

Advanced Computer Networks

MPLS

Prof. Andrzej Duda
duda@imag.fr

<http://duda.imag.fr>

1

Contents

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

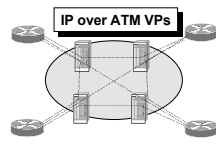
2

MPLS

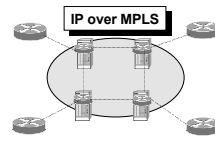
- IGP limits
 - very large routing tables in the core network
 - for every packet look up more that 100 000 entries
 - forwarding from the ISP point of view - just find the egress router
 - IP routing may ignore the real physical topology
 - ISP can put a router on the edge and use ATM/Frame Relay Virtual Path, switches in the middle
 - edge router selects the path based on the destination address
 - route look up done only once in the ISP network
 - but still scalability problems

3

Problems with ATM integration



IP over ATM VPs



IP over MPLS

- ATM cloud invisible to Layer 3 Routing
- ATM network visible to Layer 3 Routing
- Full mesh of VCs within ATM cloud - $n(n-1)/2$
- Single adjacency possible with edge router

MPLS eliminates the "n-squared" problem of IP over ATM VPs

4

MPLS

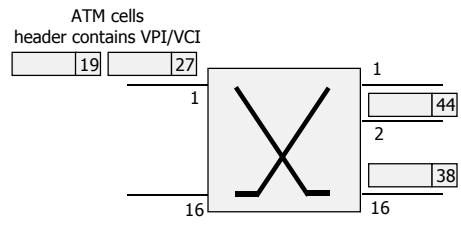
- More IGP limits
 - SPF algorithms send traffic on shared path and may ignore unloaded links
 - even if load balancing can be done in some cases (IGRP)
 - ATM or FR can use virtual circuits, but PPP/SDH/DWDM does not
- Goals of MPLS
 - increase forwarding performance
 - provide more flexibility than 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
 - ATM, FR, PPP, SDH/DWDM

5

ATM VPI/VCI switching

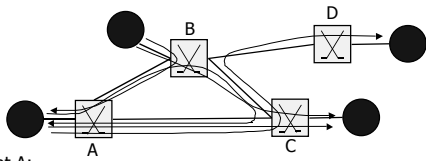
| in | VPI/VCI | out | VPI/VCI |
|----|---------|-----|---------|
| 1 | 27 | 2 | 44 |
| 1 | 19 | 16 | 38 |

ATM cells header contains VPI/VCI



6

Label swapping

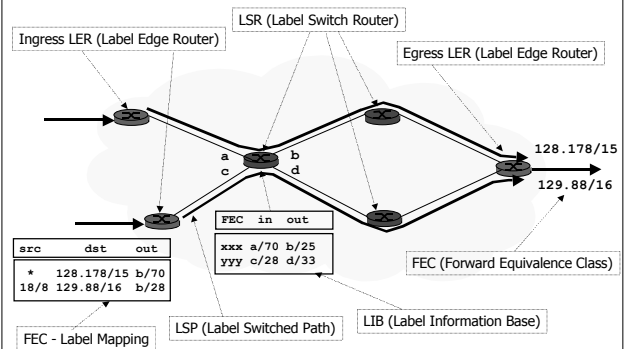


VC table at A:

| 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



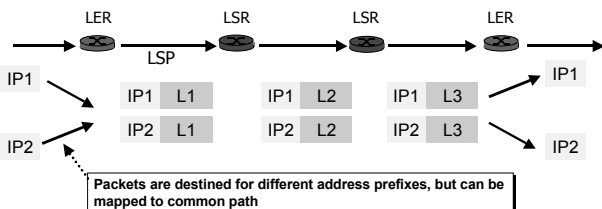
MPLS Principles

- 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

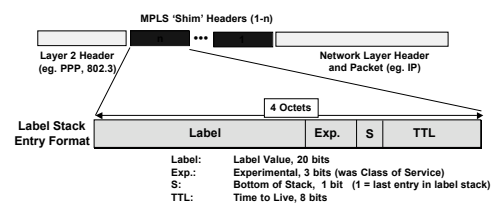
- FEC - group of IP packets
 - forwarded in the same manner, over the same path, and with the same forwarding treatment
- FEC may correspond to
 - destination IP subnet
 - source and destination IP subnet
 - traffic class that LER considers significant
- For example, all traffic with a certain value of IP precedence may constitute a FEC
- FEC in our examples
 - IP prefix

