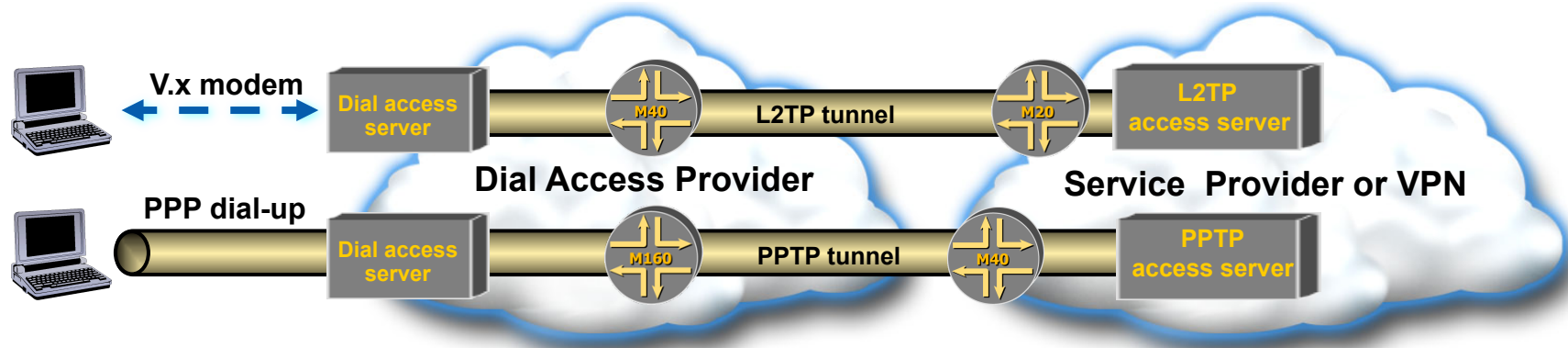


## Different Types of VPNs

- Layer 2 VPNs
  - Virtual Circuit VPN, circuit cross-connect (CCC)
  - MPLS L2 VPN
- Layer3 VPNs
  - RFC 2547bis / 4364: BGP/MPLS IP VPN
  - IPSEC VPN
  - IP-in-IP-encapsulation VPN
- End to End (CPE Based) VPNs
  - L2PT & PPTP
  - IPSEC



## End to End VPNs: L2TP and PPTP



- ❑ Application: Dial access for remote users
- ❑ Point-to-Point Tunneling Protocol (PPTP)
  - Bundled with Windows
- ❑ Layer 2 Tunneling Protocol (L2TP)
  - Open standard, RFC 2661
  - Combination of L2F and PPTP
- ❑ Both support IPsec for authentication and encryption
  - Authentication & encryption at tunnel endpoints



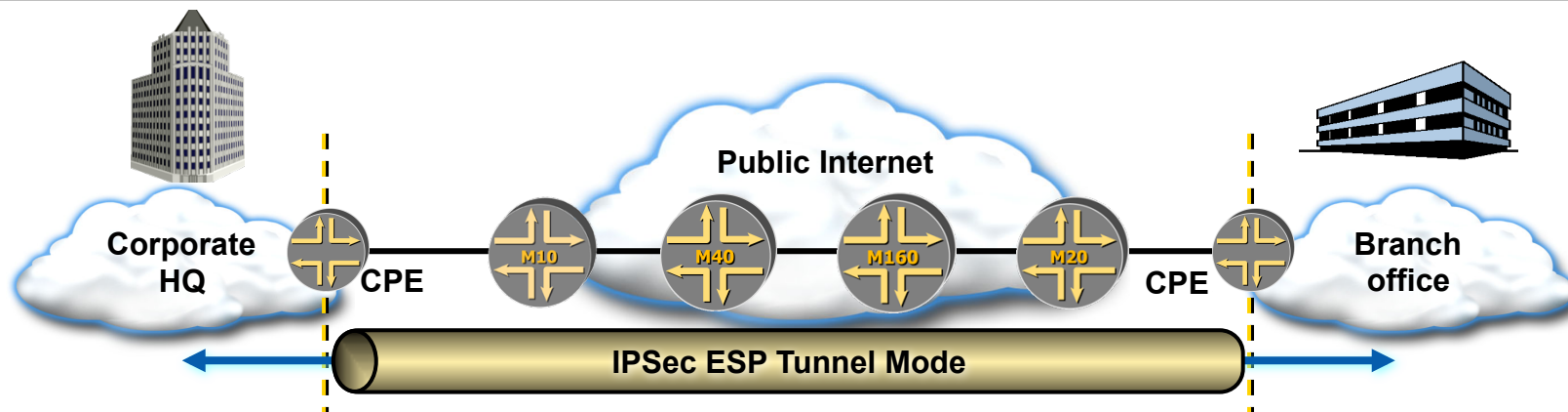
## End to End VPNs: IP Security Protocol (IPSec)

- Defines the IETF's layer 3 security architecture
- Applications:
  - Strong security requirements
  - Extend a VPN across multiple service providers
- Security services include:
  - Access control
  - Data origin authentication
  - Replay protection
  - Data integrity
  - Data privacy (encryption)
  - Key management
- Issues with IPSec include
  - Complexity; in some cases firewall traversal issues, ...

⇒ L4/L7 tunneling alternatives with DTLS / TLS / HTTP tunnels



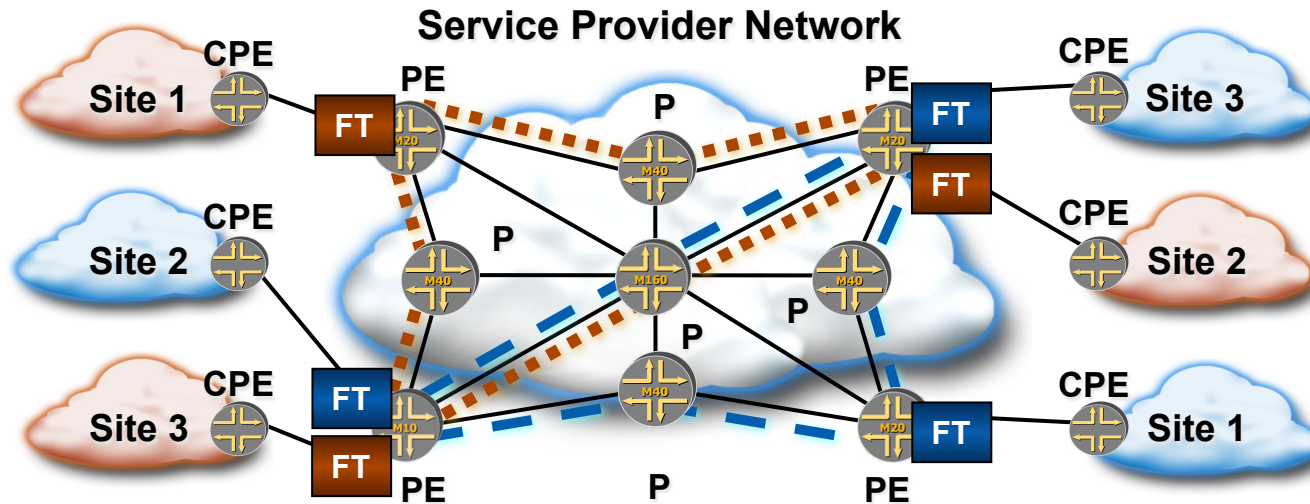
## IPSec VPNs – Example



- ❑ Routing must be performed at CPE
- ❑ Tunnels terminate on subscriber premise
  - Only CPE equipment needs to support IPSec
- ❑ ESP tunnel mode
  - Authentication insures integrity from CPE to CPE
  - Encrypts original header/payload across internet
  - Supports private address space
- ❑ Issues with IPSec VPNs include
  - Complex tunnel structure may lead to high administration effort



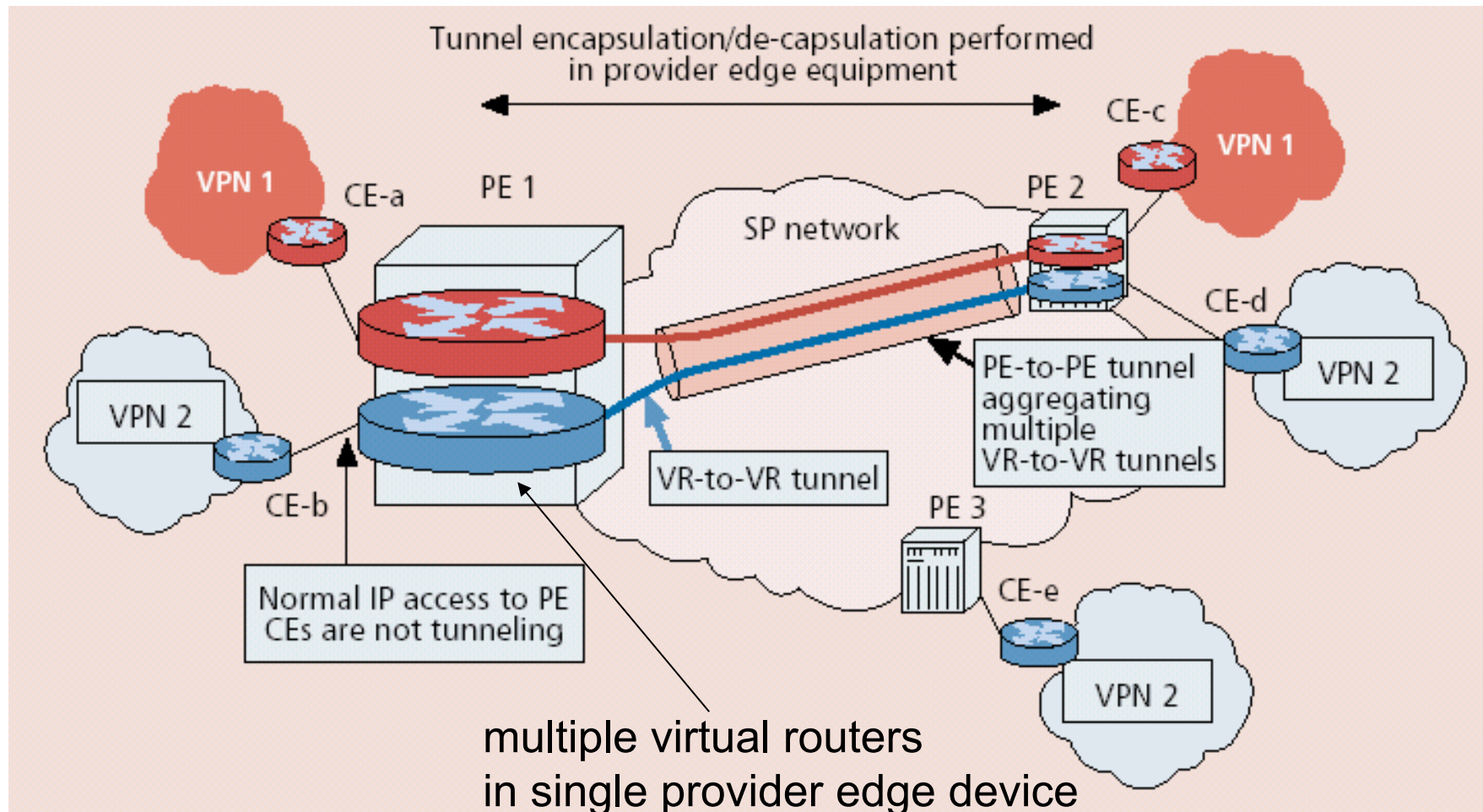
## Layer 3 VPNs: RFC 2547bis / 4364 - MPLS/BGP VPNs



