Advanced Computer Networks

MPLS

Prof. Andrzej Duda

duda@imag.fr

`http://duda.imag.fr`

Contents

- MPLS (Multi Protocol Label Switching)
  - Label swapping
  - Elements of MPLS
  - Label switching
  - Label distribution
  - Interaction with IGP
  - Traffic engineering

MPLS

- **IGP limits**
  - IP routing may ignore the real physical topology
  - OSPF algorithms send traffic on a shared path and may ignore unloaded links
    - even if load balancing can be done in some cases (Equal Cost)

- **Goals of MPLS**
  - increase forwarding performance
  - provide more flexibility than IGP routing
    - explicit routing, QoS routing
    - backup routes, load balancing, VPN
  - multiprotocol - a unifying view at 2.5 layer - a unified way of controlling the underlying Layer 2 network
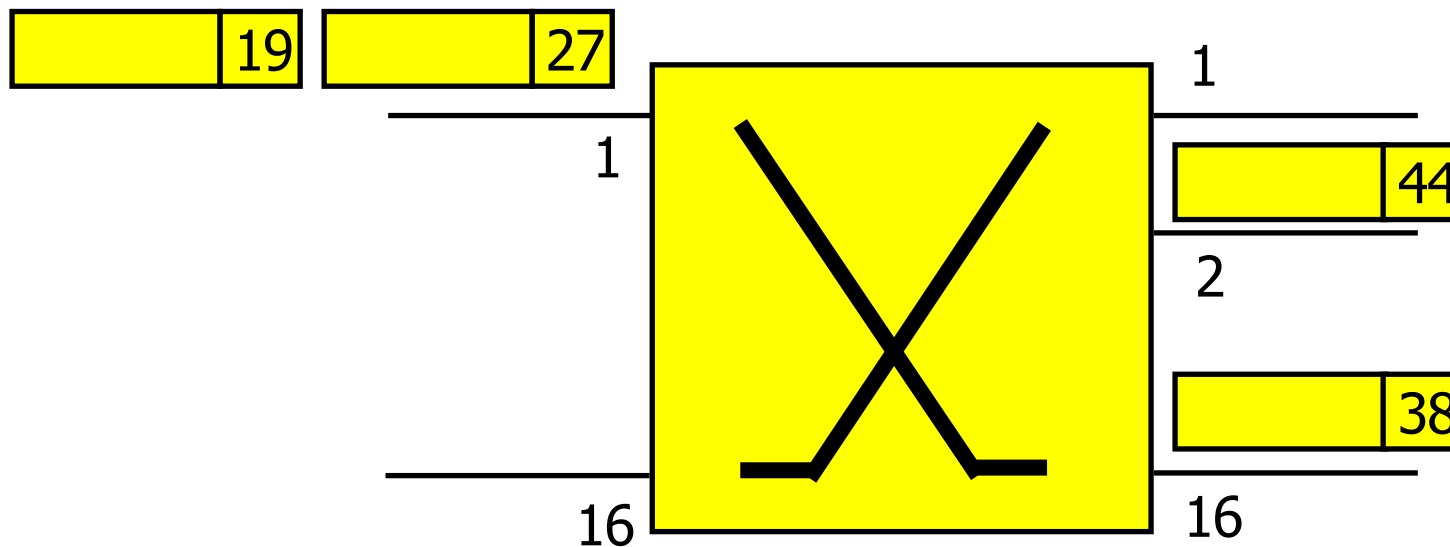   - Ethernet, PPP, SDH/DWDM