Forwarding Equivalence Classes



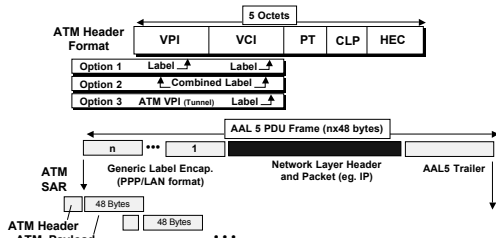
- FEC = "A subset of packets that are all treated the same way by a router"
- Conventional routing: a packet is assigned to a FEC at each hop (i.e. L3 look-up), in MPLS it is only done once at the network ingress

MPLS Encapsulation - PPP & LAN



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

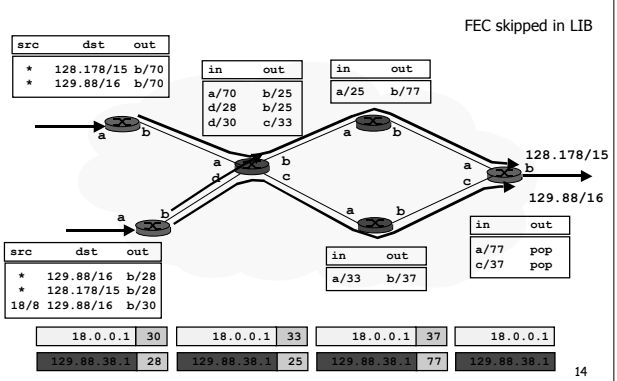
MPLS Encapsulation - ATM



- Top 1 or 2 labels are contained in the VPI/VCI fields of ATM header
 - one in each or single label in combined field, negotiated by LDP
- Further fields in stack are encoded with 'shim' header in PPP/LAN format

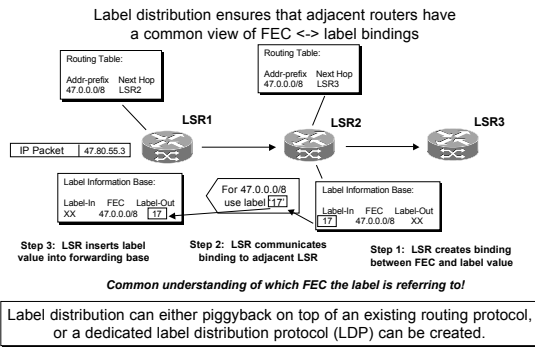
13

MPLS example



14

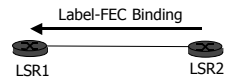
Label Distribution Protocol (LDP)



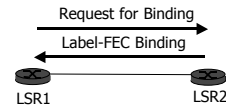
15

Label distribution

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

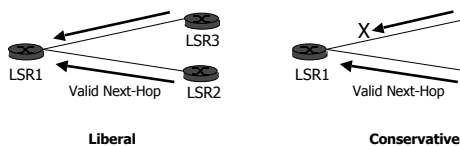


- downstream-on-demand:** upstream LSR asks for a label



16

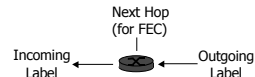
Label retention



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

17

Label control



- 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

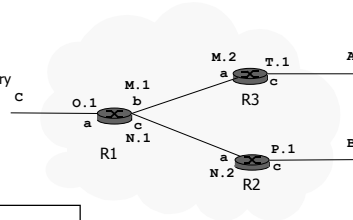
18

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, D... - interface
 a/L1-c/L3 - LIB entry