- ❑ MPLS (Multiprotocol Label Switching) is used for forwarding packets over the backbone
- ❑ BGP (Border Gateway Protocol) is used for distributing routes over the backbone
- ❑ Multiple Forwarding Tables (FT) on some edge routers, one for each VPN



# Network-based Layer 3 VPNs

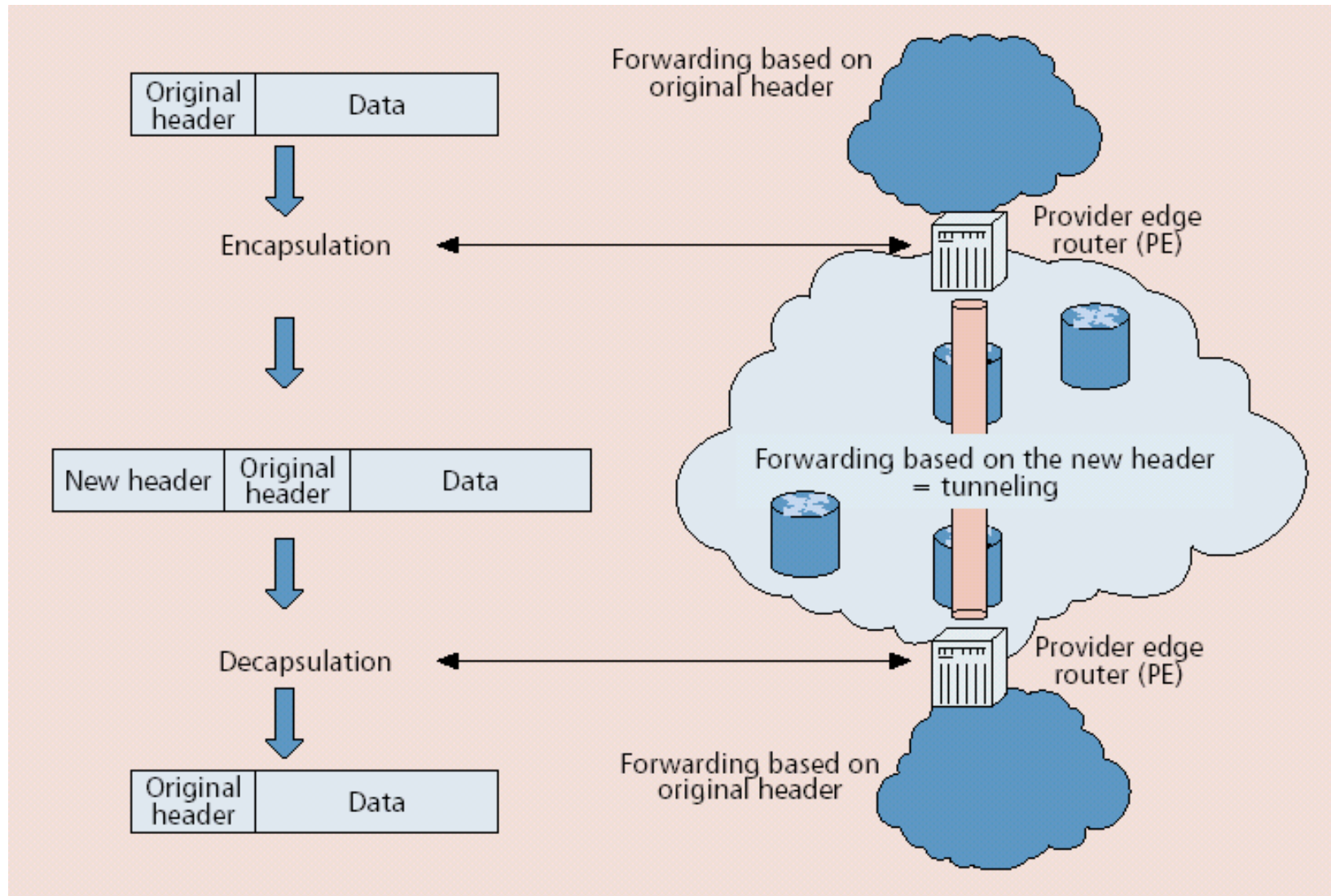


CE routers send their routes to PE routers using BGP. Routes from different VPNs remain separate in PE routers. PE routers receive IP datagrams from CE routers. Each route within a VPN is assigned a MPLS label, which is distributed by BGP.  
c.f. RFC 4364: BGP/MPLS IP Virtual Private Networks (VPNs)





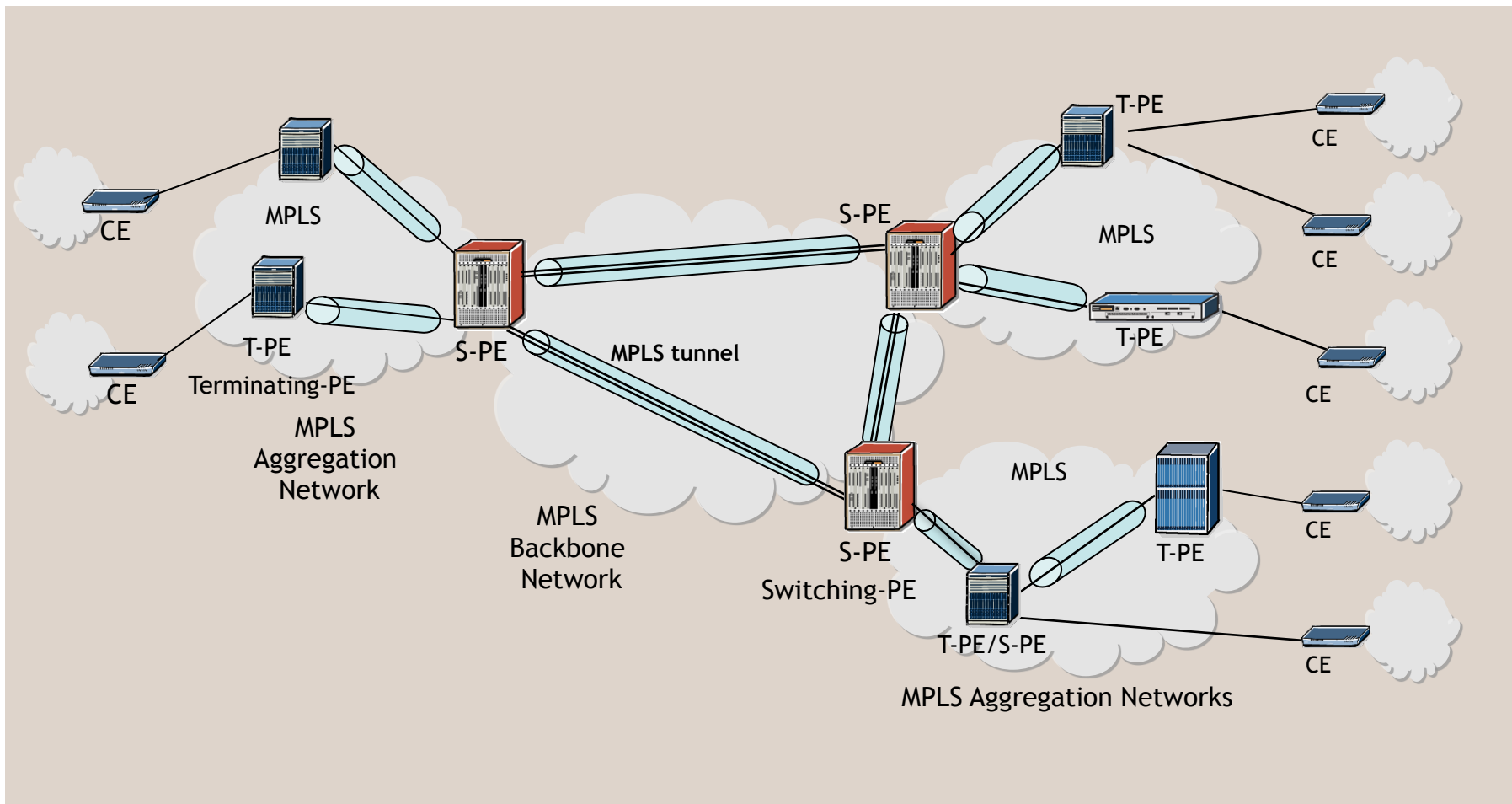
# Tunneling







## MPLS-based Layer 2 VPN



L2 VPN: Routing occurs on CE switch, which must select the appropriate circuit to send traffic. The PE switch sends it across the service provider's network to the PE switch connected to the receiving site. PE switches send data to the appropriate tunnel, and do not use customer's IP routes.



