MPLS and GMPLS

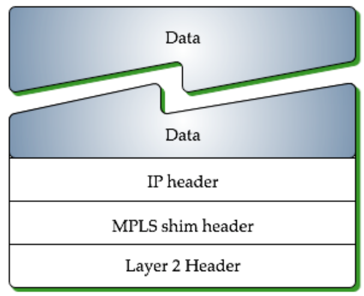
Deep Medhi & Karthik Ramasamy
October 2010
http://www.NetworkRouting.net

MPLS: a bit of history

- Initially motivated by faster IP packet processing (exact prefix matching)
  - IP switching
  - ATM switching
  - Tag switching
- Then, found its application to IP traffic engineering (late 90's)
- MPLS VPN

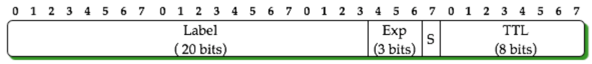
2

Where does MPLS header fit in the protocol stack?



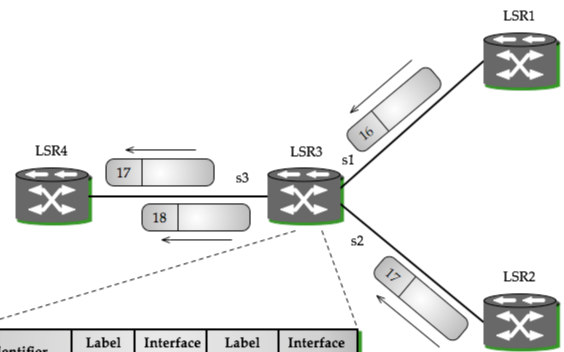
MPLS header format

- Label
- Experimental bits: can assign priority
- Time-to-Live field



Label Swapping at an MPLS router (LSR)

- I

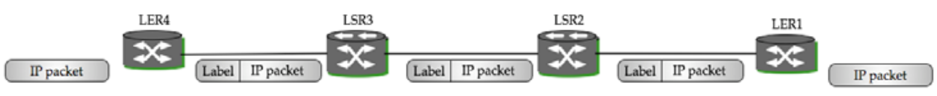


Identifier	Label In	Interface In	Label Out	Interface Out
Stream-A	16	s1	17	s3
Stream-B	17	s2	18	s3