Virtual Circuits
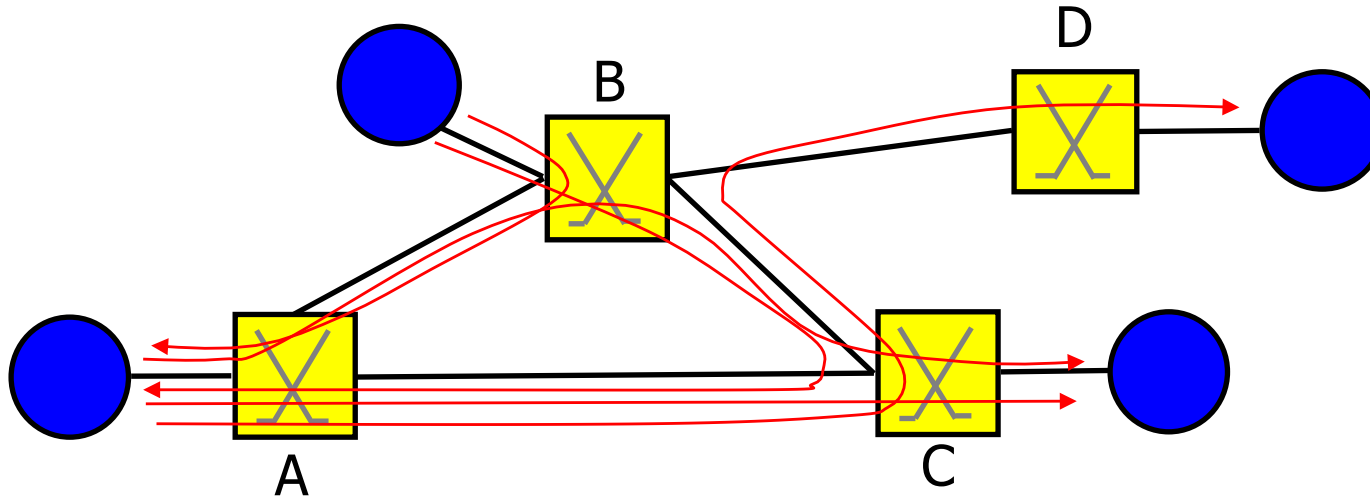# ATM VPI/VCI switching

in	VPI/VCI	out	VPI/VCI
1	27	2	44
1	19	16	38

ATM cells
header contains VPI/VCI
(Virtual Path/Virtual Circuit)
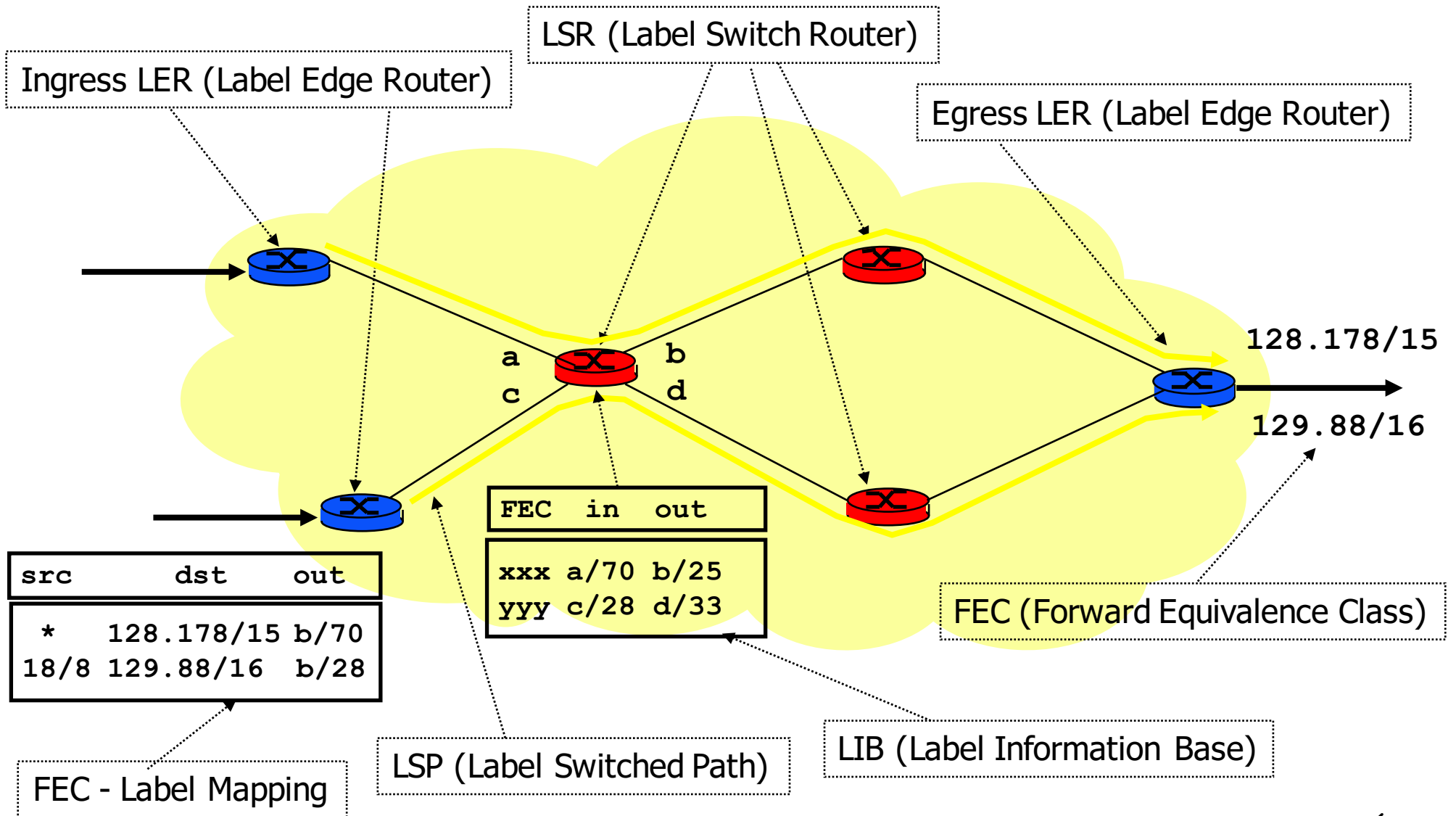
Label swapping

VC table at A:

in	VPI/VCI	out	VPI/VCI
H	0	B	0
H	1	C	0
C	1	H	2
B	1	H	3
H	4	C	2

Virtual circuits opened in the following order:

ABC

AC

BCA

BA

ACBD

MPLS elements

LSR (Label Switch Router)

Ingress LER (Label Edge Router)

Egress LER (Label Edge Router)

128.178/15

129.88/16

a b
c d

FEC	in	out
xxx	a/70	b/25
yyy	c/28	d/33

src	dst	out
*	128.178/15	b/70
18/8	129.88/16	b/28

FEC (Forward Equivalence Class)

LIB (Label Information Base)

LSP (Label Switched Path)

FEC - Label Mapping

6

MPLS Principles

- Labels

  - ingress LER classifies packets to identify FEC that determines a label; inserts the label (32 bits)

  - LSR switches based on the label

  - label swapping - label has meaning local to one LSR (requires label distribution protocol)

  - egress LER removes the label

- LSR

  - contains LIB - switching table that determines the path in the network (LSP)

  - LSP - similar to a ATM/FR virtual circuit

- Change of the forwarding paradigm

  - instead of hop by hop

  - LSP determined at entry in function of FEC, source, or other

Forwarding Equivalence Classes

- FEC - group of IP packets
  - forwarded in the same manner, over the same path, and with the same forwarding treatment

- FEC may correspond to
  - destination IP subnet
  - source and destination IP subnet
  - traffic class that LER considers significant

- For example, all traffic with a certain value of IP precedence may constitute a FEC
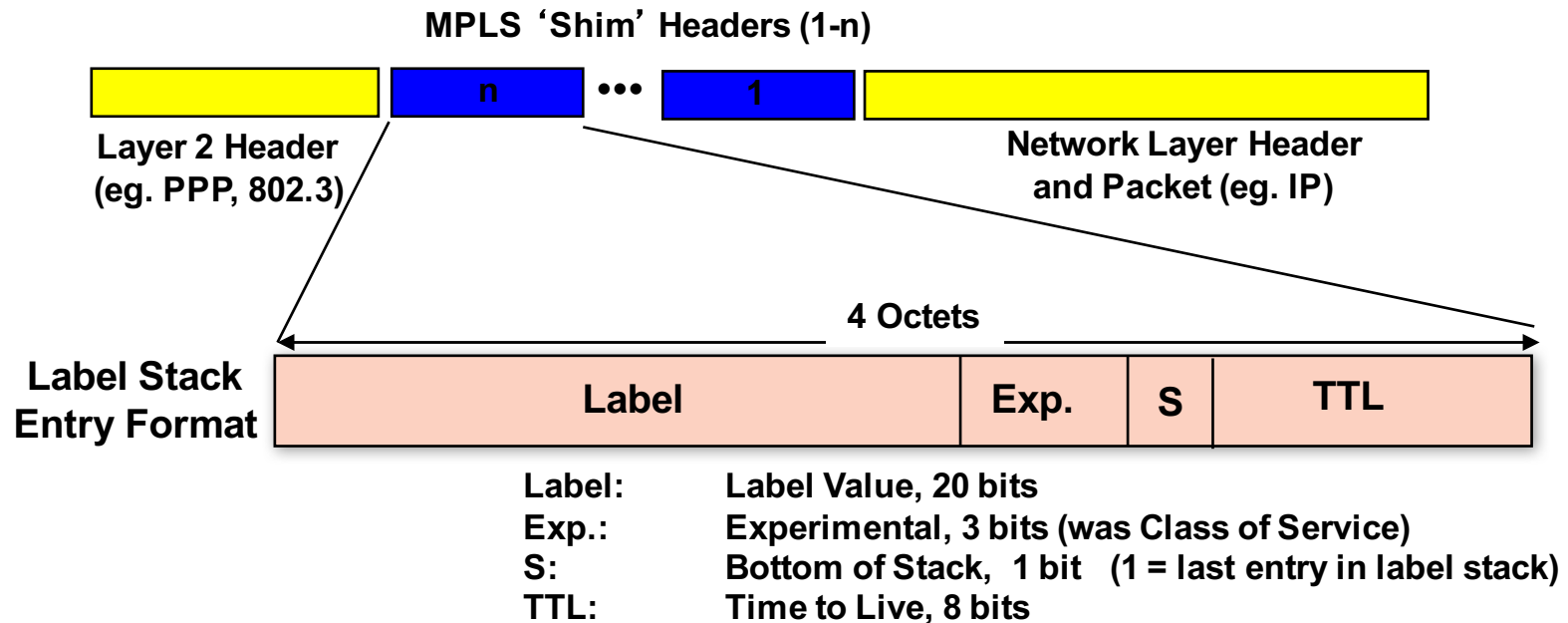
- FEC in our examples
  - IP prefix

