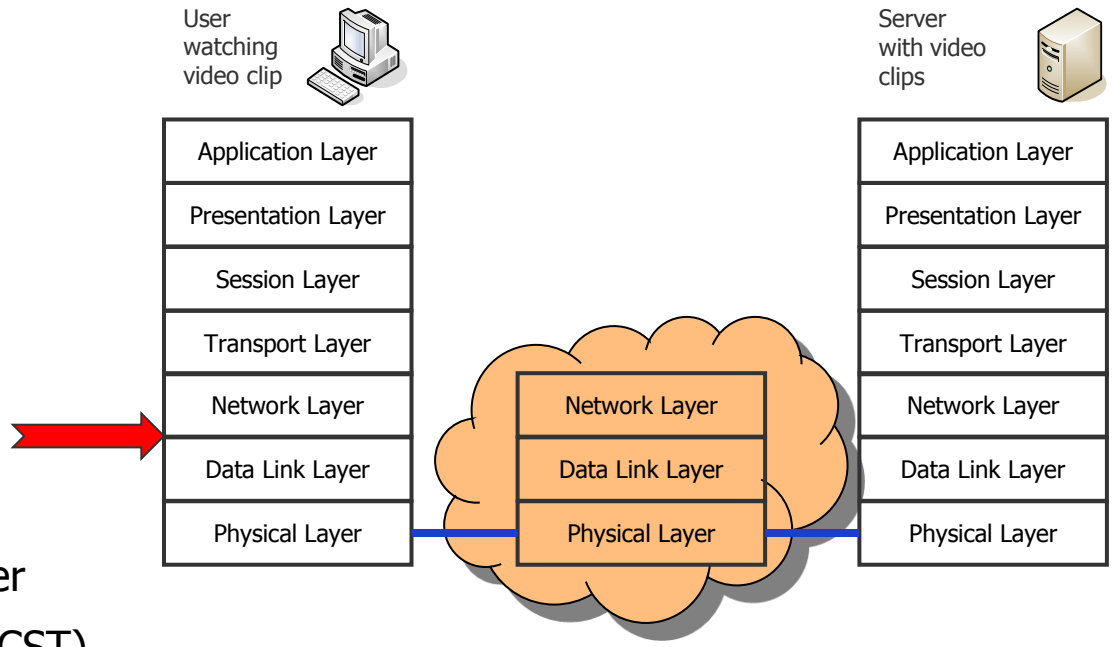
Freie Universität Berlin

Telematics
Chapter 7: MPLS



Univ.-Prof. Dr.-Ing. Jochen H. Schiller
Computer Systems and Telematics (CST)
Institute of Computer Science
Freie Universität Berlin
http://cst.mi.fu-berlin.de

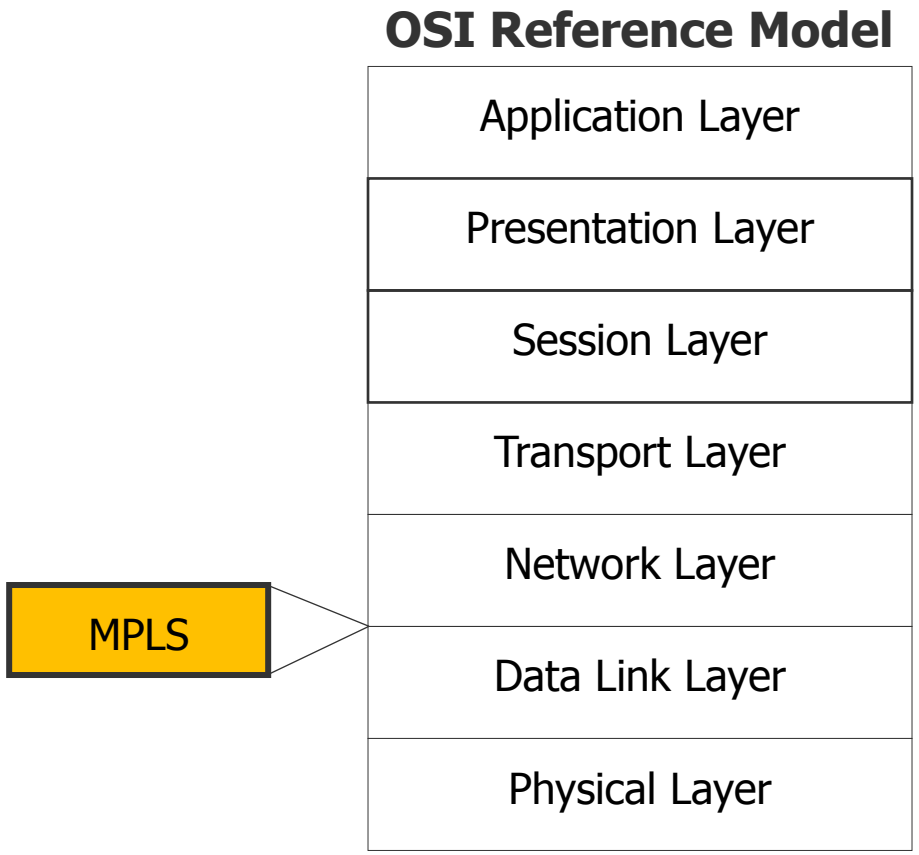
Contents

- Design issues
- Motivation
  - Convergence of IP, ATM, Frame Relay, QoS, ...
  - Traditional routing
- Overview
  - Architecture, Labels, Concepts, Components
- Examples
  - Sending of data packets, Label-distribution
- Applications
  - Fast Forwarding, Scalability, IP/ATM-Integration
  - Quality of Service
  - Virtual Private Networks
  - Voice over MPLS
  - Traffic Engineering

Design Issues

Design Issues

- Two different types of networks
  - Packet switched networks
  - Circuit switched networks
- Global network: Internet
  - The Internet is inherently connection-less on the network layer (packet switched)
  - Many applications require however QoS
  - Realize services of connection-oriented communication over a packet switched network
- MPLS is located between Layer 2 and Layer 3



Motivation

- Convergence of IP and ATM
  - ATM supports traffic management
  - ATM-Switches provide high performance and scalability
  - IP is the protocol of the Internet
  - Nearly all data traffic is carried by IP
  - Plenty of Frame-Relay networks
- New requirements/challenges to IP
  - High bandwidth, differentiated services, security, management
  - Commercial usage as VPN
  - Routing has to consider new metrics, like costs, delay, and jitter
    - Not only shortest paths (SPF, shortest path first)
- How to use existing networks with QoS- and management functionality together with IP?

MPLS

- Framework of the IETF for efficient routing, fast forwarding, and management of data flows

- Operations
  - Mechanisms for the management of data flows with different granularity
  - Independent of Layer 2 and Layer 3 protocols
  - Mechanism for **mapping** of **IP addresses** to simple (**flat**) **labels** with fix length
  - Interface to routing and signaling protocols like RSVP and OSPF
  - Supports PPP/Ethernet, ATM, Frame Relay, ...

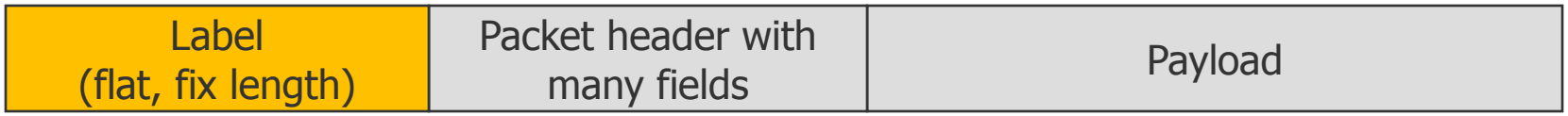
- **Forwarding of data packets is based on labels**

Traditional IP Routing vs. MPLS

- Router
  - Routing-Table
    - Mapping of address-prefix/net mask to the next router/interface
  - Routing-Algorithm
    - Method to determine the shortest/best/cheapest/... route
  - Routing-Protocol
    - Exchange of routing information, e.g., OSPF, BGP, …
  - Forwarding
    - For each incoming data packet decide based on the destination IP address to which output/neighbor to forward
    - Optionally other header fields may be considered for the decision, e.g., prioritizing data packets
- Important: Each router has to make for each packet the forwarding decision

Label

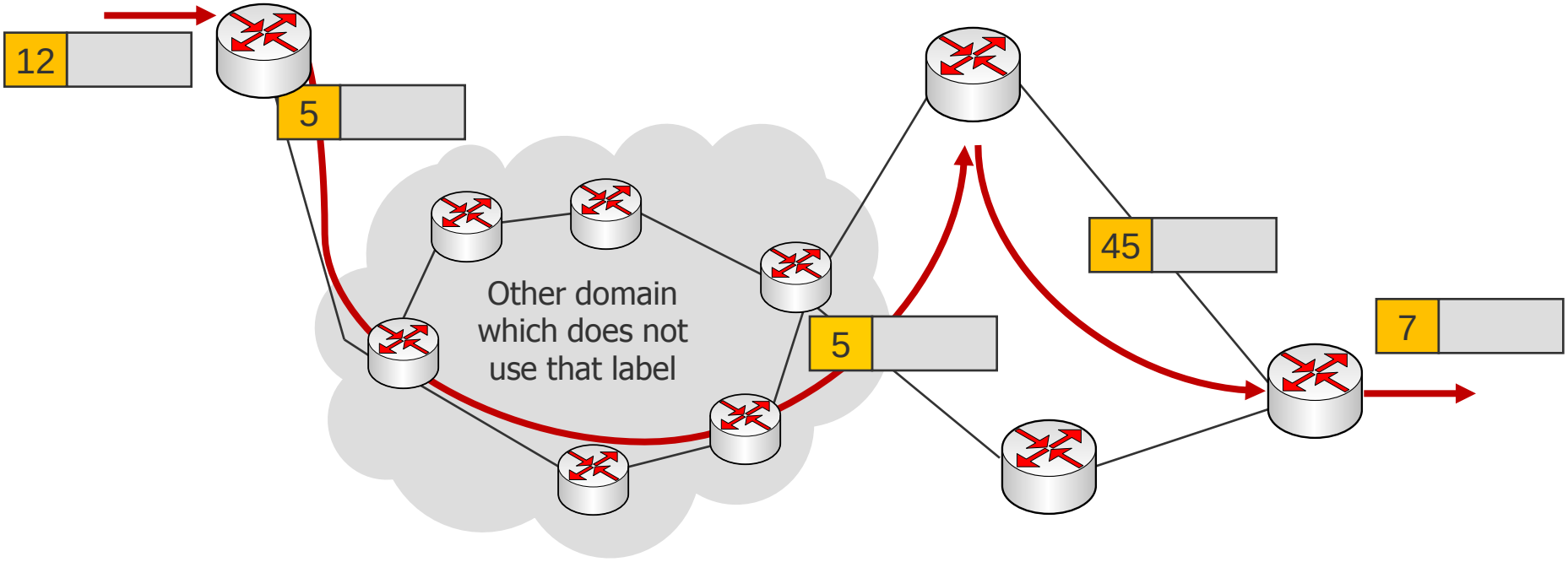
- Unique identification of packets for simple forwarding



