Deploying MPLS Traffic Engineering

Nurul Islam Roman (nurul@apnic.net)

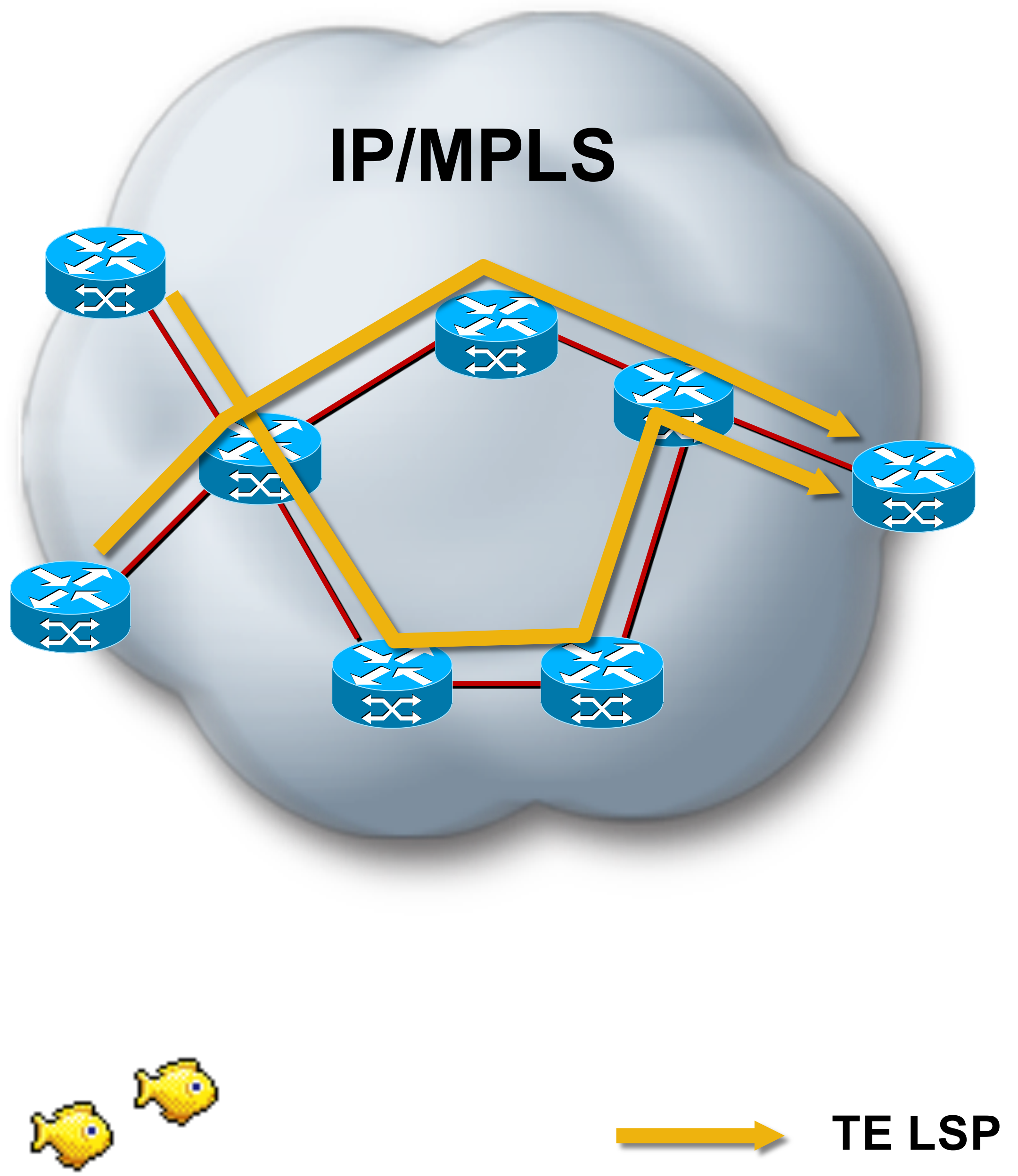
Agenda

- Technology Overview
- Bandwidth optimization
- TE for QoS
- Traffic Protection
- Inter-Domain Traffic Engineering
- General Deployment Considerations

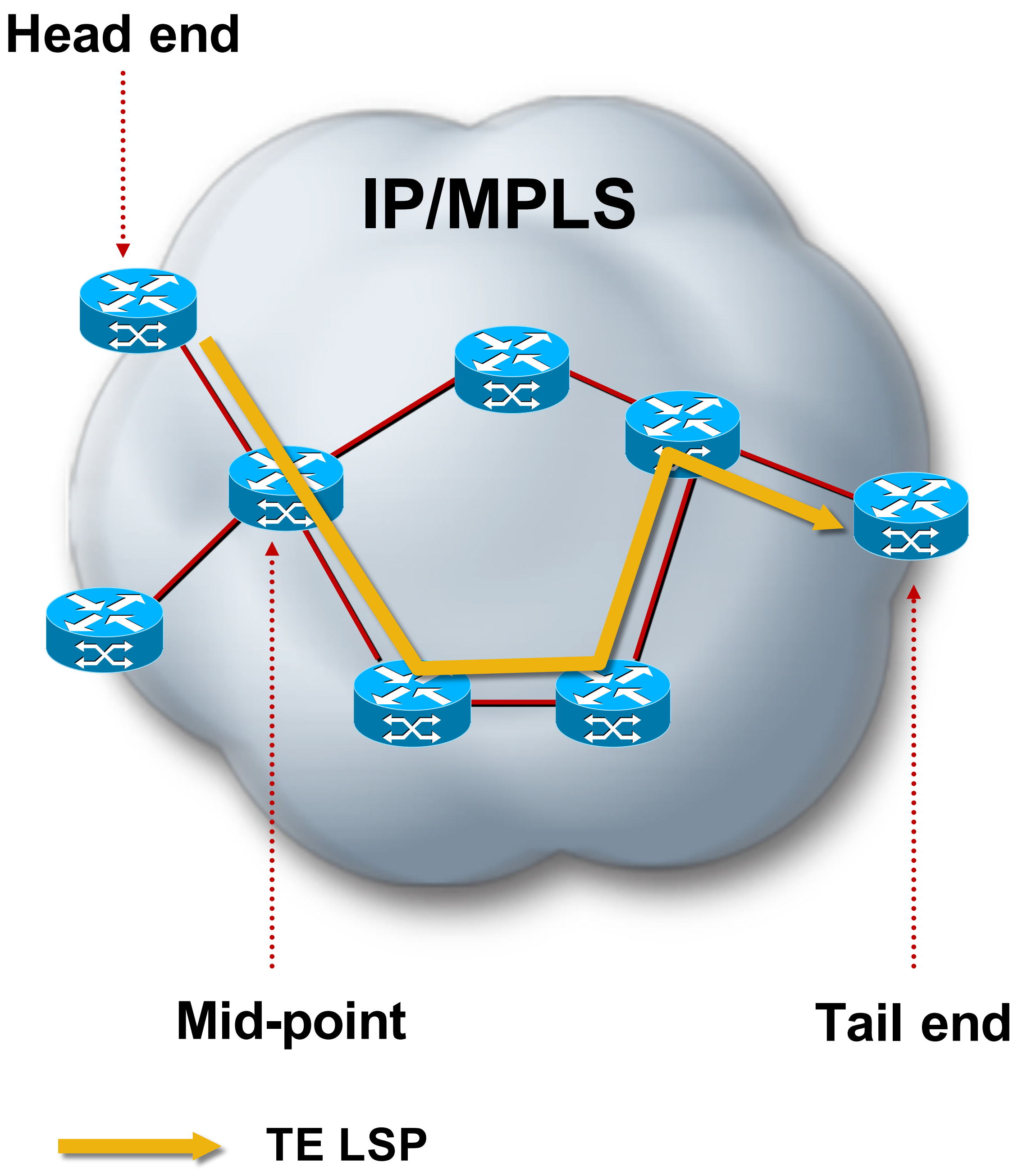
Technology Overview

MPLS TE Overview

- Introduces explicit routing
- Supports constraint-based routing
- Supports admission control
- Provides protection capabilities
- Uses RSVP-TE to establish LSPs
- Uses ISIS / OSPF extensions to advertise link attributes