Forwarding Equivalence Classes

Packets are destined for different address prefixes, but can be mapped to common path
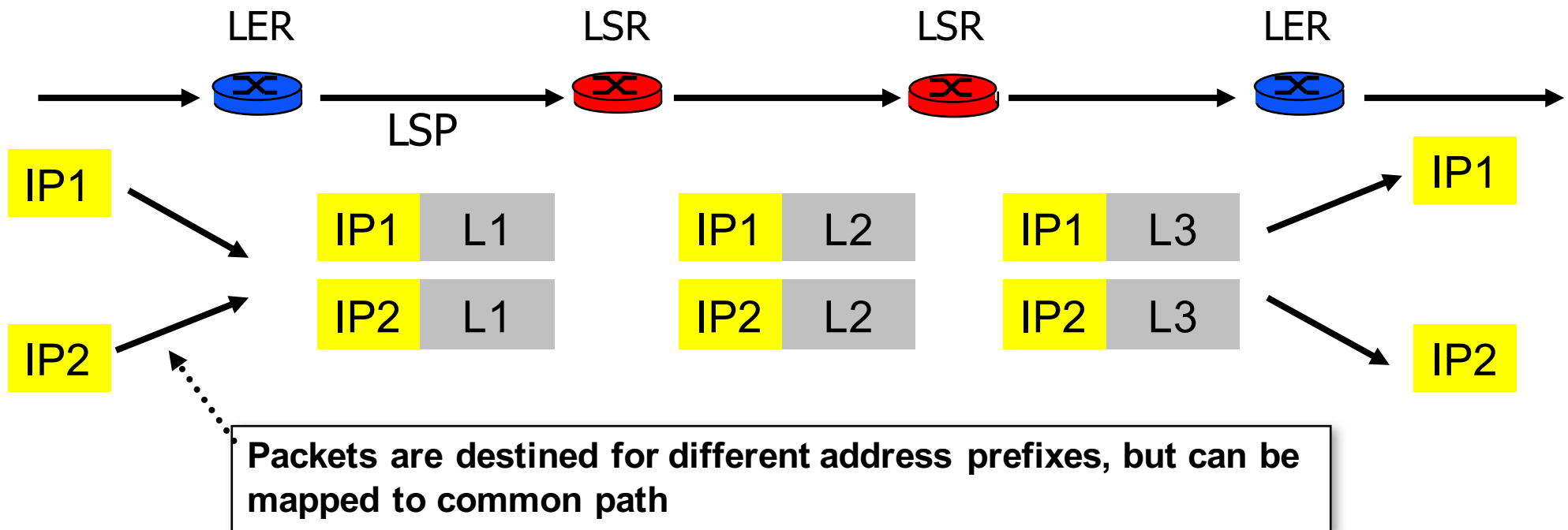
- FEC = "A subset of packets that are all treated the same way by a router"
- Conventional routing: a packet is assigned to a FEC at each hop (i.e. L3 look-up), in MPLS it is only done once at the network ingress
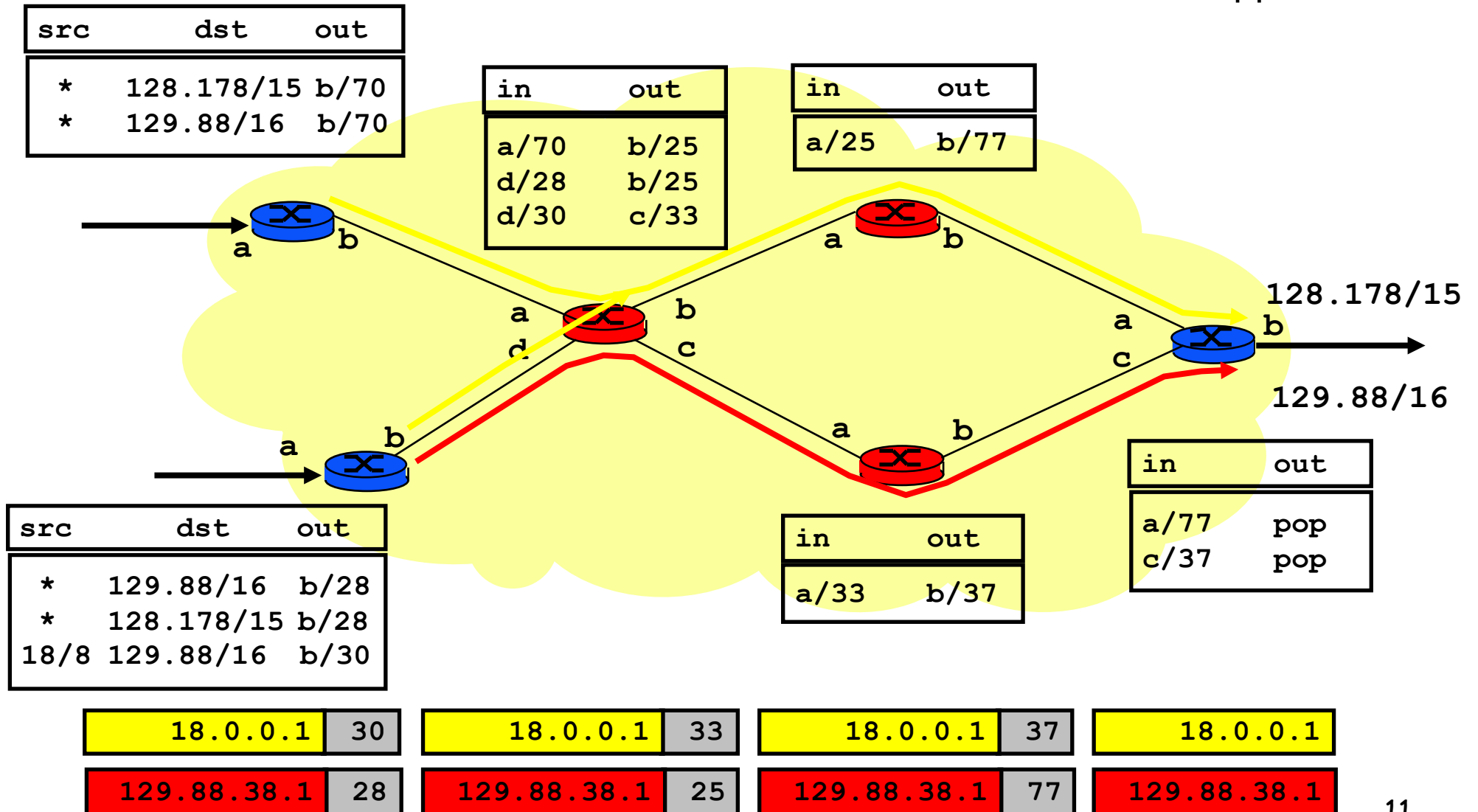
9