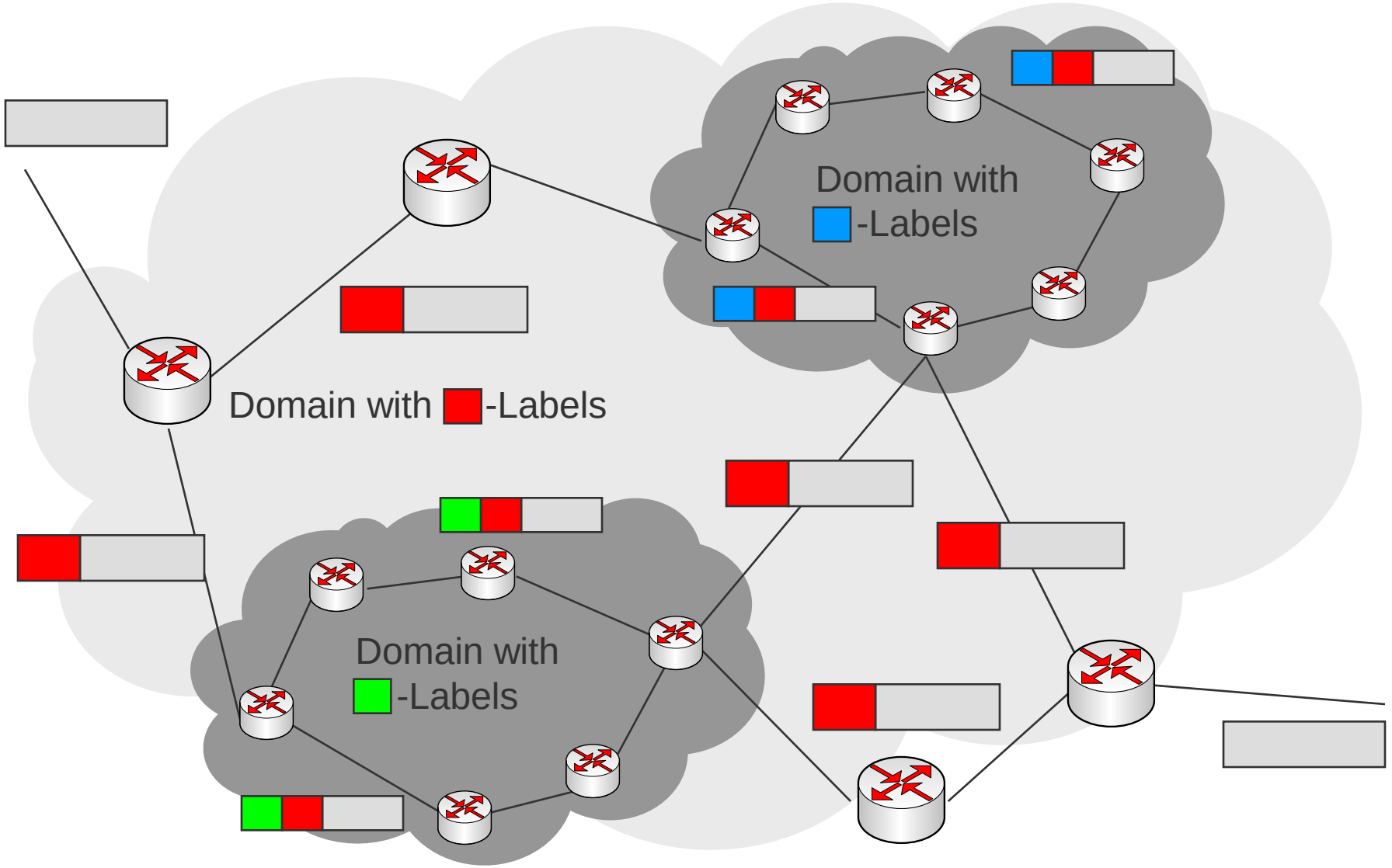
- New name for ...
  - ATM: VPI/VCI, in each cell
  - Frame Relay: DLCI – in each frame
  - STM: each time slot is implicitly a label
  - X.25: LCN is a label
  - Many proprietary protocols (Tag-Switching, ...)
  - Wave length in WDM may be interpreted as label
  - ...

Label

- A label is only locally valid
  - Local means two networks which communicate on the same level
  - Intermediate networks substitute labels (**label substitution**) and forward packets based on the label (**label switching**)



Hierarchies of Labels



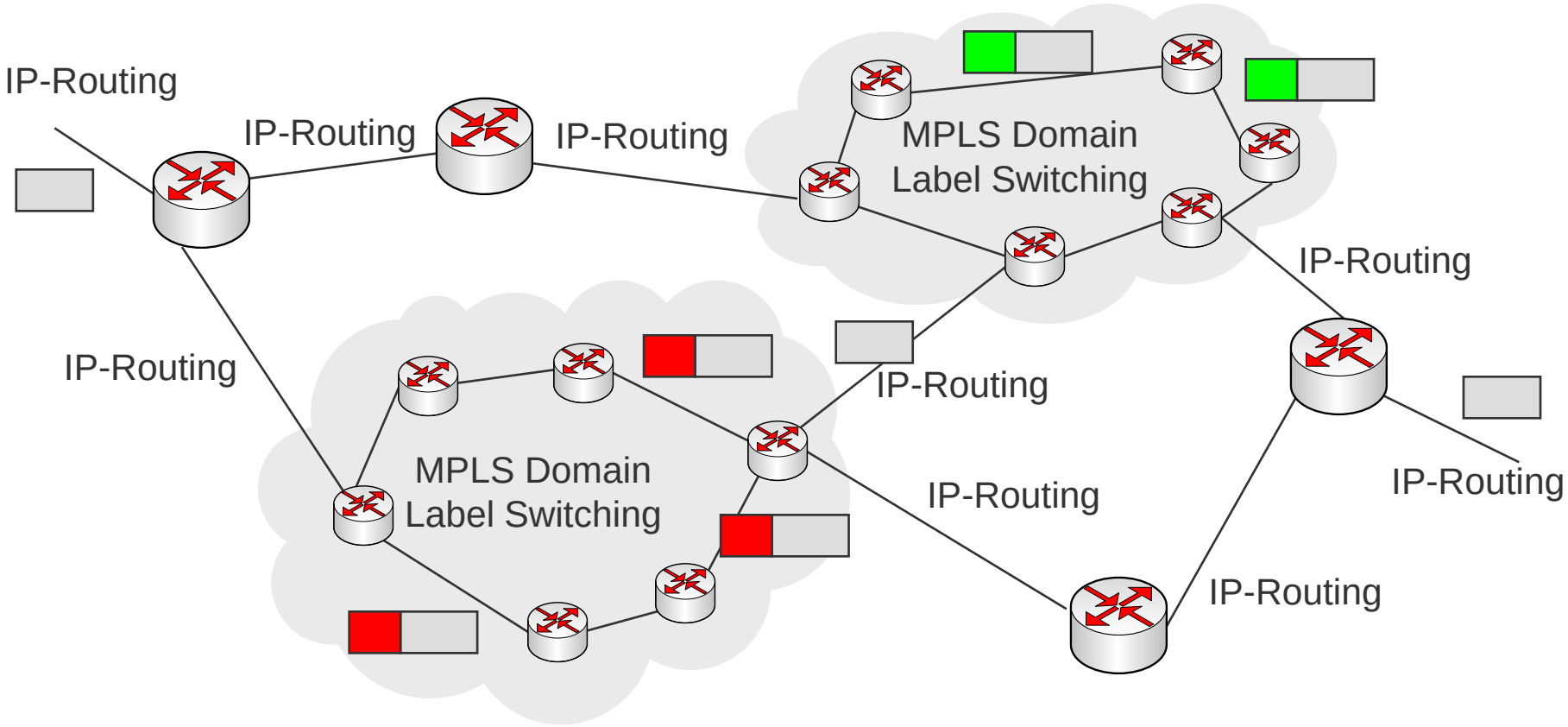
MPLS Concepts

MPLS: Concepts

- Packet forwarding is done based on labels
  - A packet is assigned a label as soon as it enters an MPLS domain
- Packet classification
  - When a packet enters an MPLS domain, it is classified according to
    - Destination address, destination network
    - Quality of Service (Security, Bandwidth, Delay, …)
    - Application
    - Virtual Private Network (VPN)
    - Multicast-Group
  - Mapping to a **Forward Equivalence Class** (**FEC**)
    - Group of packets, which has to be handled in the same manner and over the same path
    - Classification is coded into the label and is done only at the entry into the MPLS domain

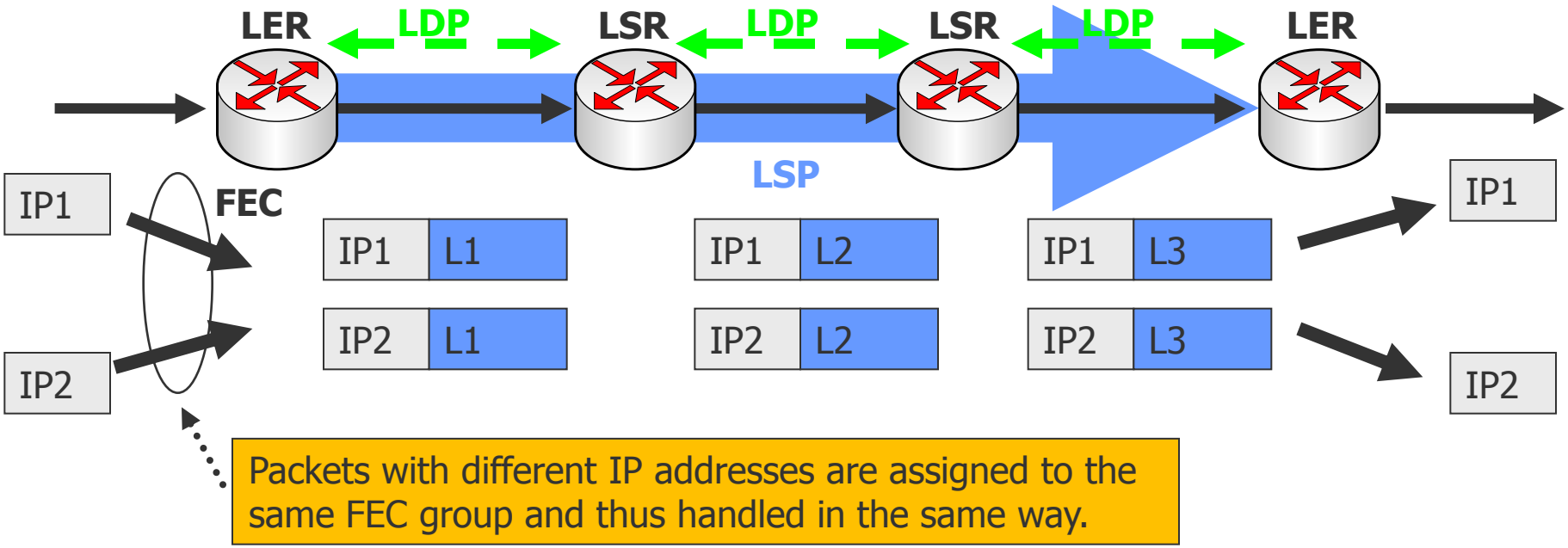
MPLS: Concepts

- IP-Routing outside of the MPLS domain and Label-Switching inside
  - An MPLS network looks to the outside like a large IP router