# MPLS

## Multi-Protocol Label Switching

Acknowledgements:

Ping Pan

Kireeti Kompella

Juniper Networks

Philip Matthews

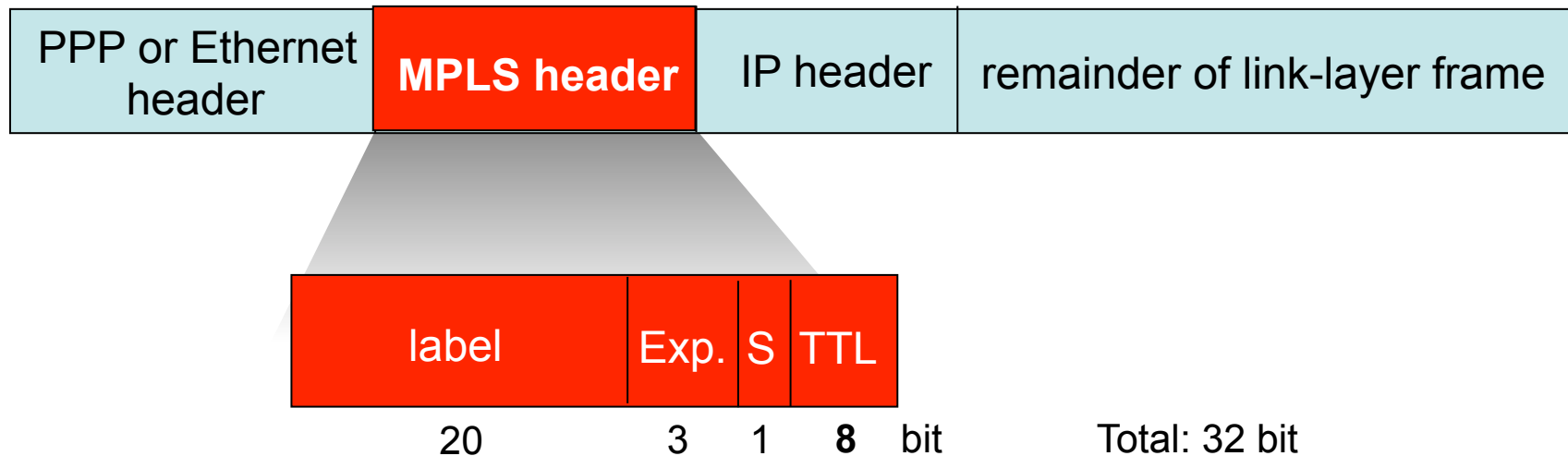
Nortel Networks





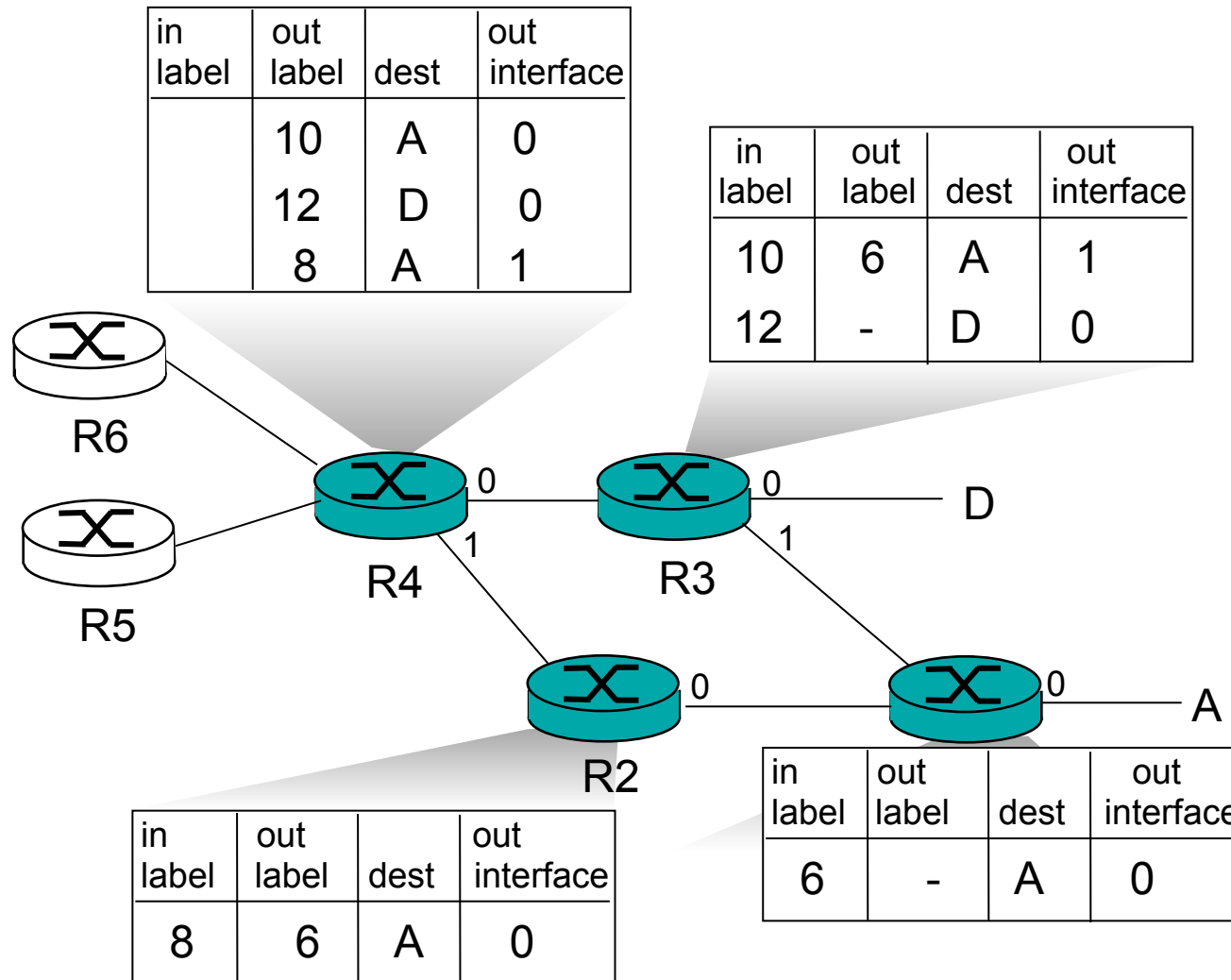
# Multiprotocol label switching (MPLS)

- Initial goal: speed up IP forwarding by using fixed length label (instead of IP address) to do forwarding
  - borrowing ideas from Virtual Circuit (VC) approach
  - IP datagram still keeps IP address
  - RFC 3032 defines MPLS header
    - Label: has role of Virtual Circuit Identifier
    - Exp: experimental usage, may specify Class of Service (CoS)
    - S: Bottom of Stack - end of series of stacked headers
    - TTL: time to live





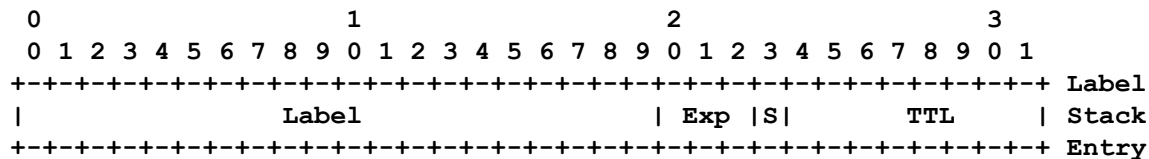
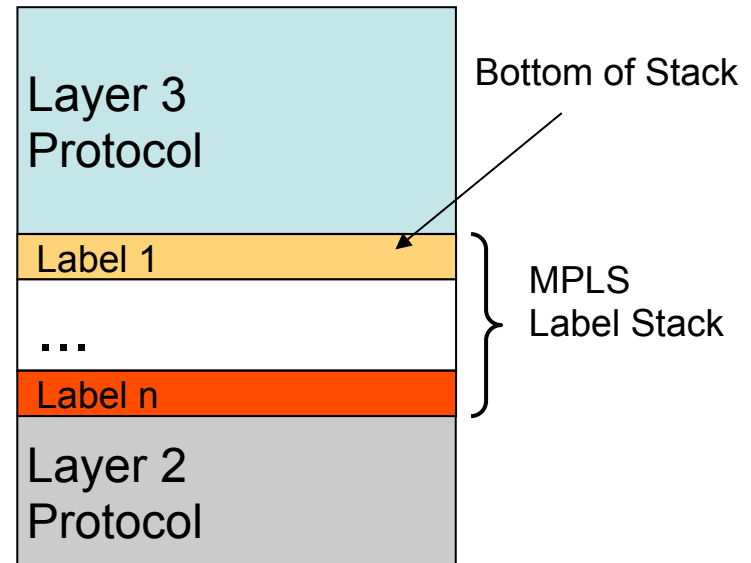
# MPLS forwarding tables





# Multi-Protocol Label Switching (MPLS)

- Properties
  - virtual connections for various protocols and technologies (at Layer 3 and Layer 2)
  - stackable labels
- Processing of Labels in LSRs (Label Switched Routers)
  - adding or dropping of labels
  - label-dependent forwarding
- Label Distribution Protocol (LDP):
  - One possible signalling protocol among LSRs



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



# MPLS

- Rationale
  - Combine IP and connection-oriented technology
  - Leverage ATM hardware
  - Fast forwarding
  - IP Traffic Engineering
  - Virtual Private Networks
  - Support Voice and Video on IP (QoS constraints)
- Two signalling variants
  - LDP - Label Distribution Protocol
    - CR-LDP: Constraint-based Label Distribution protocol
    - Label Distribution Protocol + Explicit Routes
  - RSVP = Resource Reservation Protocol
    - RSVP ext = Resource Reservation Protocol
    - + Explicit Routes + Scalability Extensions