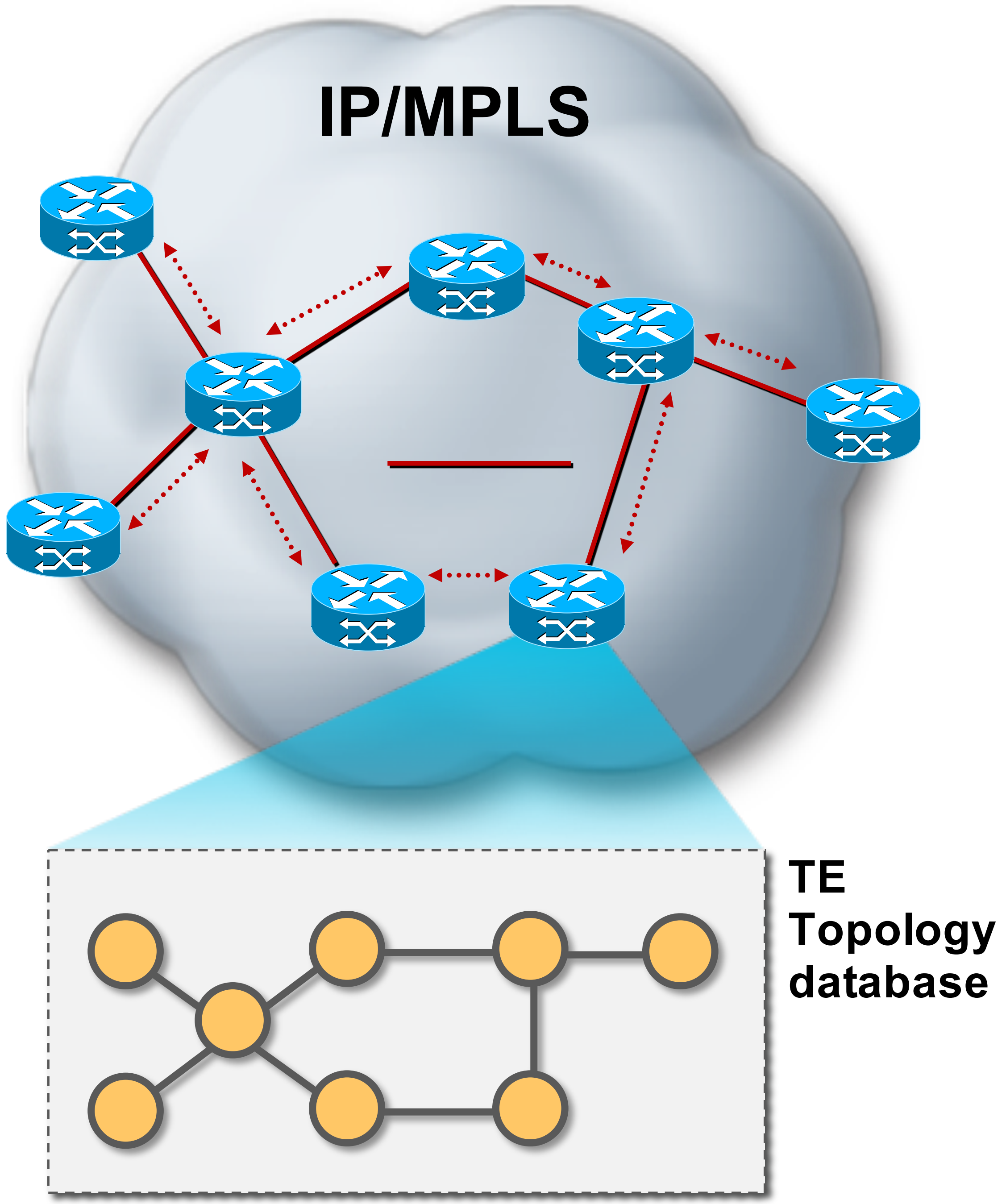
How MPLS TE Works



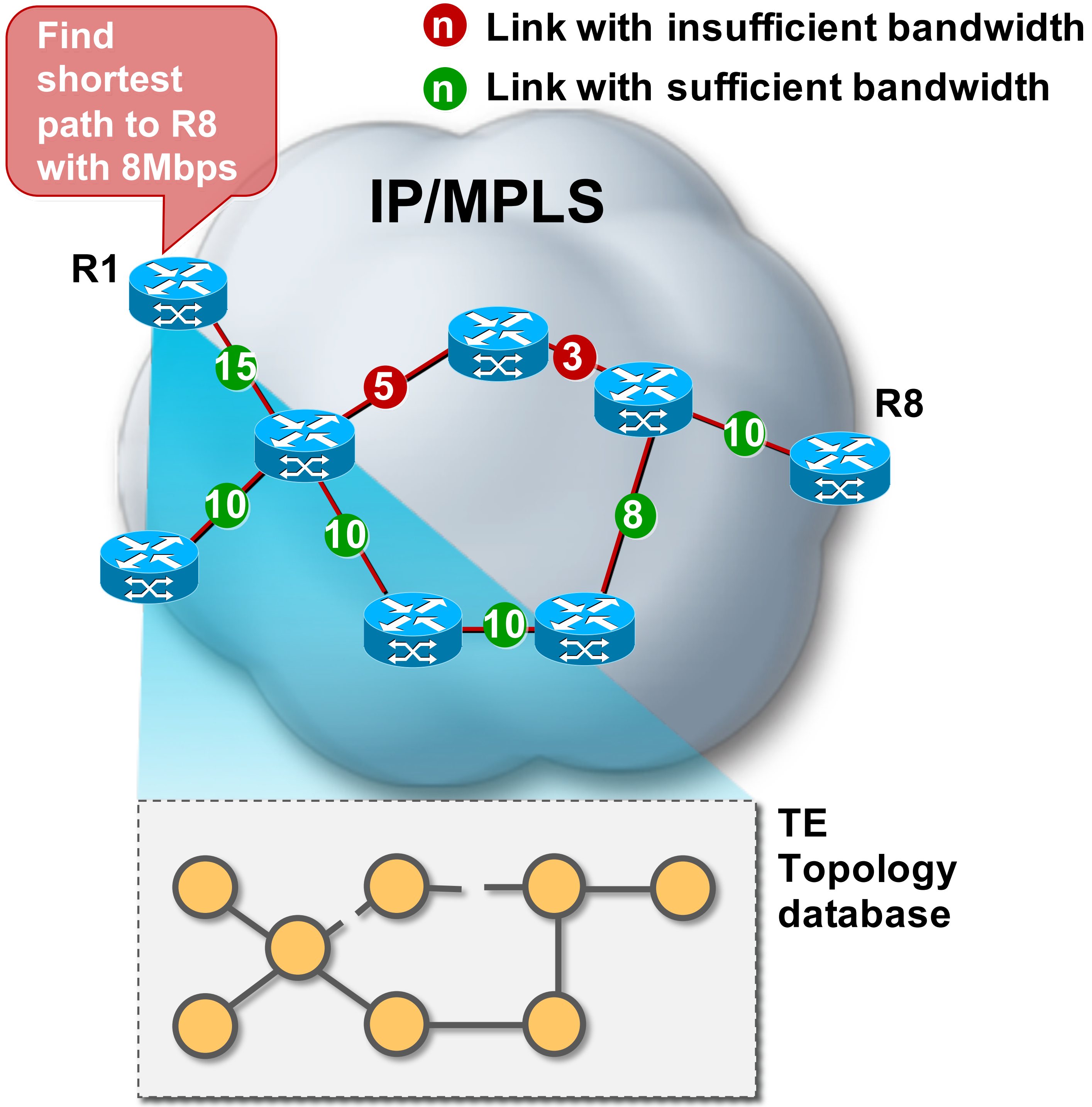
- Link information Distribution*
  - ISIS-TE
  - OSPF-TE
- Path Calculation (CSPF)*
- Path Setup (RSVP-TE)
- Forwarding Traffic down Tunnel
  - Auto-route (announce / destinations)
  - Static route
  - PBR
  - CBTS / PBTS
  - Forwarding Adjacency
  - Tunnel select

Link Information Distribution

- Additional link characteristics
  - Interface address
  - Neighbor address
  - Physical bandwidth
  - Maximum reservable bandwidth
  - Unreserved bandwidth (at eight priorities)
  - TE metric
  - Administrative group (attribute flags)
- IS-IS or OSPF flood link information
- All TE nodes build a TE topology database
- Not required if using off-line path computation



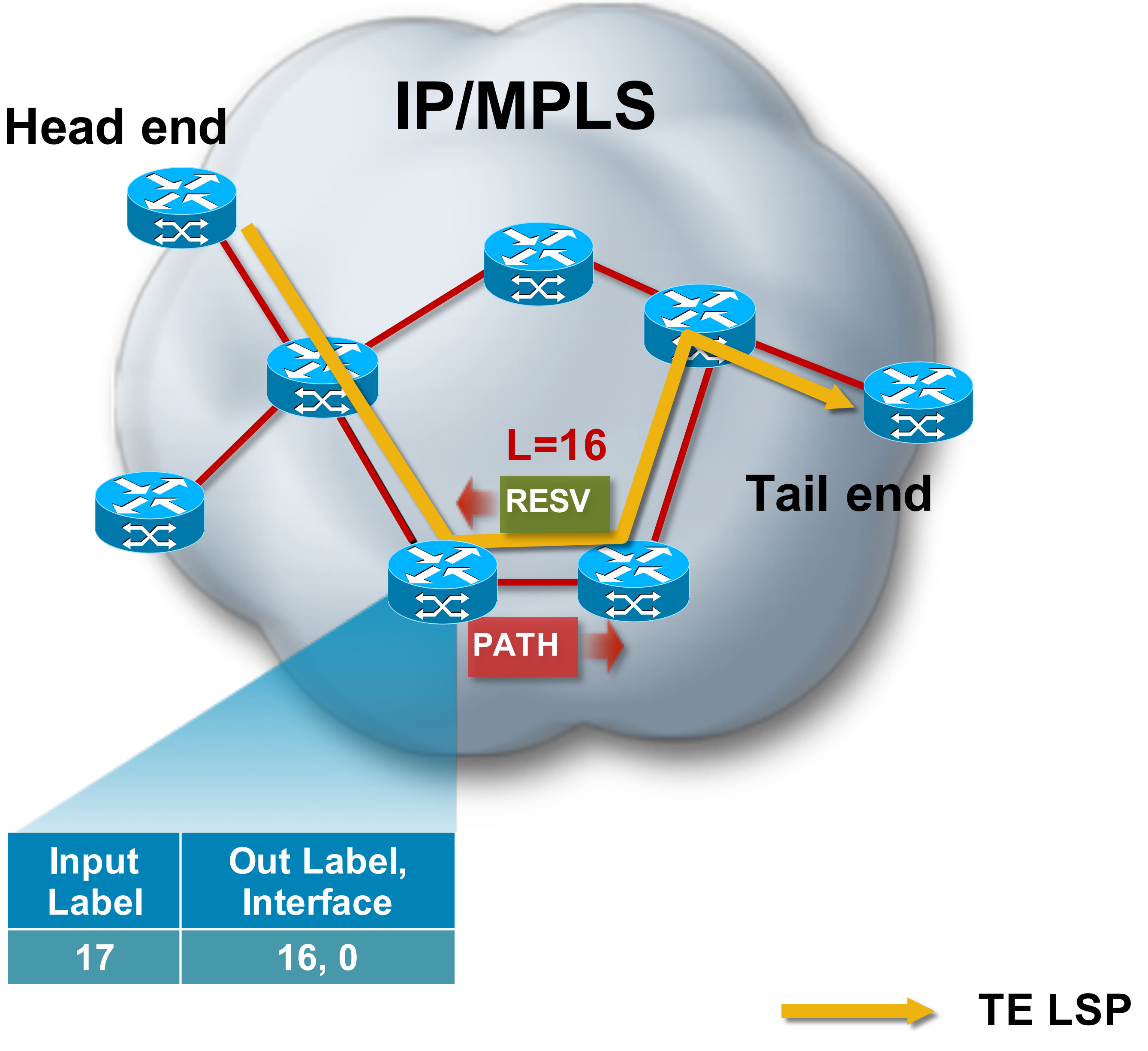
Path Calculation



- TE nodes can perform constraint-based routing
- Tunnel head end generally responsible for path calculation
- Constraints and topology database as input to path computation
- Shortest-path-first algorithm ignores links not meeting constraints
- Tunnel can be signaled once a path is found
- Not required if using offline path computation

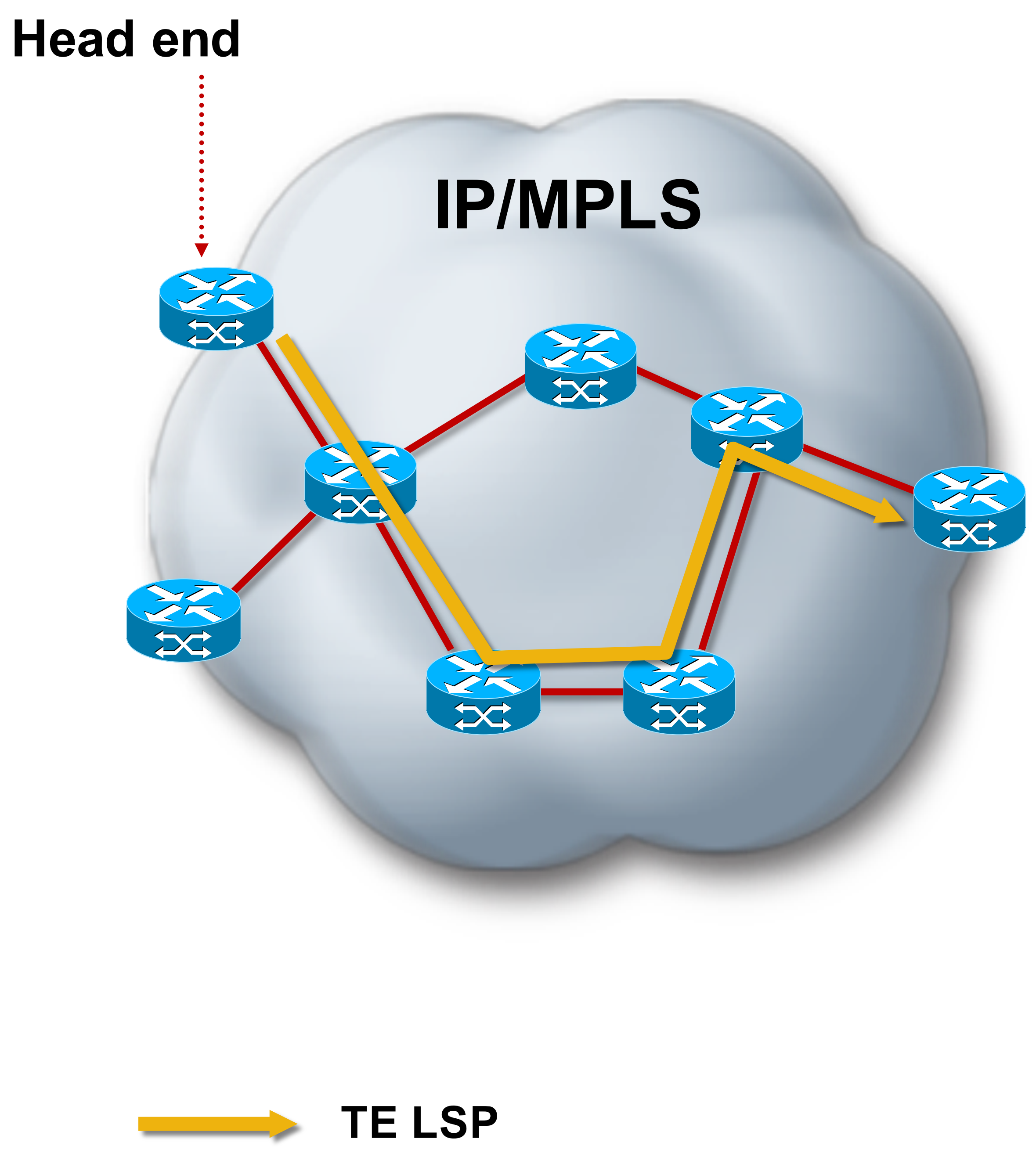
TE LSP Signaling

- Tunnel signaled with TE extensions to RSVP
- Soft state maintained with downstream PATH messages
- Soft state maintained with upstream RESV messages
- New RSVP objects
  - LABEL_REQUEST (PATH)
  - LABEL (RESV)
  - EXPLICIT_ROUTE
  - RECORD_ROUTE (PATH/RESV)
  - SESSION_ATTRIBUTE (PATH)
- LFIB populated using RSVP labels allocated by RESV messages