MPLS Encapsulation - PPP & LAN

MPLS 'Shim' Headers (1-n)

| Layer 2 Header (eg. PPP, 802.3) | n | ••• | 1 | Network Layer Header and Packet (eg. IP) |

4 Octets

Label Stack Entry Format

| Label | Exp. | S | TTL |

Label:	Label Value, 20 bits
Exp.:	Experimental, 3 bits (was Class of Service)
S:	Bottom of Stack, 1 bit (1 = last entry in label stack)
TTL:	Time to Live, 8 bits

- MPLS on PPP links and LANs uses shim header
  - inserted between layer 2 and 3 headers

10

MPLS example

FEC skipped in LIB

src	dst	out
*	128.178/15	b/70
*	129.88/16	b/70

in	out
a/70	b/25
d/28	b/25
d/30	c/33

in	out
a/25	b/77

128.178/15

129.88/16

src	dst	out
*	129.88/16	b/28
*	128.178/15	b/28
18/8	129.88/16	b/30

in	out
a/33	b/37

in	out
a/77	pop
c/37	pop

| 18.0.0.1 | 30 |

| 18.0.0.1 | 33 |

| 18.0.0.1 | 37 |

| 18.0.0.1 |

| 129.88.38.1 | 28 |

| 129.88.38.1 | 25 |

| 129.88.38.1 | 77 |

| 129.88.38.1 |

11

Label Distribution Protocol (LDP)

Label distribution ensures that adjacent routers have
a common view of FEC <-> label bindings

Routing Table:

Addr-prefix   Next Hop
47.0.0.0/8   LSR2

Routing Table:

Addr-prefix   Next Hop
47.0.0.0/8   LSR3

LSR1   **LSR2**   **LSR3**

IP Packet   47.80.55.3

Label Information Base:

Label-In   FEC   Label-Out
XX   47.0.0.0/8   17

For 47.0.0.0/8
use label 17'

Label Information Base:

Label-In   FEC   Label-Out
17   47.0.0.0/8   XX

**Step 3: LSR inserts label
value into forwarding base**

**Step 2: LSR communicates
binding to adjacent LSR**

**Step 1: LSR creates binding
between FEC and label value**

Common understanding of which FEC the label is referring to!

Label distribution can either piggyback on top of an existing routing protocol,
or a dedicated label distribution protocol (LDP) can be created.
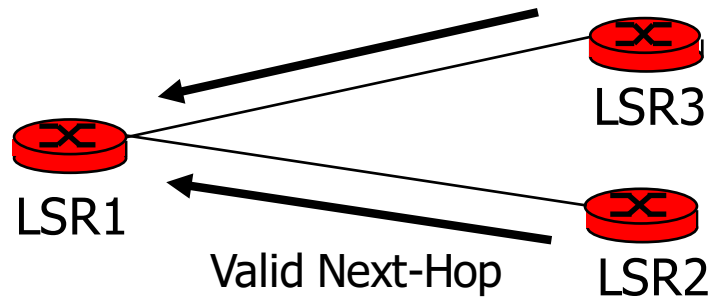
12

Label distribution

- Label distribution is always done from downstream to upstream
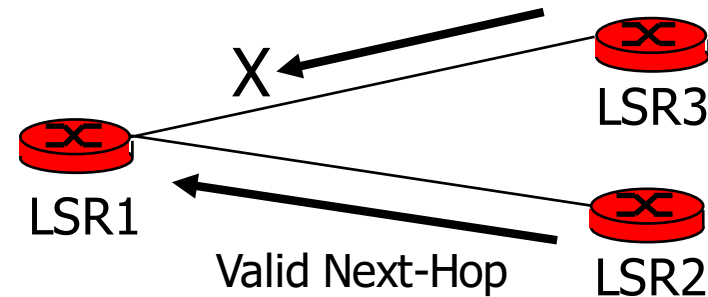   - **downstream-unsolicited**: new route => send new label

Label-FEC Binding

LSR1   LSR2
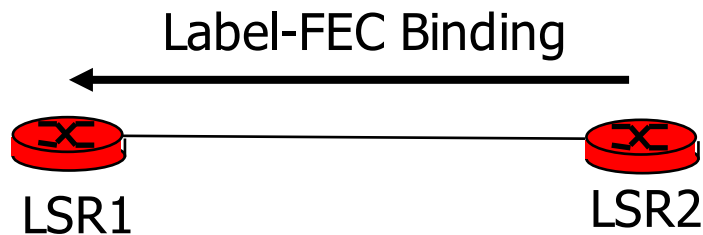
    - **downstream-on-demand**: upstream LSR asks for a label

Request for Binding

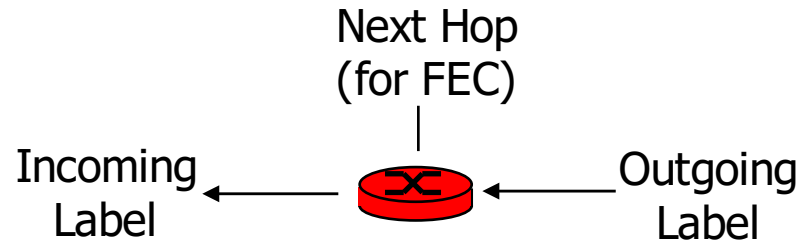Label-FEC Binding

LSR1   LSR2
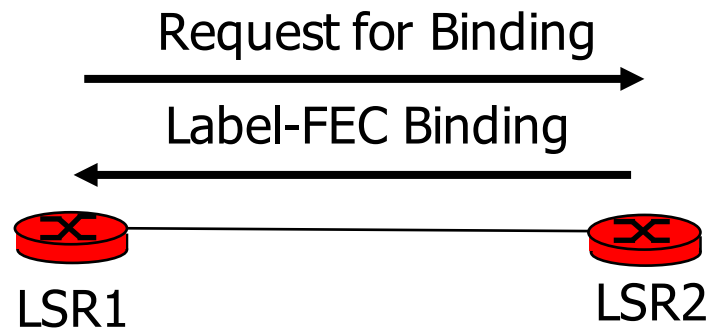
Label retention

Liberal   **Conservative**

- Label retention can be
  - **liberal**: memorize all labels from downstream LSR (faster)
  - **conservative**: memorize only selected labels (less memory)

Label control

Next Hop
(for FEC)