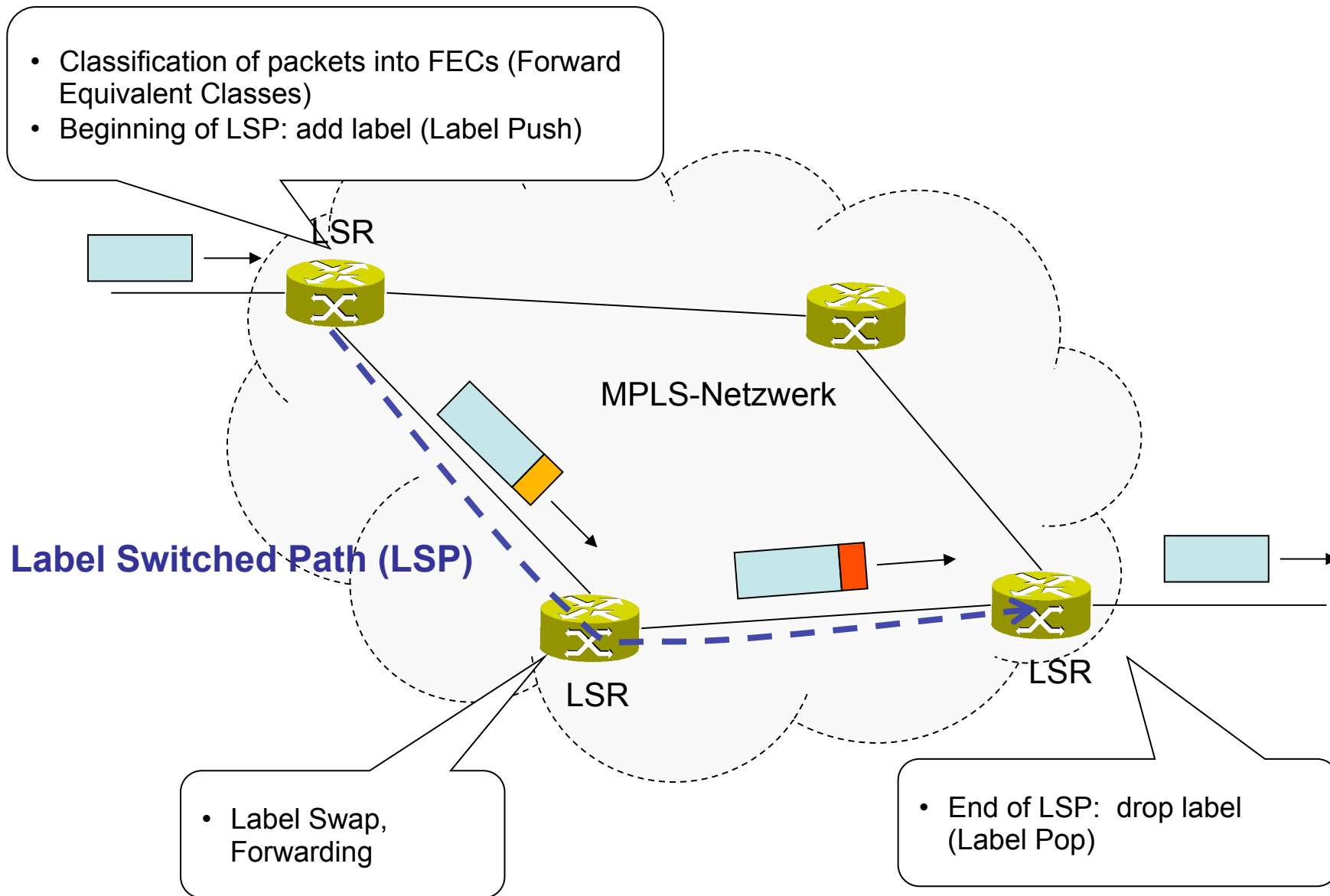


## MPLS Terminology

- ❑ LSP: Label Switched Path
  - part of a tree from every source to that destination (unidirectional)
- ❑ LDP: Label Distribution Protocol
  - builds that tree using IP forwarding tables to route the control messages
- ❑ FEC: Forwarding Equivalence Class
  - subset of packets are all treated the same by a router
  - assigned at MPLS network ingress
- ❑ LSR: Label Switching Router
- ❑ LER: Label Edge Router