Input Label	Out Label, Interface
17	16, 0

Traffic Selection



- Multiple traffic selection options
  - Auto-route
  - Static routes
  - Policy Based Routing
  - Forward Adjacency
  - Pseudowire Tunnel Selection
  - Class / Policy Based Tunnel Selection
- Tunnel path computation independent of routing decision injecting traffic into tunnel
- Traffic enters tunnel at head end

Configuring MPLS TE and Link Information Distribution Using IS-IS (Cisco IOS)

```
mpls traffic-eng tunnels
!
interface TenGigabitEthernet0/1/0
 ip address 172.16.0.0 255.255.255.254
 ip router isis
 mpls traffic-eng tunnels
 mpls traffic-eng attribute-flags 0xF
 mpls traffic-eng administrative-weight 20
 ip rsvp bandwidth 100000
!
router isis
 net 49.0001.1720.1625.5001.00
 is-type level-2-only
 metric-style wide
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng level-2
 passive-interface Loopback0
!
```

- Enable MPLS TE on this node
- Enable MPLS TE on this interface
- Attribute flags
- TE metric
- Maximum reservable bandwidth
- Enable wide metric format and TE extensions (TE Id, router level)

Configuring MPLS TE and Link Information Distribution Using OSPF (Cisco IOS)

```
mpls traffic-eng tunnels
!
interface TenGigabitEthernet0/1/0
 ip address 172.16.0.0 255.255.255.254
 mpls traffic-eng tunnels
 mpls traffic-eng attribute-flags 0xF
 mpls traffic-eng administrative-weight 20
 ip rsvp bandwidth 100000
!
router ospf 100
 log-adjacency-changes
 passive-interface Loopback0
 network 172.16.0.0 0.0.255.255 area 0
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng area 0
!
```

Enable MPLS TE on this node

Enable MPLS TE on this interface

Attribute flags

TE metric

Maximum reservable bandwidth

Enable TE extensions (TE router id and area)

Configuring MPLS TE and Link Information Distribution Using IS-IS (Cisco IOS XR)

```
router isis DEFAULT
 is-type level-2-only
 net 49.0001.1720.1625.5129.00
 address-family ipv4 unicast
  metric-style wide
  mpls traffic-eng level 2
  mpls traffic-eng router-id Loopback0
 !
 interface Loopback0
  passive
  address-family ipv4 unicast
  !
 !
 interface TenGigE0/0/0/0
  address-family ipv4 unicast
  !
 !
!
rsvp
 interface TenGigE0/0/0/0
  bandwidth 100000
 !
!
mpls traffic-eng
 interface TenGigE0/0/0/0
  admin-weight 5
  attribute-flags 0x8
 !
!
```

Enable wide metric format and TE extensions (TE Id, router level)

Configuration mode for RSVP global and interface commands

Maximum reservable bandwidth

Configuration mode for MPLS TE global and interface commands

TE metric

Attribute flags

Configuring MPLS TE and Link Information Distribution Using OSPF (Cisco IOS XR)

```
router ospf DEFAULT
 area 0
  mpls traffic-eng
  interface Loopback0
   passive
  !
  interface TenGigE0/0/0/0
  !
 mpls traffic-eng router-id Loopback0
!
rsvp
 interface TenGigE0/0/0/0
  bandwidth 100000
 !
!
mpls traffic-eng
 interface TenGigE0/0/0/0
  admin-weight 5
  attribute-flags 0x8
 !
!
```

Enable TE extensions on this area

TE router Id

Configuration mode for RSVP global and interface commands

Maximum reservable bandwidth

Configuration mode for MPLS TE global and interface commands

TE metric

Attribute flags

Configuring Tunnel at Head End (Cisco IOS)

```
interface Tunnel1
 description FROM-ROUTER-TO-DST1
 ip unnumbered Loopback0
 tunnel destination 172.16.255.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng priority 5 5
 tunnel mpls traffic-eng bandwidth  10000
 tunnel mpls traffic-eng affinity 0x0 mask 0xF
 tunnel mpls traffic-eng path-option 5 explicit name PATH1
 tunnel mpls traffic-eng path-option 10 dynamic
!
ip explicit-path name PATH1 enable
 next-address 172.16.0.1
 next-address 172.16.8.0
!
```

Destination (tunnel tail end)

TE tunnel (as opposed to GRE or others)

Setup/hold priorities

Signaled bandwidth

Consider links with 0x0/0xF as attribute flags

Tunnel path options (PATH1, otherwise dynamic)

Explicit PATH1 definition

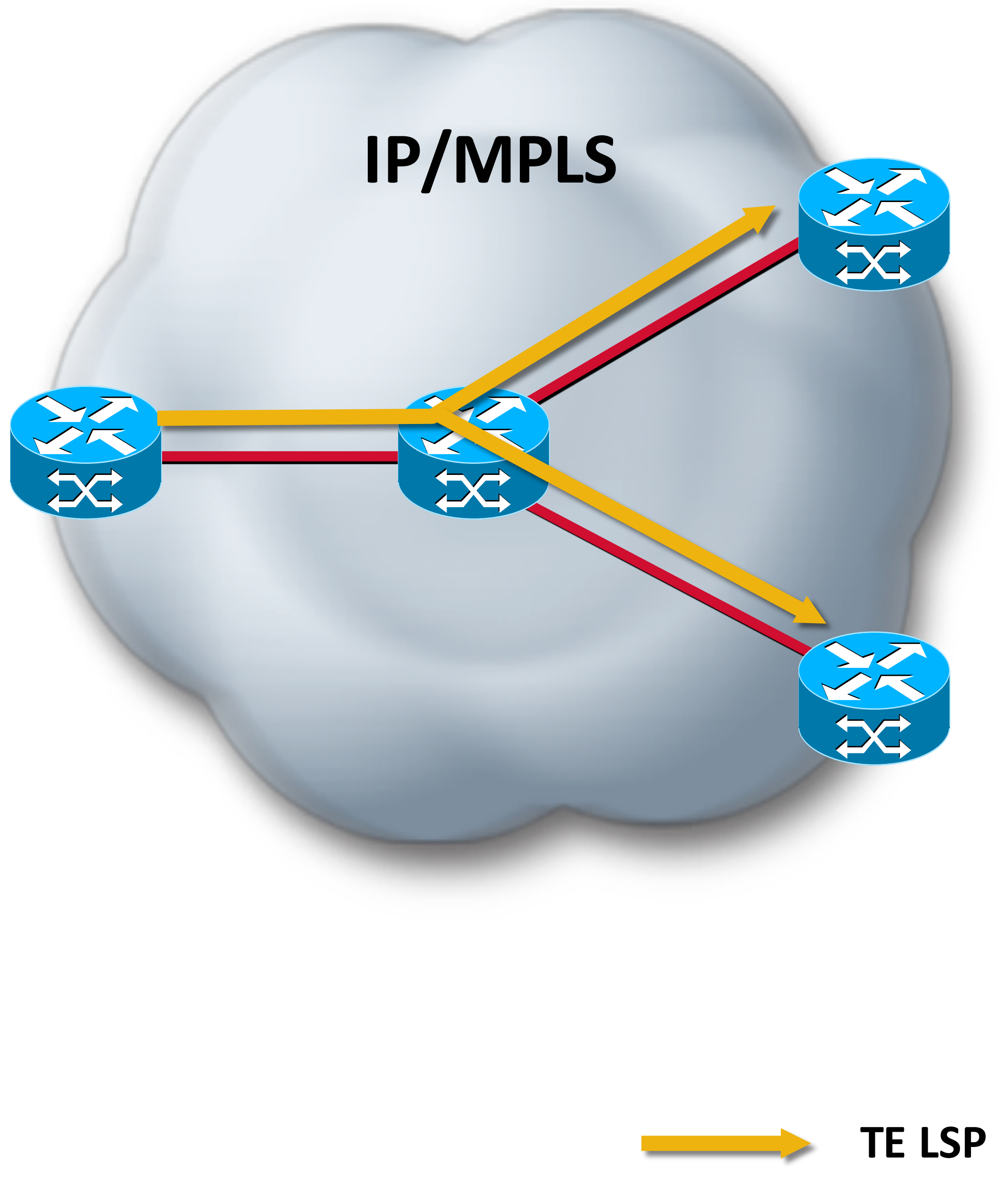
Configuring Tunnel at Head End (Cisco IOS XR)

```
explicit-path name PATH1
 index 1 next-address ipv4 unicast 172.16.0.4
 index 2 next-address ipv4 unicast 172.16.0.7
 index 3 next-address ipv4 unicast 172.16.4.2
!
interface tunnel-te1
 description FROM-ROUTER-TO-DST1
 ipv4 unnumbered Loopback0
 priority 5 5
 signalled-bandwidth 100000
 destination 172.16.255.2
 path-option 10 explicit name PATH1
 path-option 20 dynamic
 affinity f mask f
!
```

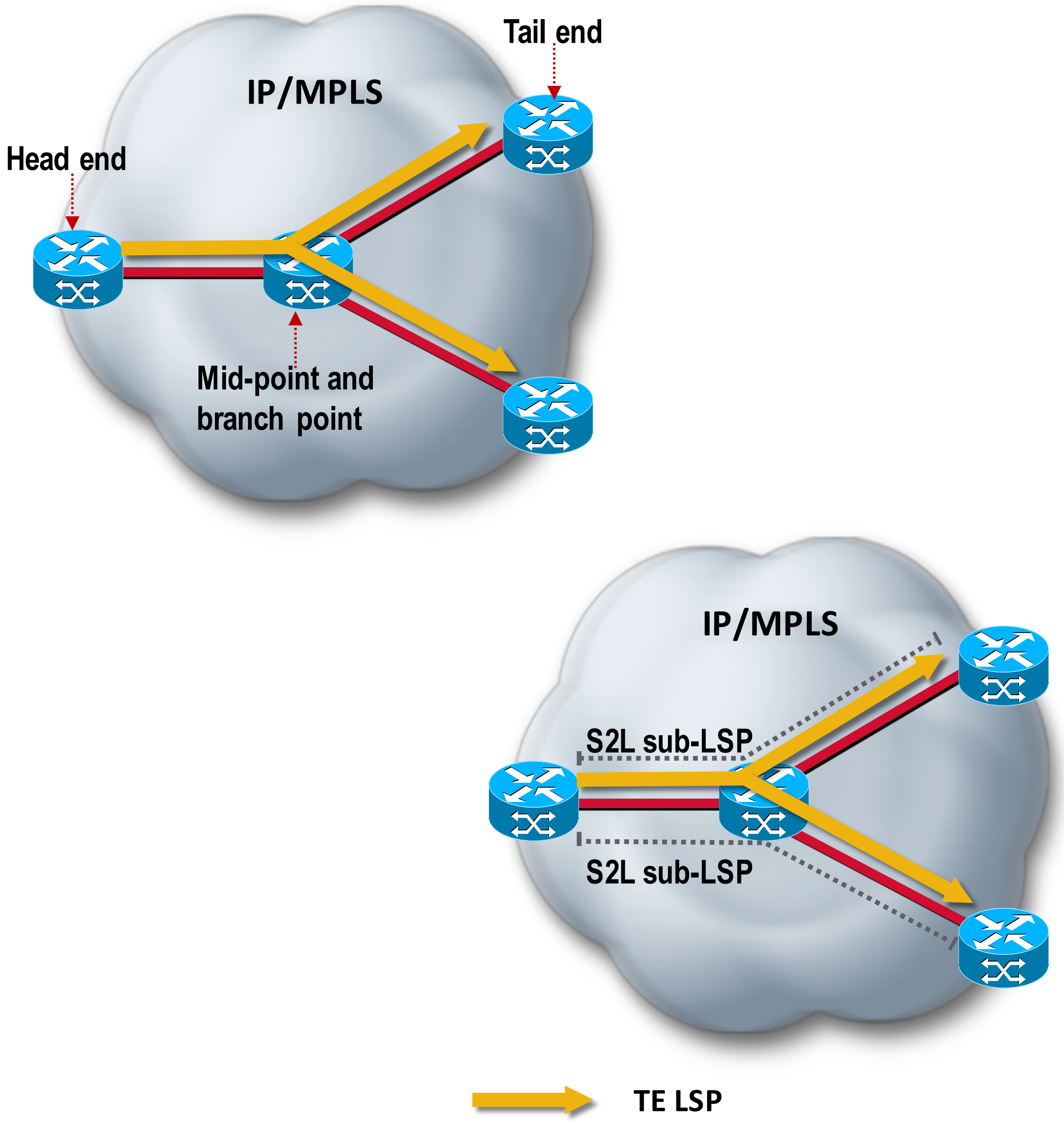
- Explicit PATH1 definition
- MPLS TE P2P tunnel
- Setup/hold priorities
- Signaled bandwidth
- Destination (tunnel tail end)
- Tunnel path options (PATH1, otherwise dynamic)
- Consider links with 0xF/0xF as attribute flags