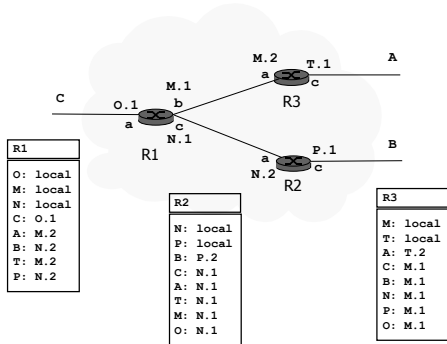


| |
|------------|
| R1 |
| C, M, N, O |

| |
|---------|
| R3 |
| A, M, T |

| |
|---------|
| R2 |
| B, N, P |

OSPF - routing tables

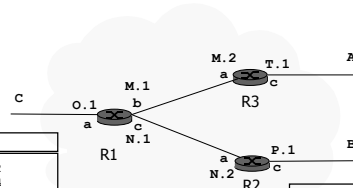


| |
|----------|
| R1 |
| O: local |
| M: local |
| N: local |
| C: O.1 |
| A: M.2 |
| B: N.2 |
| T: M.2 |
| P: N.2 |

| |
|----------|
| 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

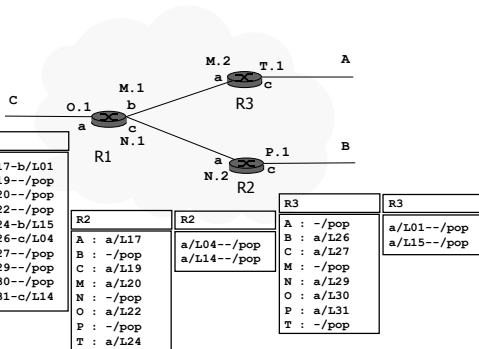


| |
|-------------------------|
| R1 |
| A : b/L01, c/L02 |
| B : b/L03, c/L04 |
| C : b/L05, c/L06, -/pop |
| M : b/L07, c/L08, -/pop |
| N : b/L09, c/L10, -/pop |
| O : b/L11, c/L12, -/pop |
| P : b/L13, c/L14 |
| T : b/L15, c/L16 |

| |
|------------------|
| 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 |

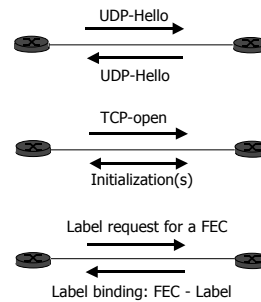
| |
|-----------|
| R2 |
| A : a/L17 |
| B : -/pop |
| C : a/L19 |
| M : a/L20 |
| N : -/pop |
| O : a/L22 |
| P : -/pop |
| T : a/L24 |

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

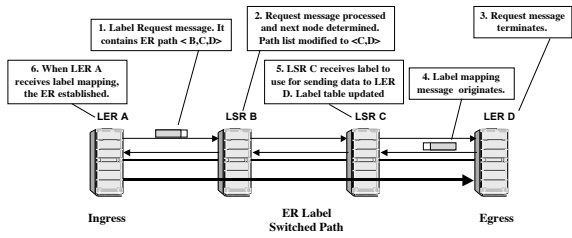
| |
|-----------|
| R3 |
| A : -/pop |
| B : a/L26 |
| C : a/L27 |
| M : -/pop |
| N : a/L29 |
| O : a/L30 |
| P : a/L31 |
| T : -/pop |

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

LDP Protocol



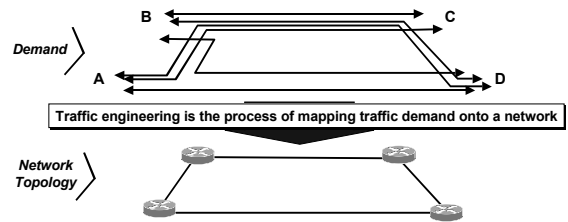
ER-LSP setup using CR-LDP



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

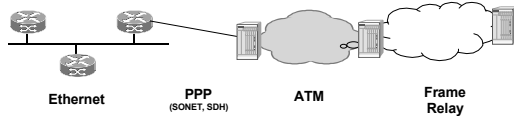
25

Traffic Engineering with MPLS



- 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 traffic₂₆

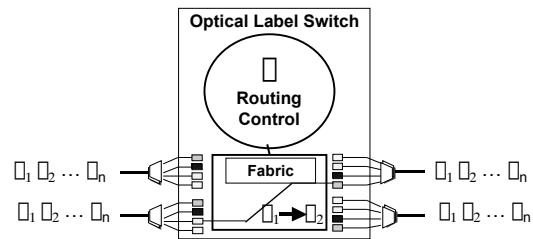
Unifying forwarding paradigm



- 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

27

GMPLS/MP \square S



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

28

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

29