# MPLS: Label Switched Path

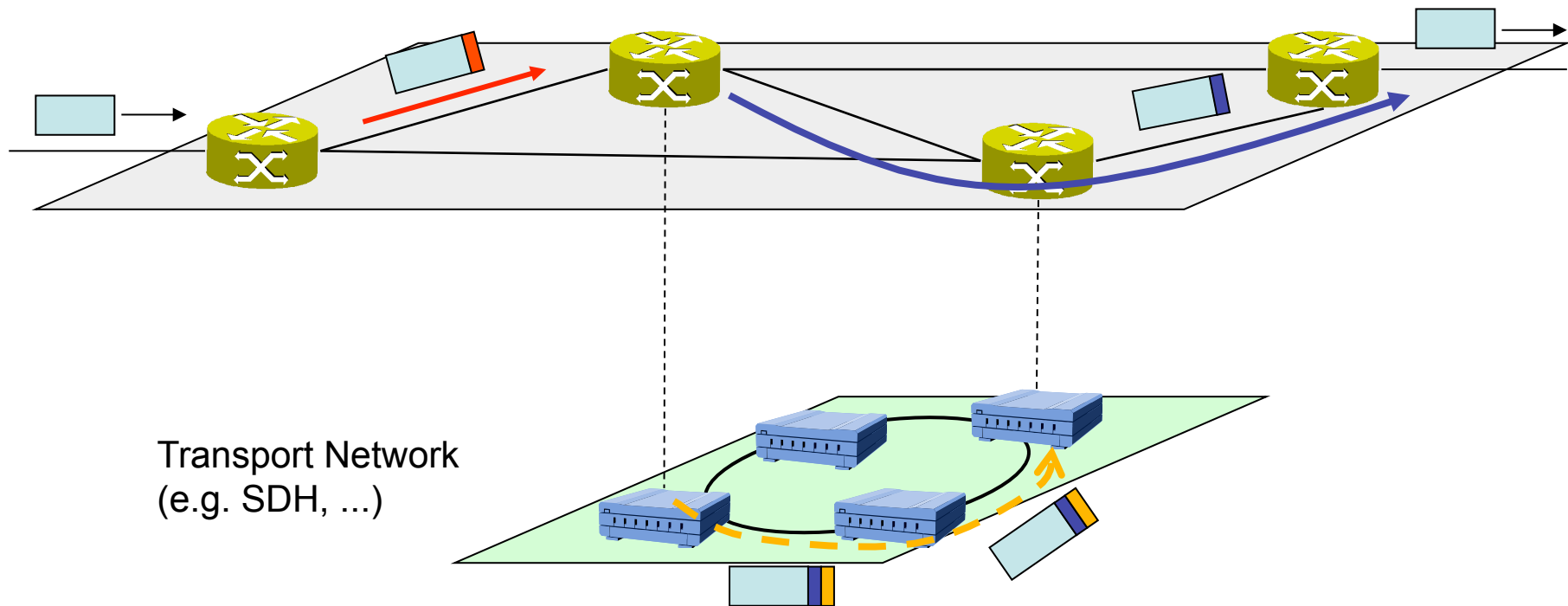






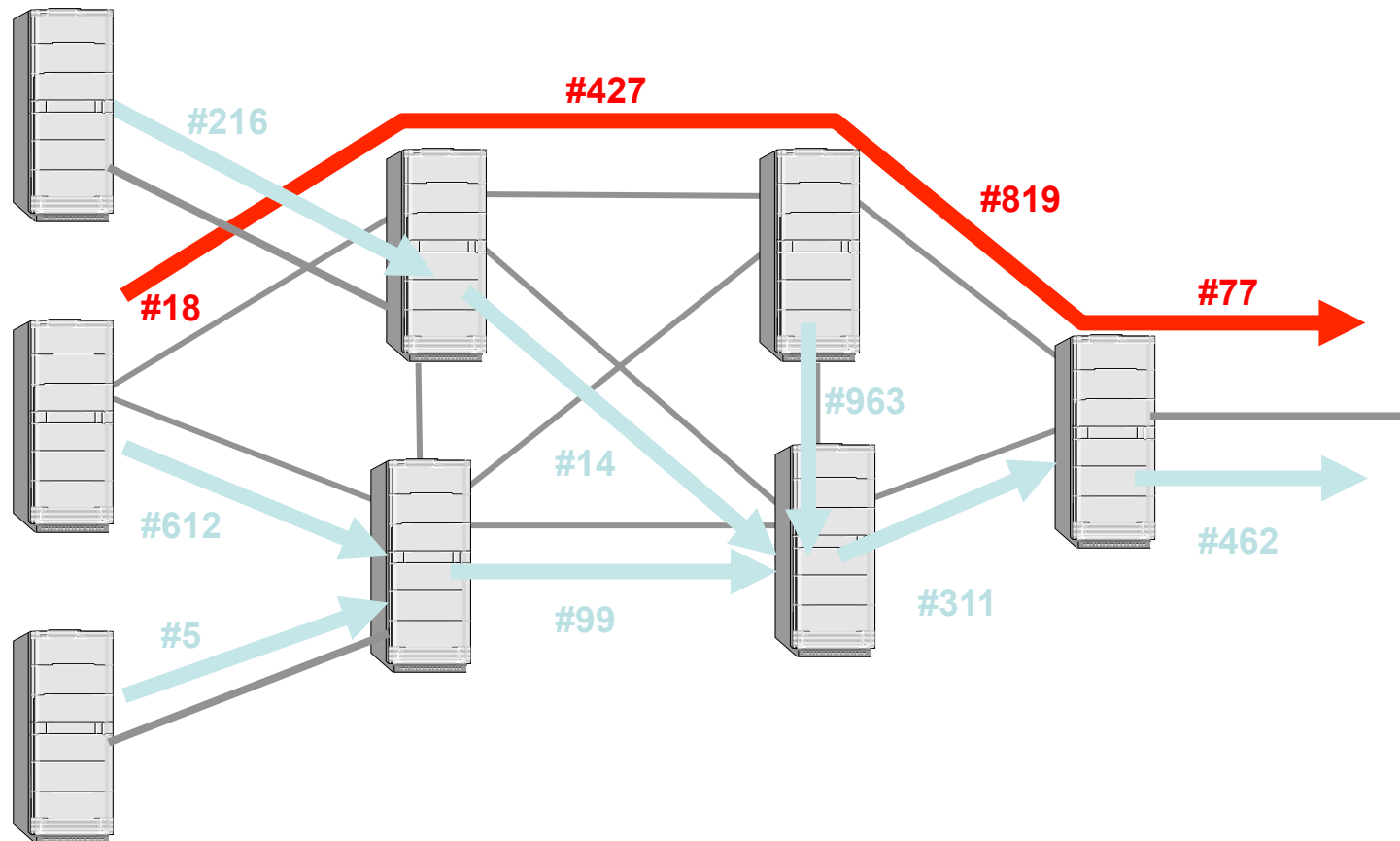
# MPLS Hierarchy

- Label-Switching on different layers







# Label Switched Path: Different Types

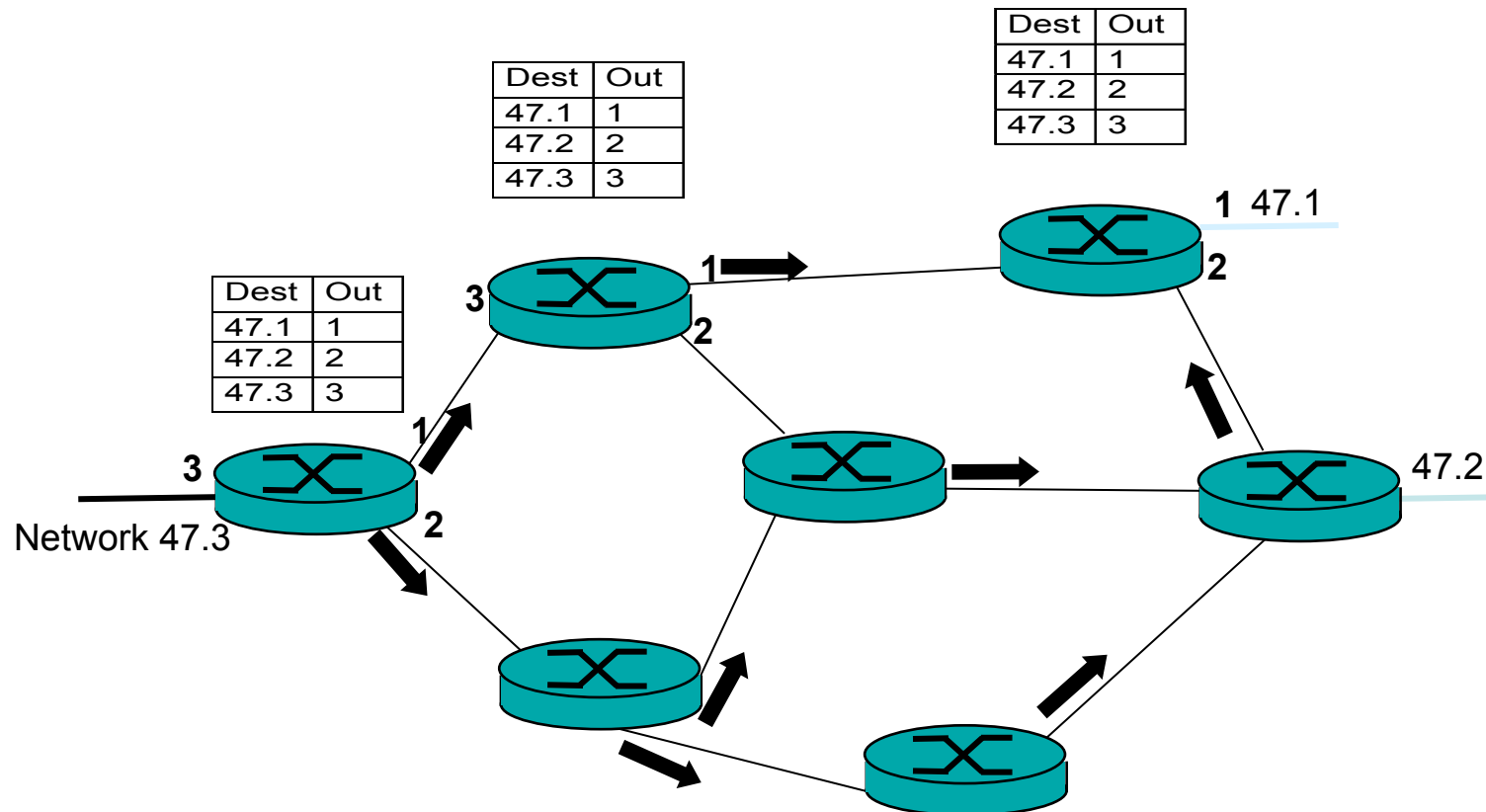


□ Two types of LDP Label Switched Paths:

- Hop by hop (“standard” LDP) 
- Explicit Routing (LDP+”ER”) 



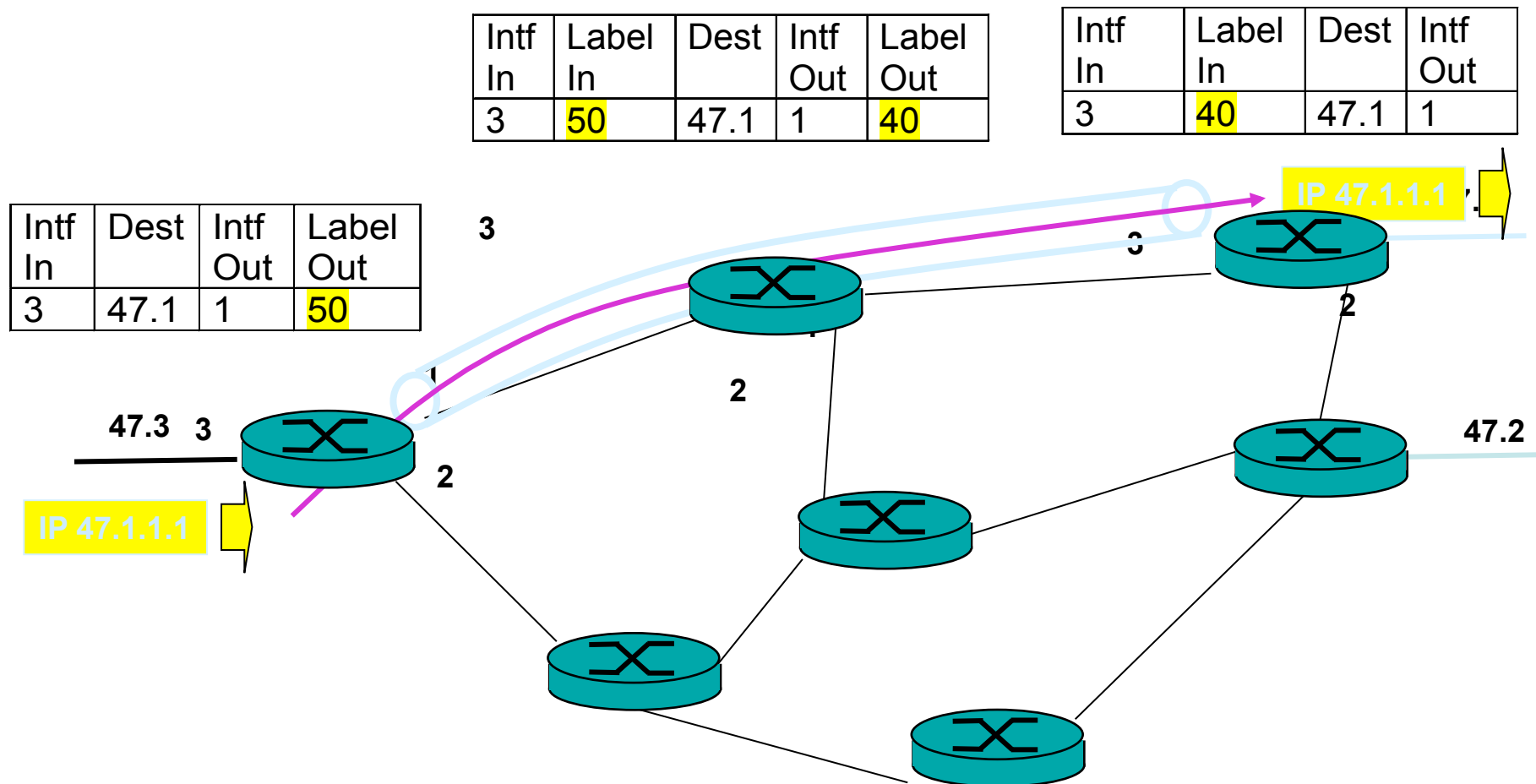
## Label Switched Path: Different Types



- ❑ A “standard” LSP creates MPLS paths for standard IP routing (from IP routing tables)
- ❑ A “standard” LSP is actually part of a tree from every source to that destination (unidirectional)
- ❑ Destination based forwarding tables as built by OSPF, IS-IS, RIP, etc.