Characteristics of P2MP TE LSP

- Unidirectional
- Explicitly routed
- One head end, but one or more tail ends (destinations)
- Same characteristics (constraints, protection, etc.) for all destinations



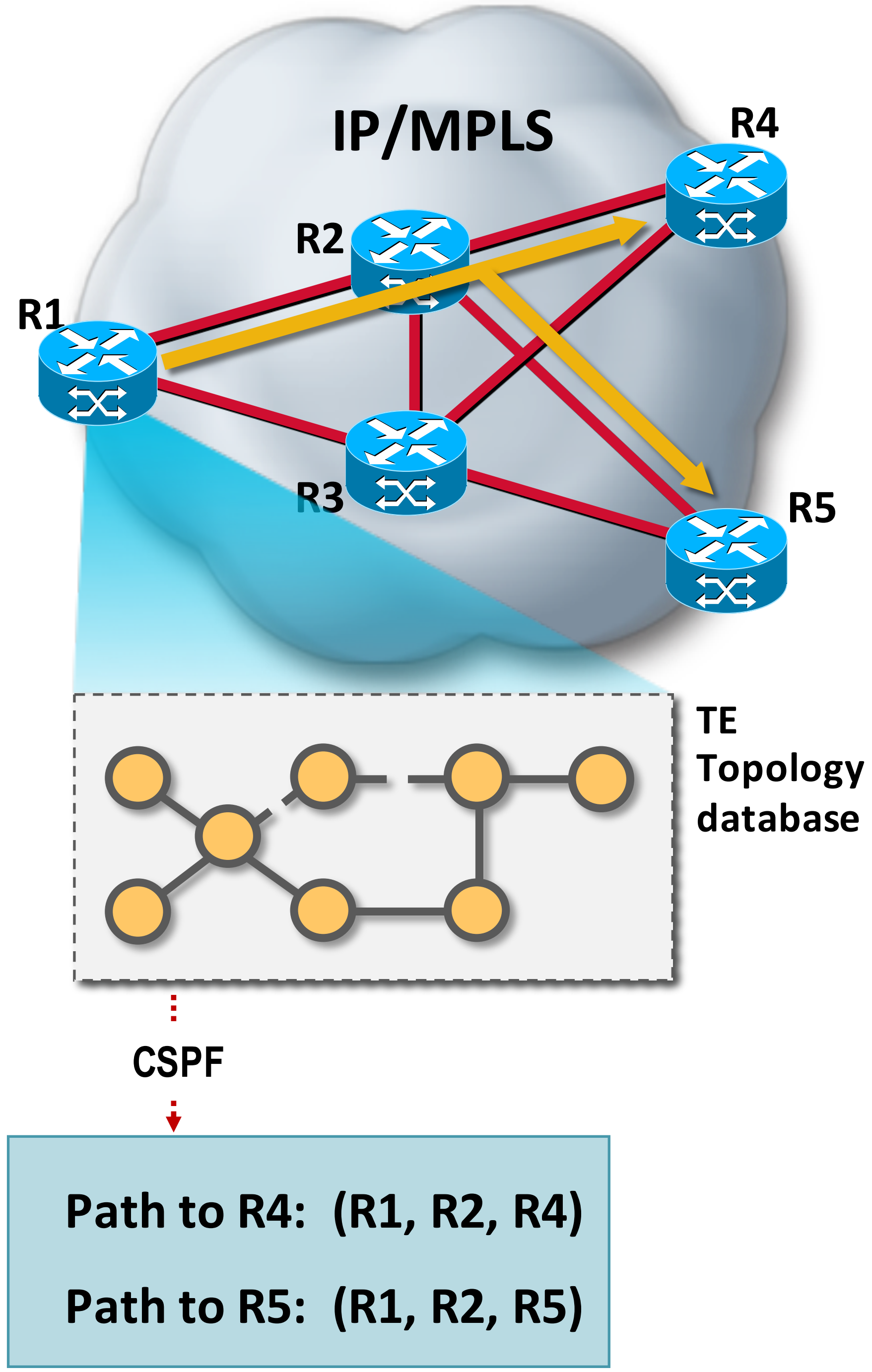
P2MP TE LSP Terminology



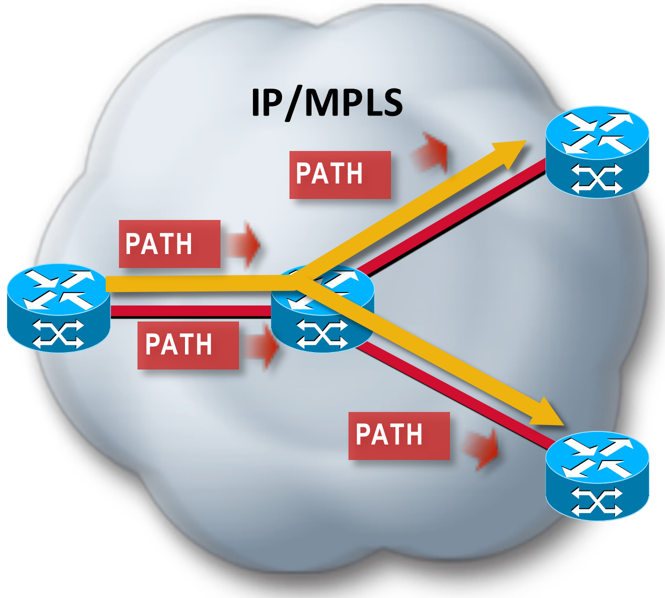
- Head-end/Source: Node where LSP signaling is initiated
- Mid-point: Transit node where LSP signaling is processed (not a head-end, not a tail-end)
- Tail-end/Leaf/destination: node where LSP signaling ends
- Branch point: Node where packet replication is performed
- Source-to-leaf (S2L) sub-LSP: P2MP TE LSP segment that runs from source to one leaf

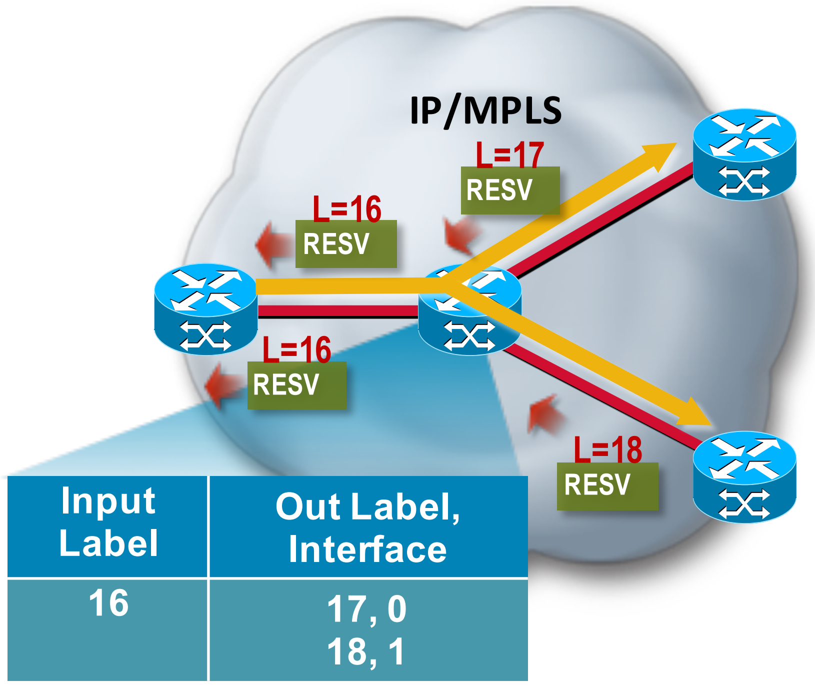
P2MP TE LSP Path Computation

- Constrained Shortest Path First (CSPF) used to compute an adequate tree
- CSPF executed per destination
- TE topology database and tunnel constraints used as input for path computation
- Path constraints may include loose, included, excluded hops
- Same constraints for all destinations (bandwidth, affinities, priorities, etc.)
- Path computation yields explicit path to each destination
- No changes to OSPF/IS-IS TE extensions
- Static paths possible with offline path computation



P2MP TE LSP Signaling

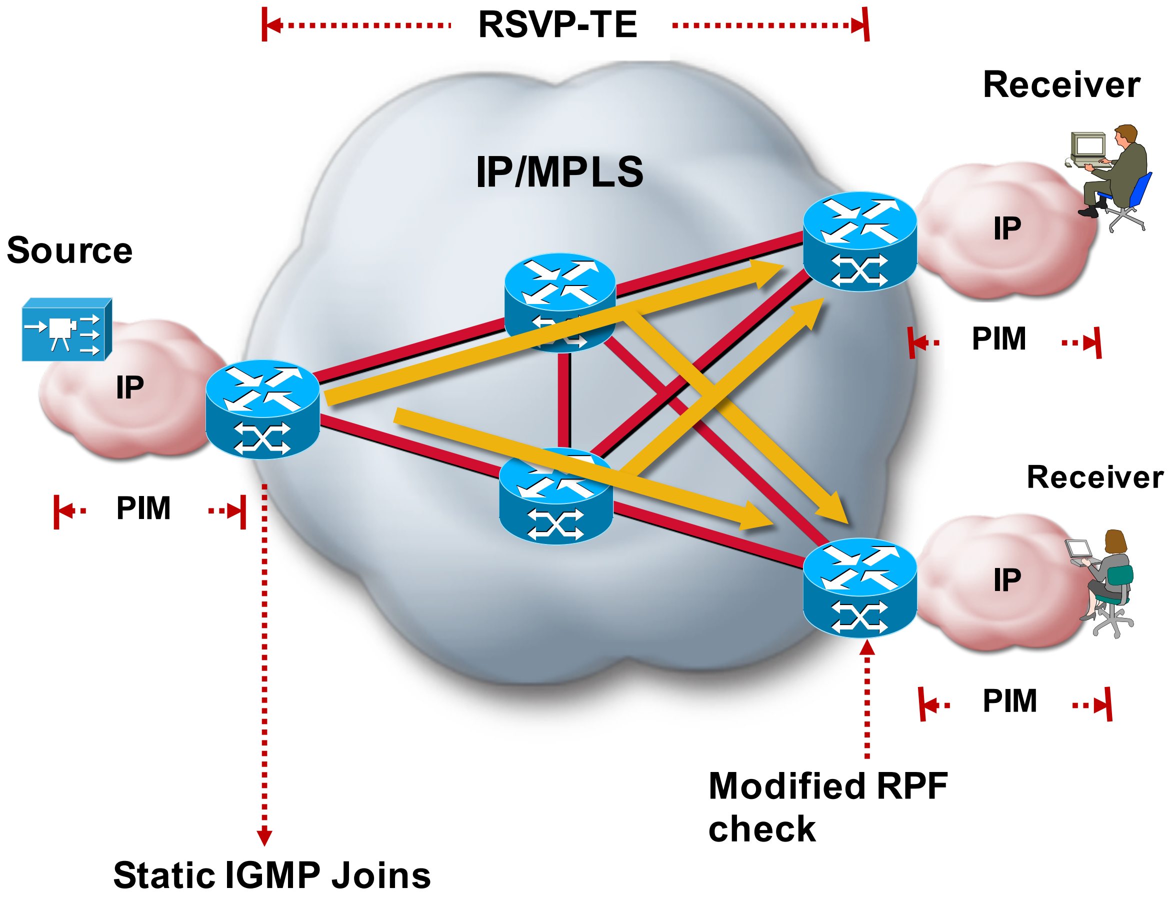




Input Label	Out Label, Interface
16	17, 0 18, 1

- Source sends unique PATH message per destination
- LFIB populated using RSVP labels allocated by RESV messages
- Multicast state built by reusing sub-LSP labels at branch points

