Deploy MPLS Traffic Engineering

APNIC Technical Workshop

APNIC

Acknowledgement

- Cisco Systems

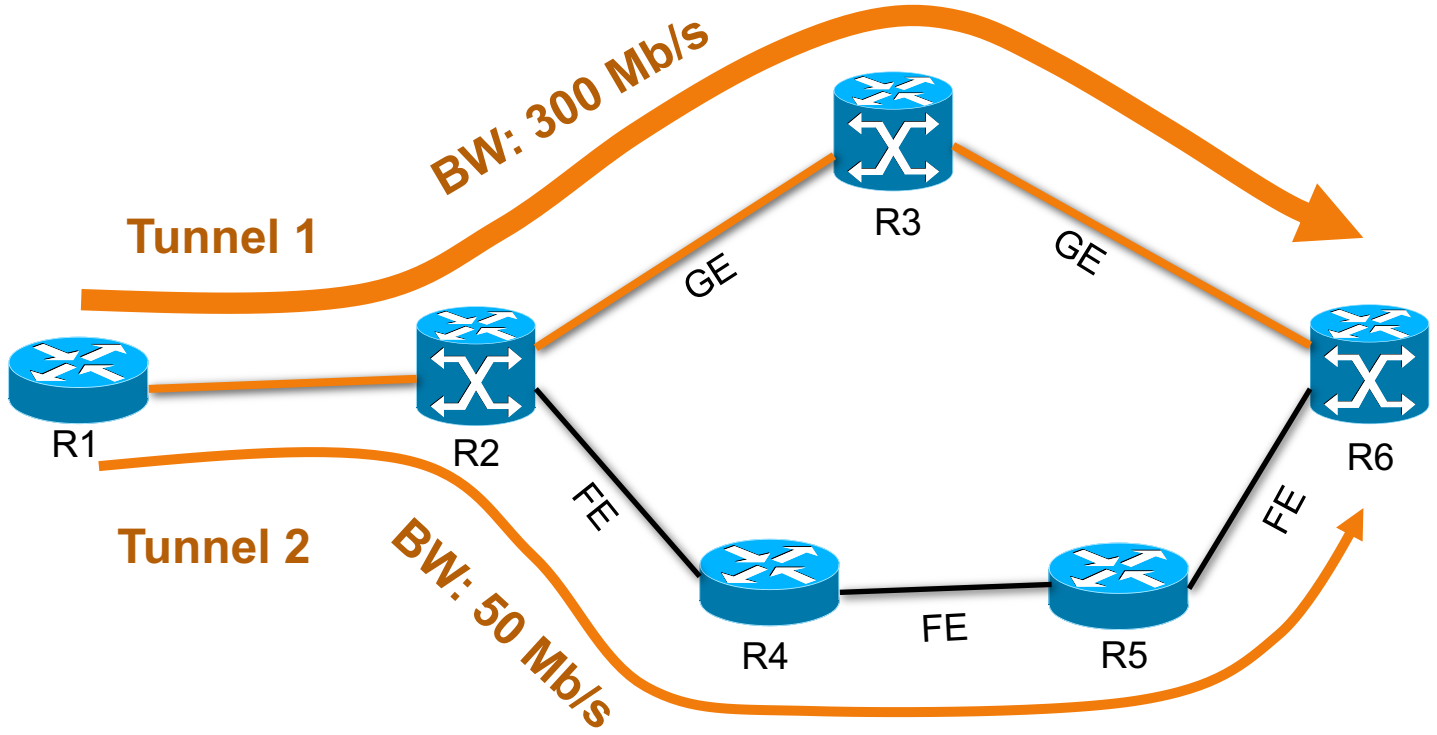
Overview of MPLS TE

APNIC

Why MPLS Traffic Engineering?

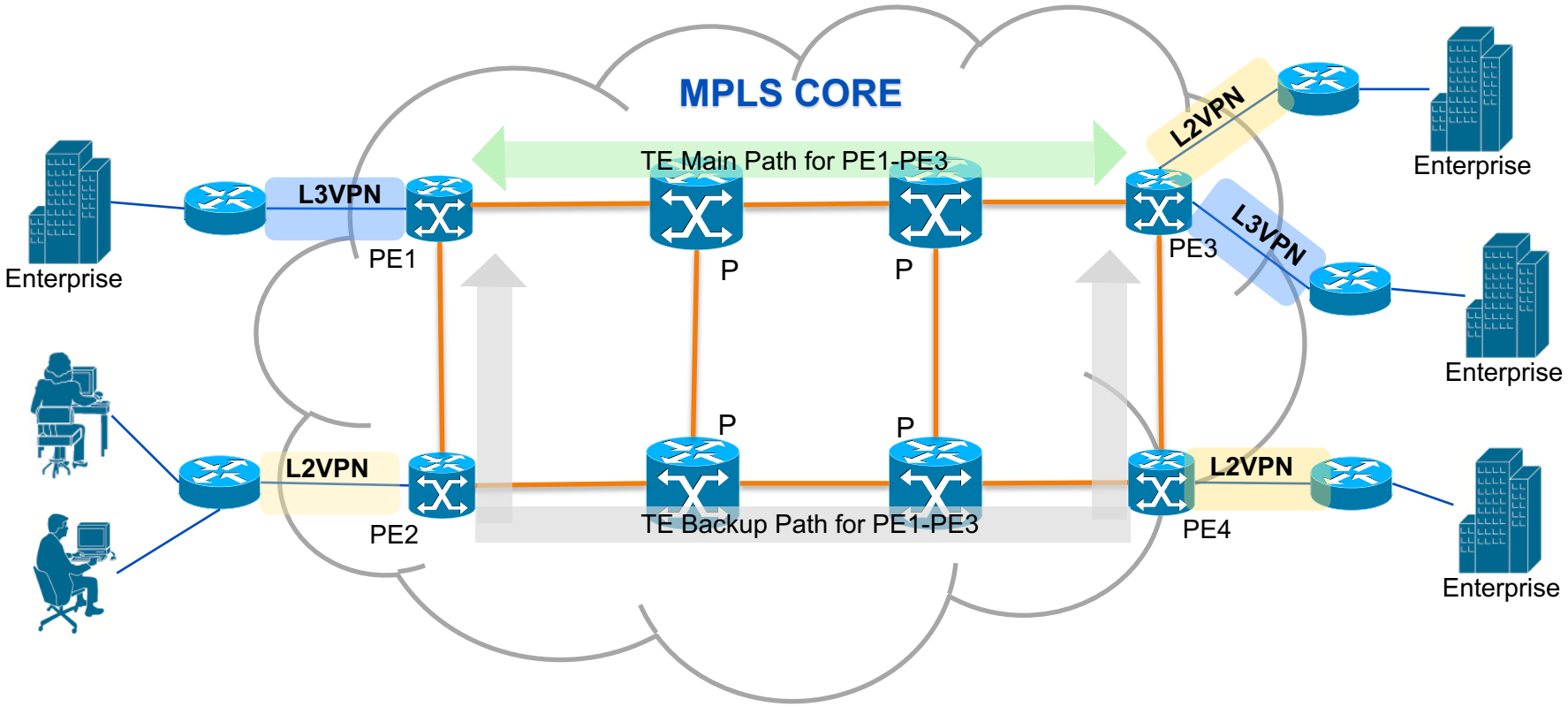
- Handling unexpected congestion
- Better utilization of available bandwidth
- Route around failed links/nodes
- Capacity planning

Optimal Traffic Engineering



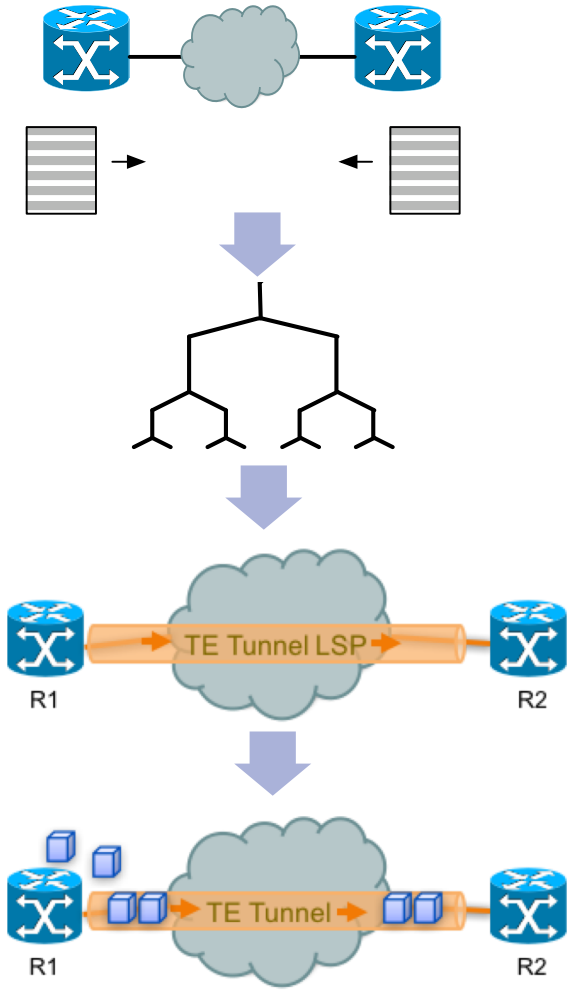
IP TE	MPLS TE
Shortest path	Determines the path at the source based on additional parameters (available resources and constraints, etc.)
Equal cost load balancing	Load sharing across unequal paths can be achieved.

MPLS Application Scenario

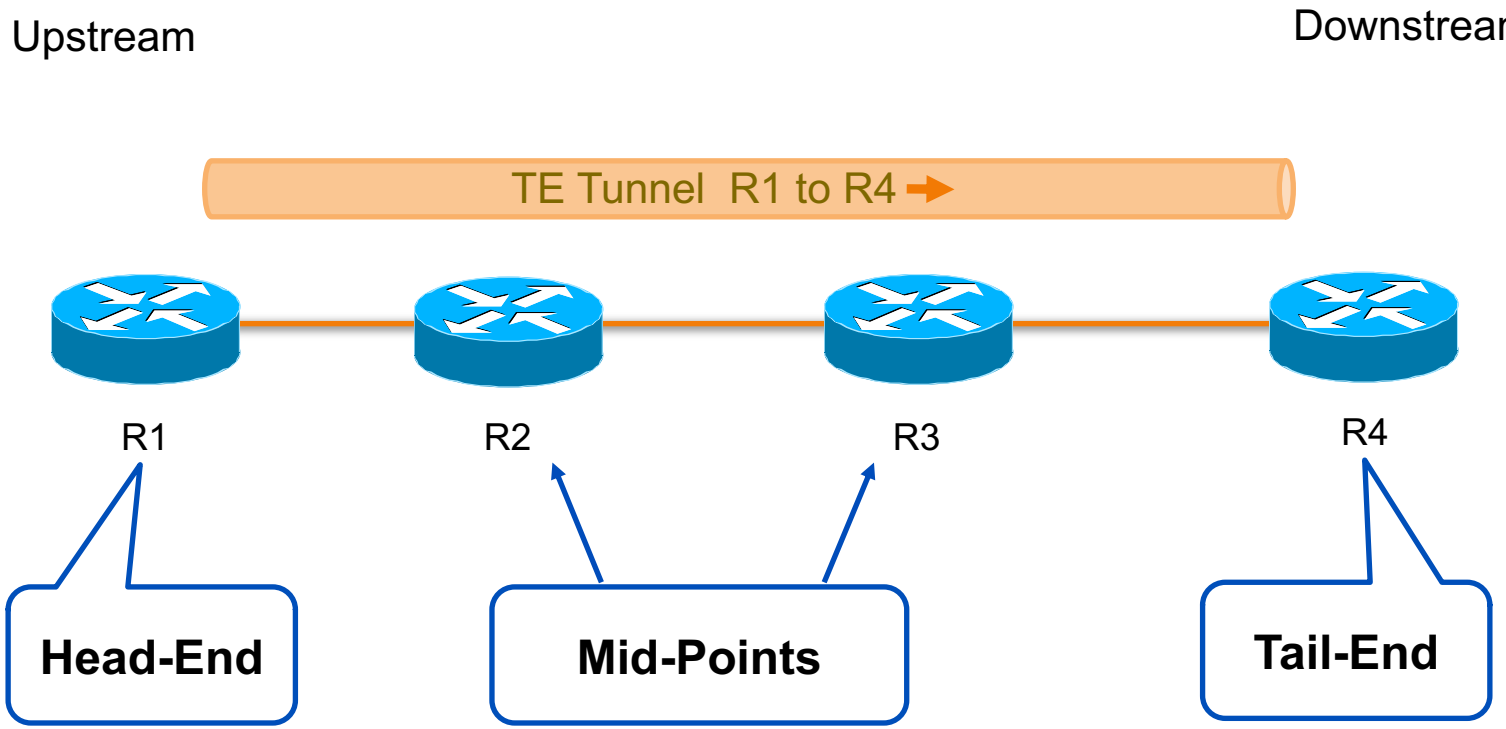


How MPLS TE Works

- **Information Distribution**
  - What is the information?
- **Path Calculation**
  - Dynamically
  - Manually
- **Path Setup**
  - RSVP-TE
  - (CR-LDP)
- **Forward the Traffic Down to the Tunnel**
  - Autoroute
  - Static
  - Policy