5

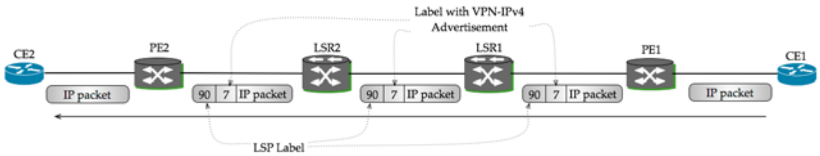
How are IP packets carried by MPLS

- Label Edge Routers, Label Switched Routers



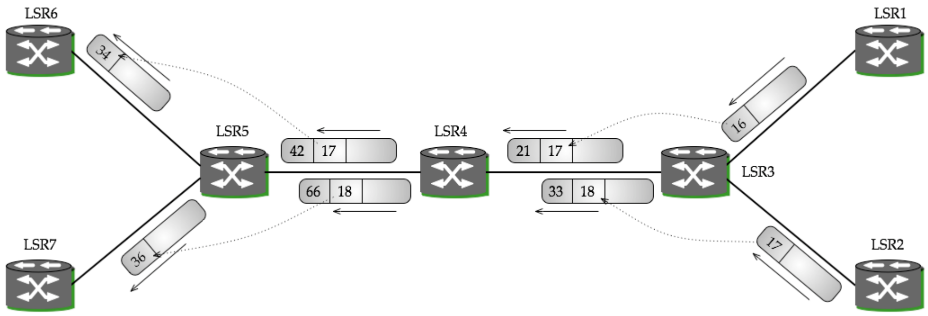
6

Label swapping with IP packets



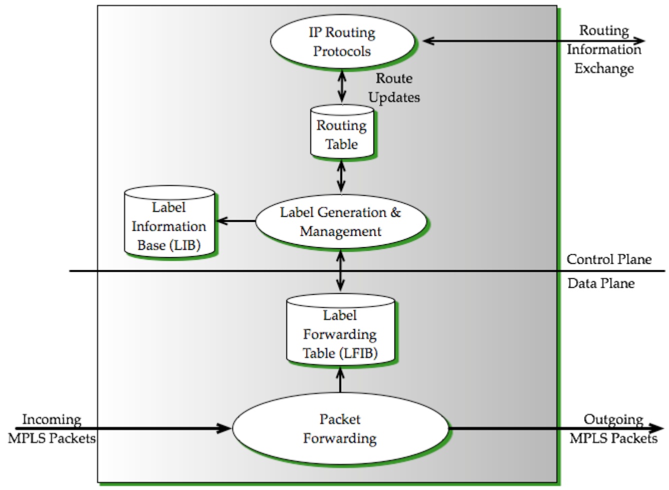
7

Two Labels…



8

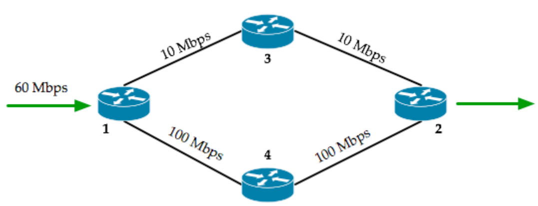
MPLS router



9

IP traffic engineering:

- Link weight adjustment
- Set up MPLS tunnels



10

How to set up and keep track of tunnel?

- Protocols for tunnel setup
  - RSVP-TE
  - CR-LDP (now "decommissioned")

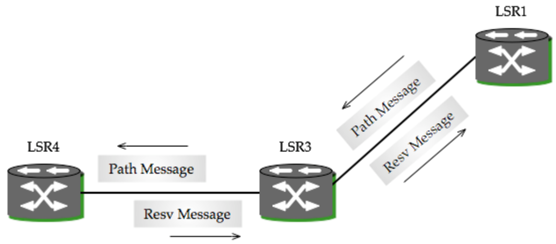
- Need to bank on Link-state routing protocol to "get" information
  - IS-IS-TE
  - OSPF-TE

11

RSVP(-TE) overview

- RSVP-TE is used for signaling between routers to set up LSP flows
- Message Types: Path, Resv, PathErr, ResvErr, PathTear, ResvTear, and ResvConf

RSVP messages for LSP set up



13

RSVP packet format

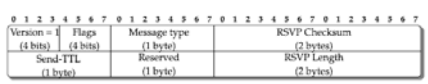


FIGURE 18.7 RSVP common header.

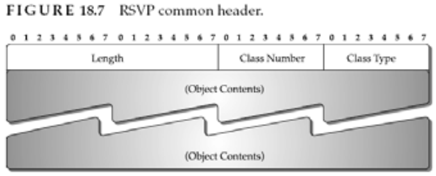


FIGURE 18.8 RSVP object format.

14

TABLE 18.1 RSVP object examples for MPLS (An up-to-date list is maintained at [318]).

Object Name	Used in	Class Number	Examples: Class type with value in parentheses (source RFC listed as [] from bibliography)
INTEGRITY	Path, Resv	4	RSVP Integrity (1) [51]
SESSION	Path, Resv	1	LSP Tunnel for IPv4 (7) [43]
RSVP-HOP	Path, Resv	3	IPv4 (1), IPv6 (2) [90]
TIME-VALUES	Path, Resv	5	Time Value (1) [90]
FILTER-SPEC	Resv	10	LSP Tunnel for IPv4 (7) [43]
SENDER-TEMPLATE	Path	11	LSP Tunnel for IP4 (7) [43]
RSVP-LABEL	Path	16	Type 1 Label (1) [43]; Generalized Label (2) [74]
LABEL-REQUEST	Path	19	No label range (1) [43]; generalized label request (4) [74]
EXPLICIT-ROUTE	Path	20	Type 1 Explicit Route (1) [43]
POLICY-DATA	Path	14	Type 1 (1) [90]
SENDER-TSPEC	Path	12	Integrated Services (1) [90]
RECORD-ROUTE	Path	21	IPv4 (1) [43]
SESSION-ATTRIBUTE	Path	207	LSP Tunnel (7) [43]
DETOUR	Path	63	IPv4 (7) [541]
FAST-REROUTE	Path	205	Type 1 (1) [541]
UPSTREAM-LABEL	Path	35	Same as in RSVP-LABEL