P2MP TE LSP Traffic Selection IP Multicast



P2MP Tunnel	Multicast Group
Tunnel1	(192.168.5.1, 232.0.0.1) (192.168.5.1, 232.0.0.2)
Tunnel2	(192.168.5.1, 232.0.0.3)

- One or more IP multicast groups mapped to a Tunnel
- Groups mapped via static IGMP join
- PIM outside of MPLS network
- Modified egress RPF check against TE LSP and tunnel head end (source address)
- Egress node may abstract TE LSP as a virtual interface (LSPVIF) for RPF purposes

Configuring P2MP Tunnel at Head End (Cisco IOS)

```
mpls traffic-eng destination list name P2MP-LIST-DST1
 ip 172.16.255.1 path-option 10 explicit name PATH1
 ip 172.16.255.2 path-option 10 dynamic
 ip 172.16.255.3 path-option 10 dynamic
 ip 172.16.255.4 path-option 10 dynamic
!
interface Tunnel1
 description FROM-ROUTER-TO-LIST-DST1
 ip unnumbered Loopback0
 ip pim passive
 ip igmp static-group 232.0.0.1 source 192.168.5.1
 ip igmp static-group 232.0.0.2 source 192.168.5.1
 tunnel mode mpls traffic-eng point-to-multipoint
 tunnel destination list mpls traffic-eng name P2MP-LIST-DST1
 tunnel mpls traffic-eng priority 7 7
 tunnel mpls traffic-eng bandwidth 1000
!
```

Destination list with one path-option per destination

Tunnel as passive PIM interface

Multicast groups mapped to tunnel

P2MP TE Tunnel

Destination list

Setup/hold priorities

Signaled bandwidth

Configuring RPF Check at P2MP Tunnel Tail End (Cisco IOS)

```
ip multicast mpls traffic-eng
ip mroute 192.168.5.1 255.255.255.255 172.16.255.5
!
```

Enable IPv4 multicast over P2MP TE LSP

Tunnel source (172.16.255.5) as next-hop for IP Multicast source (192.168.5.1) RPF check

Configuring P2MP Tunnel at Head End (Cisco IOS XR)

```
interface tunnel-mte1
 ipv4 unnumbered Loopback0
 destination 172.16.255.129
  path-option 10 explicit name PATH1
  path-option 20 dynamic
 !
 destination 172.16.255.130
  path-option 10 dynamic
 !
 priority 0 0
 signalled-bandwidth 100000
!
node-capability label-switched-multicast
multicast-routing
 address-family ipv4
  interface tunnel-mte1
   enable
  !
  interface all enable
 !
!
router igmp
 vrf default
  interface tunnel-mte1
   static-group 232.0.0.1 192.168.5.1
   static-group 232.0.0.2 192.168.5.1
  !
 !
!
```

MPLS TE P2MP tunnel

Destination with path-option list

Destination with single path-option

Setup/hold priorities

Signaled bandwidth

Enable MPLS multicast

Enable multicast forwarding over tunnel-mte1

Multicast groups mapped to tunnel-mte1

Configuring RPF Check at P2MP Tunnel Tail End (Cisco IOS XR)

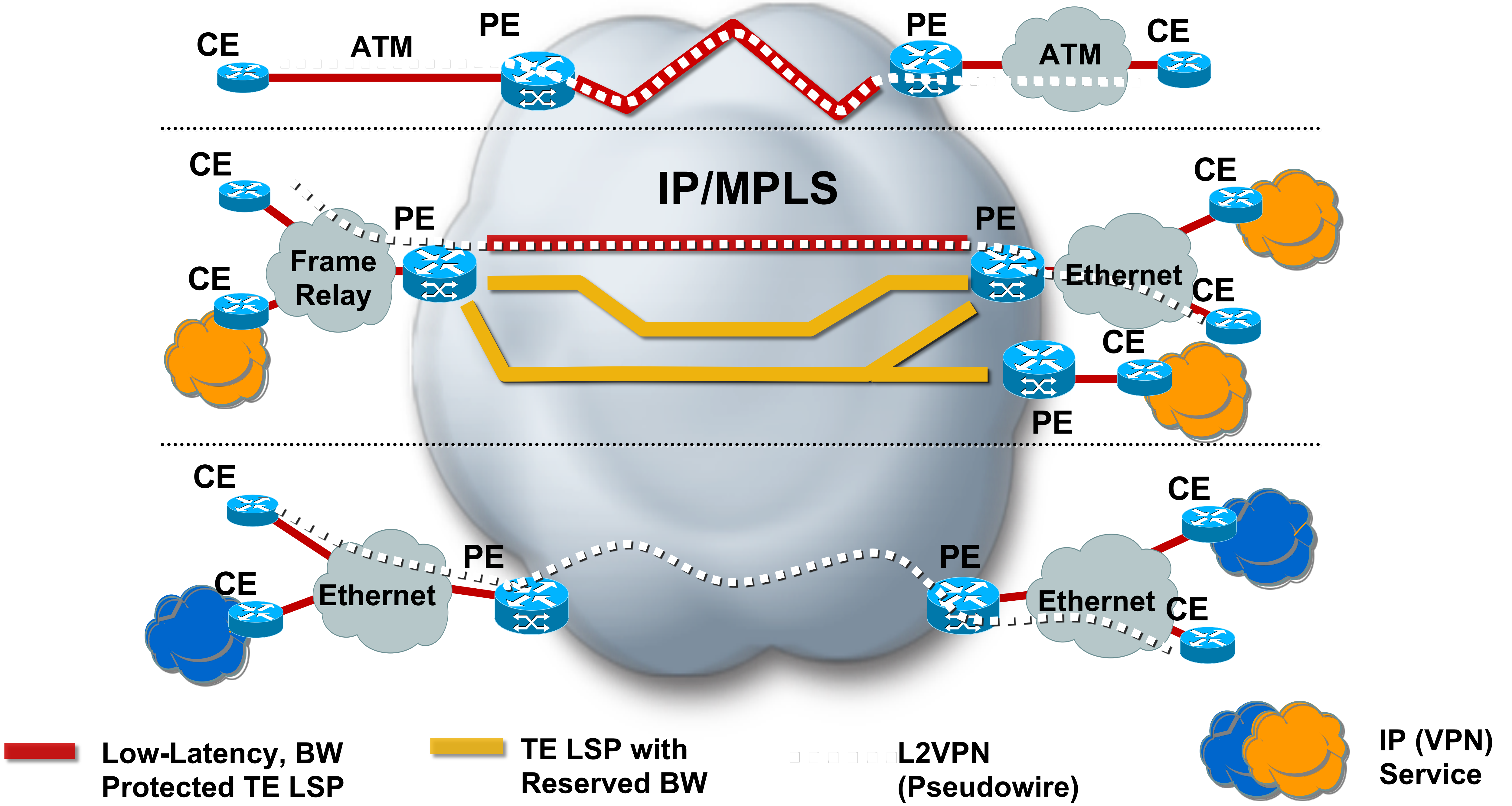
```
multicast-routing
 address-family ipv4
  core-tree-protocol rsvp-te
  static-rpf 192.168.5.1 32 mpls 172.16.255.3
  interface all enable
 !
!
```

Enable IPv4/v6 multicast over P2MP TE LSP

Tunnel source (172.16.255.3) as next-hop for IP Multicast source (192.168.5.1) RPF check

MPLS TE Integration with Network Services

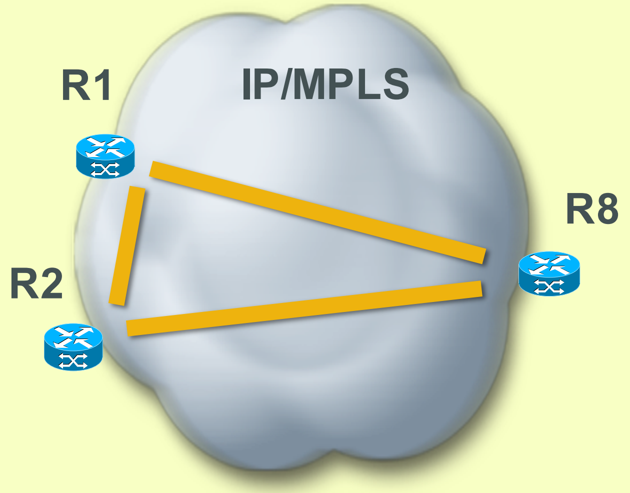
A TE LSP provides transport for different network services