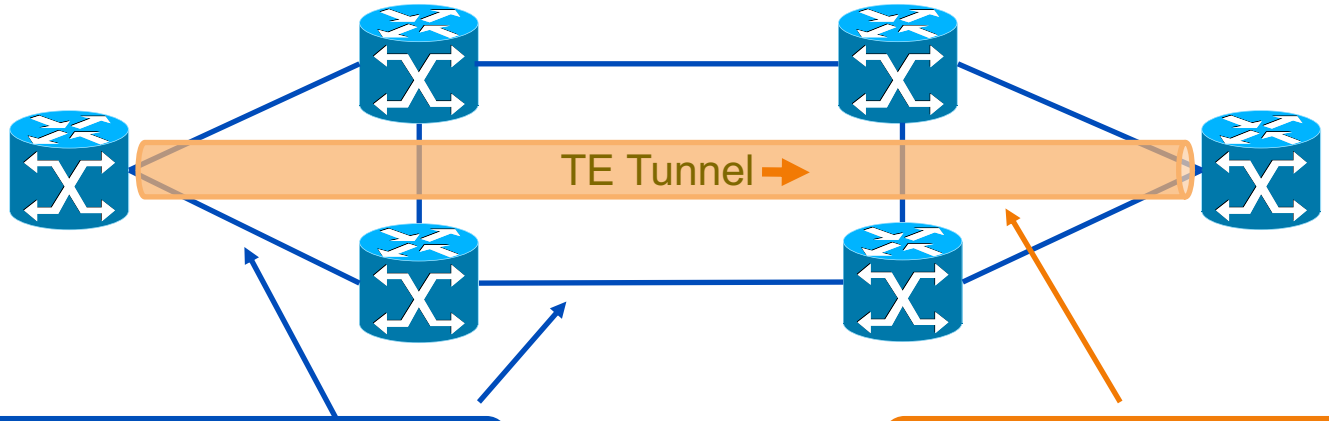
Terminology—Head, Tail, LSP



Information Distribution

APNIC

Attributes



Link Attributes

- Available Bandwidth
- Attribute flags (Link Affinity)
- Administrative weight (TE-specific link metric)

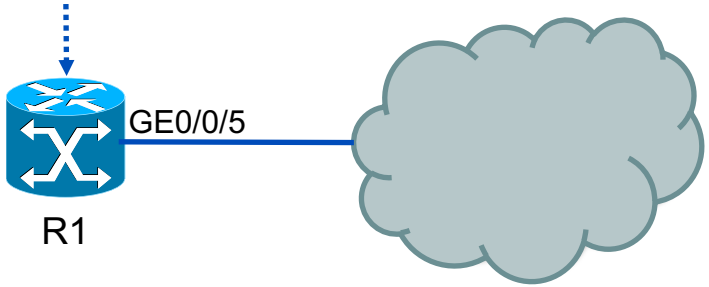
Tunnel Attributes

- Tunnel Required Bandwidth
- Tunnel Affinity & Mask
- Priority

Bandwidth on Physical Link

- Bandwidth – the amount of **reservable bandwidth** on a physical link

```
R1(Config)# interface gigabitethernet 0/0/5
R1(config-if)# mpls traffic-eng tunnels
R1(config-if)# ip rsvp bandwidth 512
```



Reserved bandwidth is 512 kbps

Bandwidth Required by Tunnel

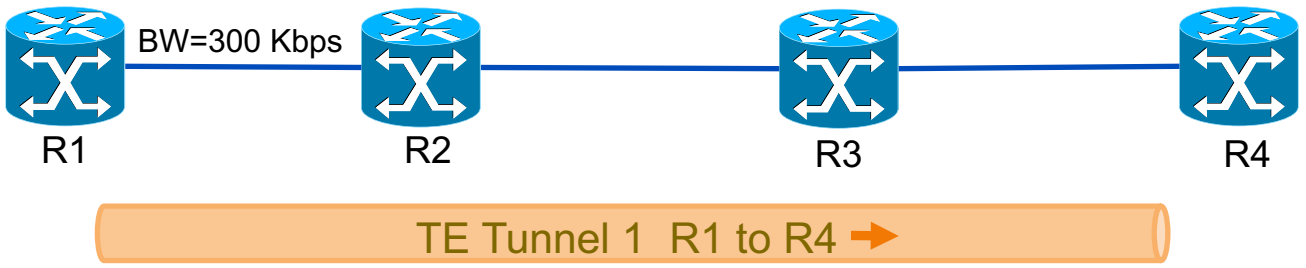
- **Bandwidth required** by the tunnel across the network

```
R1(config)# interface tunnel 1
R1(config-if)# tunnel mpls traffic-eng bandwidth 100
```

- Not a mandatory command. If not configured, tunnel is requested with zero bandwidth.

Priority

- Each tunnel has 2 priorities:
  - Setup priority
  - Holding priority
- Value: 0~7, 0 indicated the highest priority, 7 indicates the lowest priority.



BW=200 kbps Priority: S_1=3 **H_1=3**

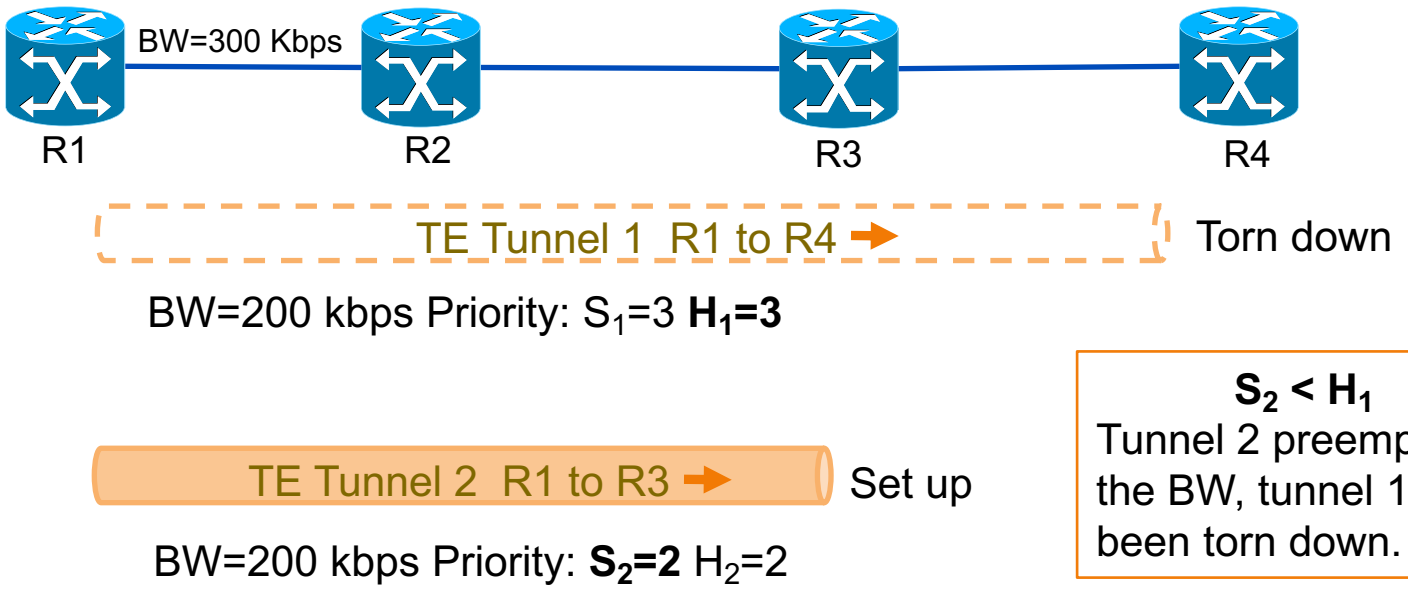
$S_2 < H_1$

TE Tunnel 2 R1 to R3

BW=200 kbps Priority: **S_2=2** H_2=2

Priority

- Each tunnel has 2 priorities:
  - Setup priority
  - Holding priority
- Value: 0~7, 0 indicated the highest priority, 7 indicates the lowest priority.



Configure Priority

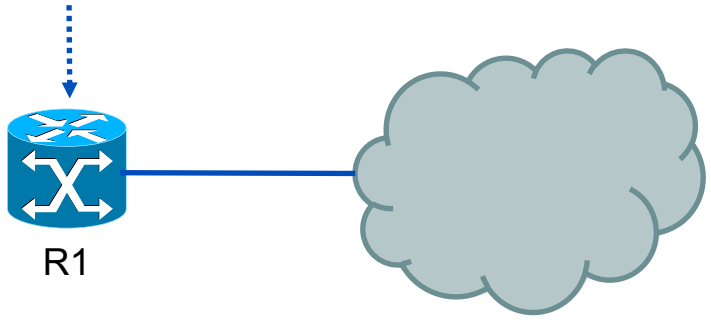
```
R1(config)# interface tunnel 1
R1(config-if)# tunnel mpls traffic-eng priority 3 3
```

- Recommended that priority S=H; if a tunnel can setup at priority “X”, then it should be able to hold at priority “X” too!
- Configuring S > H is illegal; tunnel will most likely be preempted
- Default is S = 7, H = 7

Attribute Flags

- An **attribute flag** is a 32-bit bitmap on a link that indicate the existence of up to 32 separate properties on that link, also called link affinity or administrative group.

```
R1(config)# interface ethernet 0/1
R1(config-if)# mpls traffic-eng attribute-flags 0x8
```



Tunnel Affinity & Mask

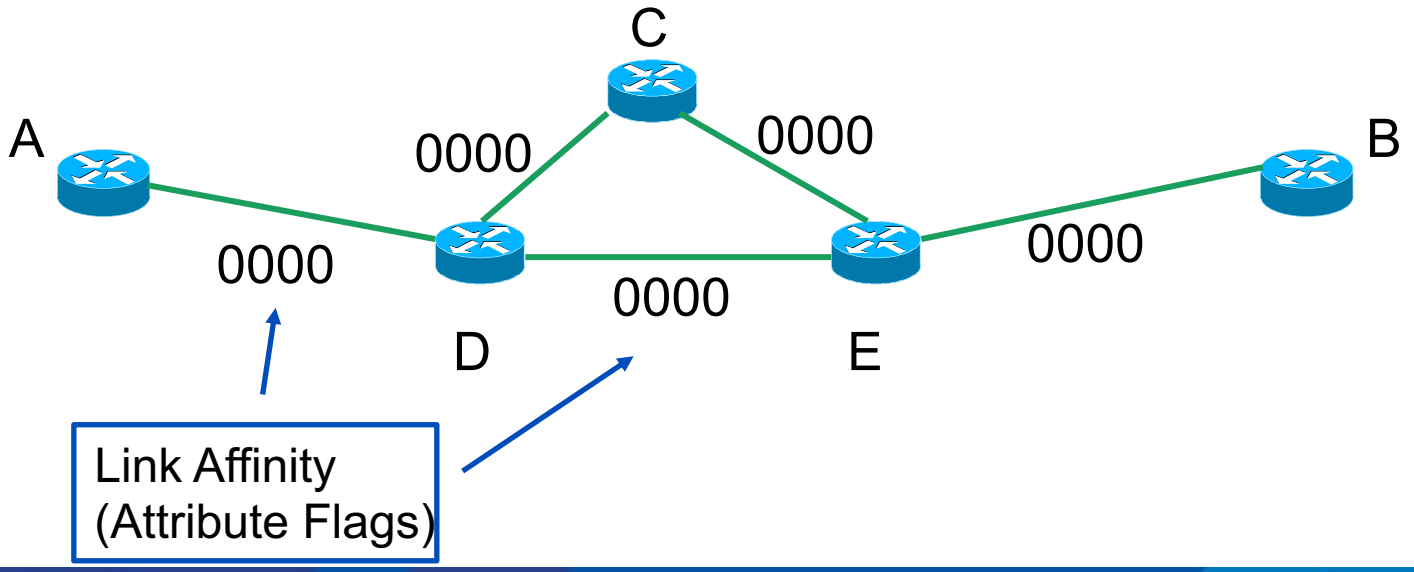
- **Tunnel affinity** helps select which tunnels will go over which links.
- For example: a network with OC-12 and Satellite links will use affinities to prevent tunnels with VoIP traffic from taking the satellite links.