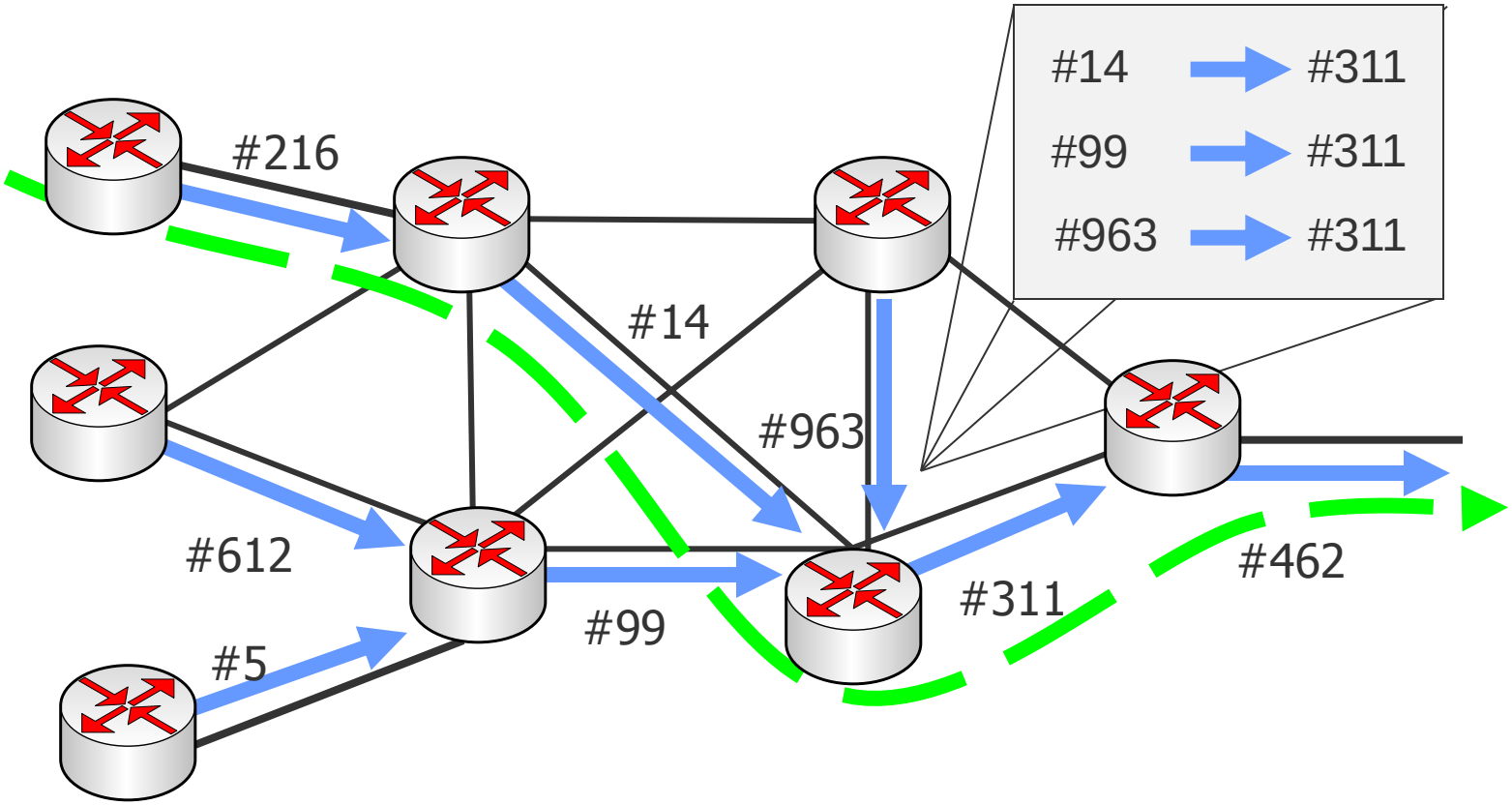


MPLS: Components

- Components of MPLS
  - FEC: Forward Equivalence Class
  - LSR: Label Switching Router
  - LER: Label Edge Router
  - LSP: Label Switched Path
  - LDP: Label Distribution Protocol



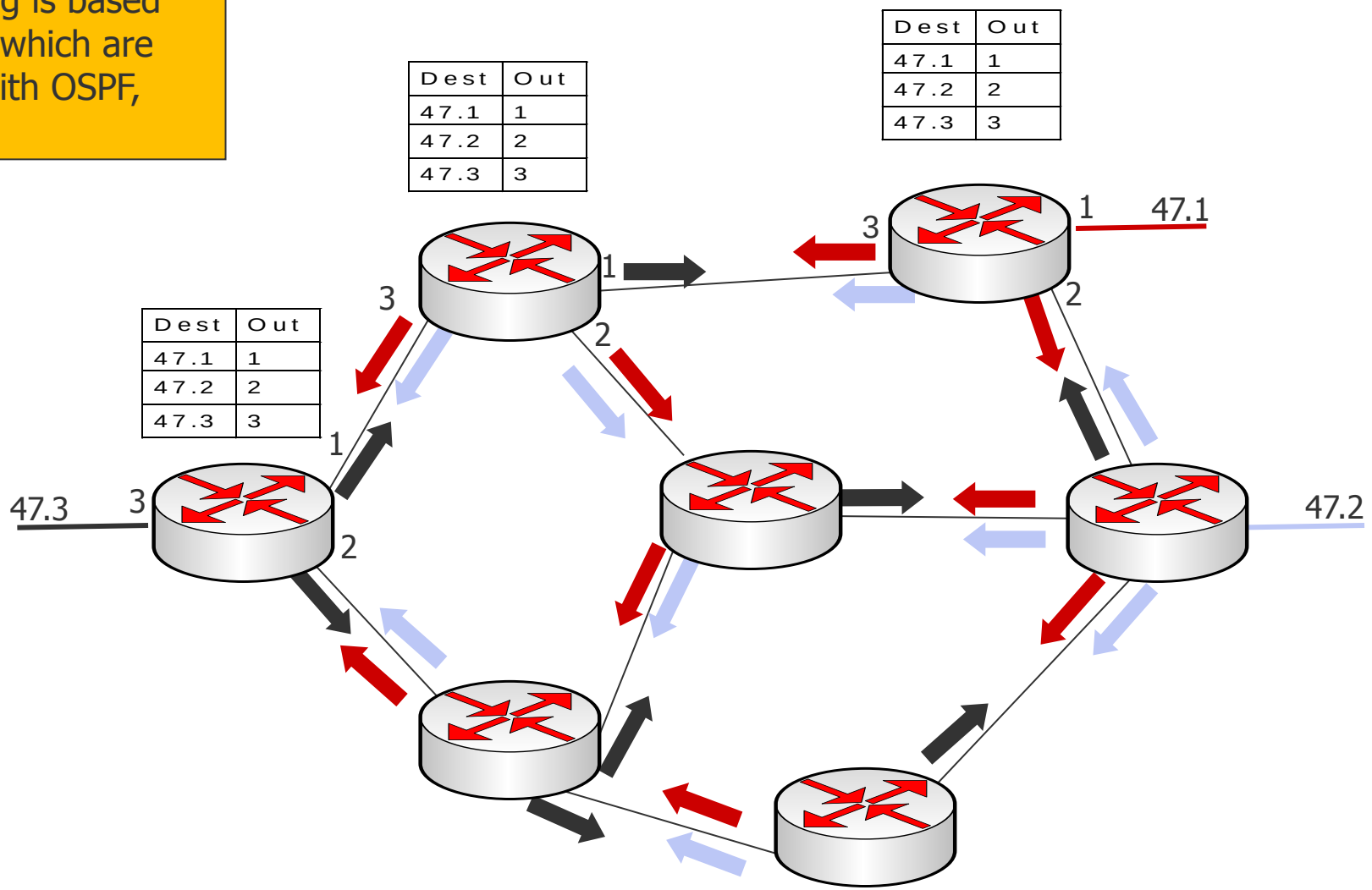
Label Switched Path (simple Variant)



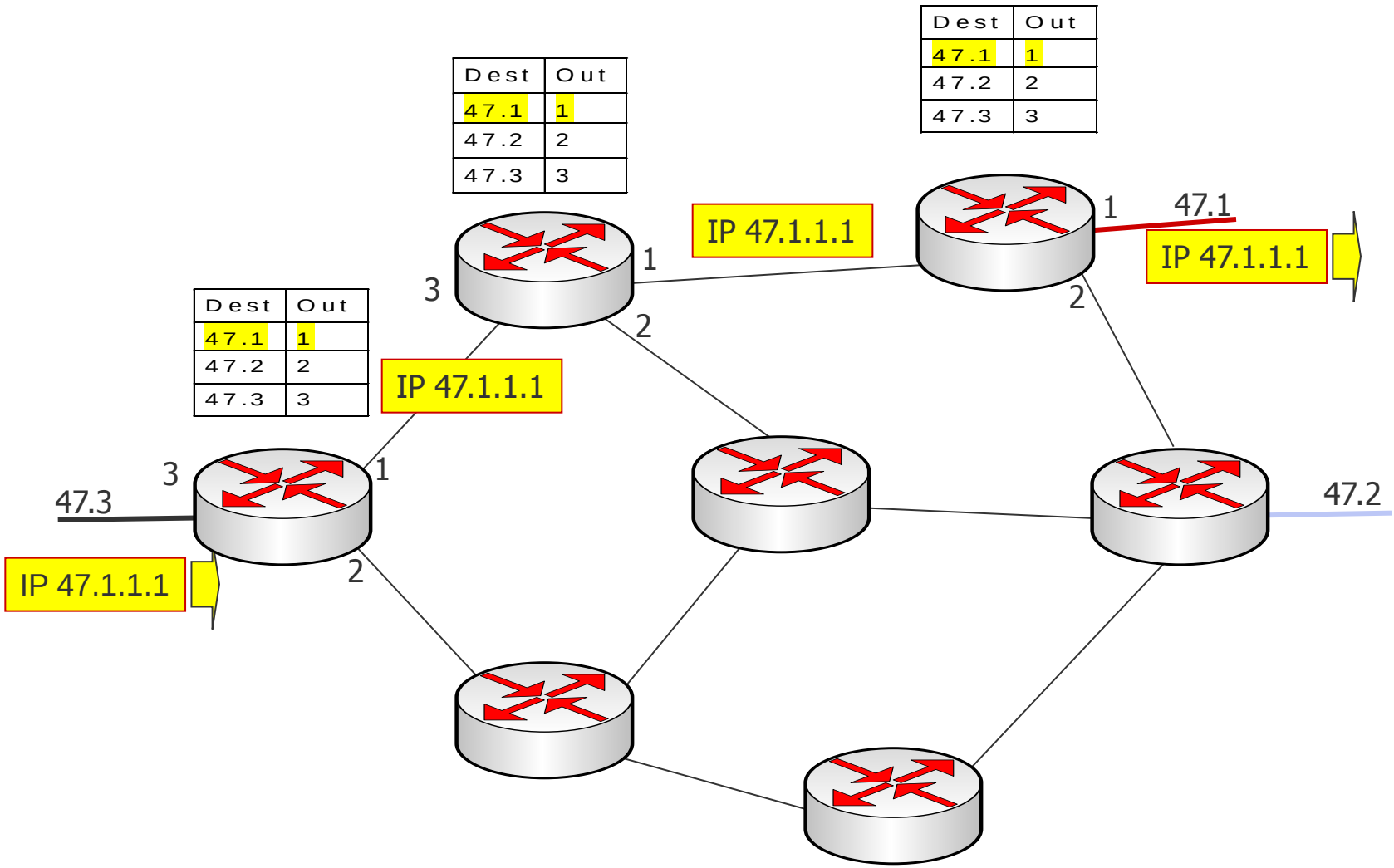
- An LSP is a part of a tree from the source to the destination
- An LDP generates that tree with the aid of existing IP routing tables

MPLS uses IP

Forwarding is based on tables which are created with OSPF, RIP, …



IP-Forwarding is done “hop-by-hop”