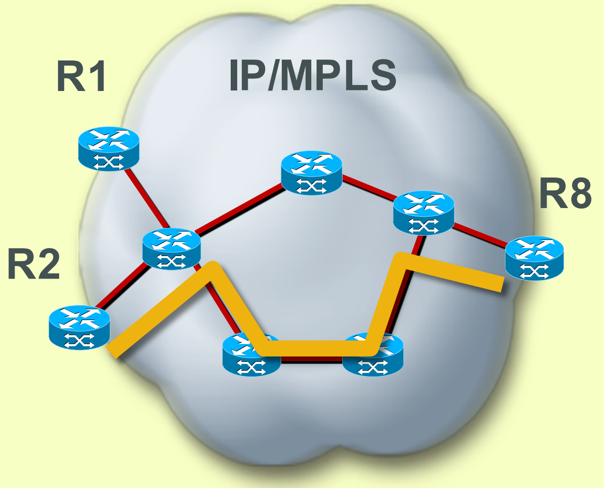
MPLS TE Deployment Models

Bandwidth Optimization

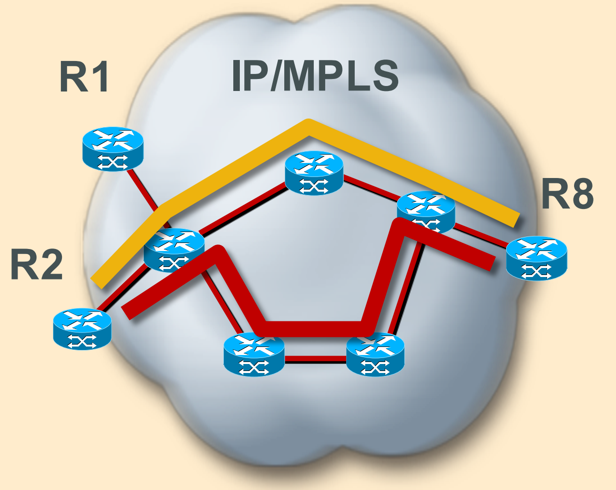
Planned



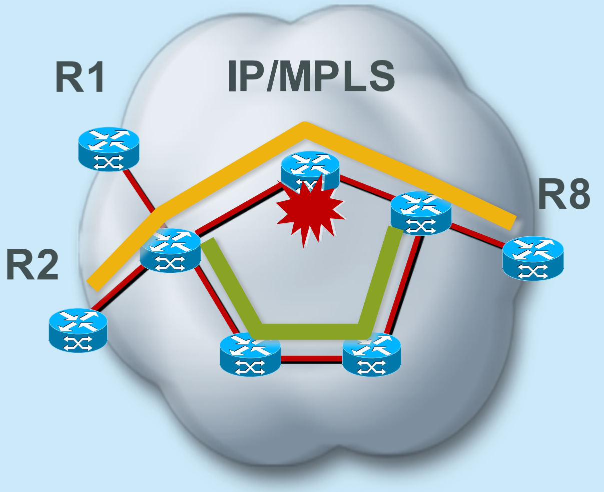
Reactive



Point-to-Point SLA



Protection



Bandwidth optimization

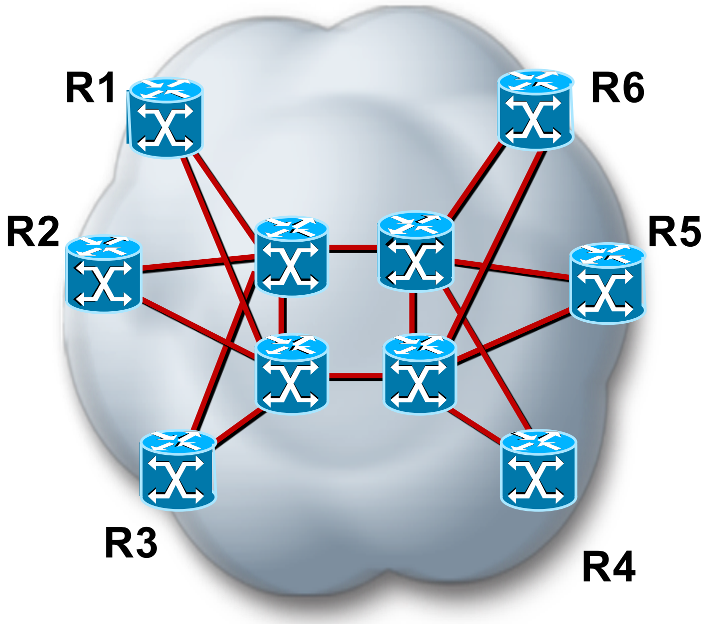
Planned Bandwidth Optimization

Traffic Matrix

	R1	R2	R3	R4	R5	R6
R1	4	7	1	5	4	5
R2	2	2	4	7	2	3
R3	1	2	9	5	5	5
R4	9	1	4	1	3	1
R5	3	7	9	2	7	7
R6	6	3	5	4	9	12

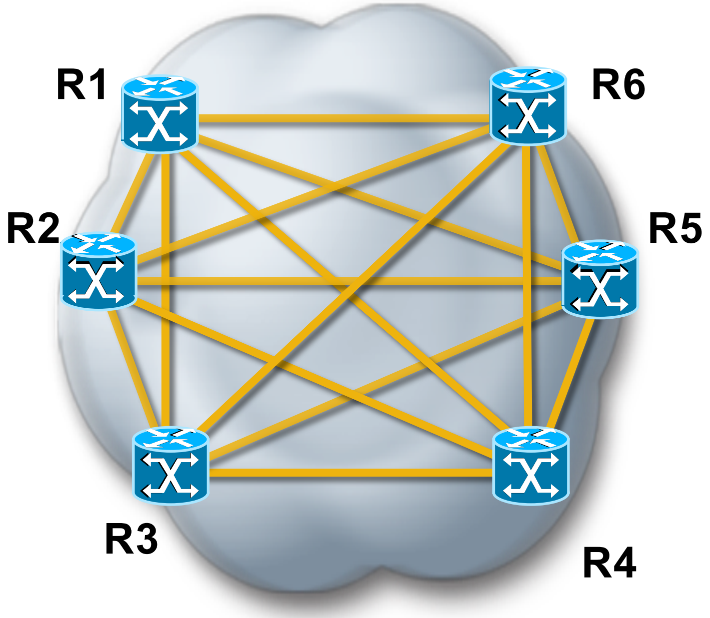
+

Physical Topology



=

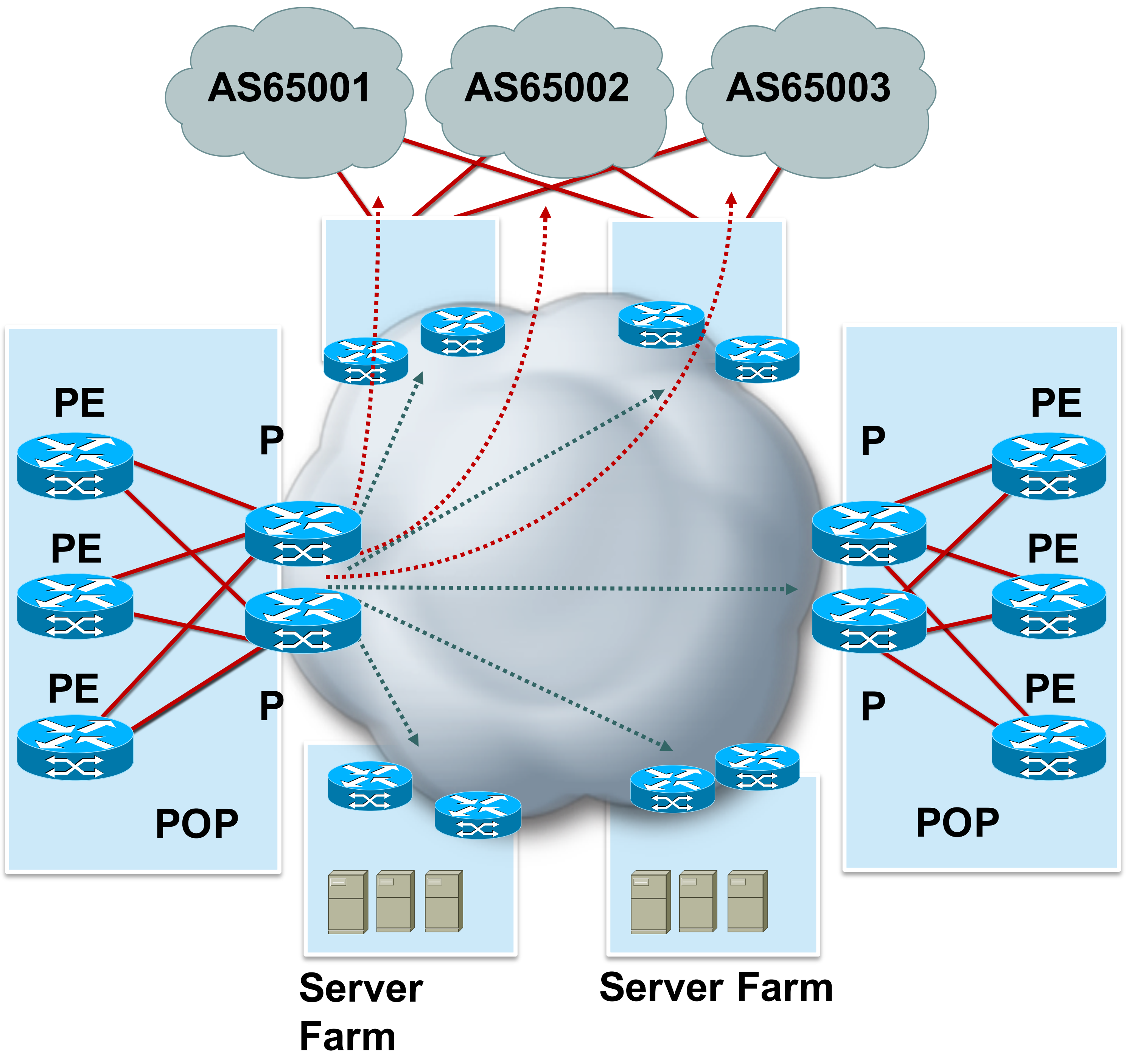
Tunnel mesh to satisfy traffic matrix



- Tries to optimize underlying physical topology based on traffic matrix
- Key goal is to avoid link over/under utilization
- On-line (CSPF) or off-line path computation
- May result in a significant number of tunnels
- Should not increase your routing adjacencies

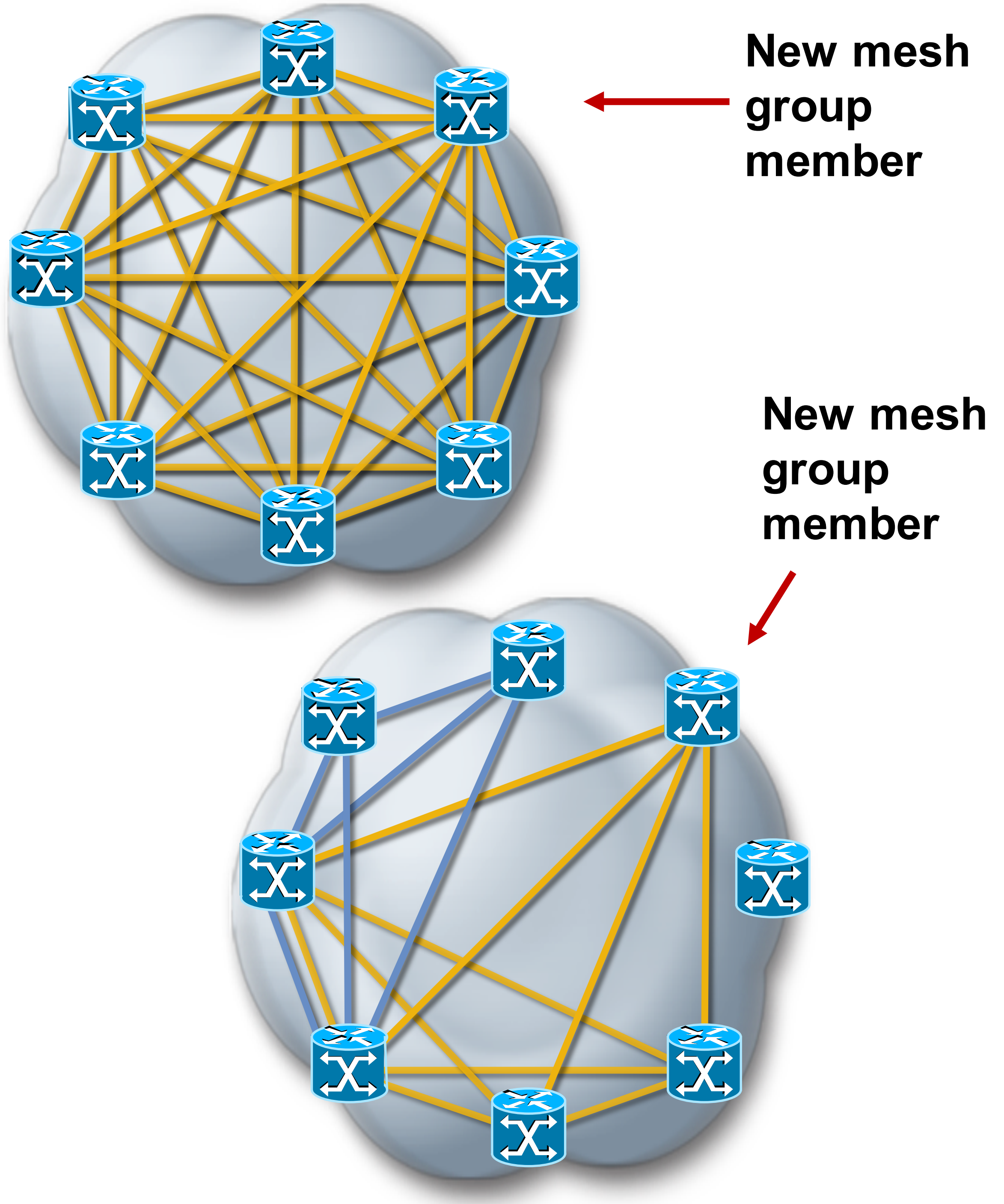
Traffic Matrix Measurement

- Interface counters on unconstrained tunnels
- Interface MIB
- MPLS LSR MIB
- NetFlow
  - NetFlow BGP Next Hop
  - MPLS-Aware NetFlow
  - Egress/Output NetFlow
- BGP policy accounting
  - Communities
  - AS path
  - IP prefix

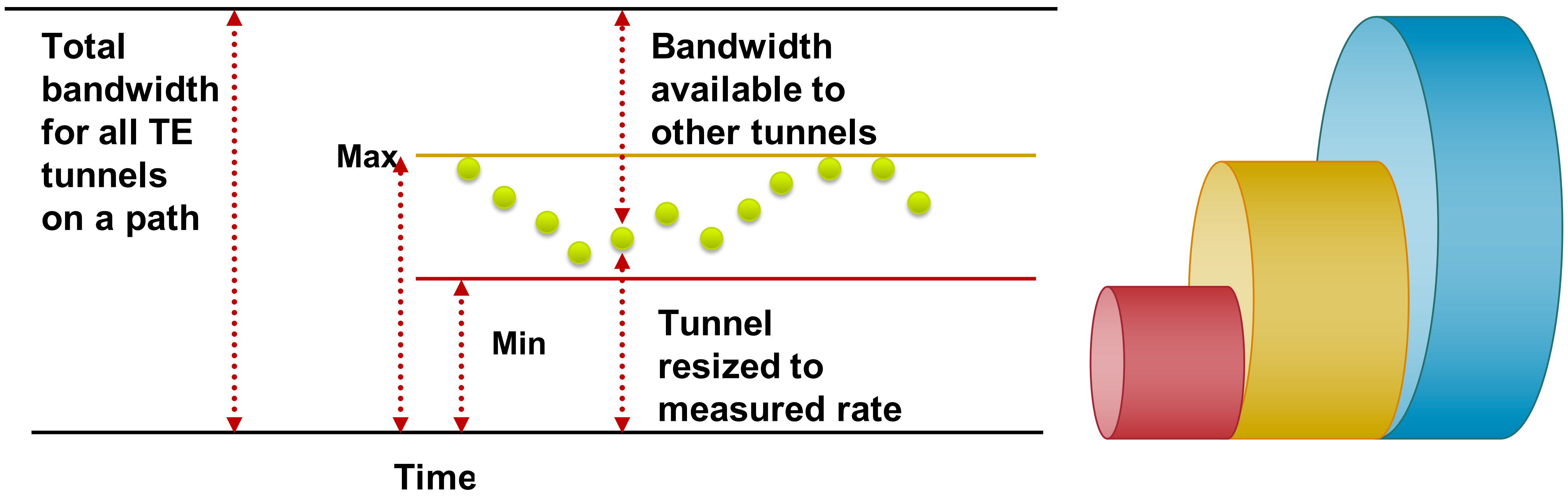


AutoTunnel Mesh

- Mesh group: LSRs to mesh automatically
- Membership identified by
  - Matching TE Router ID against ACL
  - IGP mesh-group advertisement
- Each member automatically creates tunnel upon detection of a member
- Tunnels instantiated from template
- Individual tunnels not displayed in router configuration