# Label Switched Path

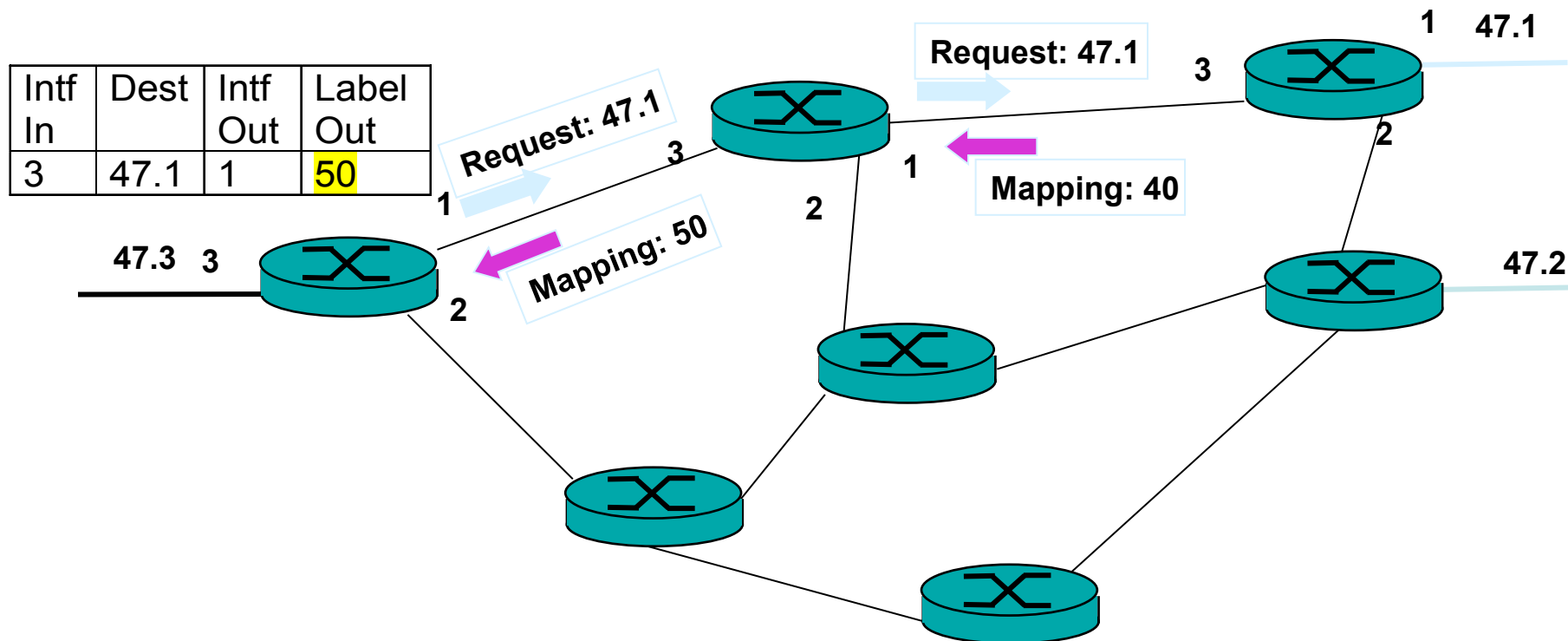




# MPLS Label Distribution

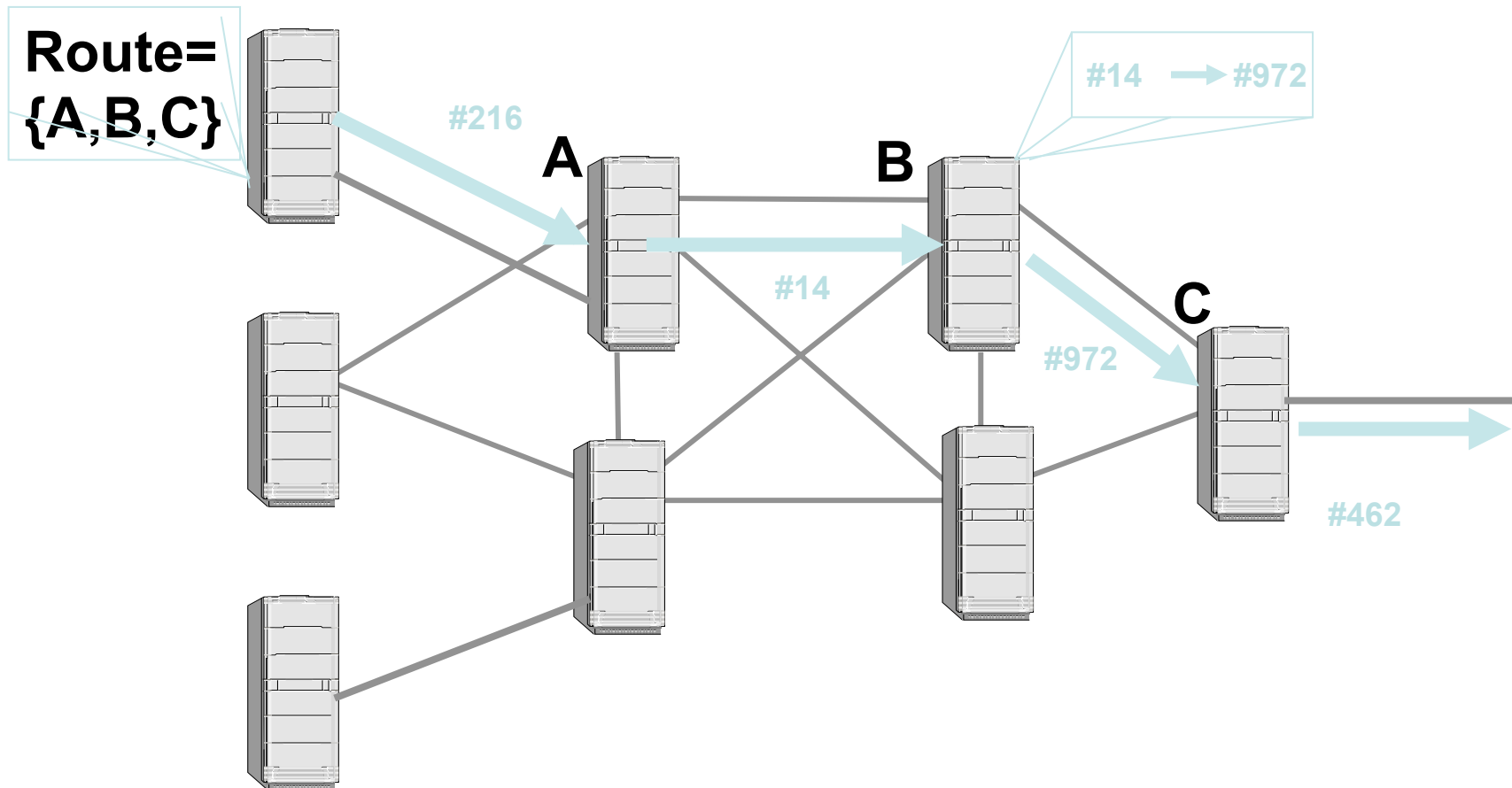
Intf In	Label In	Dest	Intf Out	Label Out
3	50	47.1	1	40

Intf In	Label In	Dest	Intf Out
3	40	47.1	1





## Explicitly Routed LSP (ER-LSP)



- ER-LSP follows the route that source chooses, i.e. control message to establish the LSP (label request) is ***source routed***



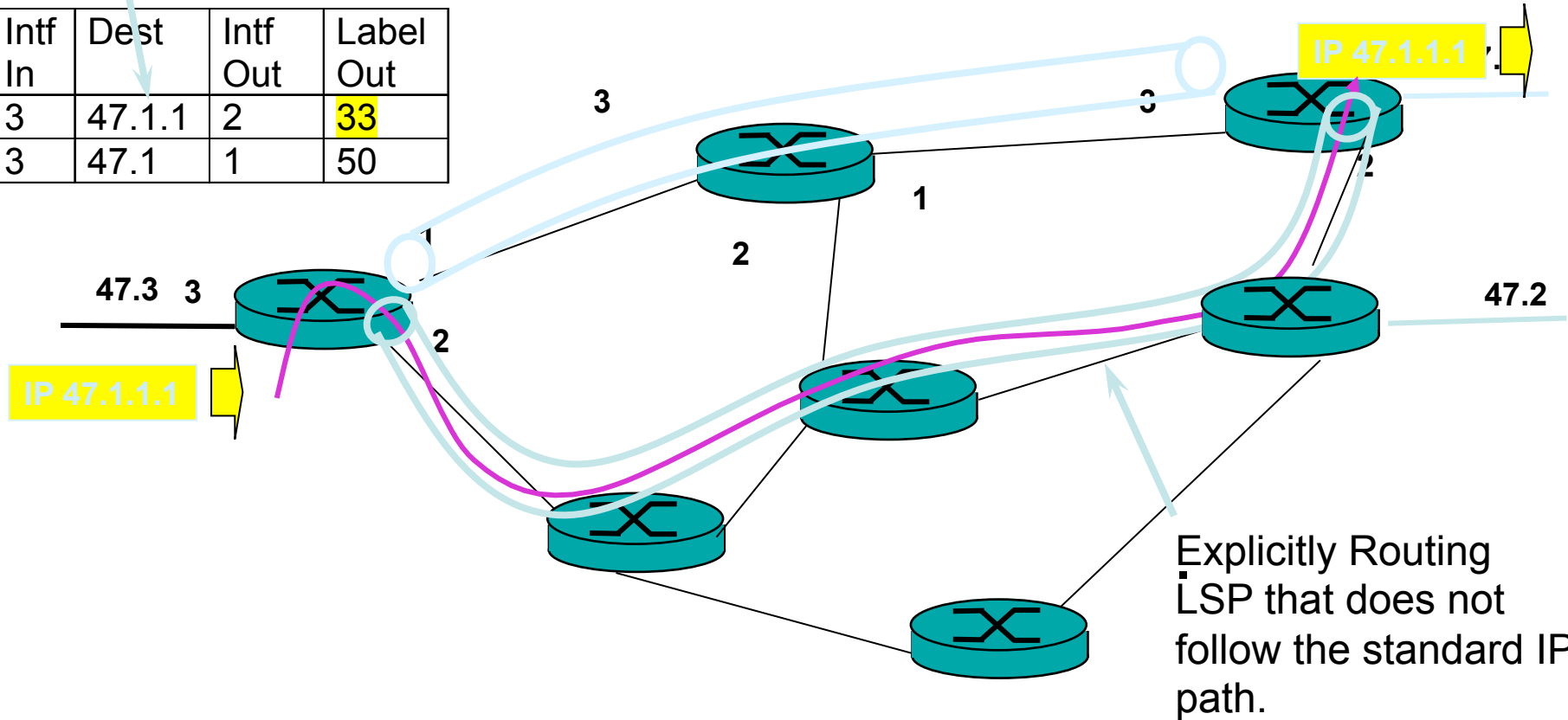
# Explicitly Routed LSP (ER-LSP)

This entry gives the longest prefix match.