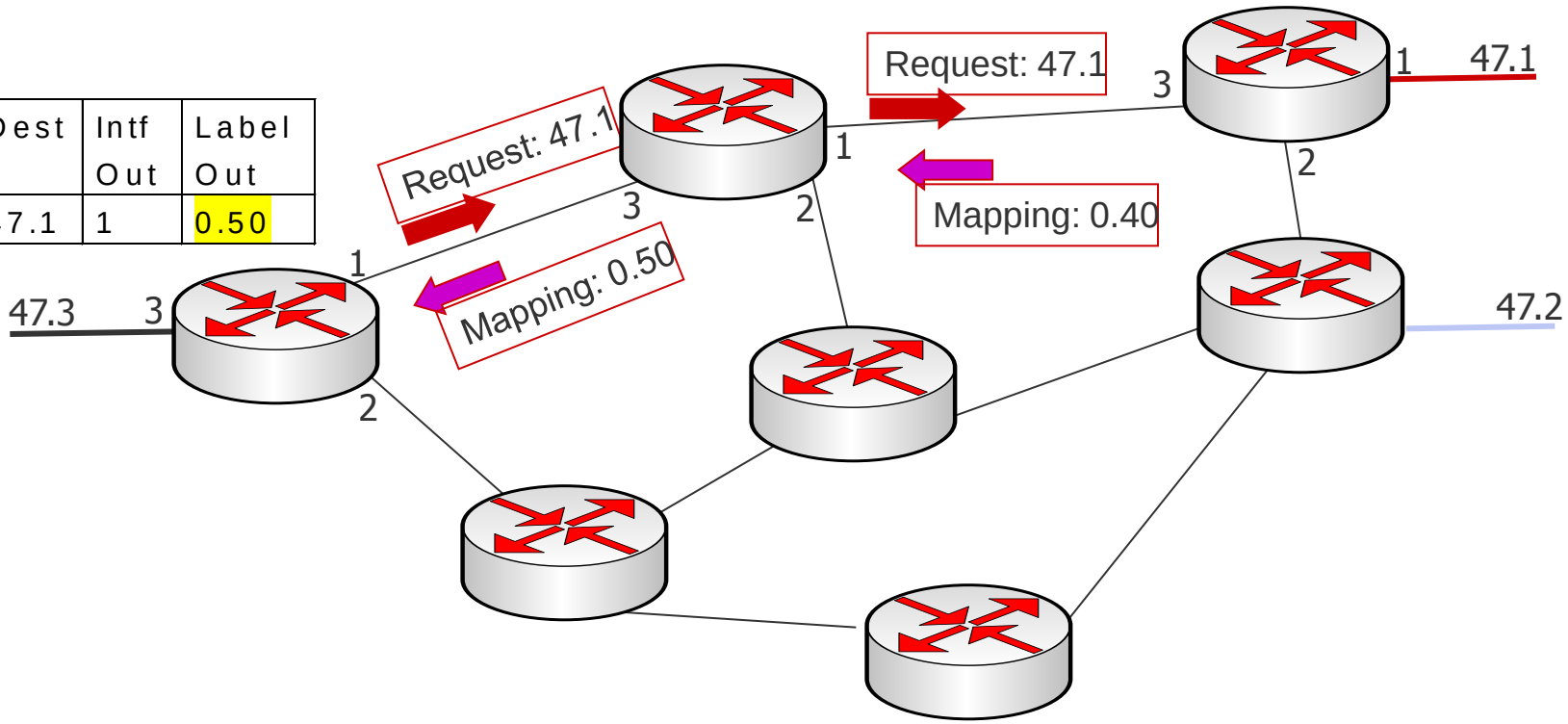


MPLS-Label-Distribution (simple Variant)

Intf In	Label In	Dest	Intf Out	Label Out
3	0.50	47.1	1	0.40

Intf In	Label In	Dest	Intf Out
3	0.40	47.1	1

Intf In	Dest	Intf Out	Label Out
3	47.1	1	0.50

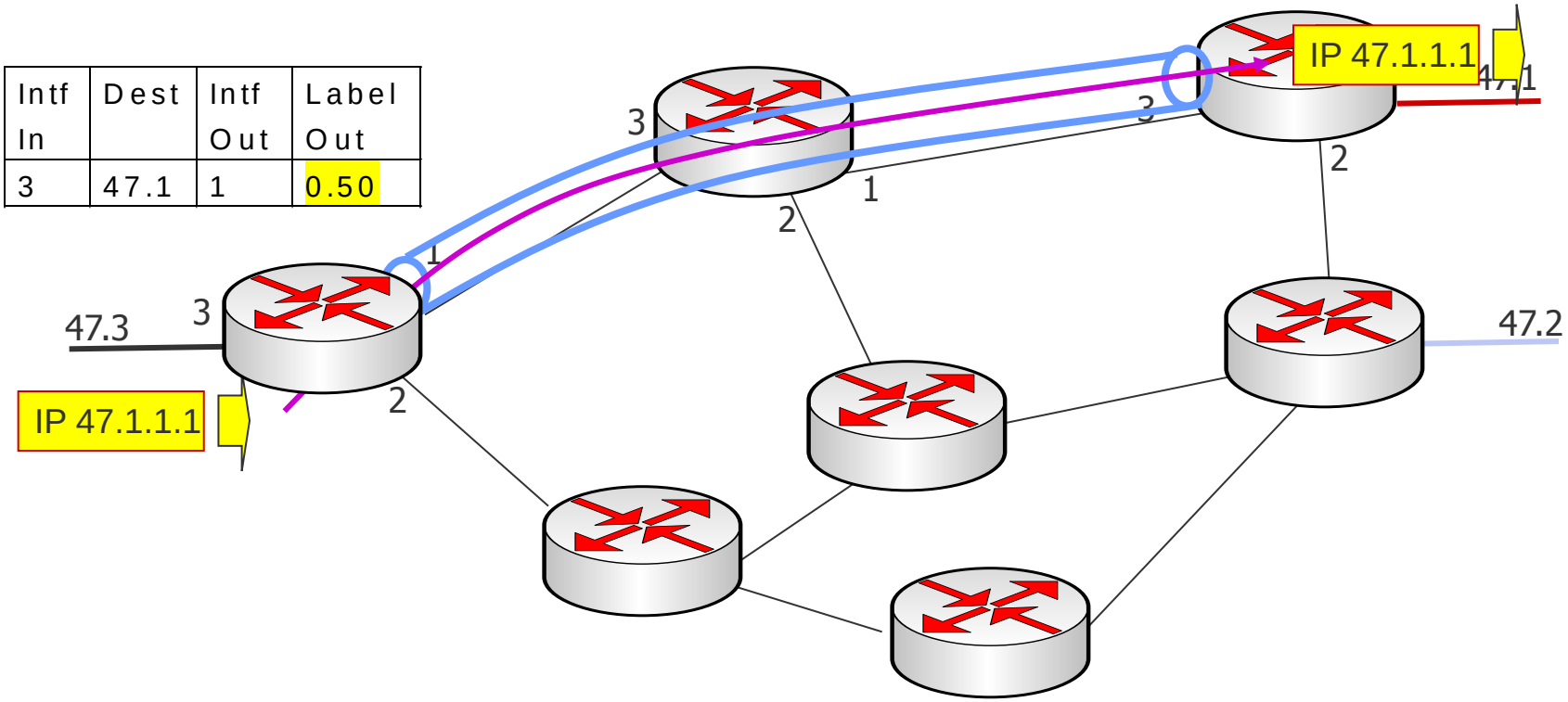


Label Switched Path (LSP)

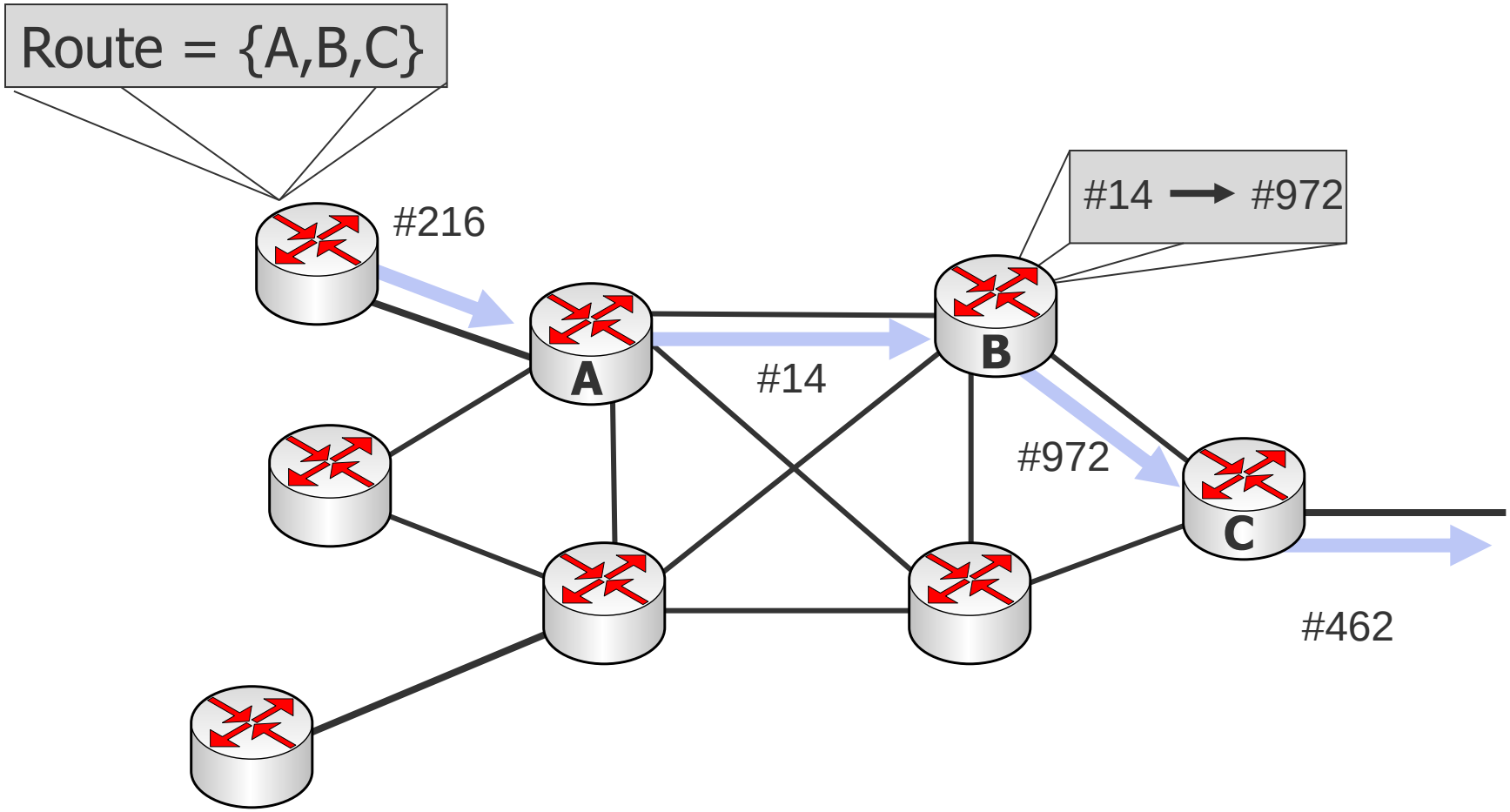
Intf In	Label In	Dest	Intf Out	Label Out
3	0.50	47.1	1	0.40