Tunnel can only go over a link if
(Tunnel Mask) && (Attribute Flags) == Tunnel Affinity

```
R1(config)# interface tunnel 1
R1(config-if)# tunnel mpls traffic-eng affinity 0x80 mask 0x80
```

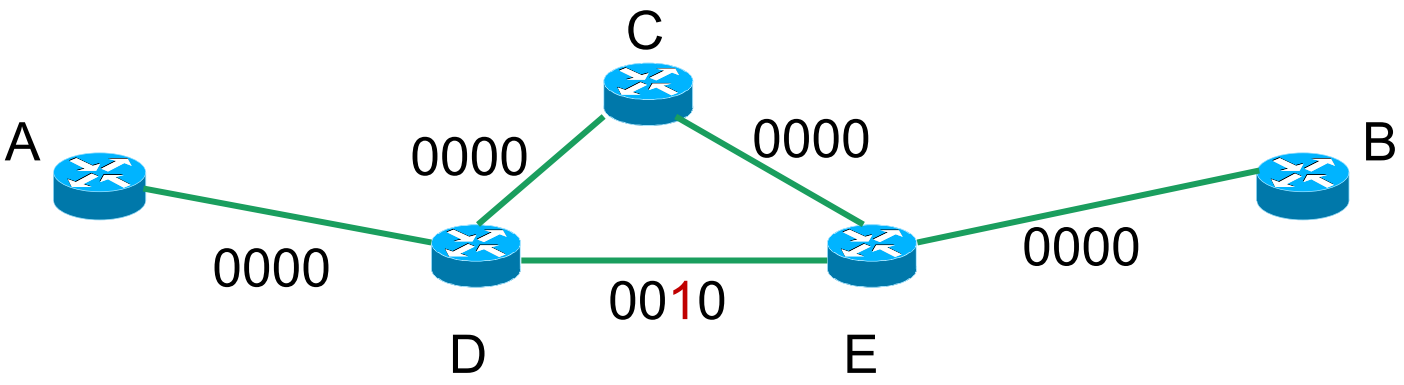
Example0: 4-bit string, default

- Traffic from A to B:
  - tunnel affinity= 0000, t-mask = 0011
- ADEB and ADCEB are possible



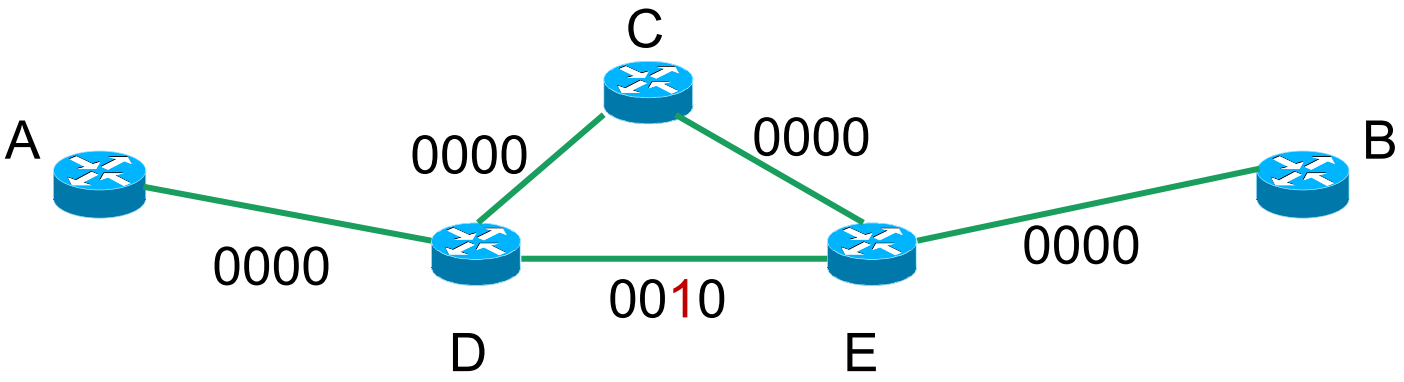
Example1a: 4-bit string

- Setting a link bit in the lower half drives all tunnels off the link, except those specially configured
- Traffic from A to B:
  - tunnel = 0000, t-mask = 0011
- Only ADCEB is possible



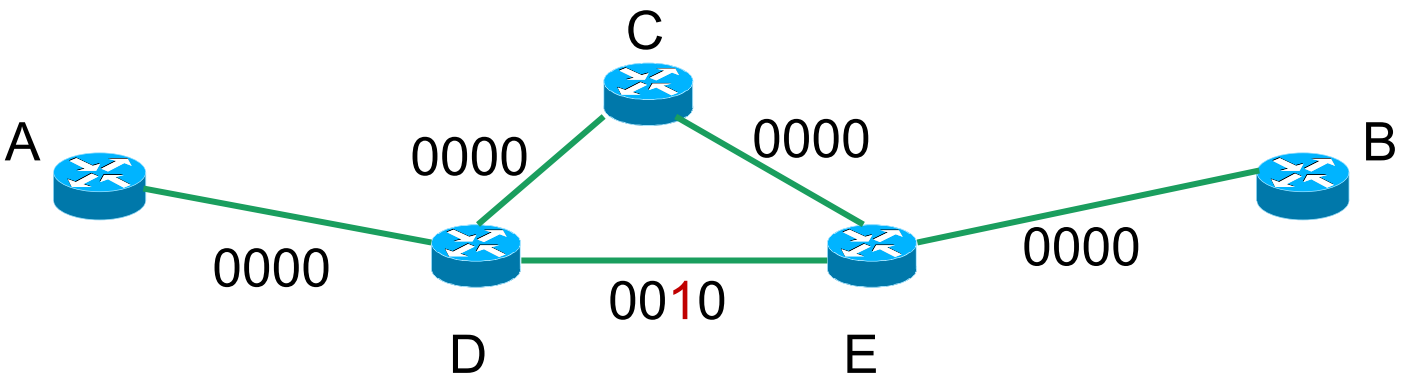
Example1b: 4-bit string

- A specific tunnel can then be configured to allow such links by clearing the bit in its affinity attribute mask
- Traffic from A to B:
  - tunnel = 0000, t-mask = 0001
- Again, ADEB and ADCEB are possible



Example1c: 4-bit string

- A specific tunnel can be restricted to only such links by instead turning on the bit in its affinity attribute bits
- Traffic from A to B:
  - tunnel = 0010, t-mask = 0011
- No path is possible



Administrative Weight (TE Metric)

- Two costs are associated with a link
  - TE cost (Administrative weight)
  - IGP cost
- The default TE cost on a link is the same as the IGP cost. We can also configure as following:

```
Router(config)# interface ethernet 0/1
Router(config-if)# mpls traffic-eng administrative-weight 20
```

Link-State Protocol Extensions/ IGP Flooding

- TE finds paths other than shortest-cost. To do this, TE must have more info than just per-link cost
- OSPF and IS-IS have been extended to carry additional information
  - Physical bandwidth
  - RSVP configured bandwidth
  - RSVP available bandwidth
  - Link TE metric
  - Link affinity

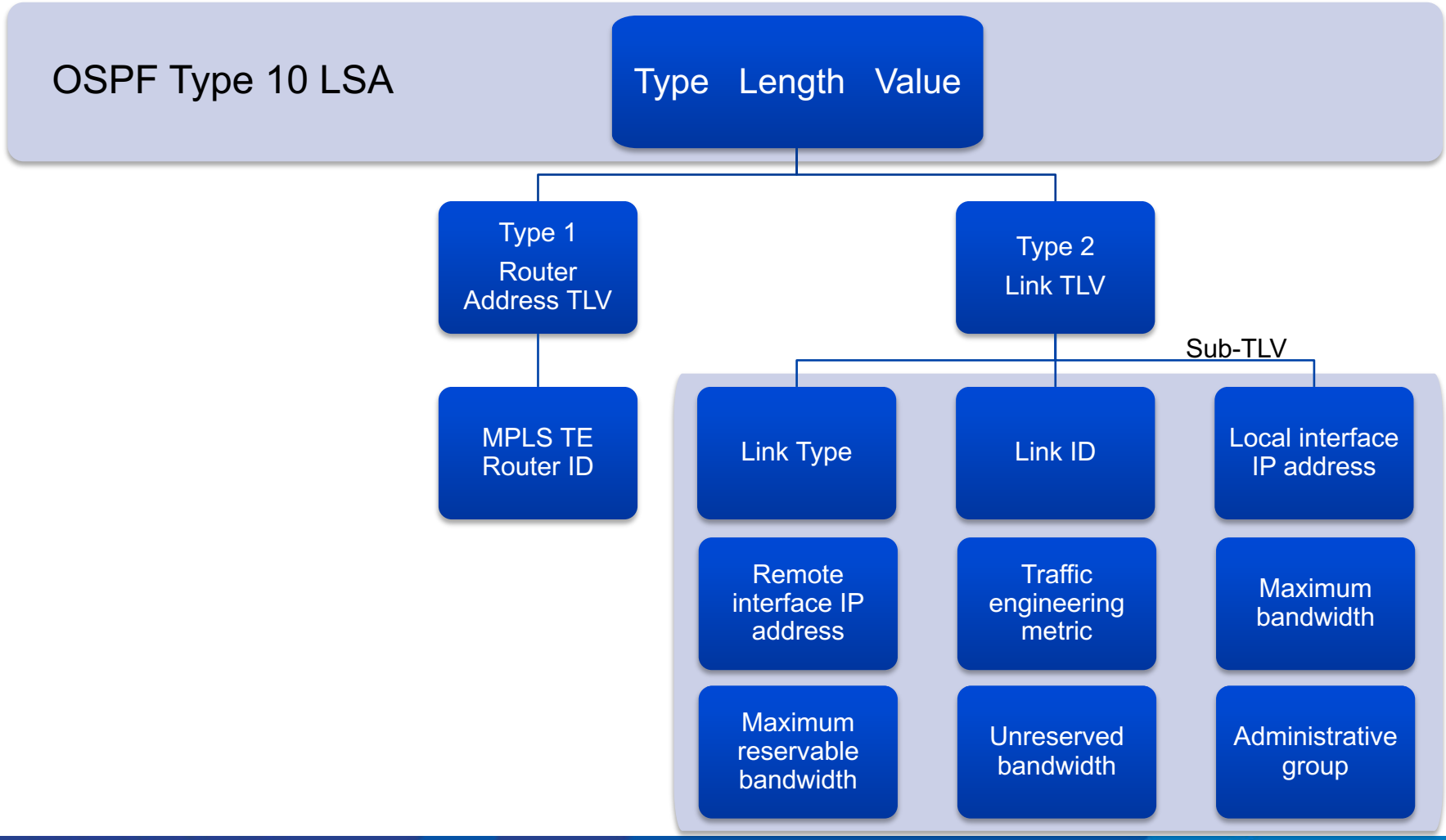
OSPF Extensions

- OSPF
  - Uses Type 10 (Opaque Area-Local) LSAs

- Enable OSPF TE:

```
Router(config)# router ospf 100
Router(config-router)# mpls traffic-eng router-id loopback 0
Router(config-router)# mpls traffic-eng area 0
```