15

Why OSPF-TE

- New attributes of a links needed for MPLS
  - Max link bandwidth
  - Max reserved bandwidth
  - Unreserved bandwidth at different priority levels
  - Traffic engineering metric

OSPF-TE through LSA extention

- In OSPFv2, several link state advertisement (LSA) types: one of them: *opaque LSA*
  - The intended use of opaque LSA is to allow a general LSA feature so that it might be useful for any future extension.
- Opaque LSA: three link state types have been presented for the scope of flooding
  - type 9, type 10, and type 11 for local subnet flooding, intra-area flooding, and flooding in the entire autonomous systems, respectively.

17

OSPF-TE (cont'd)

- For MPLS traffic engineering, opaque LSA type 10 is used
  - This limits flooding to an intra-area of an OSPF domain; it is known as a traffic engineering LSA.
- TE LSA contains a standard header that includes information such as link state age, advertising router, and link state sequence number; in addition, it uses nested TLV to contain information needed for TE LSA.
- At the top level, there are two TLVs: (1) a router address TLV and (2) a Link TLV.
- Link TLV contains several sub-TLVs: these are used for the new attributes of a link (such as bandwidth etc)

IS-IS-TE

- Similar to OSPF-TE, IS-IS is also extended to allow TE features

MPLS traffic enginering

- Basic Steps:
  - Determine points to be connected:
    - End points: ingress-egress point
    - Size of the tunnel (e.g. 6Mbps)
  - Determine bandwidth availability
    - Using ISIS-TE, OSPF-TE
  - Set up paths using RSVP-TE (like a telephone call)

20

How to select

- Constrained Shortest-path
- Network flow modeling

ALGORITHM 19.1 Constrained shortest path first algorithm: from node i to node v, for bandwidth constraint, computed at time t

1. Network $\mathcal{N}$ and cost of link $d^i_{km}(t)$ and available bandwidth on $b^i_{km}(t)$ on link k–m, as known to node i at the time of computation, t.
2. For link k–m, if available bandwidth, $b^k_{km}(t)$, is smaller than bandwidth request $\bar{b}$, then set link cost temporarily to infinity, i.e., $d^i_{km}(t) = \infty$.
3. Initially, consider only source node i in the list of nodes considered ("permanent list"), i.e., $\mathcal{S} = \{i\}$; mark the list with all the rest of the nodes as $\mathcal{S}'$ ("tentative list"). Initialize

   $$\underline{D}_{ij}(t) = d^i_{ij}(t), \quad \text{for all} \quad j \in \mathcal{S}'.$$

4. Identify a neighboring node (intermediary) k not in the current list $\mathcal{S}$ with the minimum-cost path from node i, i.e., find $k \in \mathcal{S}'$ such that $\underline{D}_{ik}(t) = \min_{m \in \mathcal{S}'} \underline{D}_{im}(t)$

   if k is the same as destination v, *stop*.

   Add k to permanent list $\mathcal{S}$, i.e., $\mathcal{S} = \mathcal{S} \cup \{k\}$,

   Drop k from tentative list $\mathcal{S}'$, i.e., $\mathcal{S}' = \mathcal{S}' \backslash \{k\}$.

   If $\mathcal{S}'$ is empty, *stop*.

5. Consider neighboring nodes $\mathcal{N}_k$ of the intermediary k (but do not consider nodes already in permanent list $\mathcal{S}$) to check for improvement in the minimum-cost path, i.e.,

   for $j \in \mathcal{N}_k \cap \mathcal{S}'$

   $$\underline{D}_{ij}(t) = \min\{\underline{D}_{ij}(t), \underline{D}_{ik}(t) + d^i_{kj}(t)\} \tag{19.2.1}$$