Intf In	Label In	Dest	Intf Out
3	0.40	47.1	1

Intf In	Dest	Intf Out	Label Out
3	47.1	1	0.50



Explicitly Routed LSP (ER-LSP)



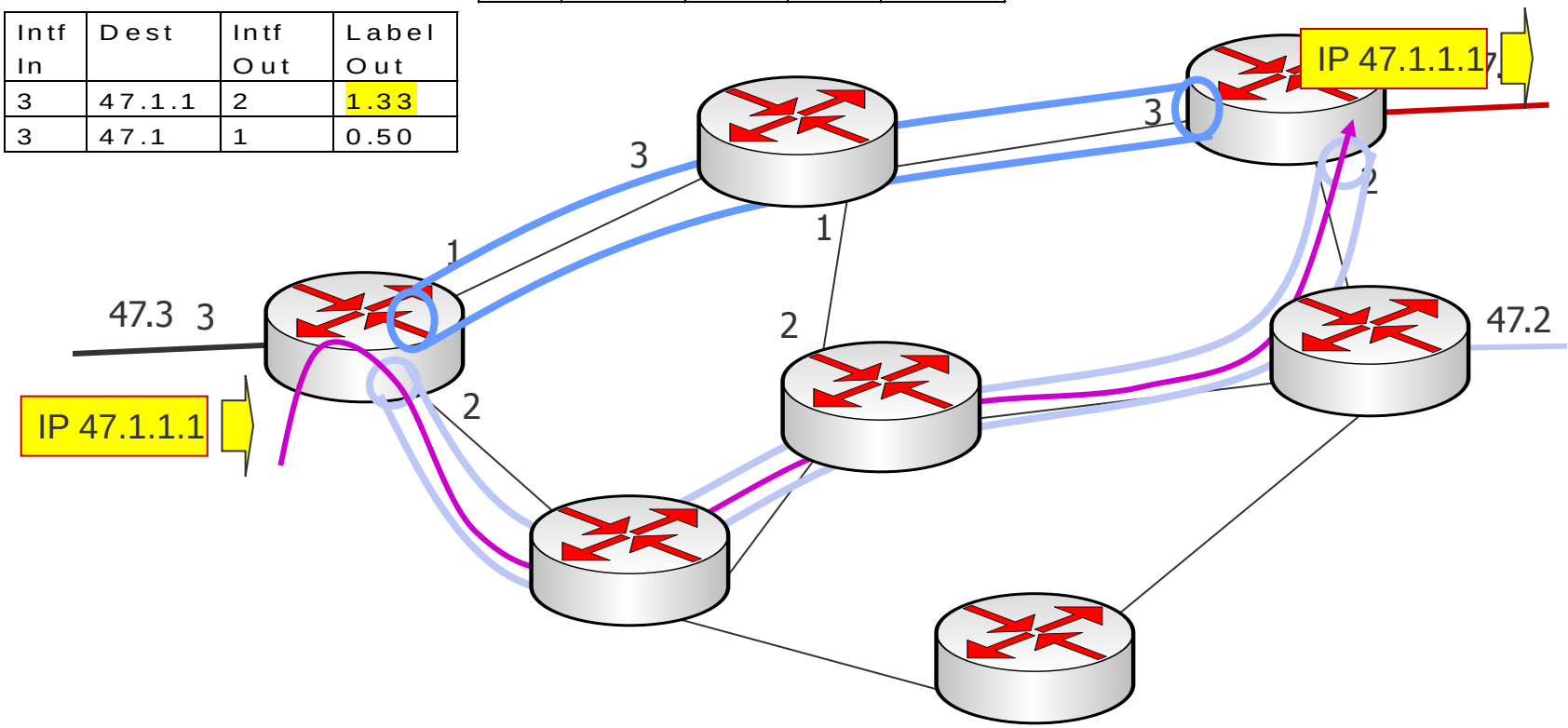
A ER-LSP is selected by the source. A label request is send from the source (source routing).

Explicitly Routed LSP (ER-LSP)

Intf In	Label In	Dest	Intf Out	Label Out
3	0.50	47.1	1	0.40

Intf In	Label In	Dest	Intf Out
3	0.40	47.1	1

Intf In	Dest	Intf Out	Label Out
3	47.1.1	2	1.33
3	47.1	1	0.50



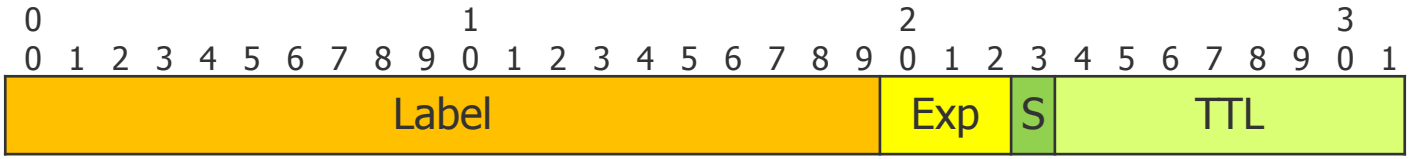
Advantages of ER-LSP

- The network provider controls the route selection
  - Based on individual decisions
    - QoS, Costs, Load, Policy, ...
  - Thus, not only shortest path routes are taken
  - Support of Traffic Engineering
- Several protocols may be used
  - CR-LDP (Constraint Routing): LDP + ER-Extensions
  - RSVP ext: RSVP + Extensions for scalability + ER-Extensions
  - ...

MPLS Labels

MPLS Label

- MPLS labels can be contained in different headers
  - In the Layer 2 protocol, e.g., ATM and Frame Relay
  - In a **shim header**, i.e., between Layer 2 and Layer 3
  - In the Layer 2 protocol as well as in the shim header
- Labels may have various formats
  - Negotiated by the peers
  - Depends on the Layer 2 protocol
  - Specified for Frame Relay, ATM, PPP,…
- Labels can be stacked
  - Only the top label is considered
  - Bottom label is marked in particular way
- General format: 4 bytes
  - 20 bit label
  - 3 bit experimental
    - Used for Quality of Service (QoS)
  - 1 bit bottom label
  - 8 bit TTL



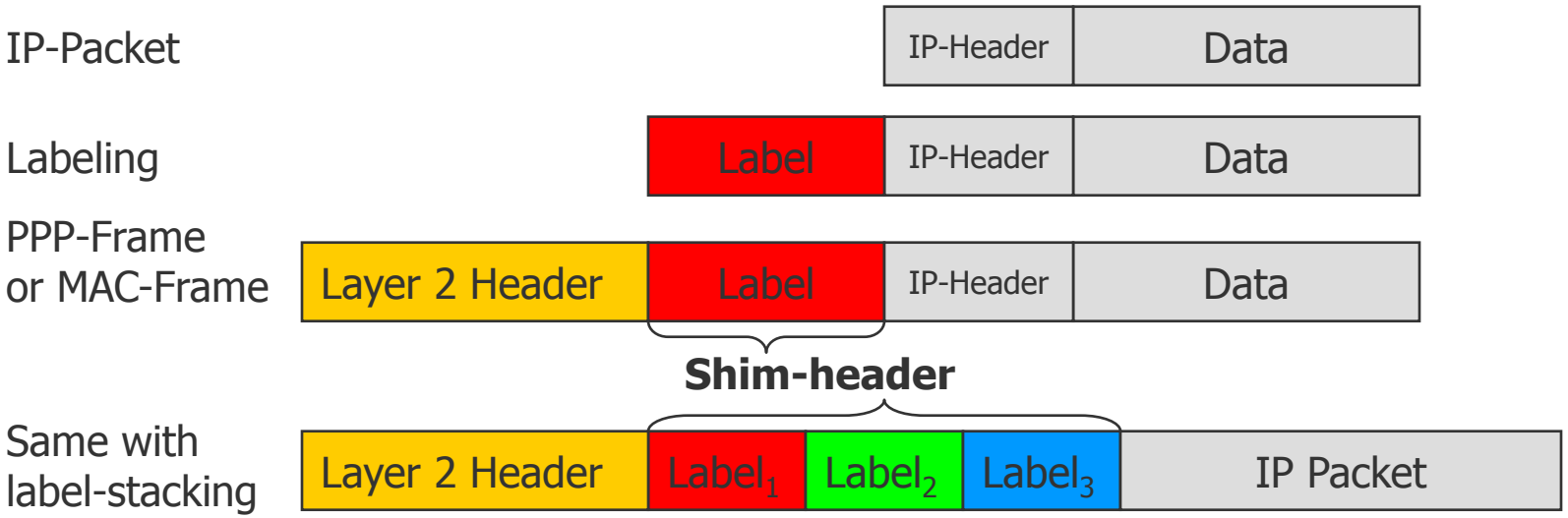
MPLS Label

- Label stacking

Label	Exp	0	TTL
Label	Exp	0	TTL
...			
Label		1	TTL

MPLS Label with PPP/LAN-MAC

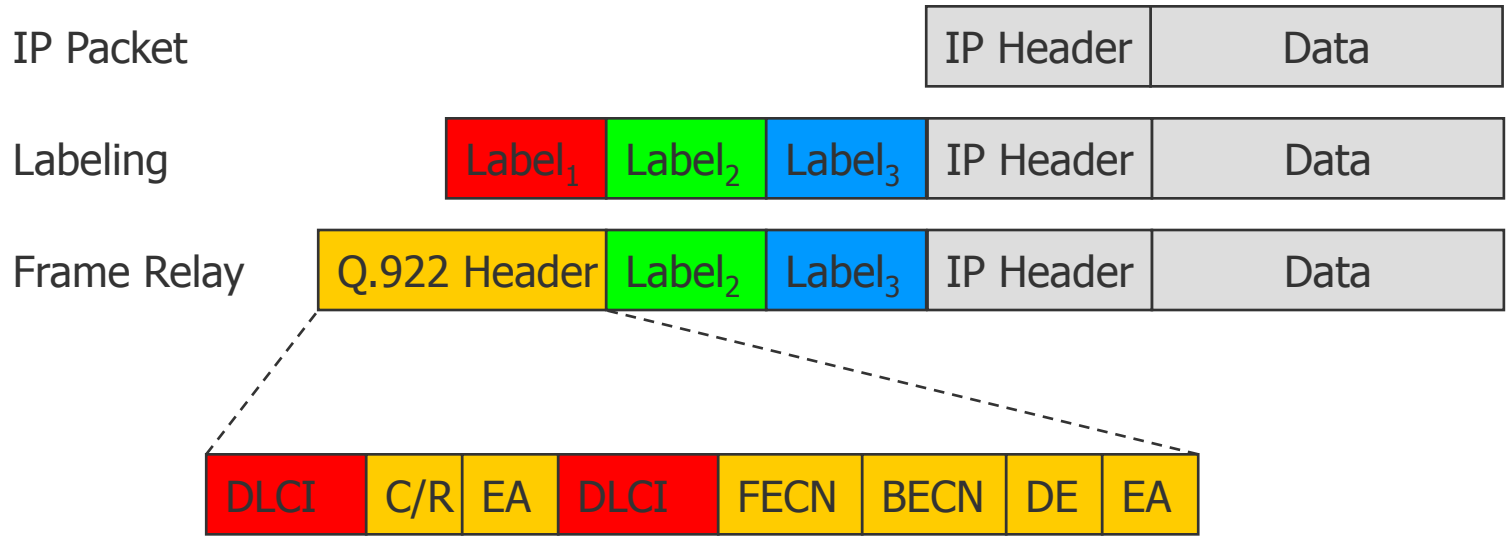
- Layer 3 protocol has to be defined at the bottom label
  - Similar to the type field in Ethernet, which may refer to IP
- Label TTL
  - At the first labeling the Label-TTL is set to IP-TTL
  - When the last label is removed the IP-TTL is set to Label-TTL
- Packet length
  - Due to multiple labeling the packet grows ➡ max. IP datagram size for labeling