OSPF Type 10 LSA



IS-IS Extensions

- IS-IS
  - Uses Type 22 TLVs to carrry MPLS TE link information
  - Also uses Type 134 and Type 135 TLV

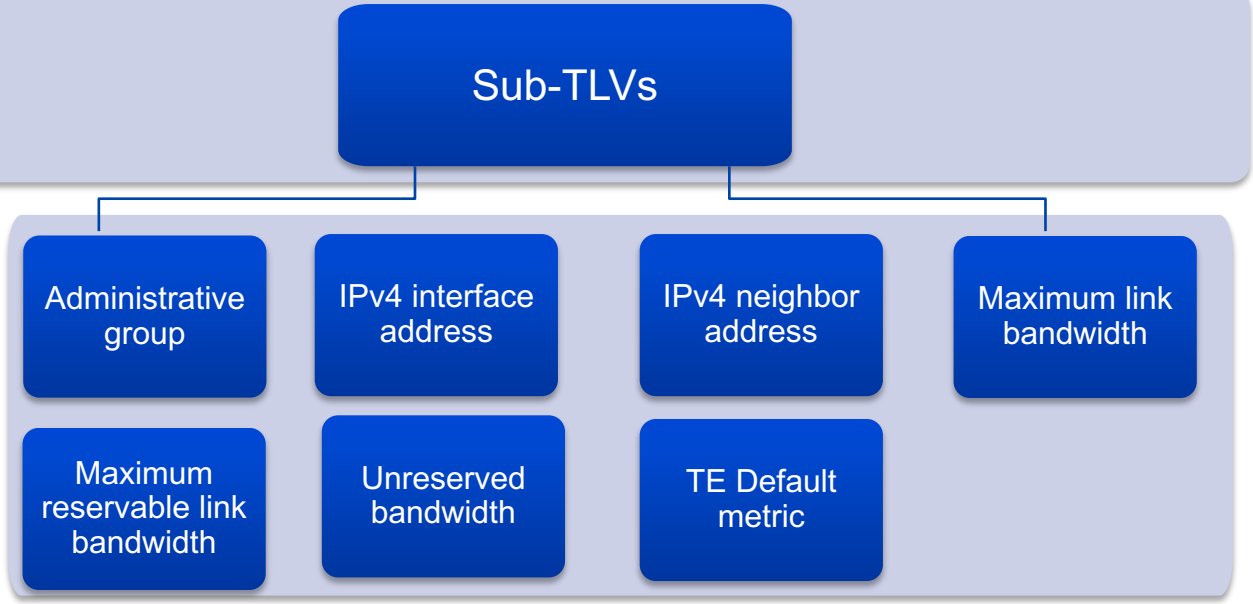
- Enable IS-IS TE

```
Router(config)# router isis
Router(config-router)# mpls traffic-eng level-2
Router(config-router)# mpls traffic-eng router-id loopback0
Router(config-router)# metric-style wide
```

- Support for wide metrics must be enabled.

IS-IS Extensions for TE

IS-IS Type 22 TLV



IS-IS Type 134 TLV — Advertise a 32-bit router ID

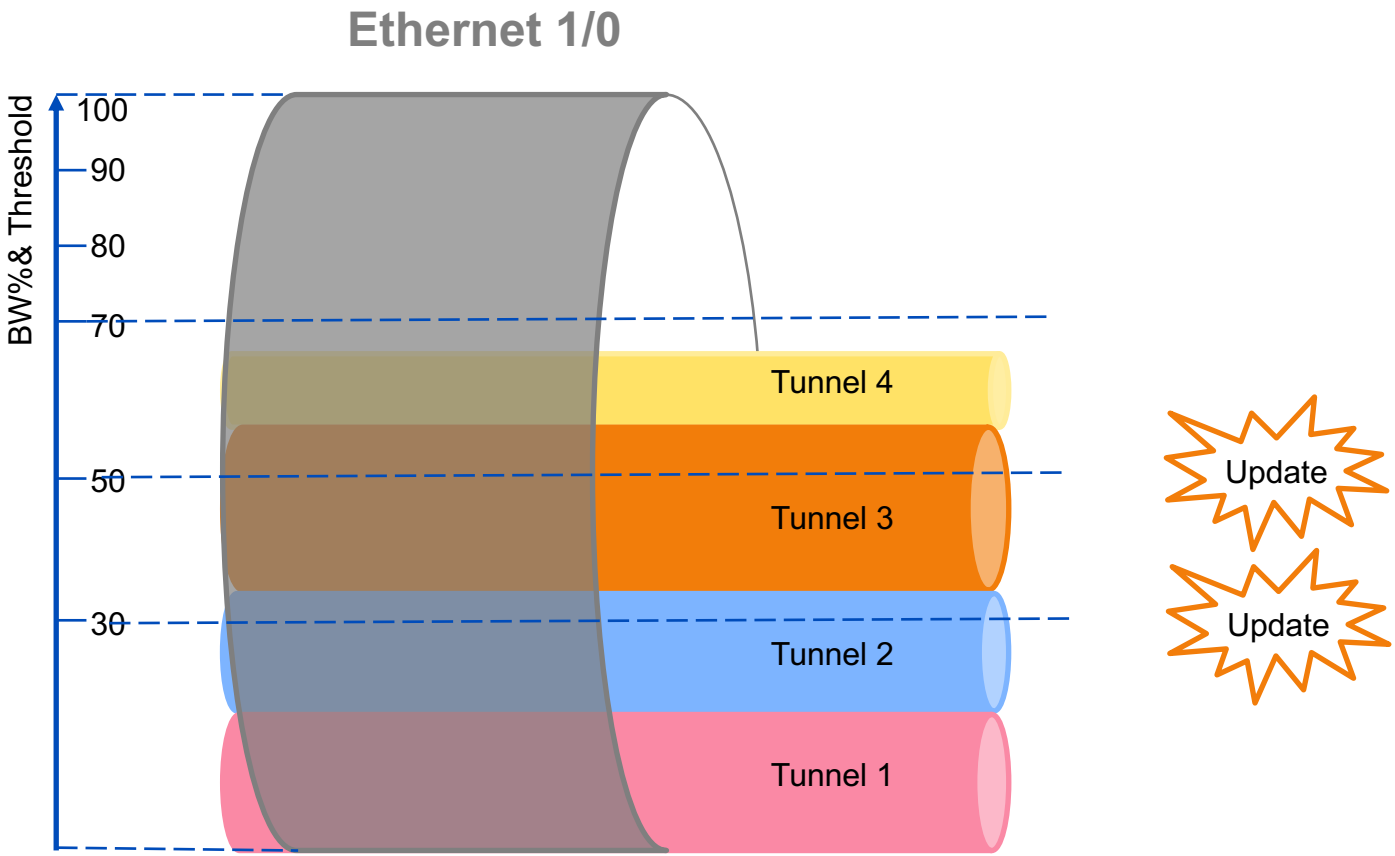
IS-IS Type 135 TLV — Expand the link metrics and path metric

When to Flood the Information

- When a link goes up or down
- When a link's configuration is changed
- Periodically reflood the router's IGP information
- When link bandwidth changes **significantly**

Bandwidth Significant Change

- Each time a threshold is crossed, an update message is sent.



Default Threshold

- We can view the current threshold using the command `show mpls traffic-eng link-management badwidth-allocation`

```
Router1#show mpls traffic-eng link bandwidth-allocation ethernet 1/0
...(Omitted)
      Up Thresholds:        15 30 45 60 75 80 85 90 95 96 97 98 99 100 (default)
      Down Thresholds:      100 99 98 97 96 95 90 85 80 75 60 45 30 15 (default)
...(Omitted)
```

  - Denser population as utilization increases
  - Different thresholds for Up and Down

Path Calculation and Setup

APNIC

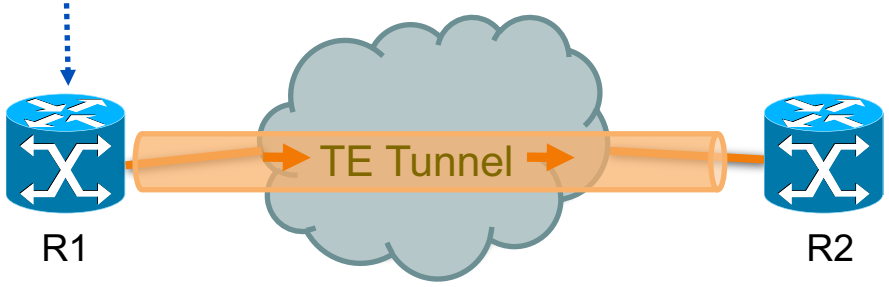
Tunnel Path Selection

- Tunnel has two path options
  1. Dynamic
  2. Explicit
- Path is a set of next-hop addresses (physical or loopbacks) to destination
- This set of next-hops is called Explicit Route Object (ERO)

Dynamic Path Option

- Dynamic = router calculates path using TE topology database
- Router will take best IGP path that meets BW requirements, also called CSPF algorithm.

```
R1(config)# interface tunnel 1
R1(config-if)# tunnel mpls traffic-eng path-option 10 dynamic
```



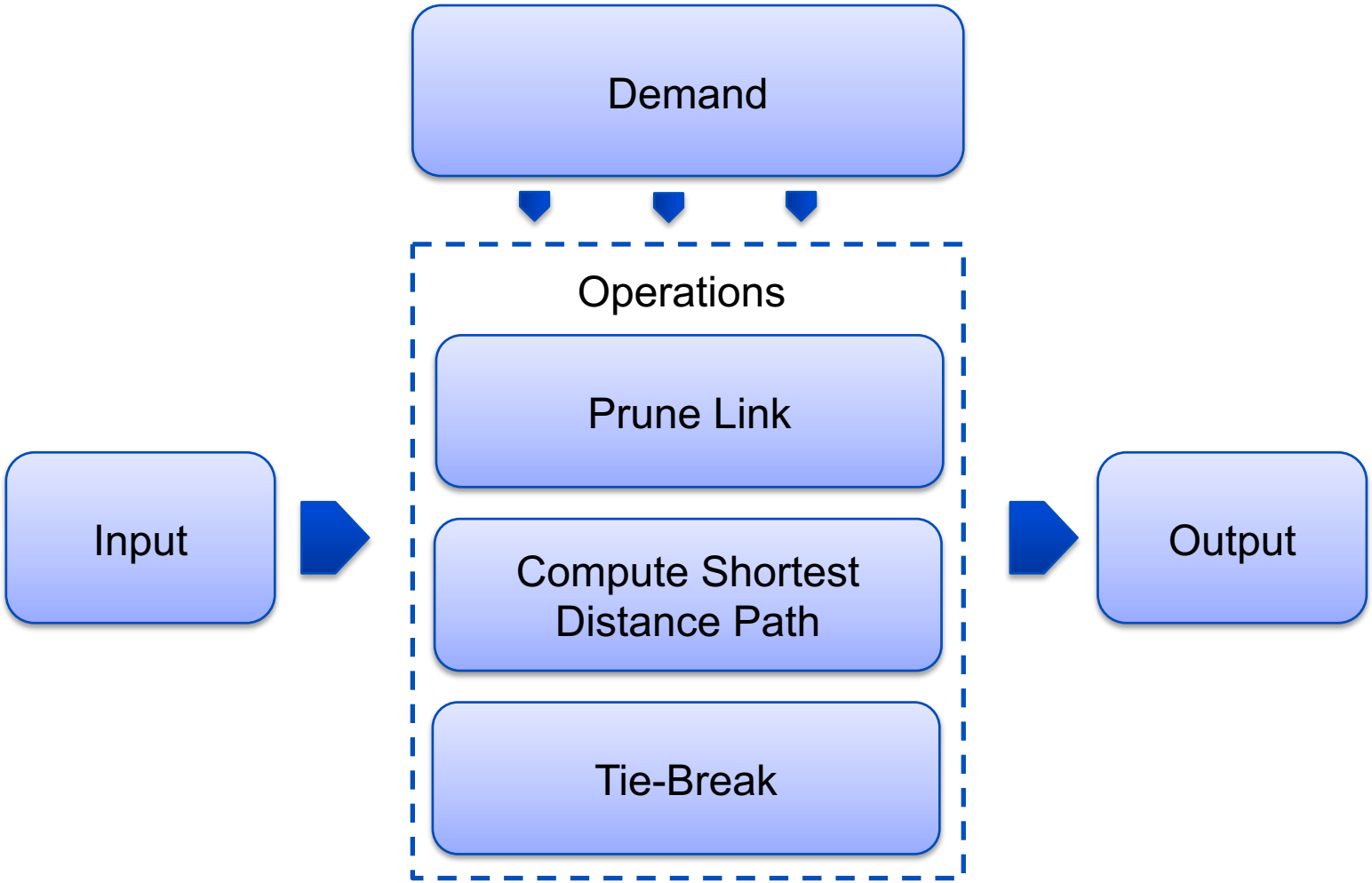
Path Calculation

- Modified Dijkstra
- Often referred to as CSPF
  - Constrained SPF
- …or PCALC (path calculation)
- Final result is explicit route meeting desired constraint