Intf In	Dest	Intf Out	Label Out
3	47.1.1	2	33
3	47.1	1	50

Intf In	Label In	Dest	Intf Out	Label Out
3	50	47.1	1	40

Intf In	Label In	Dest	Intf Out
3	40	47.1	1





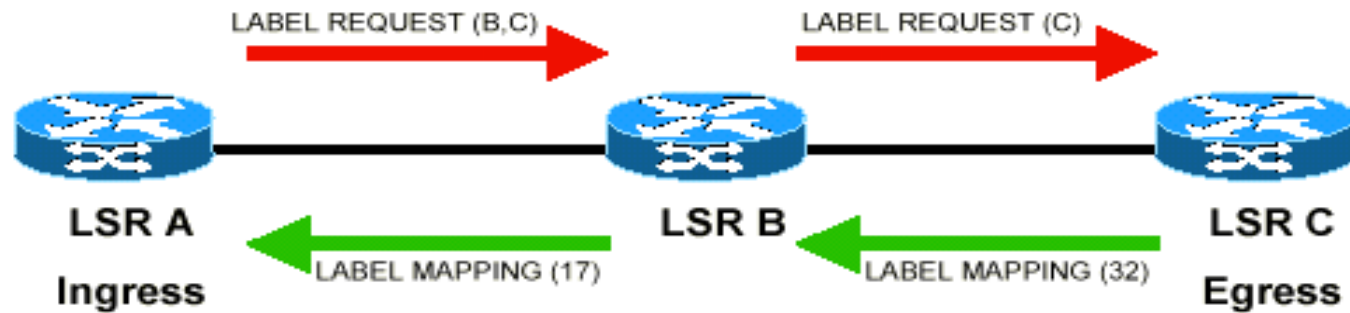
## Constraint-based Routing LDP (CR-LDP)

- ❑ Constraint-Based Routing is one method of Traffic Engineering
- ❑ RFC 2702: Requirements for Traffic Engineering Over MPLS
  - Strict & Loose ER
  - Specification of QoS
  - Specification of Traffic Parameters
  - Route Pinning
  - Preemption
  - Failure Recovery





## Signalling Protocol CR-LDP



**LSP Setup Flow**

- ❑ Hard State Protocol
- ❑ UDP used for peer discovery
- ❑ TCP used for session, advertisement, notification, and LDP messages
- ❑ Supports Diffserv and Operator configurable QOS classes
- ❑ Failure reported using TCP

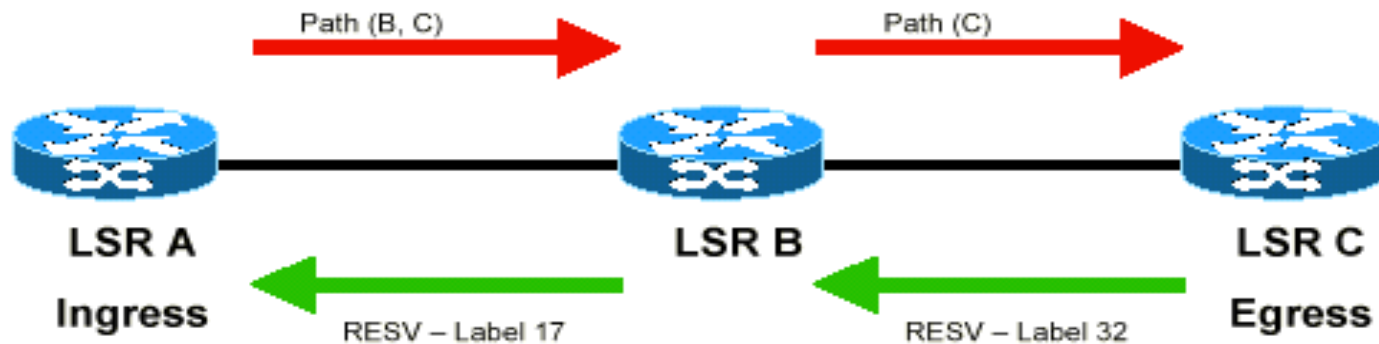


## ER-LSP setup using RSVP-TE

- TE (Traffic Engineering) extensions to RSVP
- Built on RSVP messages over IP
  - In RSVP, a source requests resources along a path
  - Then the source regularly sends refresh messages to keep the reservations active
- Extensions to RSVP
  - Explicit Route Object
  - Label Request
  - Label Object
  - Session Attribute
  - Record Route Object
- Defines a set of constraints for LSP computation and admission
  - Expectation and Allocation of resources: Uses Inserv-style reservations
  - Preemption Level: Setup and Holding Priority with respect to other LSPs



## Signalling Protocol Extended RSVP



- Extension of the classical connectionless RSVP
- **Path** and **Resv** messages used with
  - Label\_Request Object
  - Explicit\_Route Object
  - Label Object
- Aggregation of flows to reduce state information in routers
- Soft State Control and scalability concerns



## MPLS TTL

- MPLS TTL
  - is 8 bits long, used as a control mechanism to prevent packets looping in the network.
  - TTL value is decremented by 1 at ingress LSR/eLER. A packet with TTL of 0 is not transmitted.
- Two approaches to TTL handling on ingress to the MPLS network
  - Pipe Mode
    - iLER decrements the IP TTL value and sets MPLS TTL to a value different than IP TTL.
    - At eLER MPLS TTL is decremented, IP TTL is decremented when IP packet is processed.
  - Uniform Mode
    - iLER decrements the IP TTL value and copies resulting value to the MPLS TTL field.
    - At eLER MPLS TTL is decremented and copied to IP TTL.



## MPLS TTL Processing

### c.f. RFC 3032 - MPLS Label Stack Encoding

- Protocol-independent rules
  - "outgoing TTL" of a labeled packet is either
    - a) one less than the incoming TTL, or b) zero
  - Packets with TTL=0 are discarded
- IP-dependent rules
  - When an IP packet is first labeled, the TTL field of the label stack is set to the value of the IP TTL field
  - If the IP TTL field needs to be decremented, as part of the IP processing, it is assumed that this has already been done
  - When a label is popped, and the resulting label stack is empty, then the value of the IP TTL field **SHOULD BE** replaced with the outgoing MPLS TTL value
  - A network administration may prefer to decrement the IPv4 TTL by one as it traverses an MPLS domain



## ICMP

- ❑ When a router receives an IP datagram that it can not forward, it sends an ICMP message to the datagram's originator
- ❑ The ICMP message indicates why the datagram could not be delivered, e.g., Time Expired, Destination Unreachable
- ❑ The ICMP message also contains the IP header and at least leading 8 octets of the original datagram
  - RFC 1812 - Requirements for IP Version 4 Routers extends this to “as many bytes as possible”
  - Historically, every ICMP error message has included the IP header and at least leading 8 octets
  - Including only the first 8 data bytes of the datagram that triggered the error frequently is no longer adequate, due to use e.g. of IP-in-IP tunneling
  - Therefore ICMP datagram SHOULD contain as much of original datagram as possible (max. ICMP length 576 bytes)



## ICMP in presence of MPLS

- ❑ When a Label Switched Router (LSR) receives an MPLS encapsulated datagram that it can not deliver
  - It removes entire MPLS labels stack
  - It sends an ICMP message to datagram's originator
- ❑ ICMP message indicates why the datagram could not be delivered (e.g., time expired, destination unreachable)
- ❑ ICMP message also contains IP header and leading 8 octets of the original datagram
  - RFC 1812 extends this to “as many bytes as possible”



## ICMP in Presence of MPLS

### Issue

- ❑ The ICMP message contains no information regarding the MPLS stack that encapsulated the datagram when it arrived at the LSR
- ❑ This is a significant omission because:
  - The LSR tried to forward the datagram based upon that label stack
  - Resulting ICMP message may be confusing





## ICMP in Presence of MPLS

### Issue

- ❑ ICMP Destination Unreachable
  - Message contains IP header of original datagram
  - Original datagram couldn't be delivered because MPLS forwarding path was broken
- ❑ ICMP Time Expired
  - Message contains IP header of original datagram
  - TTL value in IP header is greater than 1
  - TTL expired on MPLS header. ICMP Message contains IP header of original datagram



## ICMP with MPLS

c.f. RFC 4950 - ICMP Extensions for Multiprotocol Label Switching

- ❑ defines an ICMP extension object that permits LSR to append MPLS information to ICMP messages.
- ❑ ICMP messages include the MPLS label stack, as it arrived at the router that is sending the ICMP message.
- ❑ equally applicable to ICMPv4 [RFC792] and ICMPv6 [RFC4443]
- ❑ sample output from an enhanced TRACEROUTE:

```
> traceroute 192.0.2.1
```

```
traceroute to 192.0.2.1 (192.0.2.1), 30 hops max, 40 byte packets
```

```
1 192.0.2.13 (192.0.2.13) 0.661 ms 0.618 ms 0.579 ms
```

```
2 192.0.2.9 (192.0.2.9) 0.861 ms 0.718 ms 0.679 ms  
   MPLS Label=100048 Exp=0 TTL=1 S=1
```

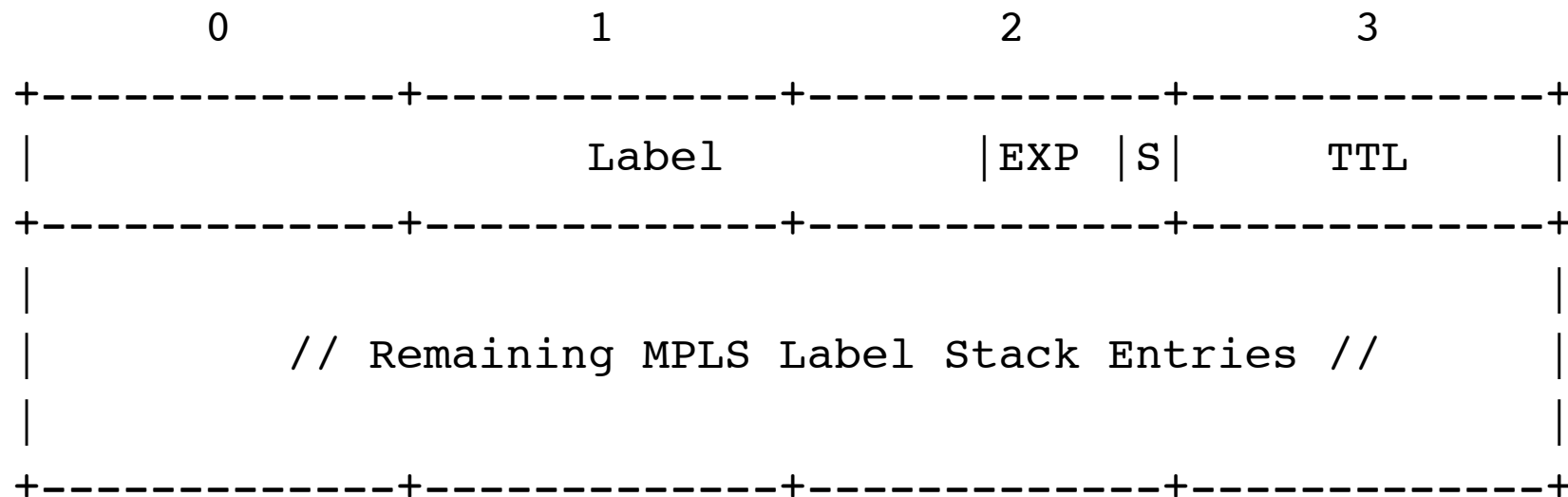
```
3 192.0.2.5 (192.0.2.5) 0.822 ms 0.731 ms 0.708 ms  
   MPLS Label=100016 Exp=0 TTL=1 S=1
```

```
4 192.0.2.1 (192.0.2.1) 0.961 ms 8.676 ms 0.875 ms
```