MPLS Label with Frame Relay

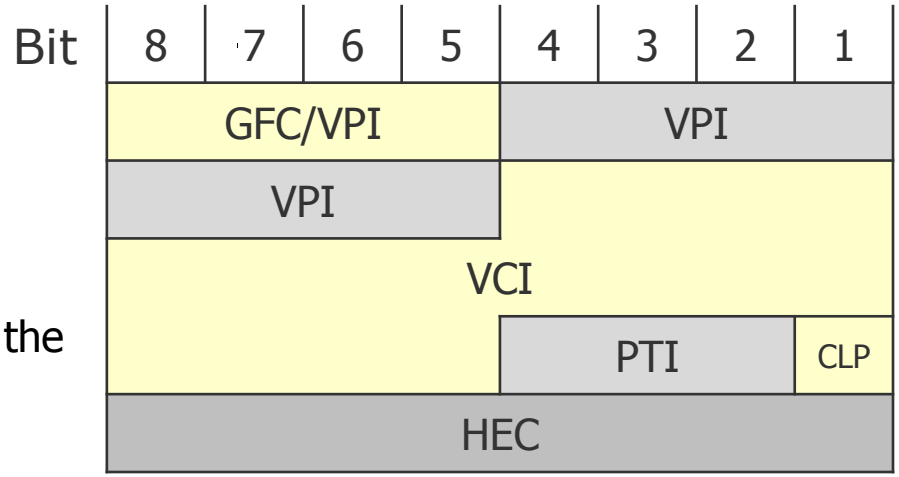
- Current Label
  - Is transported in the DLCI-Field of Frame Relay
  - May use 2 or 4 byte addresses according to Q.922
  - Other labels may be contained in the PPP/LAN format

- Example



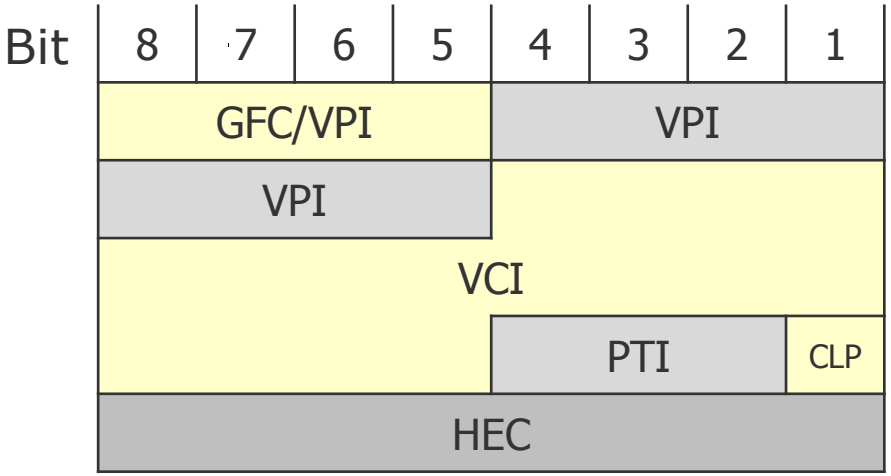
MPLS Label with ATM

- Two ATM header formats:
  - Communication between switches and endpoints: User-Network Interface (UNI)
  - Communication between two switches: Network-Network Interface (NNI)
- Header Fields
  - Generic Flow Control (GFC)
    - Only with UNI, for local control of the transmission of data into the network. Typically unused.
    - With NNI these bits are used to increase the VPI field.
  - Payload Type Identifier (PTI)
    - Describes content of the data part, e.g. user data or different control data
  - Cell Loss Priority (CLP)
    - If the bit is 1, the cell can be discarded within overload situations.
  - Header Error Control (HEC)
    - CRC for the first 4 bytes; single bit errors can be corrected.



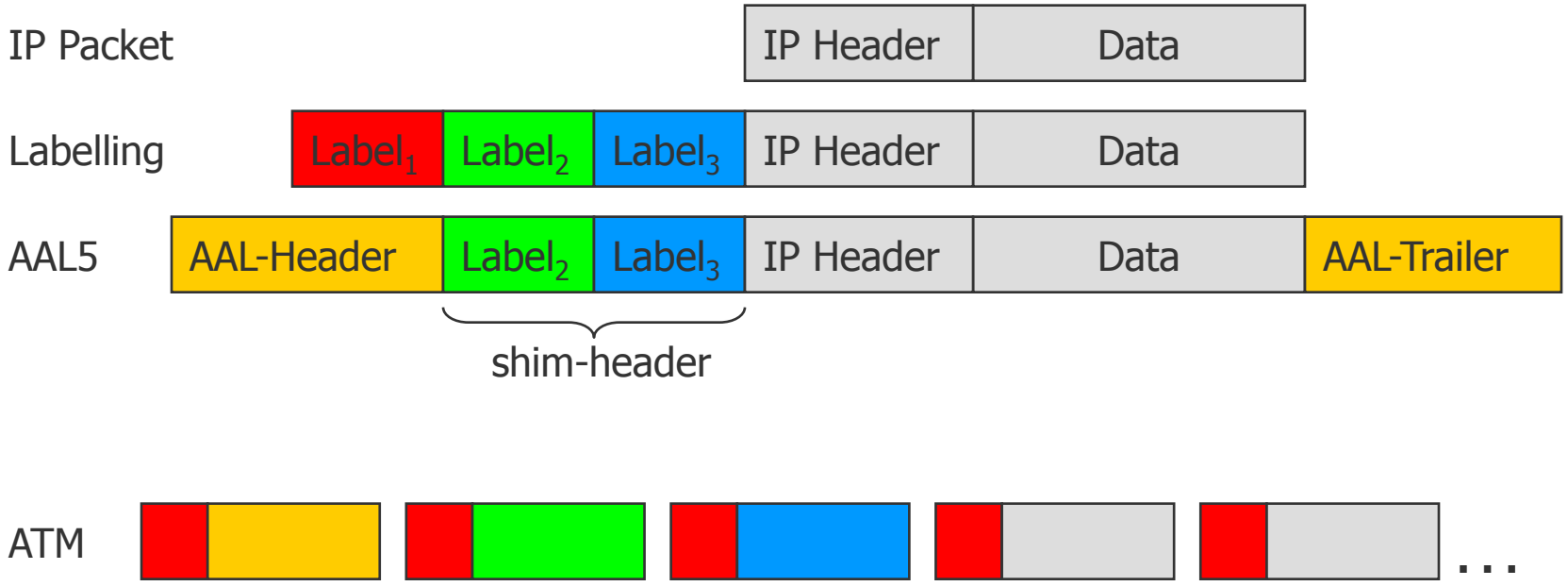
MPLS Label with ATM

- Option 1: SVC Encoding (Switched Virtual Circuits, SVC)
  - (Top) Label is stored in VPI/VCI field
    - Works with all networks
    - Only one label in the ATM header
- Option 2: SVP Encoding (Switched Virtual Paths, SVP)
  - Top label in VPI field, second label in VCI field
    - Supporting of "VP switching"
- Option 3: SVP Multipoint Encoding
  - Top label in VPI field, second label as part of the VCI field. Remainder of the VCI field represents the LSP ingress LSR (aids for label merging)



MPLS Label with ATM

- Label Stacking (multiple hierarchical labels) with ATM?
  - Other labels and fields which do not fit in VPI/VCI have to be stored in the shim header in the PPP/LAN format
- Example



MPLS Label Distribution

Label Distribution

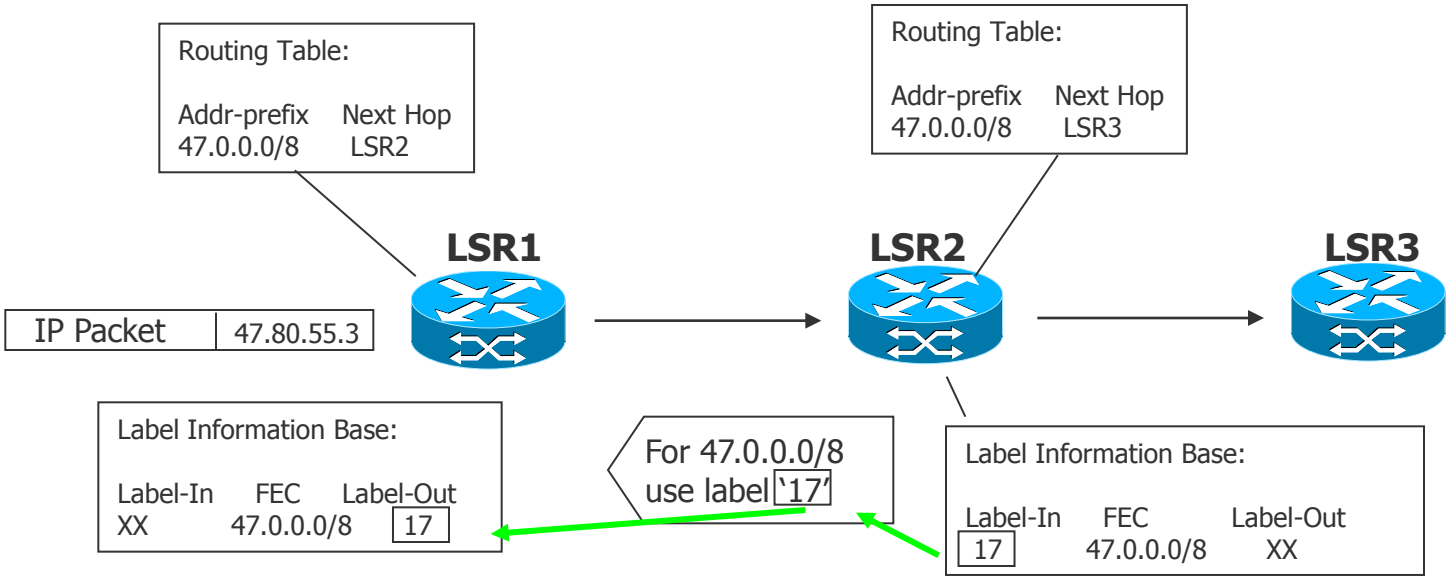
- MPLS do not specify a particular way of label distribution
  - Piggyback on existing IP routing protocol
  - Separate protocol to distribute labels

- Piggyback on existing IP routing protocol
  - Extension of protocols for particular deployment of resources
    - Constraint Route Label Distribution Protocol (CR-LDP)
    - Resource Reservation Protocol (RSVP-TE)
  - Protocol Independent Multicast (PIM)
  - BGP: Forwarding of external labels, e.g., for VPNs

- Label Distribution Protocol (LDP)
  - New protocol of the IETF developed with MPLS for the distribution and maintenance of labels

Label Distribution Protocol (LDP)

The distribution of labels guarantees that neighbored LSRs have the same set of FECs



Step 3: LSR1 inserts the label into its forwarding table

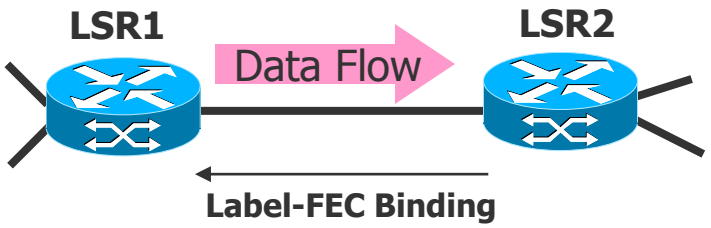
Step 2: LSR2 sends the mapping to neighbored LSRs