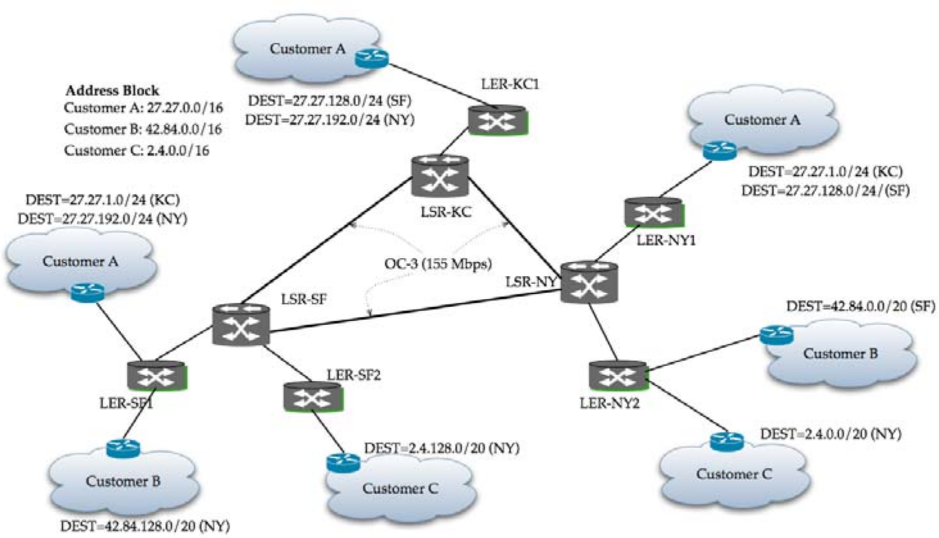
   go to Step-4.

TABLE 19.1 Customer demand matrix.

Customer ID	Locations between		Bandwidth Requirement
Customer A (27.27.0.0/16)	Kansas City (27.27.1.0/24)	San Francisco (27.27.128.0/24)	45 Mbps
	Kansas City (27.27.1.0/24)	New York (27.27.192.0/24)	60 Mbps
	San Francisco (27.27.128.0/24)	New York (27.27.192.0/24)	20 Mbps
Customer B (42.84.0.0/16)	San Francisco (42.84.0.0/20)	New York (42.84.128.0/20)	80 Mbps
Customer C (2.4.0.0/16)	San Francisco (2.4.0.0/20)	New York (2.4.128.0/20)	100 Mbps