Auto Bandwidth



- Dynamically adjust bandwidth reservation based on measured traffic
- Optional minimum and maximum limits
- Sampling and resizing timers
- Tunnel resized to largest sample since last adjustment
- Actual resizing can be subject to adjustment threshold and overflow/underflow detection

Configuring AutoTunnel Mesh (Cisco IOS)

```
mpls traffic-eng tunnels
mpls traffic-eng auto-tunnel mesh
!
interface Auto-Template1
 ip unnumbered Loopback0
 tunnel destination mesh-group 10
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng autoroute announce
 tunnel mpls traffic-eng path-option 10 dynamic
 tunnel mpls traffic-eng auto-bw frequency 3600
!
router ospf 16
 log-adjacency-changes
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng area 0
 mpls traffic-eng mesh-group 10 Loopback0 area 0
 passive-interface Loopback0
 network 172.16.0.0 0.0.255.255 area 0
!
```

Enable Auto-tunnel Mesh

Tunnel template

Template cloned for each member of mesh group 10

Dynamic (CSPF) path to each mesh group member

Tunnels will adjust bandwidth reservation automatically

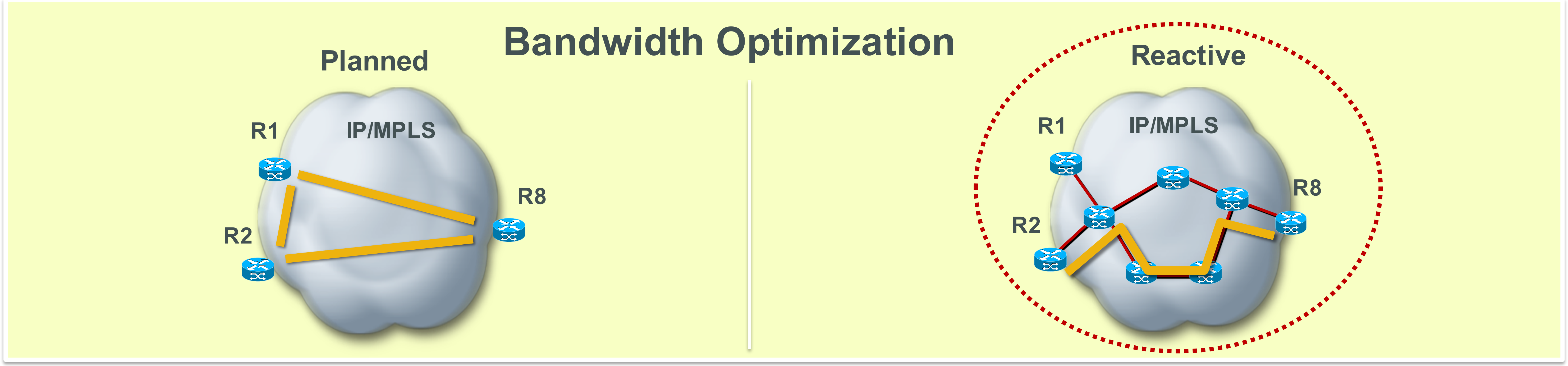
Advertise mesh group 10 membership in area 0

Configuring AutoTunnel Mesh (Cisco IOS XR)

```
mpls traffic-eng
 auto-tunnel mesh
  group 10
   attribute-set 10
   destination-list DST-RID-ACL
  !
  tunnel-id min 1000 max 2000
 !
 attribute-set auto-mesh 10
  autoroute announce
  auto-bw collect-bw-only
 !
!
```

- Enable Auto-tunnel Mesh
- Mesh group 10
- Attribute set to use
- ACL matching matching TE router ids associated with mesh
- Range of dynamically created tunnel interfaces
- Attribute set definition

Reactive Bandwidth Optimization



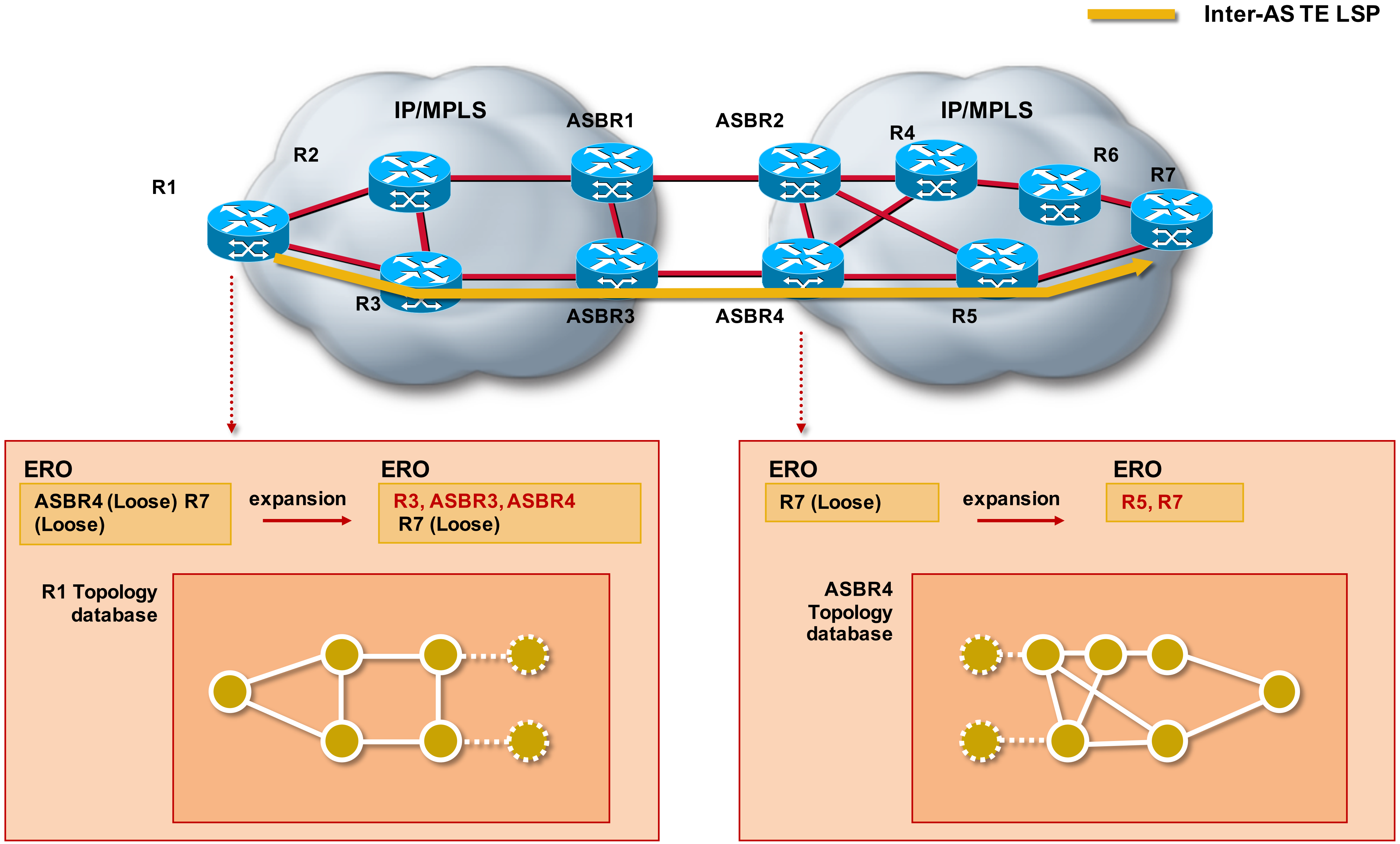
- Selective deployment of tunnels when highly-utilized links are identified
- Generally, deployed until next upgrade cycle alleviates congested links

Inter-Domain Traffic Engineering

Inter-Domain Traffic Engineering: Introduction

- Domain defined as an IGP area or autonomous system
- Head end lacks complete network topology to perform path computation in both cases
- Two path computation approaches
  - Per-domain (ERO loose-hop expansion)
  - Distributed (Path Computation Element)

Per-Domain Path Computation Using ERO Loose-hop Expansion



Configuring Inter-Area Tunnels (Cisco IOS)

```
mpls traffic-eng tunnels
!
interface Tunnel1
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 172.16.255.7
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng path-option 10 explicit name LOOSE-PATH
!
ip route 172.16.255.7 255.255.255.255 Tunnel1
!
ip explicit-path name LOOSE-PATH enable
 next-address loose 172.16.255.3
 next-address loose 172.16.255.5
!
```

Loose-hop path

Static route mapping IP traffic to Tunnel1

List of ABRs as loose hops

Configuring Inter-Area Tunnels with Autoroute Destinations (Cisco IOS)

```
interface Tunnel1
 ip unnumbered Loopback0
 tunnel mode mpls traffic-eng
 tunnel destination 172.16.255.7
 tunnel mpls traffic-eng autoroute destination
 tunnel mpls traffic-eng path-option 10 explicit name LOOSE-PATH
!
ip explicit-path name LOOSE-PATH enable
 next-address loose 172.16.255.3
 next-address loose 172.16.255.5
!
```

Create static route to tunnel destination (172.16.255.7)

Loose-hop path

List of **ABRs** as loose hops

Configuring Inter-Area Tunnels (Cisco IOS XR)