Step 1: LSR2 creates a mapping from FEC to a Label

The label distribution is done either via "piggybacking" on existing routing protocols or with a particular **Label Distribution Protocol** (**LDP**)

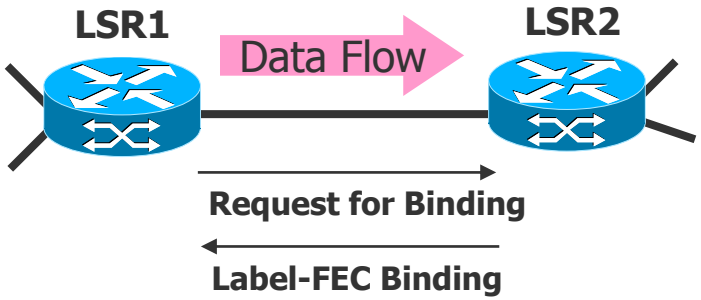
Alternatives for Label Distribution

- Downstream "unsolicited" Label Distribution



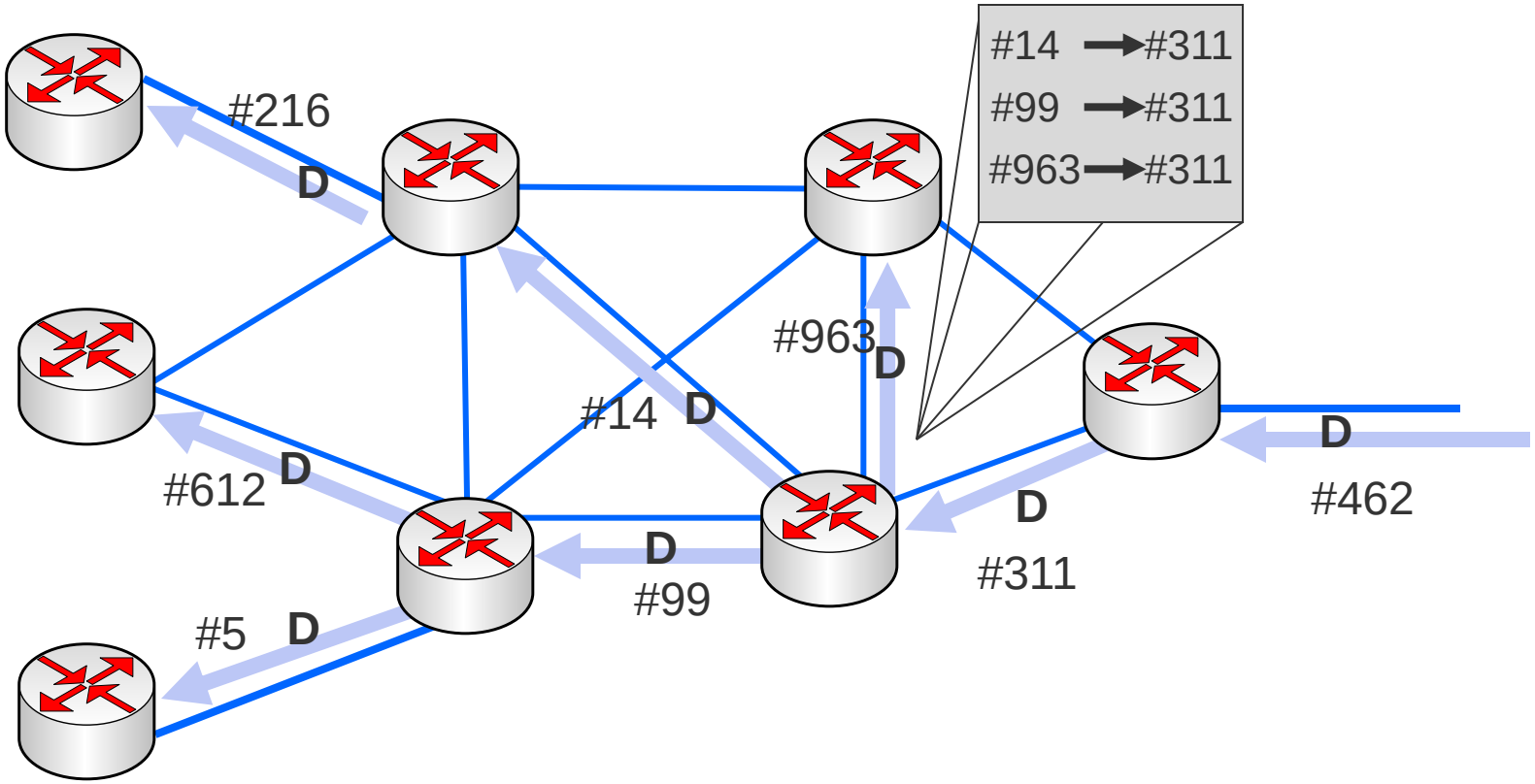
  - LSR2 is the "downstream" LSR for LSR1
  - LSR2 detects a "next hop" for a FEC
  - LSR2 creates a label for the FEC and sends the mapping to LSR1
  - LSR1 inserts the mapping into its table
  - Is LSR2 the "next hop" for a FEC, then LSR1 can use the label

- Downstream-on-Demand Label Distribution

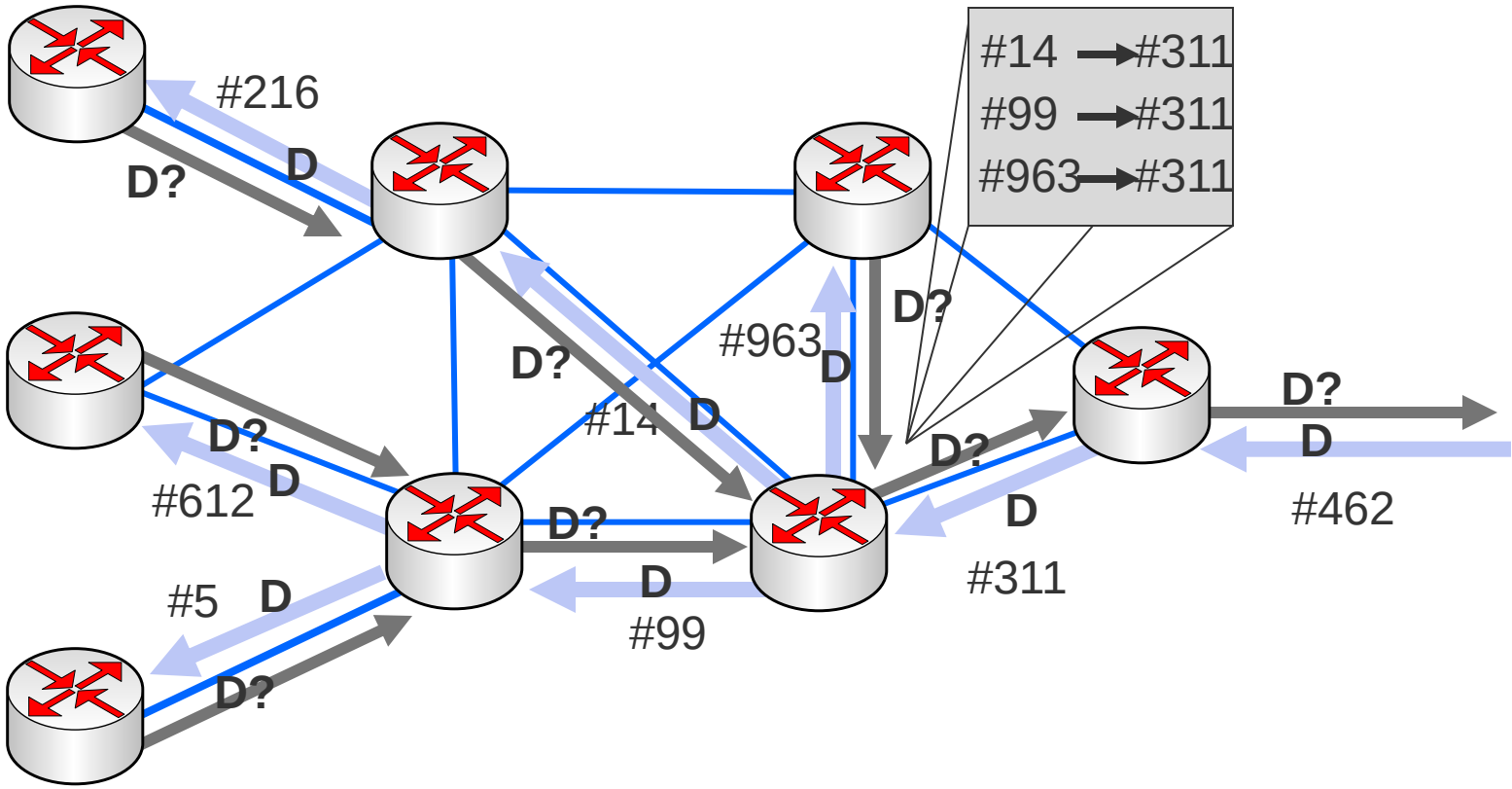


  - LSR1 detects LSR2 as one "next hop" for a FEC
  - Request to LSR2 for a mapping from FEC to a label
  - When LSR 2 recognizes the FEC and has a "next hop", then it creates a mapping and sends it to LSR1
  - Both LSRs have the same mapping for the FEC

Downstream-Mode generates a Shortest-Path-First Tree



In Analogous with Downstream-on-Demand



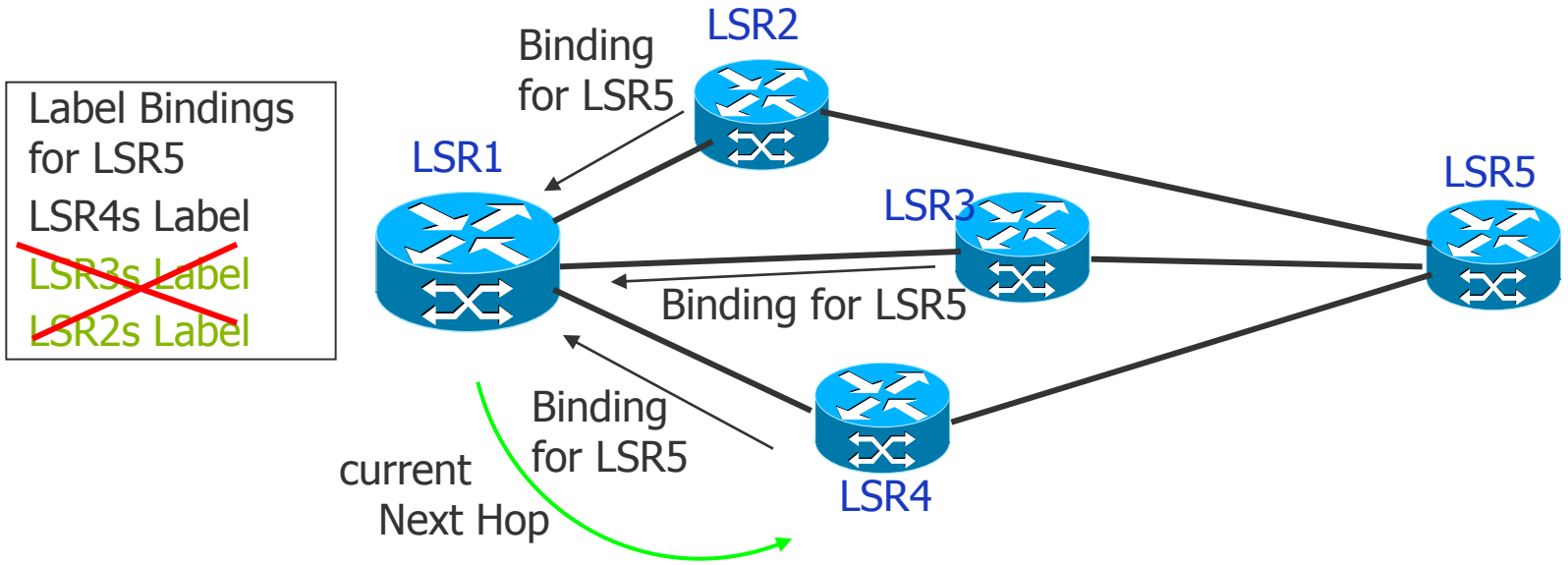
Control of Label Distribution

- Independent LSP-Control
  - Each LSR creates independent mappings of FEC-to-Label
  - Forwarding as soon as "next hop" determined
  - LSP is created by connecting in/out label
- Advantages
  - Label-Generation and exchange faster
  - Independent of availability of Egress-LSRs
- Disadvantages
  - Inconsistent availability of labels
  - Loop prevention required