## ICMP with MPLS

- ❑ MPLS Label Stack Object: can be appended to ICMP Time Exceeded and Destination Unreachable messages



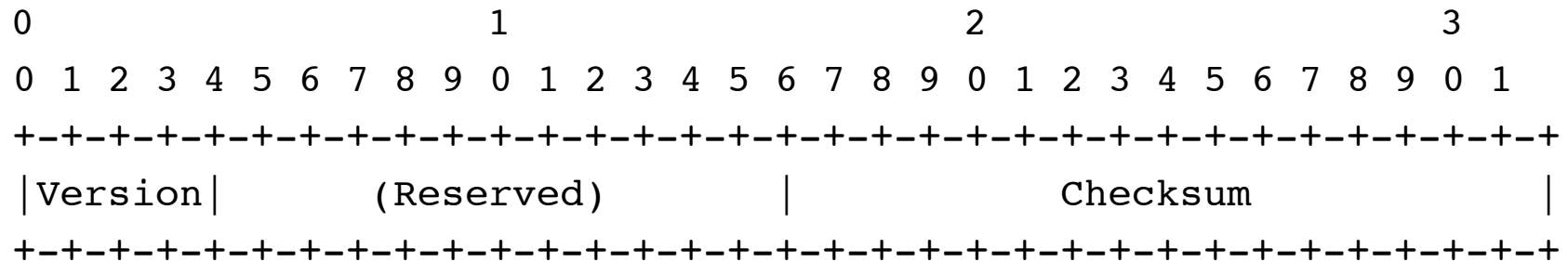
- ❑ Must be preceded by an ICMP Extension Structure Header and an ICMP Object Header, defined in [RFC4884].



# Multi-Part ICMP Messages - RFC 4884

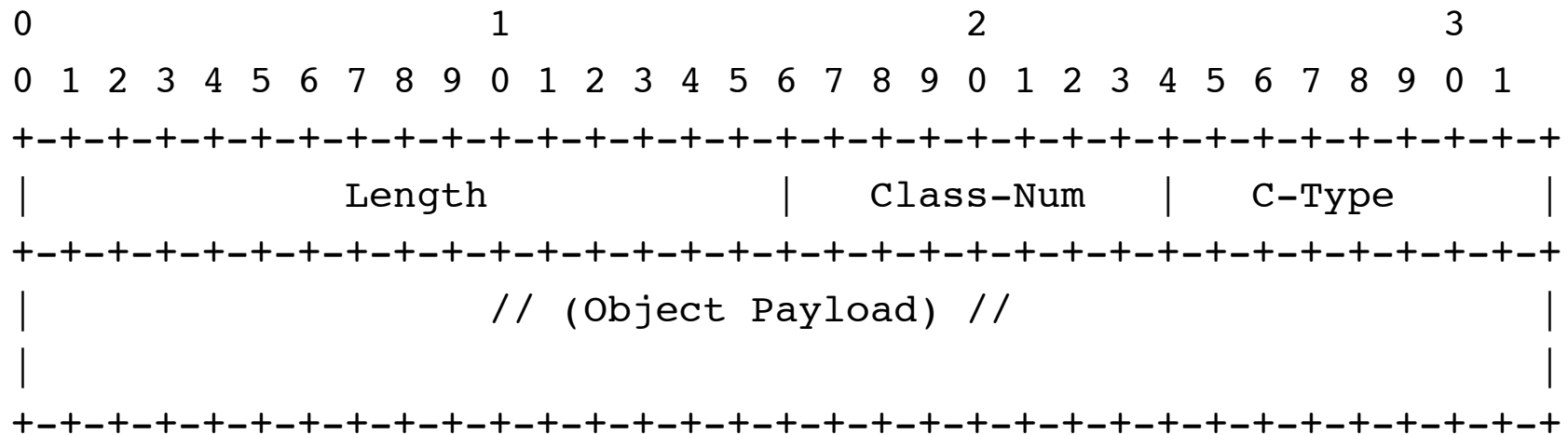
- ❑ ICMP Extension Structure may be appended to ICMP v4 / v6 Destination Unreachable and Time Exceeded messages

- ❑ ICMP Extension Structure Header



ICMP extension version number: 2

- ❑ ICMP Object Header and Object Payload





## MPLS for Linux

# The work of James Leu (last updated July 2011):

<https://sourceforge.net/projects/mpls-linux/>

Discussions:

[http://sourceforge.net/mailarchive/forum.php?forum\\_name=mpls-linux-devel](http://sourceforge.net/mailarchive/forum.php?forum_name=mpls-linux-devel)

# Bug fixes of Jorge Boncompte:

<http://mpls-linux.git.sourceforge.net/git/gitweb.cgi?p=mpls-linux/net-next;a=shortlog;h=refs/heads/net-next-mpls>

# Additional bug fixes by Igor Maravić:

<https://github.com/i-maravic/MPLS-Linux>

<https://github.com/i-maravic/iproute2>

# MPLS for Linux Labs

by Irina Dumitrascu and Adrian Popa: graduation project with purpose of teaching MPLS to university students, at Limburg Catholic University College

<http://ontwerpen1.khlim.be/~lrutten/cursussen/comm2/mpls-linux-docs/>

includes e.g. Layer 2 VPN with MPLS, Layer 3 VPN with MPLS



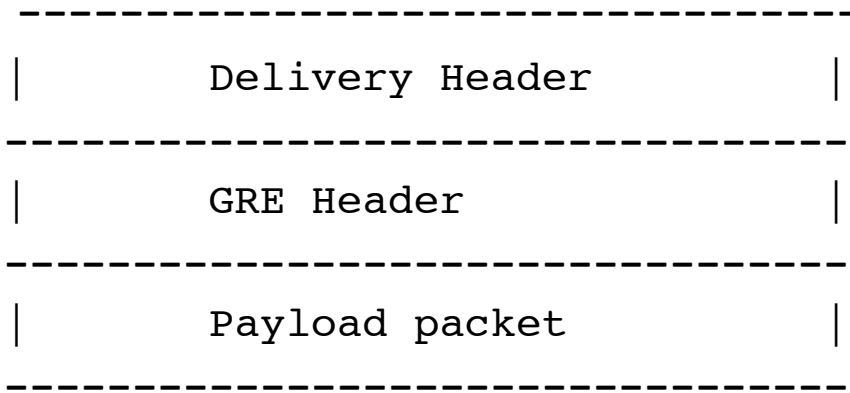
## MPLS Assessment

- Tunnels using MPLS vs. tunnels using IP
- MPLS: LDP – automated tunnel setup, following IP routing
- IP
  - IP-in-IP
  - GRE (Generic Routing Encapsulation Protocol)
  - IPSec

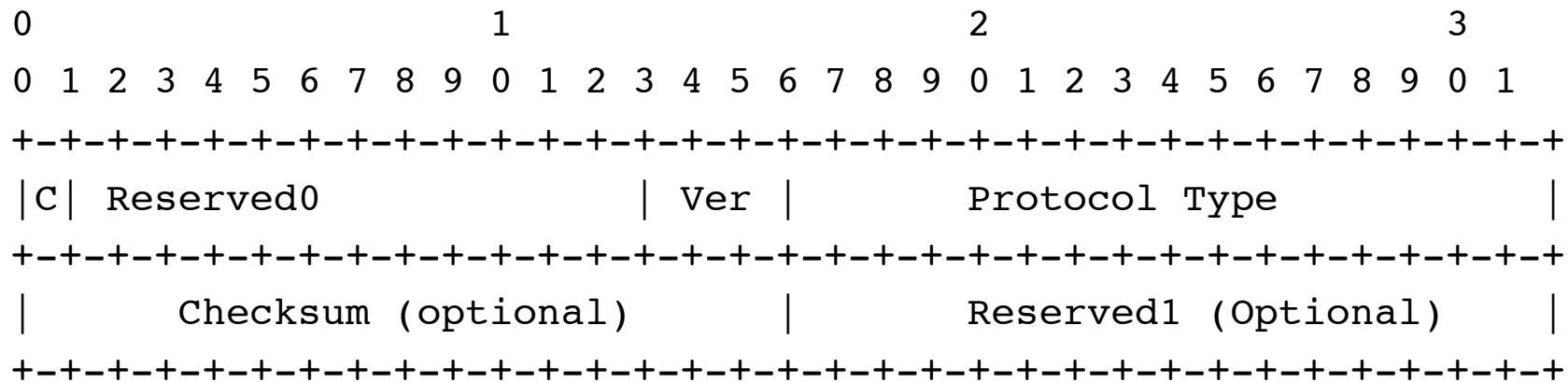


# GRE: Generic Routing Encapsulation Protocol

- RFC 2784
- Structure of a GRE Encapsulated Packet



- The GRE packet header has the form





## Tunnel Comparison

### MPLS tunnels

- ❑ Small header
- ❑ Label stacking
- ❑ Signaling for tunnel setup
- ❑ **MPLS-specific routing**
- ❑ Harder to spoof
- ❑ No data security

### IP tunnels

- ❑ Big header
- ❑ No stacking (typically)
- ❑ Configured tunnels (typically)
- ❑ **IP-only routing**
- ❑ Spoofable
- ❑ IPSec