Incoming
Label

Outgoing
Label

- Label control can be
  - **ordered**: LSR only binds and advertise a label for a particular FEC if
    - it is the egress LSR for that FEC or
    - it has already received a label binding from its next-hop
    - LSP formation 'flows' from egress to ingress
  - **independent**: LSR binds a Label to a FEC independently, whether or not the LSR has received a label from the next-hop for the FEC
    - LSR then advertises the label to its neighbor
    - LSP is formed as incoming and outgoing labels are spliced together
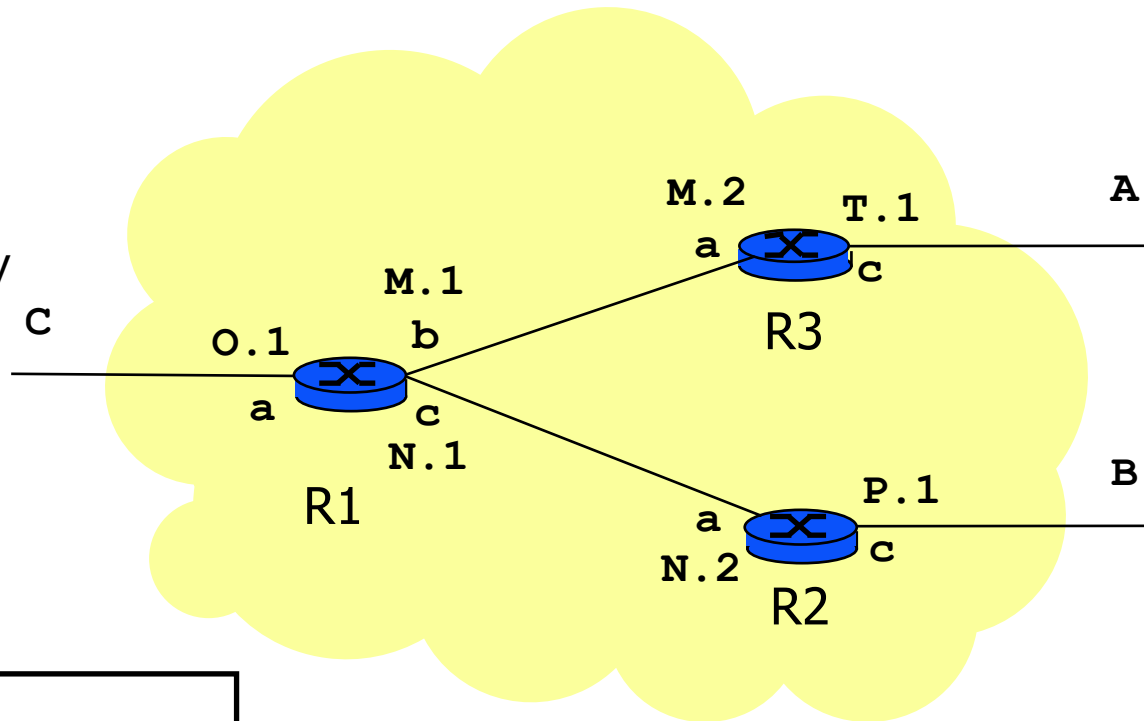
Label distribution

- Different label distribution protocols

- LDP (Label Distribution Protocol)

  - defined for MPLS

- Extension of BGP

- Extension of RSVP

  - RSVP-TE: traditional RSVP + Explicit Route

- CR-LDP (Constraint-Based LDP)

  - LDP + Explicit Route

Label distribution example - OSPF

A, B… - IP prefix
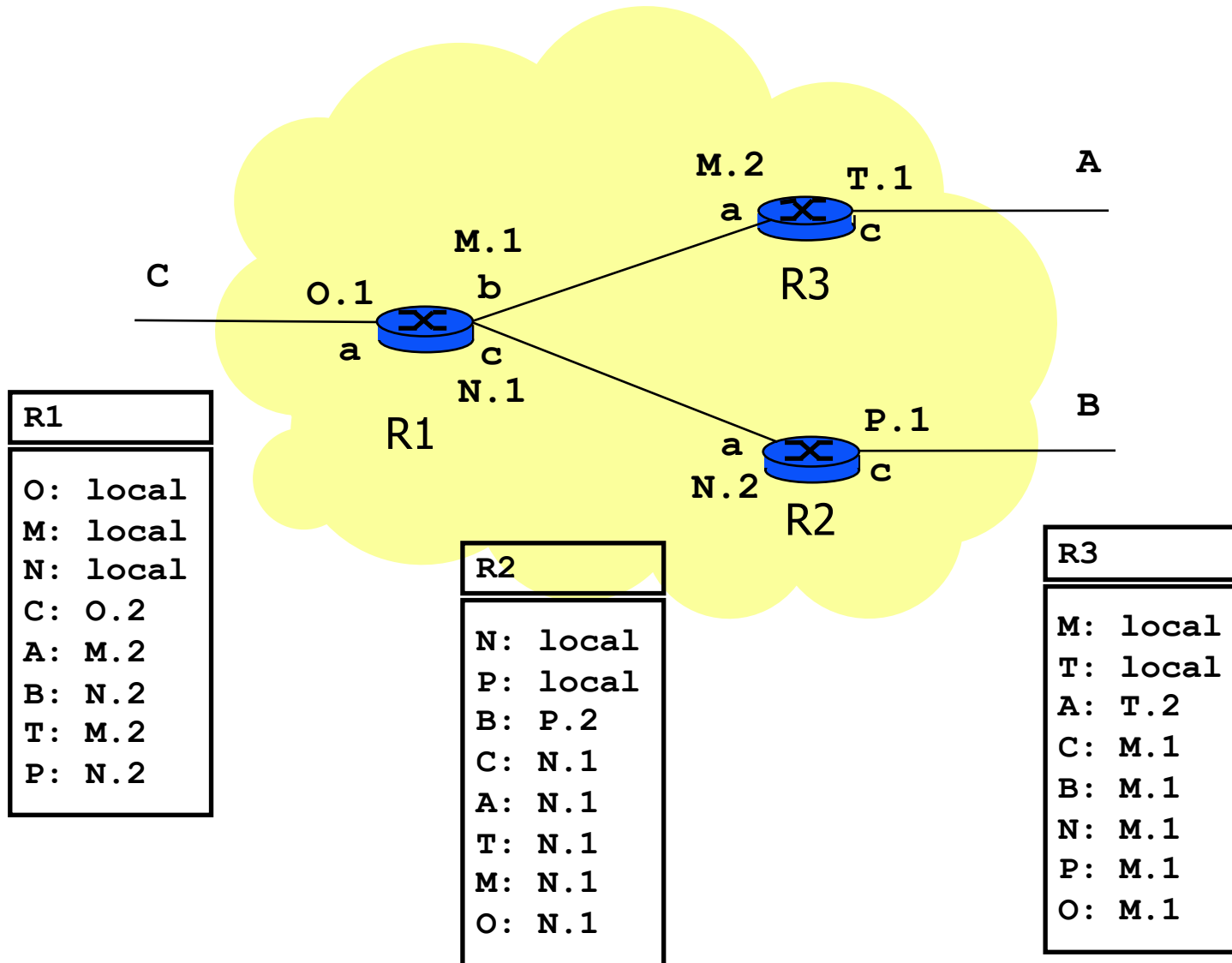M.1… - IP address
a, b… - interface
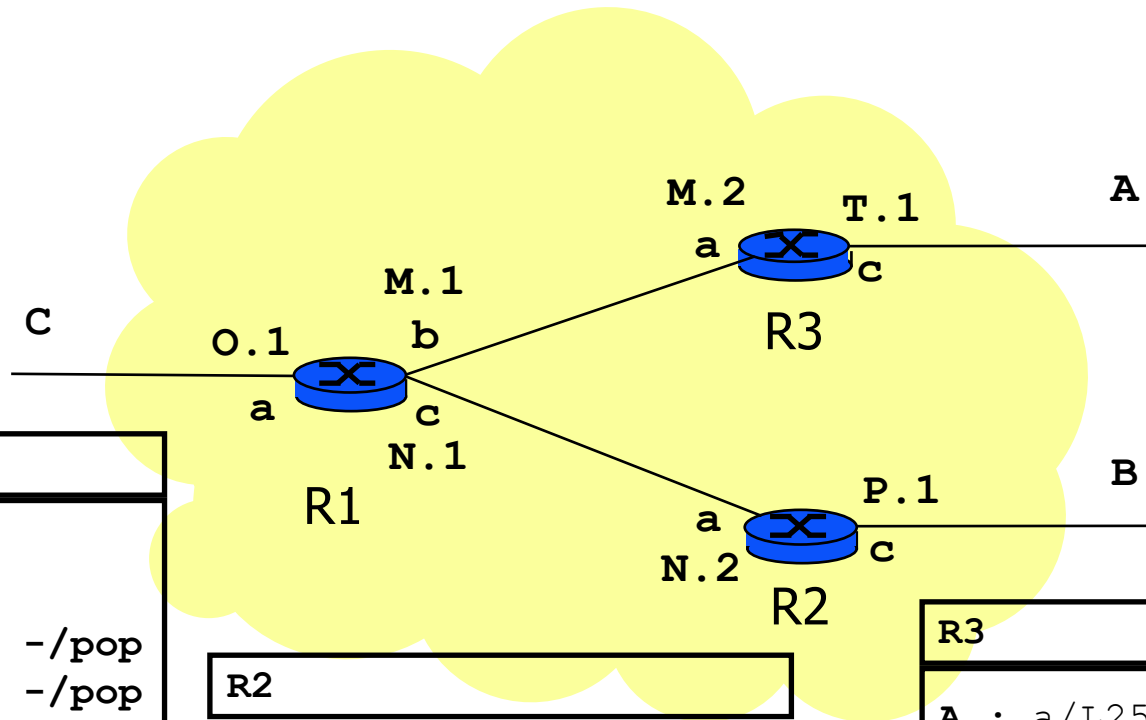a/L1-c/L3 - LIB entry