- Ordered LSP-Control
  - One router is responsible for distributing labels
    - Typically the egress LSR
- Advantages
  - Consistent and loop free
  - Deployment for explicit route selection and multicast
- Disadvantages
  - Higher delay for LSP creation
  - Depends on egress LSR (bottleneck)

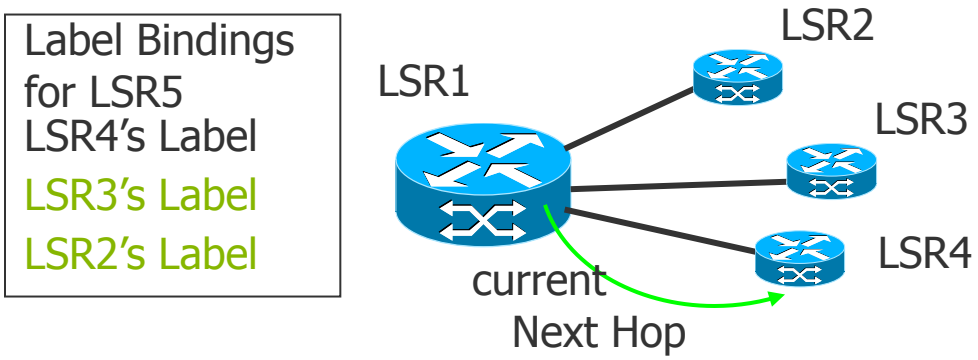
Retention of Labels

An LSR can receive mappings from several LSRs



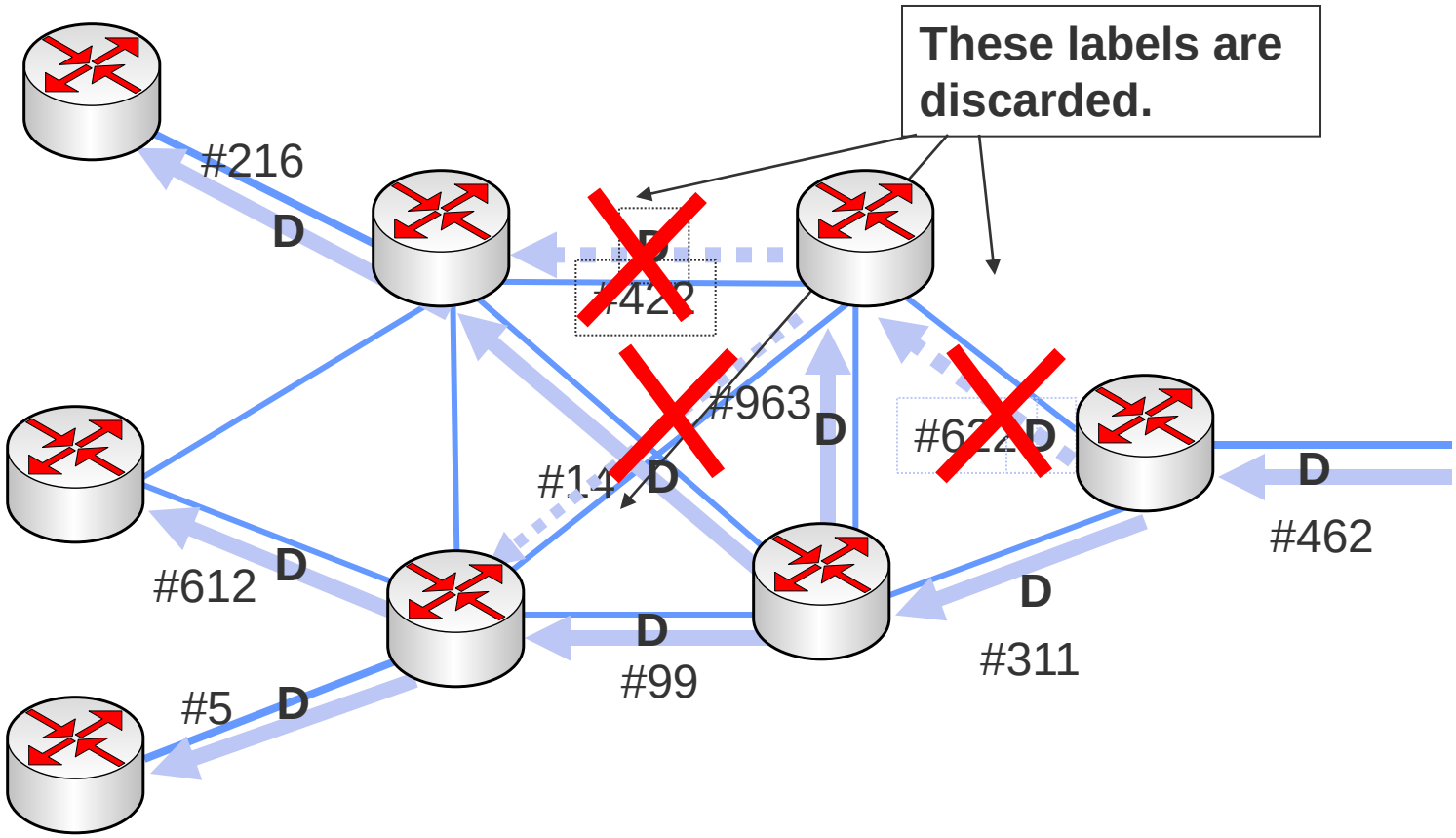
- Functions
  - An LSR keeps only the mapping which it receives from the valid "next hop"
  - When the "next hop" changes, then a new mapping from the new "next hop" has to be requested
  - Restricted adaption to changes
  - An LSR has to maintain less mappings

Liberal Retention of Labels



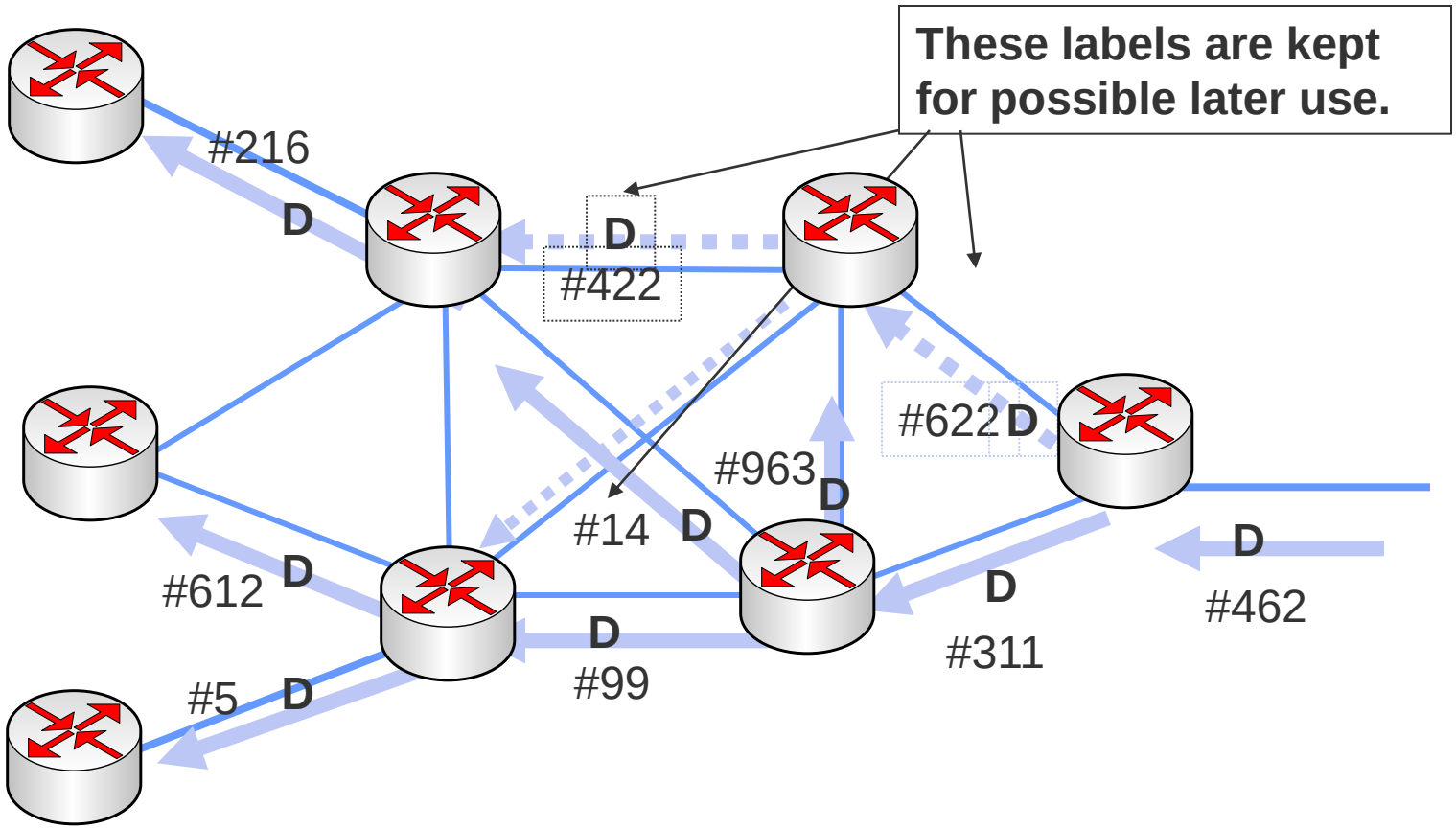
- Function
  - An LSR maintains also mappings which are received from others than the "next hop" neighbor
  - When the "next hop" changes, the stored mappings can be used directly
    - Fast adaptation to changes
  - Requires the maintenance of many labels
- The selection of the retention mode is a trade-off between adaptability and memory.

Conservative Retention of Labels



Keep only the labels which are used for forwarding and discard others.

Liberal Retention of Labels



Summary

- Key advantages
  - Leverage growth of MPLS deployment in core and edge networks
  - Removes primary bottleneck of TCP termination
  - Realization in standard off-the-shelf switch hardware
  - Implements sophisticated request routing functions
- Requirements
  - Assign some request-routing functionality to proxies
  - MPLS-aware proxies at the network edges
  - Implementation of control protocol for label distribution