C-SPF

- Shortest-cost path is found that meets administrative constraints
- These constraints can be
  - bandwidth
  - link attribute (aka color, resource group)
  - priority
- The addition of constraints is what allows MPLS-TE to use paths other than *just* the shortest one

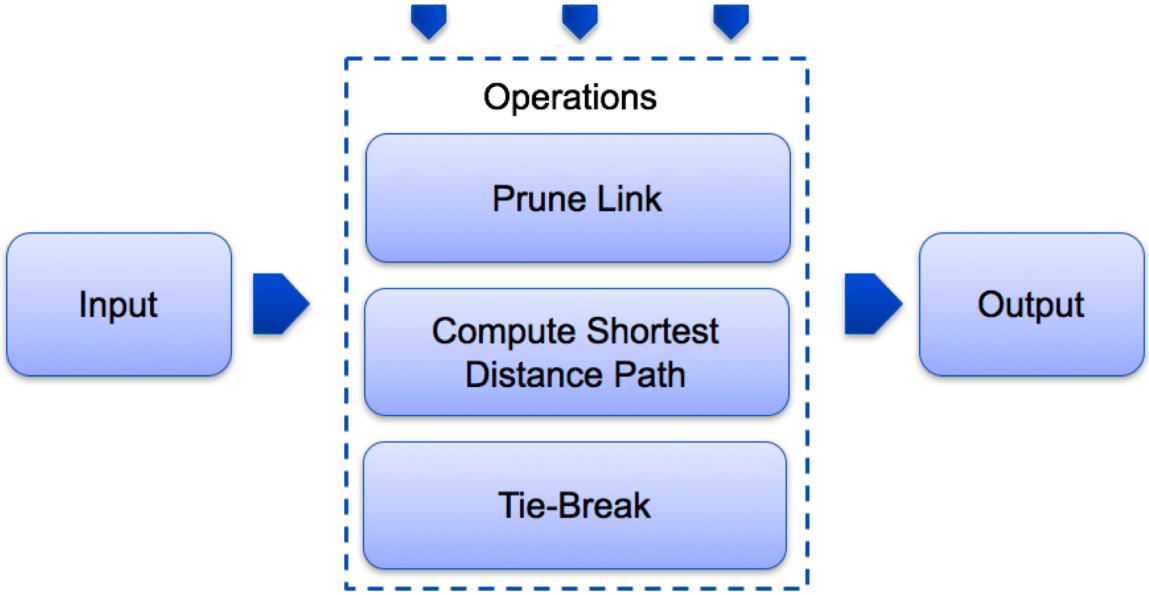
Path Computation



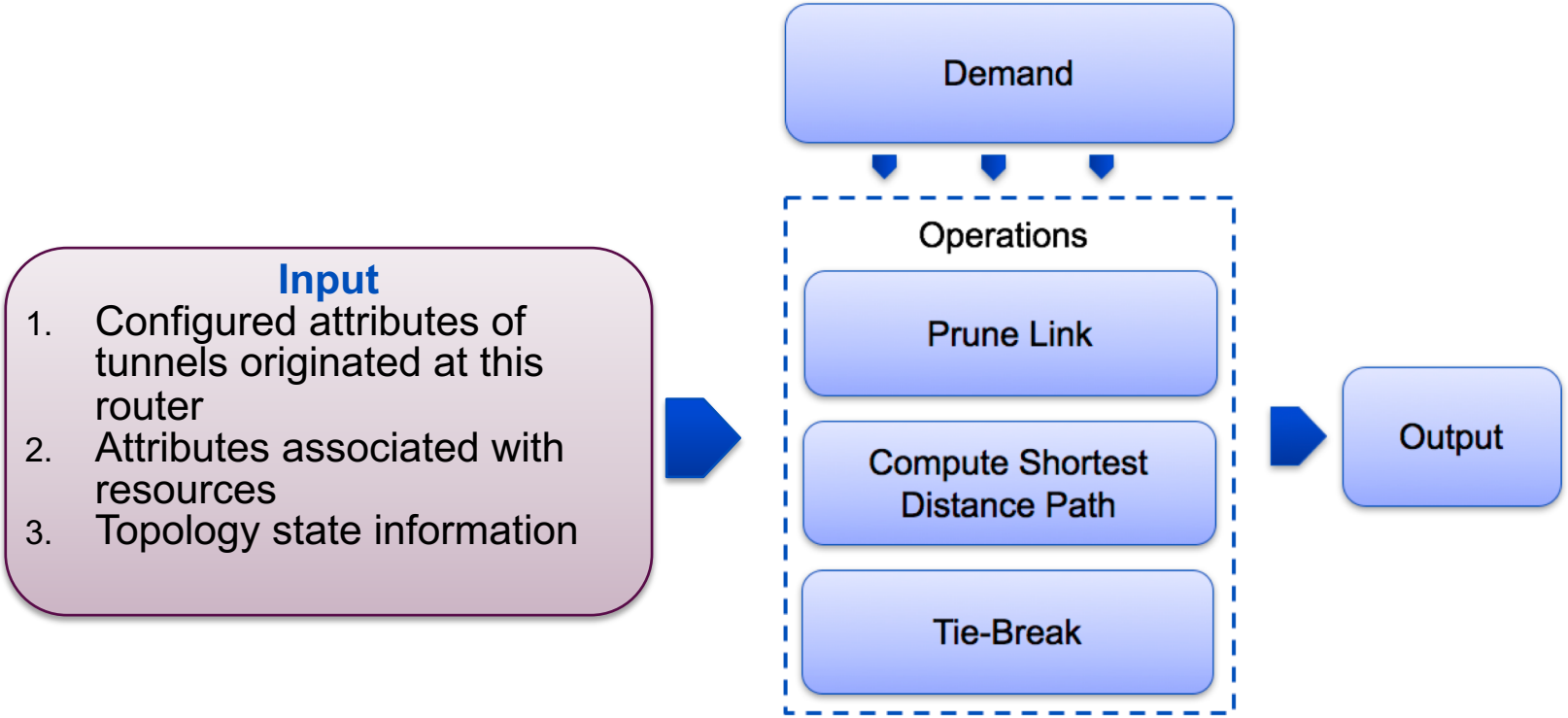
Path Computation - Demand

"On demand" by the tunnel's head-end:

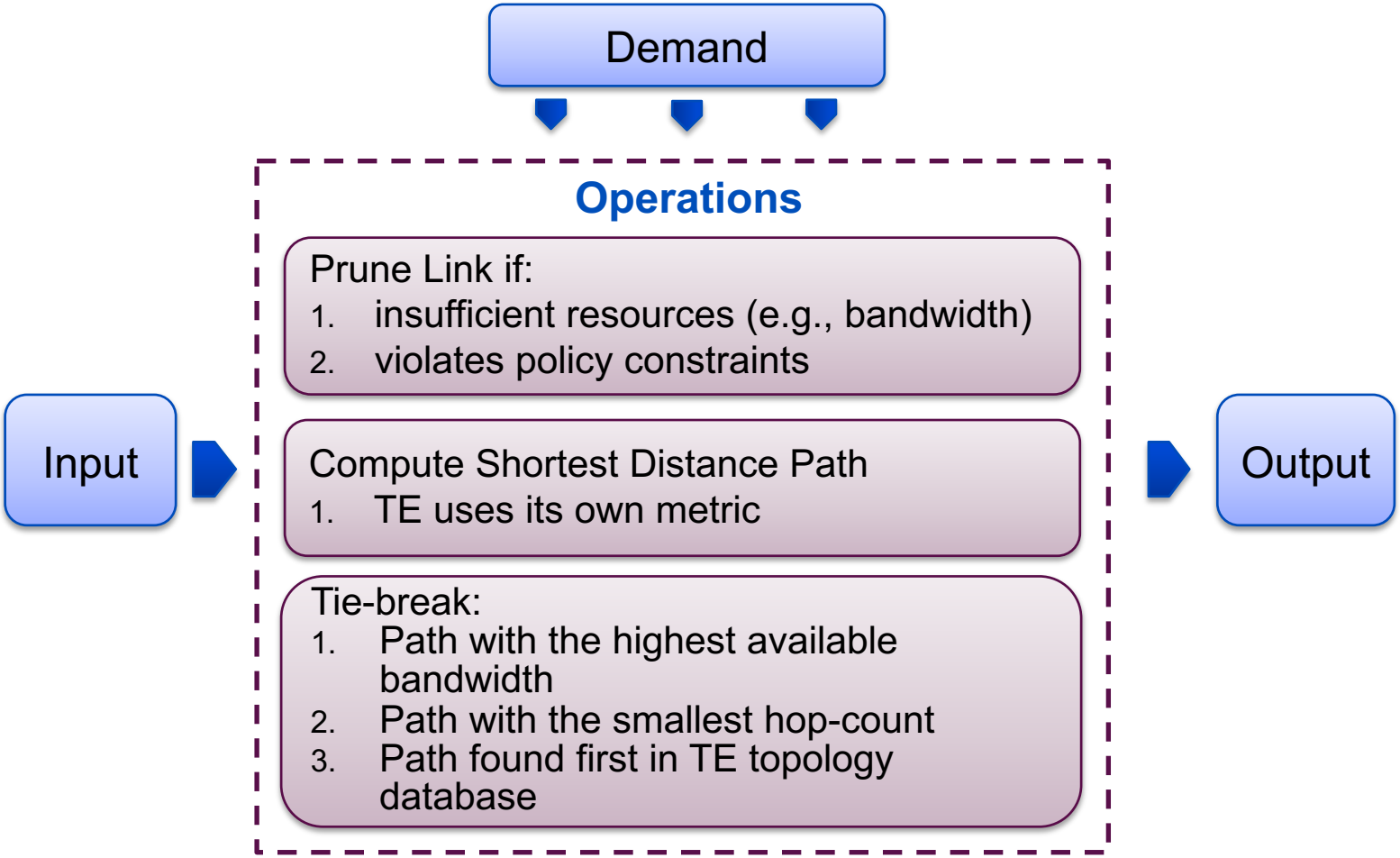
1. for a new tunnel
2. for an existing tunnel whose (current) LSP failed
3. for an existing tunnel when doing re-optimization