R1
C, M, N, O

R3
A, M, T

R2
B, N, P

OSPF - routing tables

C

O.1

M.1
b

a

c

N.1

R1

M.2

a

R3

T.1

c

A

P.1

a

c

N.2

R2

B

R1

O: local
M: local
N: local
C: O.2
A: M.2
B: N.2
T: M.2
P: N.2

R2

N: local
P: local
B: P.2
C: N.1
A: N.1
T: N.1
M: N.1
O: N.1

R3

M: local
T: local
A: T.2
C: M.1
B: M.1
N: M.1
P: M.1
O: M.1

Label bindings

R1

```
A : b/L01, c/L02
B : b/L03, c/L04
C : b/L05, c/L06, -/pop
M : b/L07, c/L08, -/pop
N : b/L09, c/L10, -/pop
O : b/L11, c/L12, -/pop
P : b/L13, c/L14
T : b/L15, c/L16
```

R2

```
A : a/L17
B : a/L18, -/pop
C : a/L19
M : a/L20
N : a/L21, -/pop
O : a/L22
P : a/L23, -/pop
T : a/L24
```

R3

```
A : a/L25, -/pop
B : a/L26
C : a/L27
M : a/L28, -/pop
N : a/L29
O : a/L30
P : a/L31
T : a/L32, -/pop
```

Switching tables
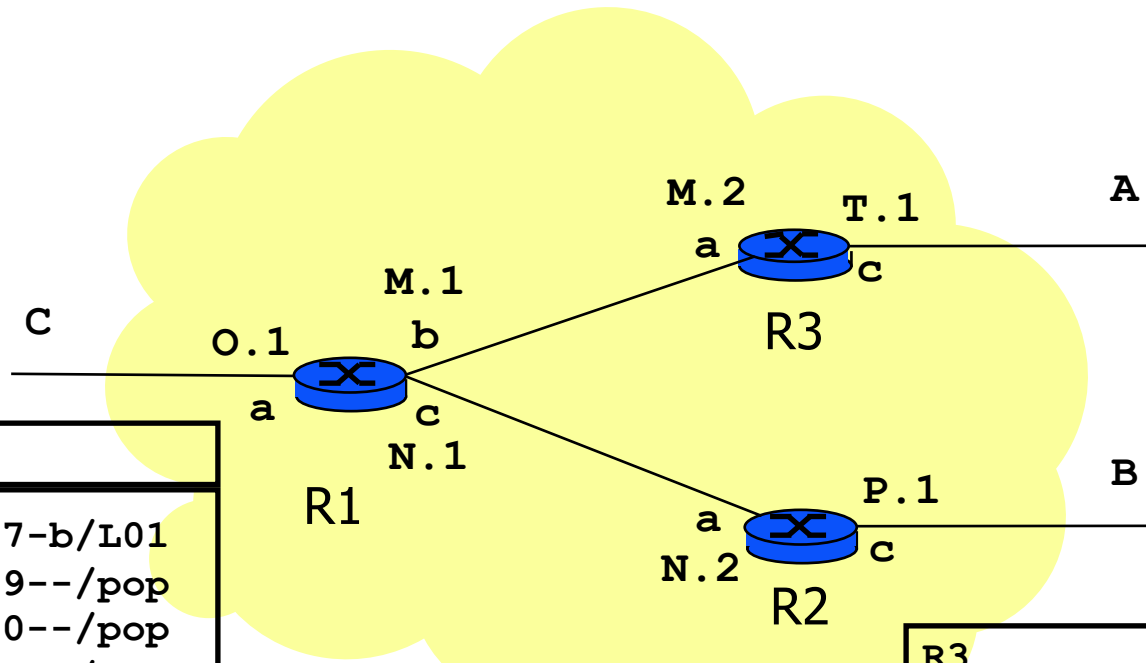
R1	
A : b/L01	
B : c/L04	
C : -/pop	
M : -/pop	
N : -/pop	
O : -/pop	
P : c/L14	
T : b/L15	