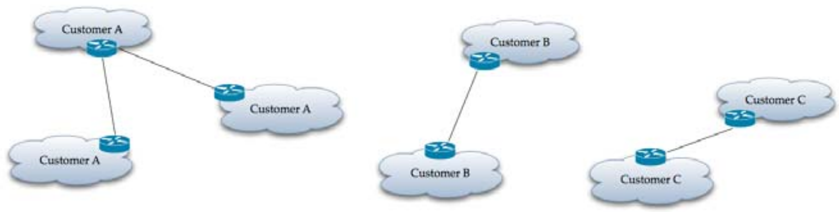
23

MPLS-VPN Traffic Engineering

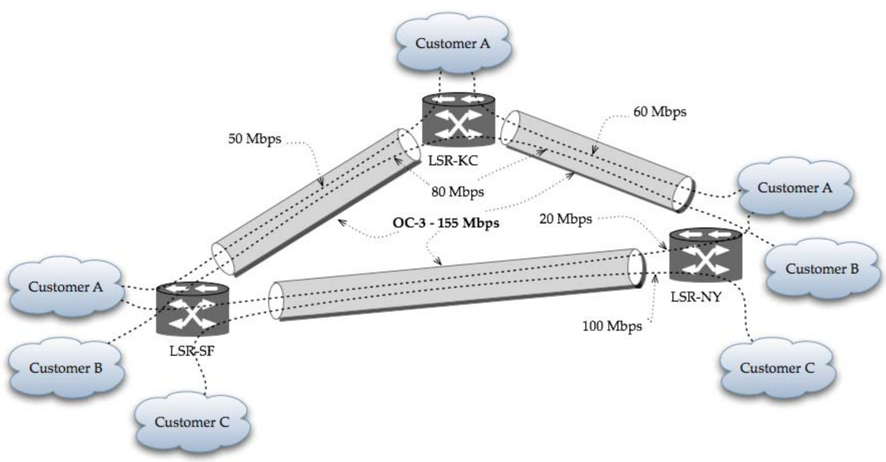


24

Customer's logical view



VPN routing view



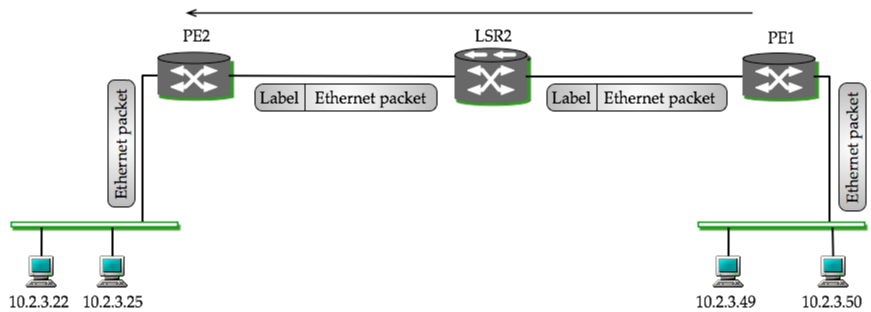
1. Set up TE 100-Mbps tunnel for Customer C (at LSR-SF from SF to NY, and reverse)
   Available link bandwidth: SF-NY: 55 Mbps; SF-KC: 155 Mbps; KC-NY: 155 Mbps
2. Set up TE 80-Mbps tunnel for Customer B (at LSR-SF from SF to NY, and reverse)
   Available link bandwidth: SF-NY: 55 Mbps; SF-KC: 75 Mbps; KC-NY: 75 Mbps
3. Set up TE 20-Mbps tunnel for Customer A (at LSR-SF from SF to NY, and reverse)
   Available link bandwidth: SF-NY: 35 Mbps; SF-KC: 75 Mbps; KC-NY: 75 Mbps
4. Set up TE 45-Mbps tunnel for Customer A (at LSR-SF from SF to KC, and reverse)
   Available link bandwidth: SF-NY: 35 Mbps; SF-KC: 30 Mbps; KC-NY: 75 Mbps
5. Set up TE 60-Mbps tunnel for Customer A (at LSR-KC from KC to NY, and reverse)
   Available link bandwidth: SF-NY: 35 Mbps; SF-KC: 30 Mbps; KC-NY: 15 Mbps

1′. Set up TE 80-Mbps tunnel for Customer B (at LSR-SF from SF to NY, and reverse)
Available link bandwidth: SF-NY: 75 Mbps; SF-KC: 155 Mbps; KC-NY: 155 Mbps

2′. Set up TE 100-Mbps tunnel for Customer C (at LSR-SF from SF to NY, and reverse)
Available link bandwidth: SF-NY: 75 Mbps; SF-KC: 55 Mbps; KC-NY: 55 Mbps

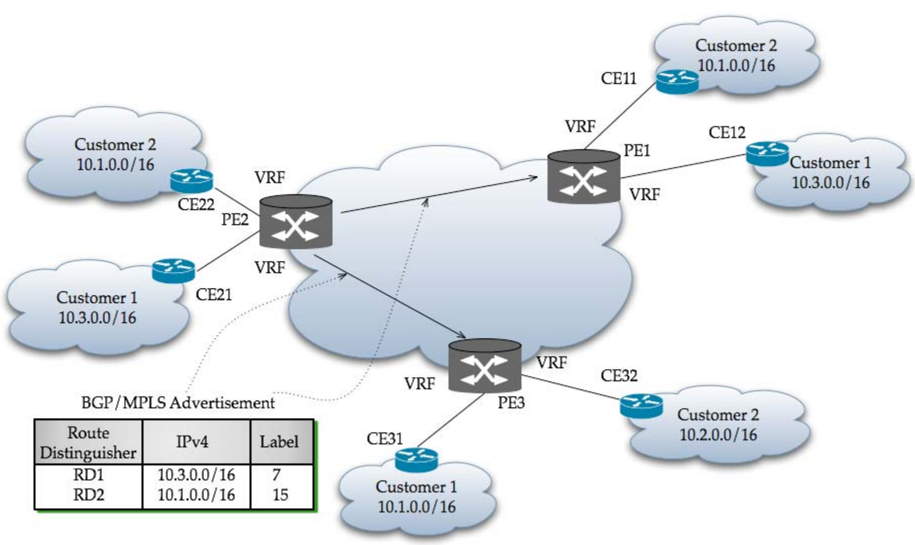
3. Set up TE 20-Mbps tunnel for Customer A (at LSR-SF from SF to NY, and reverse)
   Available link bandwidth: SF-NY: 55 Mbps; SF-KC: 55 Mbps; KC-NY: 55 Mbps
4. Set up TE 45-Mbps tunnel for Customer A (at LSR-SF from SF to KC, and reverse)
   Available link bandwidth: SF-NY: 55 Mbps; SF-KC: 10 Mbps; KC-NY: 55 Mbps