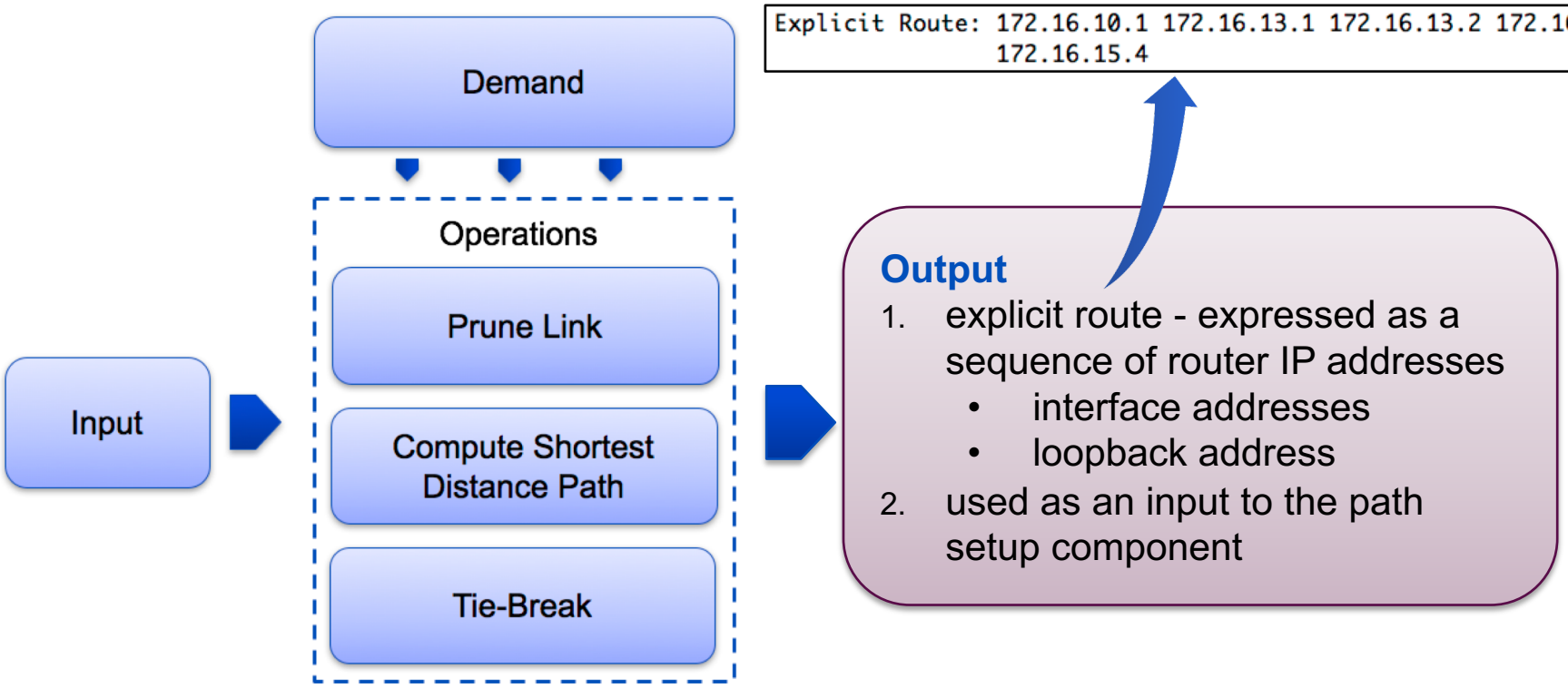
Path Computation - Input



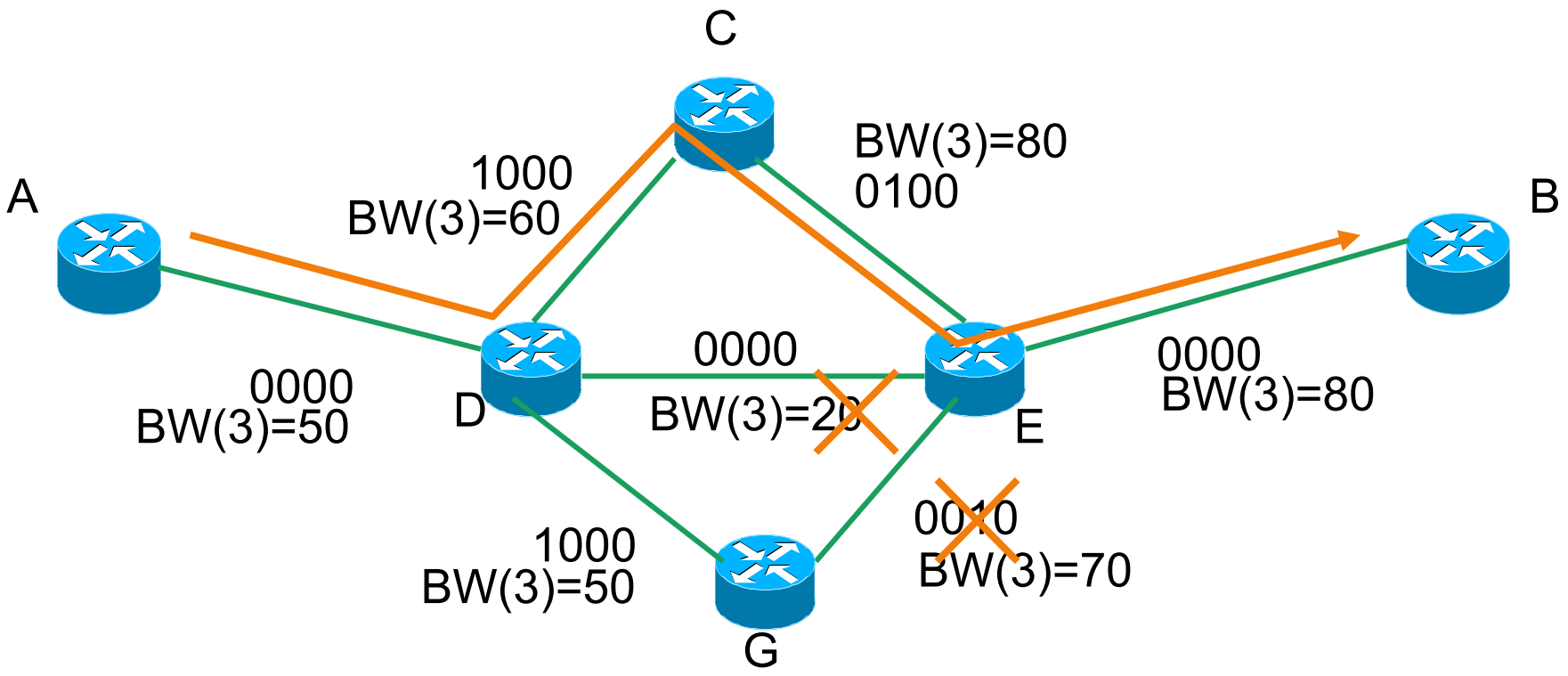
Path Computation - Operation



Path Computation - Output



BW/Policy Example

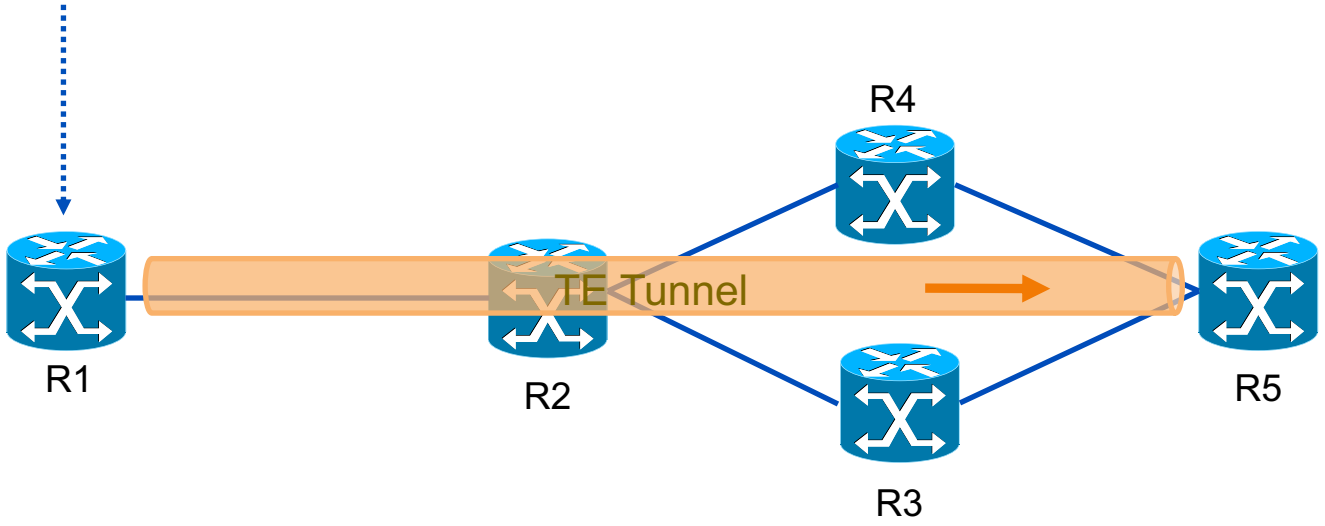


- Tunnel's request:
  - Priority 3, BW = 30 units,
  - Policy string: 0000, mask: 0011

Explicit Path Option

- explicit = take specified path.
- Router sets up path you specify.

```
R1(config)# interface tunnel 1
R1(config-if)# tunnel mpls traffic-eng path-option 10 explicit
name R1toR5
```

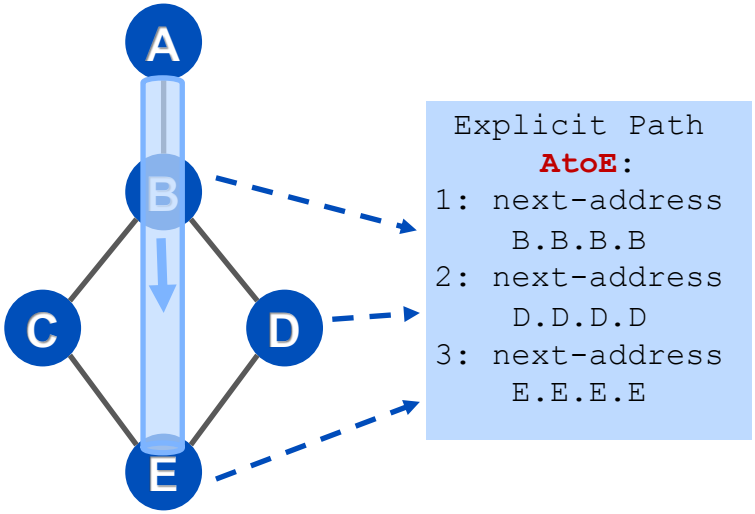


Strict and Loose Path

- Paths are configured manually. Each hop is a physical interface or loopback.

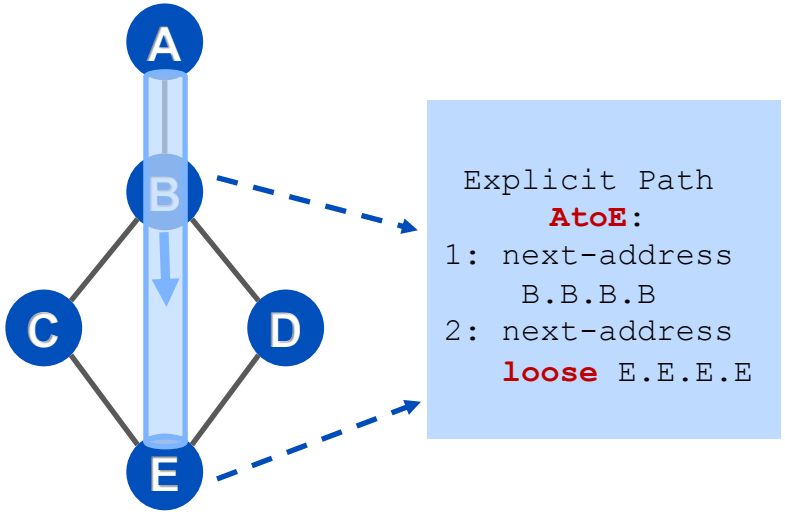
Strict Path

A network node and its preceding node in the path must be adjacent and directly connected.

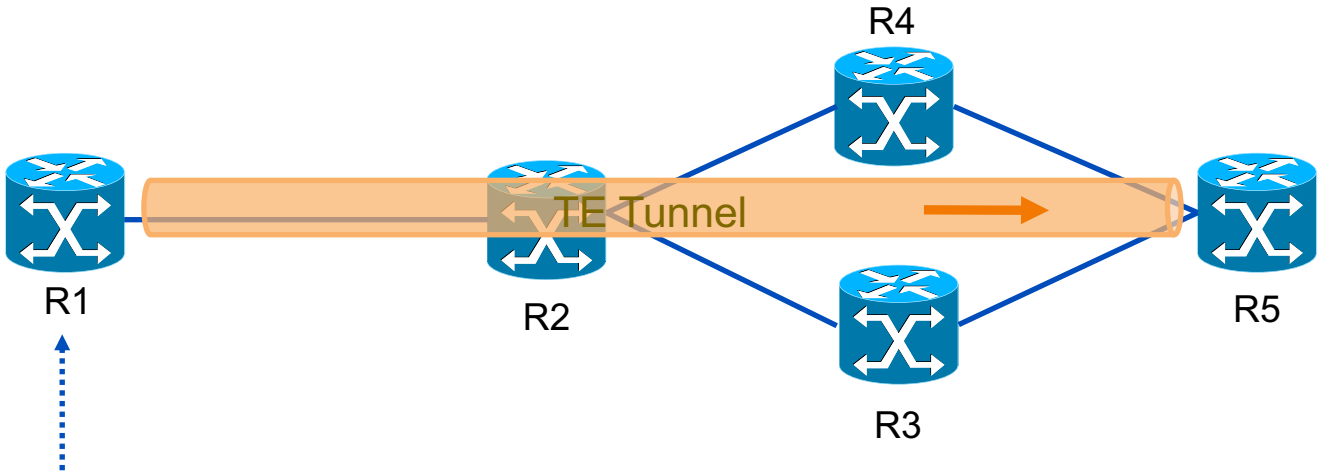


Loose Path

A network node must be in the path but is not required to be directly connected to its preceding node.