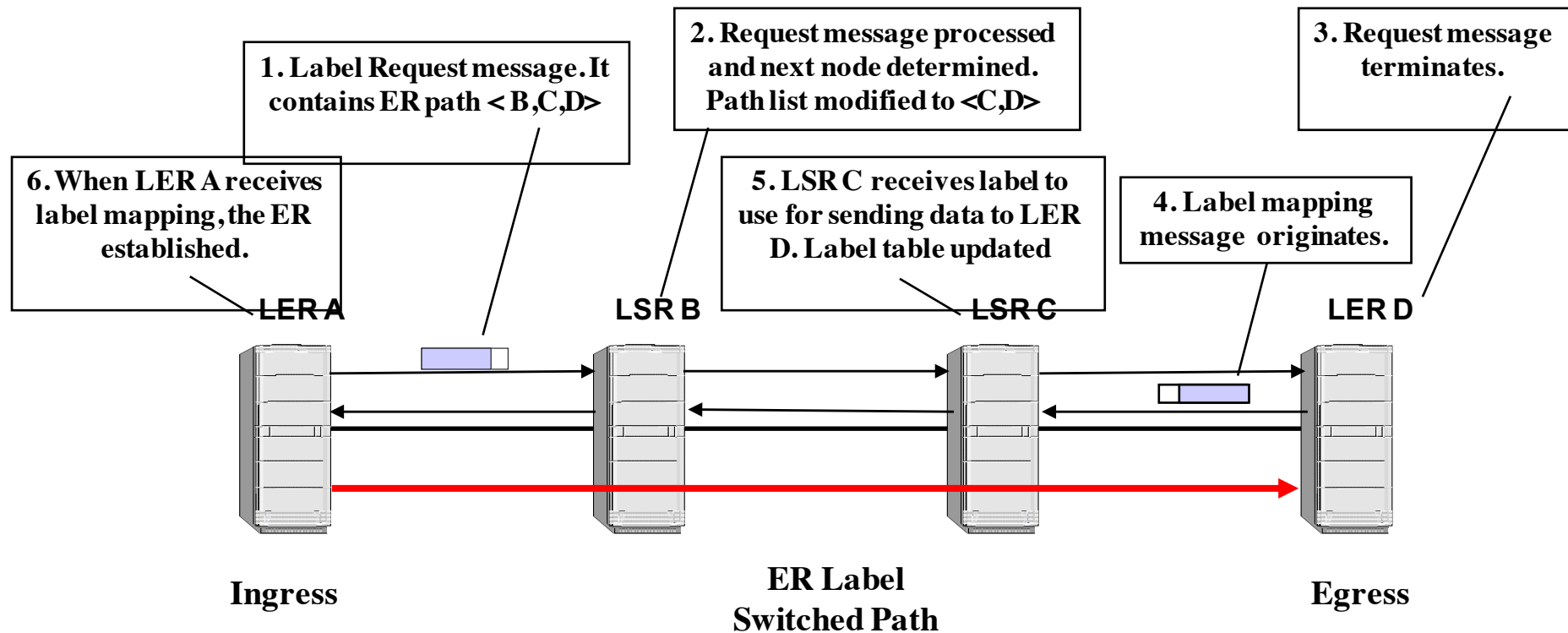| R1 |
|---|
| c/L17-b/L01 |
| c/L19--/pop |
| c/L20--/pop |
| c/L22--/pop |
| c/L24-b/L15 |
| b/L26-c/L04 |
| b/L27--/pop |
| b/L29--/pop |
| b/L30--/pop |
| b/L31-c/L14 |

R2
A : a/L17
B : -/pop
C : a/L19
M : a/L20
N : -/pop
O : a/L22
P : -/pop
T : a/L24

R2
a/L04--/pop
a/L14--/pop

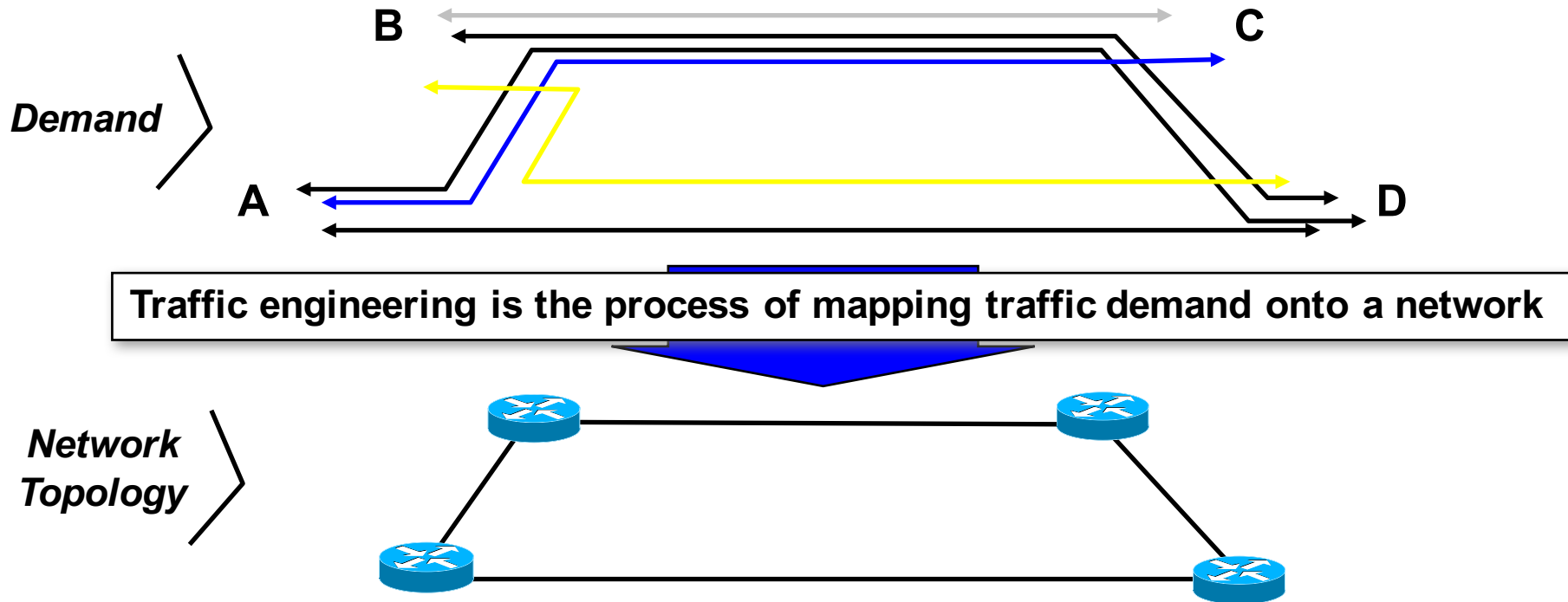| R3 |
|---|
| A : -/pop |
| B : a/L26 |
| C : a/L27 |
| M : -/pop |
| N : a/L29 |
| O : a/L30 |
| P : a/L31 |
| T : -/pop |

R3
a/L01--/pop
a/L15--/pop

ER-LSP setup using CR-LDP

1. Label Request message. It contains ER path < B,C,D>

2. Request message processed and next node determined. Path list modified to <C,D>

3. Request message terminates.

6. When LER A receives label mapping, the ER established.

5. LSR C receives label to use for sending data to LER D. Label table updated

4. Label mapping message originates.

LER A   **LSR B**   **LSR C**   **LER D**

Ingress

ER Label Switched Path

Egress

- Explicit Route setup
- CR-LDP can also be used to reserve resources
  - peak rate, committed rate, burst size