Virtual LAN service ("emulated Ethernet Services) over MPLS



29

BGP/MPLS IP VPN architecture



BGP/MPLS route distinguisher

Type 0

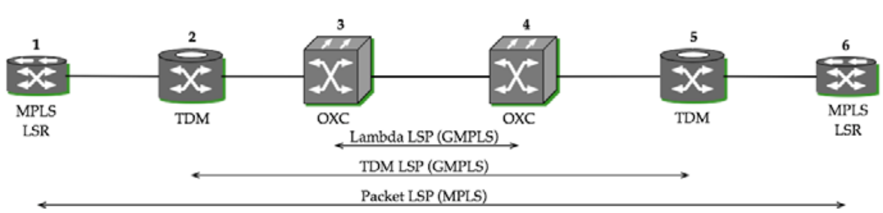
Type Field (2 bytes)	Administrator Subfield (2 bytes)	Assigned Number Subfield (4 bytes)	IPv4 Address (4 bytes)

Type 1 and Type 2

Type Field (2 bytes)	Administrator Subfield (4 bytes)	Assigned Number Subfield (2 bytes)	IPv4 Address (4 bytes)

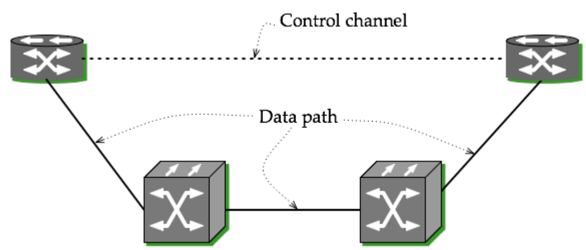
31

MPLS and GMPLS labels



32

Separation of Data path and control path in GMPLS



Scalability issues/Manageability issues

- How many tunnels in the networks?
- How often to change them?
  - Bandwidth adjustment
  - Route adjustment

0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7

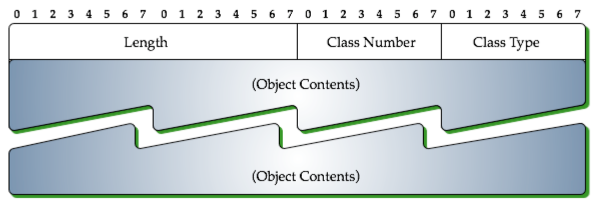
Waveband Identifier (4 bytes)
Start Label (4 bytes)
End Label (4 bytes)

35

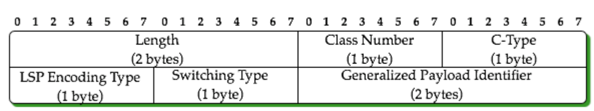
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7

Version = 1 (4 bits)	Flags (4 bits)	Message type (1 byte)	RSVP Checksum (2 bytes)
Send-TTL (1 byte)		Reserved (1 byte)	RSVP Length (2 bytes)

36


0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
Length
Class Number
Class Type
(Object Contents)
(Object Contents)

37


0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
Length (2 bytes)
Class Number (1 byte)
C-Type (1 byte)
LSP Encoding Type (1 byte)
Switching Type (1 byte)
Generalized Payload Identifier (2 bytes)

38

0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7
Length (2 bytes)		Class Number (1 byte)	C-Type (1 byte)
LSP Encoding Type (1 byte)	Switching Type (1 byte)	Generalized Payload Identifier (2 bytes)	

39