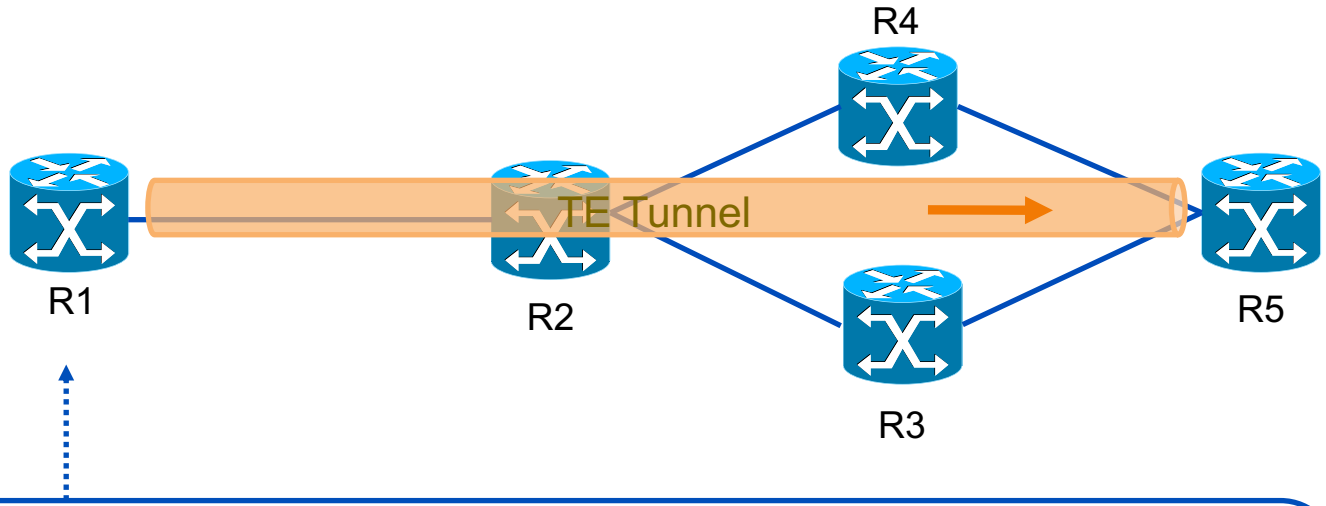


Configure Strict Explicit Path




```
R1(config)# ip explicit-path name R1toR5
R1(cfg-ip-expl-path)# next-address strict 2.2.2.2
Explicit Path name R1toR5:
    1: next-address 2.2.2.2
R1(cfg-ip-expl-path)# next-address strict 4.4.4.4
Explicit Path name R1toR5:
    1: next-address 2.2.2.2
    2: next-address 4.4.4.4
R1(cfg-ip-expl-path)# next-address strict 5.5.5.5
Explicit Path name R1toR5:
    1: next-address 2.2.2.2
    2: next-address 4.4.4.4
    3: next-address 5.5.5.5
```

Configure Loose Explicit Path




```
R1(config)# ip explicit-path name R1toR5
R1(cfg-ip-expl-path)# next-address 2.2.2.2
Explicit Path name R1toR5:
    1: next-address 2.2.2.2
R1(cfg-ip-expl-path)# next-address loose 5.5.5.5
Explicit Path name R1toR5:
    1: next-address 2.2.2.2
    2: next-address loose 5.5.5.5
```

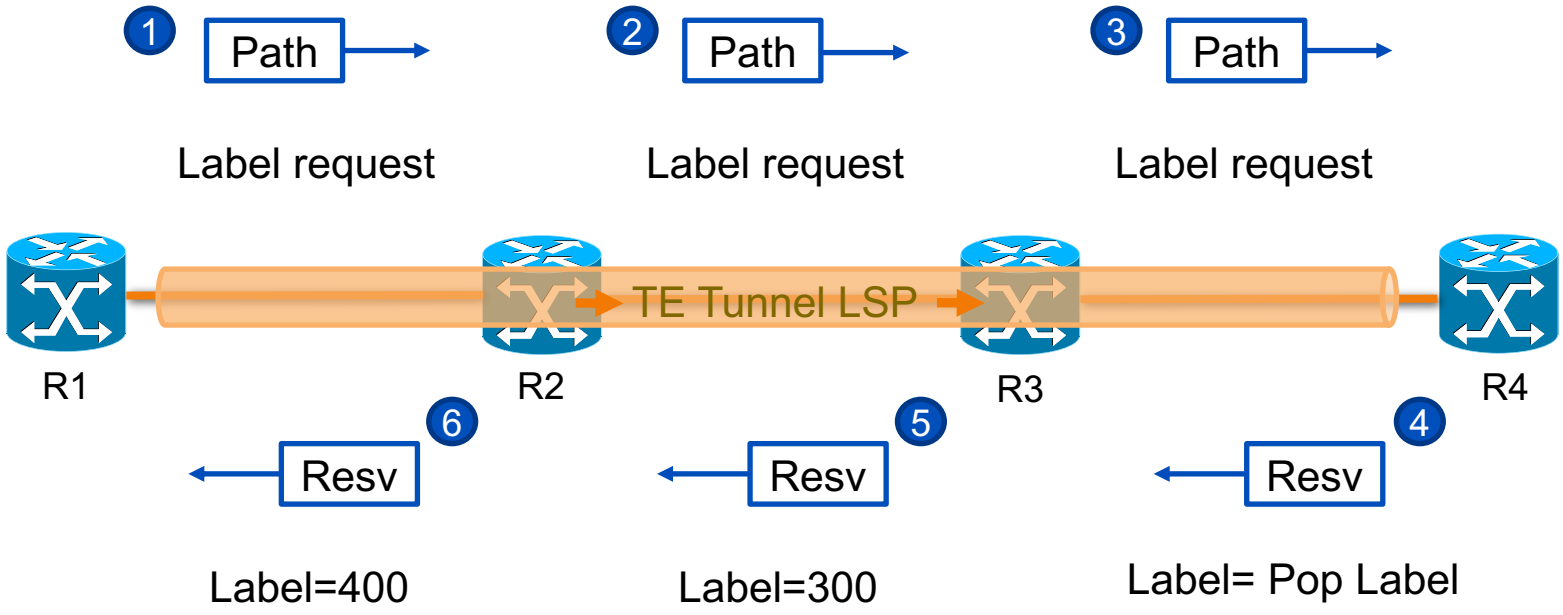
By default, it is strict.

RSVP-TE

- After calculating the path, tunnel will be set up by using RSVP-TE.
- RSVP has three basic functions:
  - Path setup and maintenance
  - Path teardown
  - Error signalling

Setup of TE LSP

- In following topo, R1 will set up a TE tunnel from R1 to R4:



Check RSVP Status

- show ip rsvp interface

```
Router2# show ip rsvp interface
interface    rsvp       allocated  i/f max  flow max sub max  VRF
Fa0/0        ena        0          256K     256K     0
Fa0/1        ena        100K       256K     256K     0
Et1/0        ena        100K       256K     256K     0
Et1/1        ena        0          256K     256K     0
```

- show ip rsvp neighbor

```
Router2# show ip rsvp neighbor
Neighbor         Encapsulation  Time since msg rcvd/sent
172.16.10.2      Raw IP         00:00:11   00:00:06
172.16.13.2      Raw IP         00:00:08   00:00:14
```

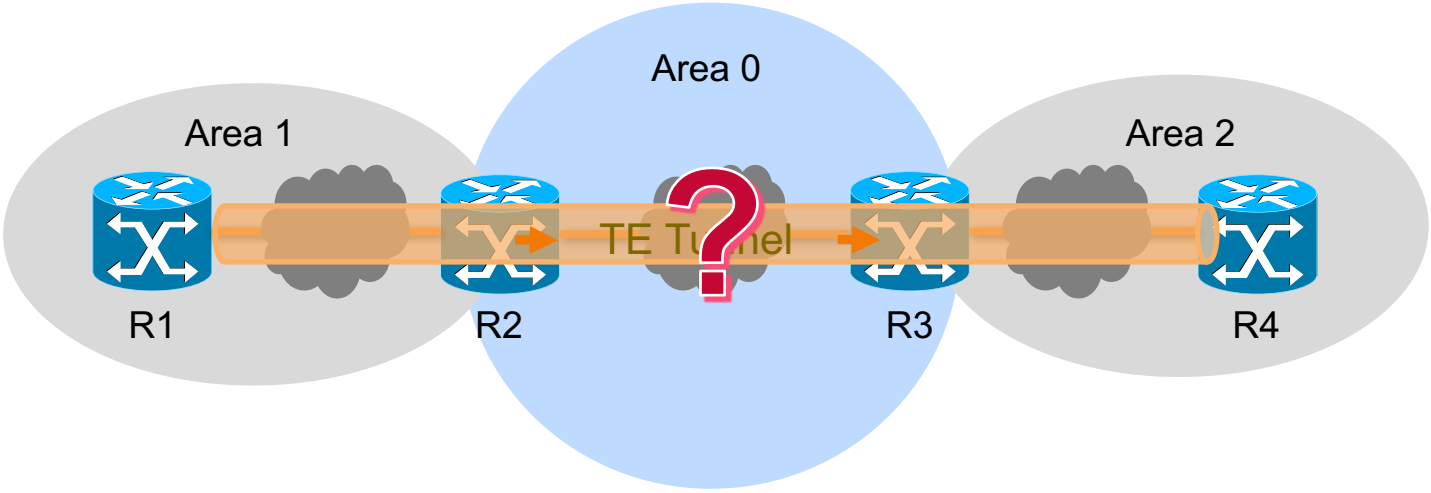
Check RSVP Reservation

- show ip rsvp reservation detail

```
Router2# show ip rsvp reservation detail
Reservation:
  Tun Dest:   172.16.15.1  Tun ID: 10  Ext Tun ID: 172.16.15.4
  Tun Sender: 172.16.15.4  LSP ID: 3
  Next Hop: 172.16.10.2 on Ethernet1/0
  Label: 3 (outgoing)
  Reservation Style is Shared-Explicit, QoS Service is Controlled-Load
  Resv ID handle: 02000413.
  Created: 20:27:28 AEST Thu Dec 22 2016
  Average Bitrate is 100K bits/sec, Maximum Burst is 1K bytes
  Min Policed Unit: 0 bytes, Max Pkt Size: 1500 bytes
  Status:  Policy: Accepted. Policy source(s): MPLS/TE
Reservation:
  Tun Dest:   172.16.15.4  Tun ID: 10  Ext Tun ID: 172.16.15.1
  Tun Sender: 172.16.15.1  LSP ID: 814
  Next Hop: 172.16.13.2 on FastEthernet0/1
  Label: 16 (outgoing)
  Reservation Style is Shared-Explicit, QoS Service is Controlled-Load
  Resv ID handle: 02000408.
  Created: 19:56:58 AEST Thu Dec 22 2016
  Average Bitrate is 100K bits/sec, Maximum Burst is 1K bytes
  Min Policed Unit: 0 bytes, Max Pkt Size: 1500 bytes
  Status:  Policy: Accepted. Policy source(s): MPLS/TE
```