Traffic Engineering with MPLS

Demand

B   C

A   D

Traffic engineering is the process of mapping traffic demand onto a network
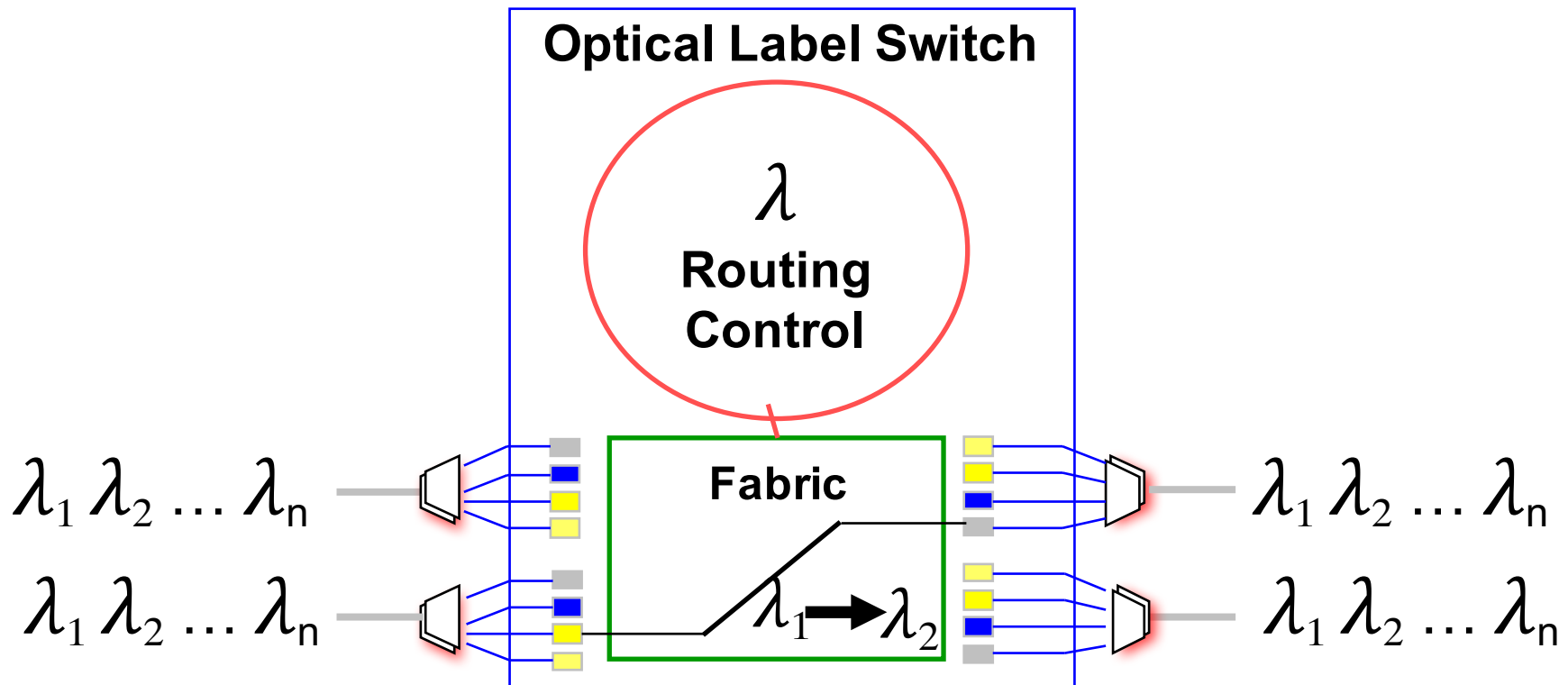
Network Topology

- Maximize utilization of links and nodes
- Engineer links to achieve required delay
- Spread the network traffic across network links, minimize impact of single failure
- Ensure available spare link capacity for re-routing traffic

Unifying forwarding paradigm

Ethernet

PPP
(SONET, SDH)

ATM

Frame
Relay
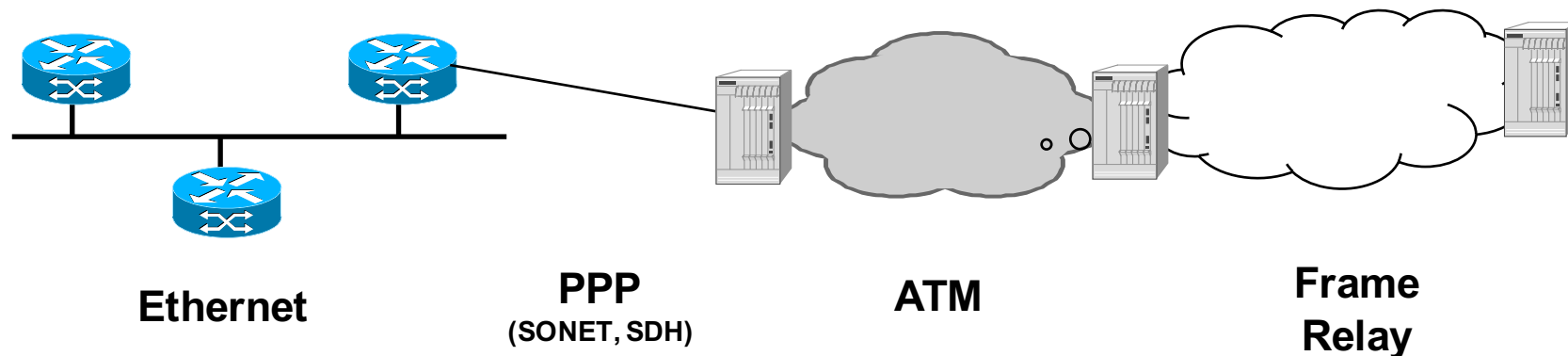
- MPLS offers an end-to-end unifying forwarding paradigm
  - MPLS is "multiprotocol" below (link layer) as well as above (network layer)
  - provides for consistent operations, engineering across multiple technologies
  - allows operators to control different technologies in a unified manner

GMPLS/MPλS

Optical Label Switch

λ
Routing
Control

Fabric

$\lambda_1 \rightarrow \lambda_2$

$\lambda_1 \lambda_2 \ldots \lambda_n$

$\lambda_1 \lambda_2 \ldots \lambda_n$

$\lambda_1 \lambda_2 \ldots \lambda_n$

$\lambda_1 \lambda_2 \ldots \lambda_n$

- Optical networks such as DWDM
  - λ is just another label to distribute - no new protocols required

Facts to remember

- MPLS allows flexible packet classification and network resources optimization
- Labels are distributed by different protocols
  - LDP, RSVP, BGP
- Labels have local (LSR) significance
  - no need for global (domain) wide label allocation/numbering
- Different link layer protocols may co-exist in the same LSR