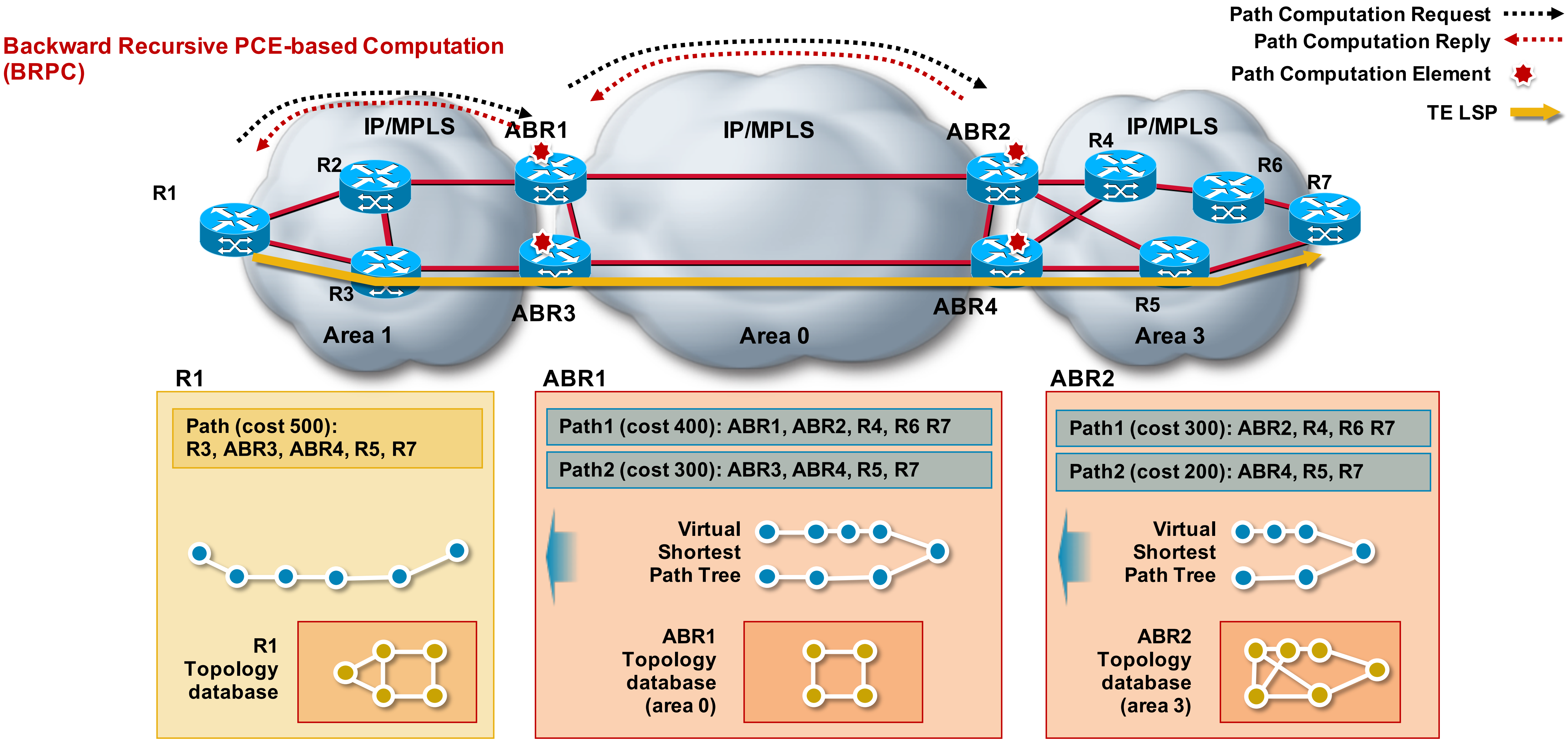
```
explicit-path name LOOSE-PATH
 index 1 next-address loose ipv4 unicast 172.16.255.129
 index 2 next-address loose ipv4 unicast 172.16.255.131
!
interface tunnel-te1
 description FROM-ROUTER-TO-DST3
 ipv4 unnumbered Loopback0
 destination 172.16.255.2
 path-option 10 explicit name LOOSE-PATH
!
router static
 address-family ipv4 unicast
  172.16.255.2/32 tunnel-te1
 !
```

List of ABRs as loose hops

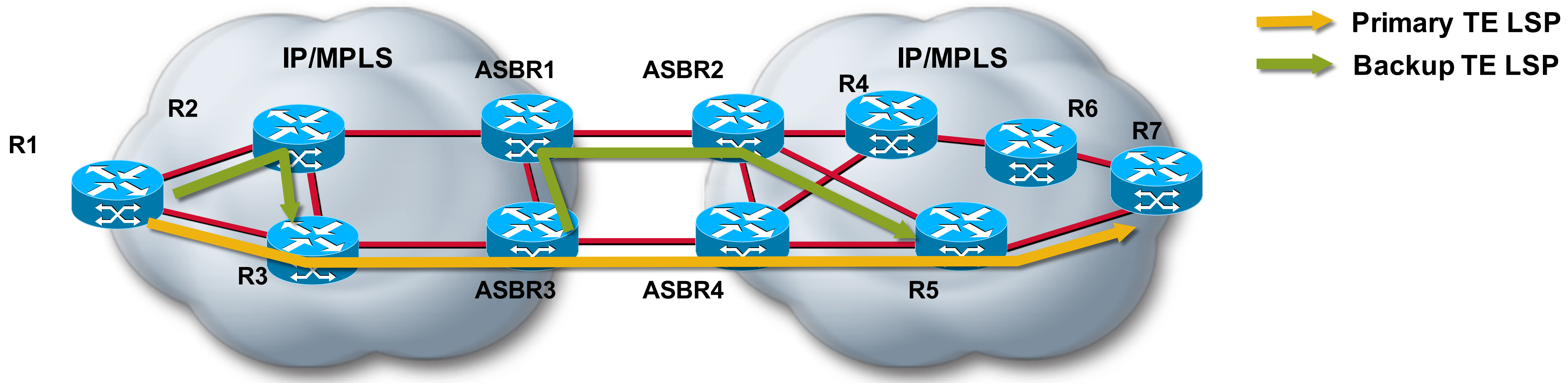
Loose-hop path

Static route mapping IP traffic to tunnel-te1

Distributed Path Computation using Path Computation Element



Inter-Domain TE – Fast Re-route



- Same configuration as single domain scenario
- Support for node-id sub-object required to implement ABR/ASBR node protection
- Node-id helps point of local repair (PLR) detect a merge point (MP)

Inter-Domain TE
Take into Account before Implementing

- Semantics of link attributes across domain boundaries
- Semantics of TE-Classes across domain boundaries for DS-TE
- Auto-route destinations creates a static route to tunnel destination and facilitates traffic selection
- Auto-route announce not applicable for traffic selection

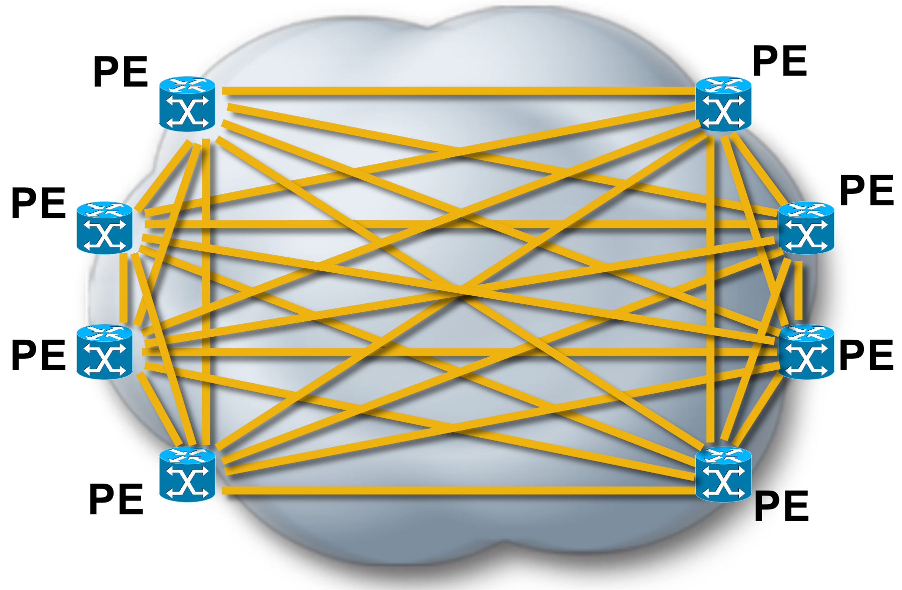
General Deployment Considerations

Should RSVP-TE and LDP be Used Simultaneously?

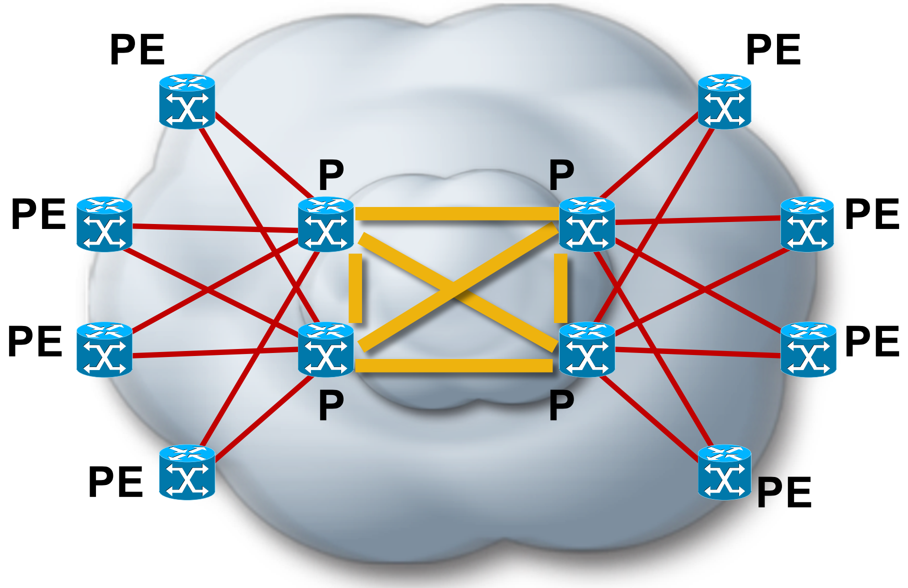
- Guarantees forwarding of VPN traffic if a TE LSP fails
- May be required if full mesh of TE LSPs not in use
- Increased complexity

How Far should Tunnels Span?

56 TE LSP



12 TE LSP



- PE-to-PE Tunnels

  More granular control on traffic forwarding

  Larger number of TE LSPs

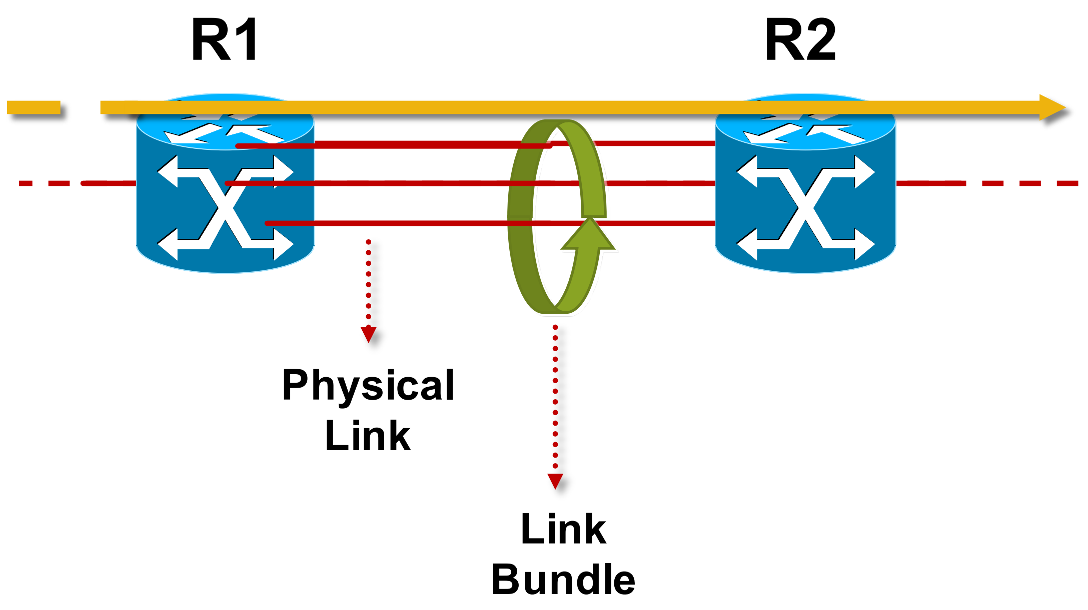
- P-to-P Tunnels

  Requires IP tunnels or LDP over TE tunnels to carry VPN traffic (deeper label stack)

  Fewer TE LSPs

  May be extended with PE-P tunnels

MPLS TE on Link Bundles



- Different platforms support different link bundles
  - Ethernet
  - POS
  - Multilink PPP
- Bundles appear as single link in topology database
- Same rules for link state flooding
- LSP preemption if bundle bandwidth becomes insufficient
- Configurable minimum number of links to maintain bundle active
- Bundle failure can act as trigger for FRR

Scaling Signaling (Refresh Reduction)

SRefresh Message

	MSG_Id	Path State
LSP1	22	…
LSP2	62	…
:	:	…
LSPn	94	…

	MSG_Id	Resv State
LSP1	43	…
LSP2	37	…
:	:	…
LSPn	29	…

- RSVP soft state needs to be refreshed periodically
- Refresh reduction extensions use message Identifier associated with Path/Resv state
- Summary Refresh (SRefresh) message refreshes state using a message_id list
- SRefresh only replaces refresh Path/Resv messages

Configuring Refresh Reduction (Cisco IOS)

```
mpls traffic-eng tunnels
!
interface TenGigabitEthernet0/1/0
 ip address 172.16.0.0 255.255.255.254
 mpls traffic-eng tunnels
 ip rsvp bandwidth 100000
!
router ospf 100
 log-adjacency-changes
 passive-interface Loopback0
 network 172.16.0.0 0.0.255.255 area 0
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng area 0
!
ip rsvp signalling refresh reduction
!
```

Enable refresh reduction

*** Enabled by default in Cisco IOS XR**

BUILT FOR THE HUMAN NETWORK

CISCO