Interarea Tunnel

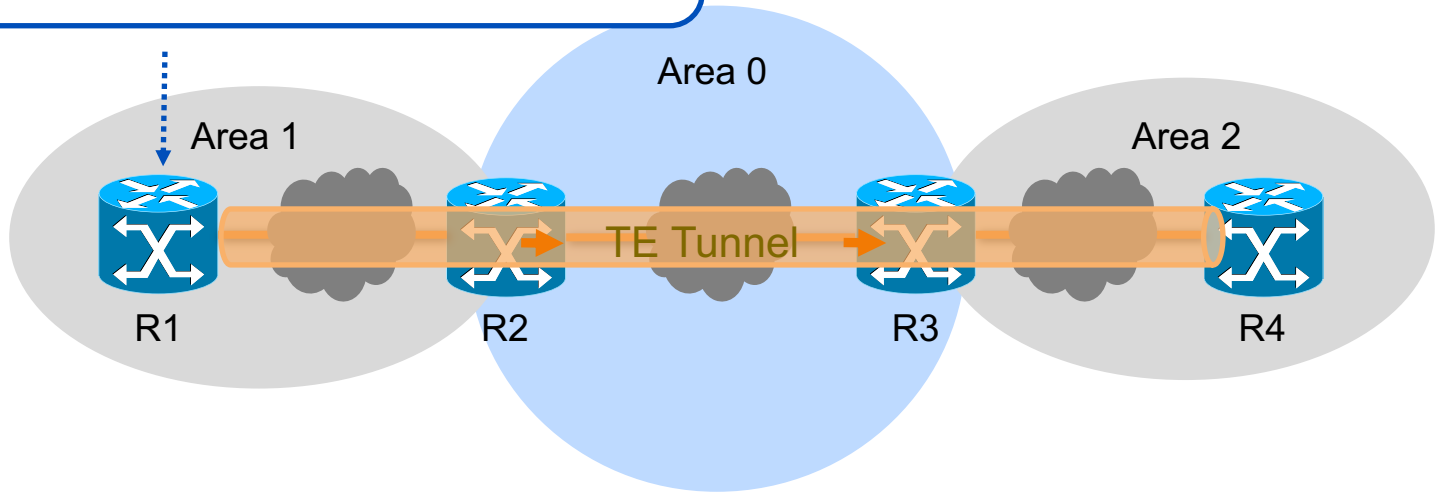
- Remember?
  - OSPF Type 10 LSA's flooding scope is local area.
- The TE LSP path is computed for single IGP area only due to lack of visibility of the topology in other IGP areas.
- How can we build tunnels between areas?



Define Contiguous Path

- An **explicit path** identifying the **ABR**s is required.
- The head-end router and the ABRs along the specified explicit path expand the loose hops, each computing the path segment to the next ABR or tunnel destination.

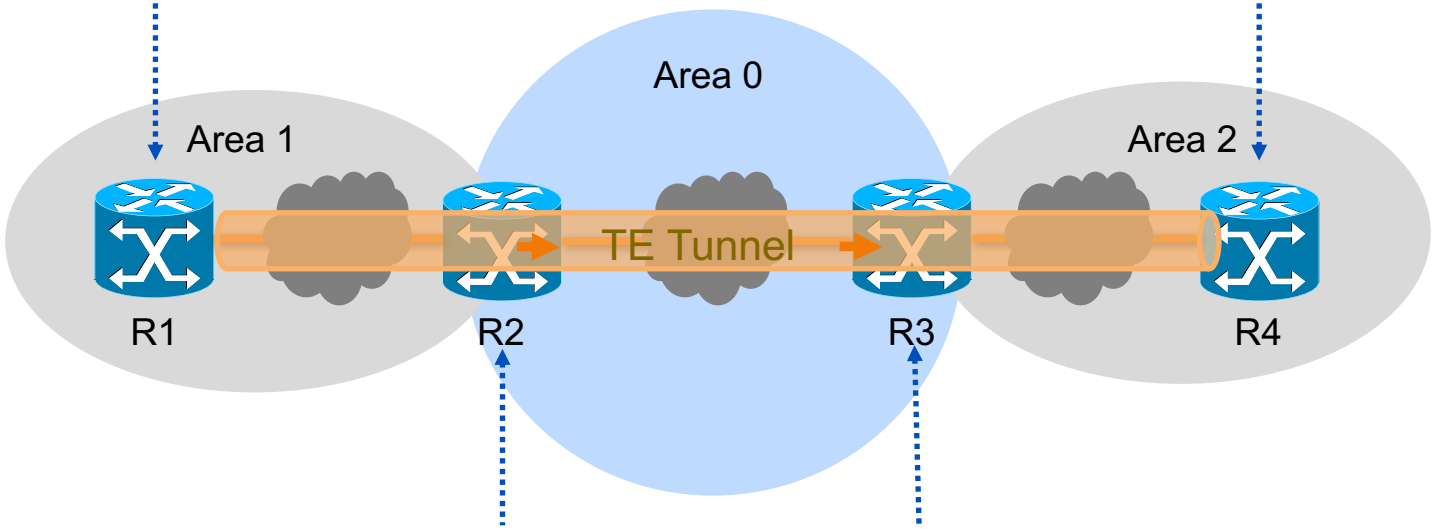
```
ip explicit path name R1toR4
  next-address loose 2.2.2.2
  next-address loose 3.3.3.3
  next-address loose 4.4.4.4
```



Enable MPLS TE for Areas

```
router ospf 1
mpls traffic-eng router-id Loopback0
mpls traffic-eng area 1
```

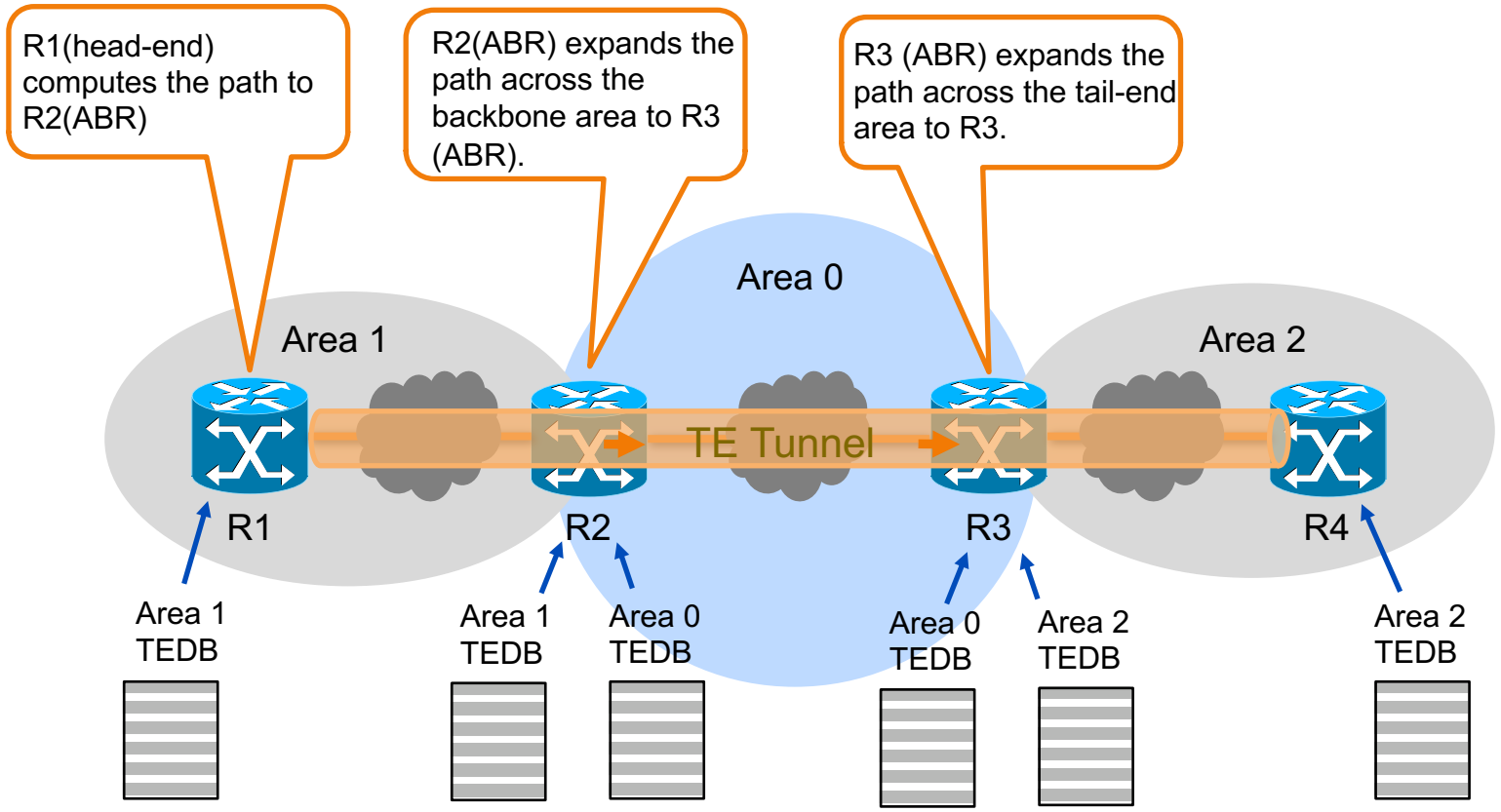
```
router ospf 1
mpls traffic-eng router-id Loopback0
mpls traffic-eng area 2
```




```
router ospf 1
mpls traffic-eng router-id Loopback0
mpls traffic-eng area 0
mpls traffic-eng area 1
```

```
router ospf 1
mpls traffic-eng router-id Loopback0
mpls traffic-eng area 0
mpls traffic-eng area 2
```

Setup Interarea Tunnel



ABR is maintaining multiple TEDB.
One TEDB for one area.

Forwarding Traffic Down Tunnels

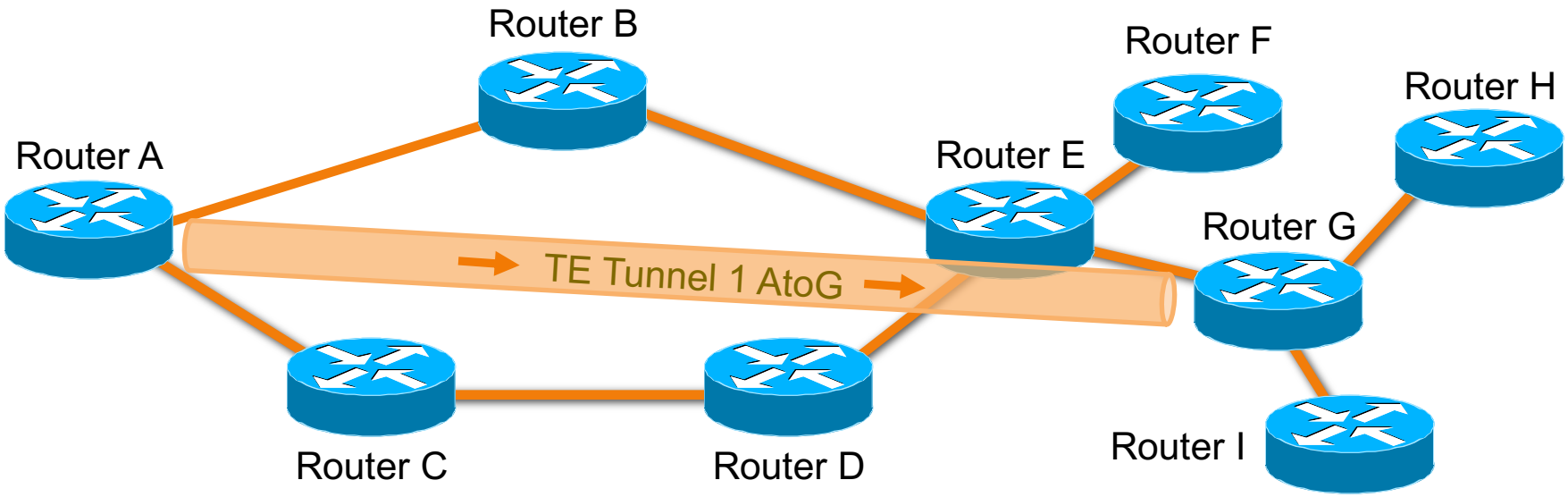
APNIC

Routing Traffic Down a Tunnel

- Once the tunnel is established and operational, it's ready to forward data traffic.
- However, no traffic will enter the tunnel unless the IP routing tables and FIB tables are modified.
- How to get traffic down the tunnel?
  1. Static route
  2. Auto route
  3. Policy routing

Example Topology

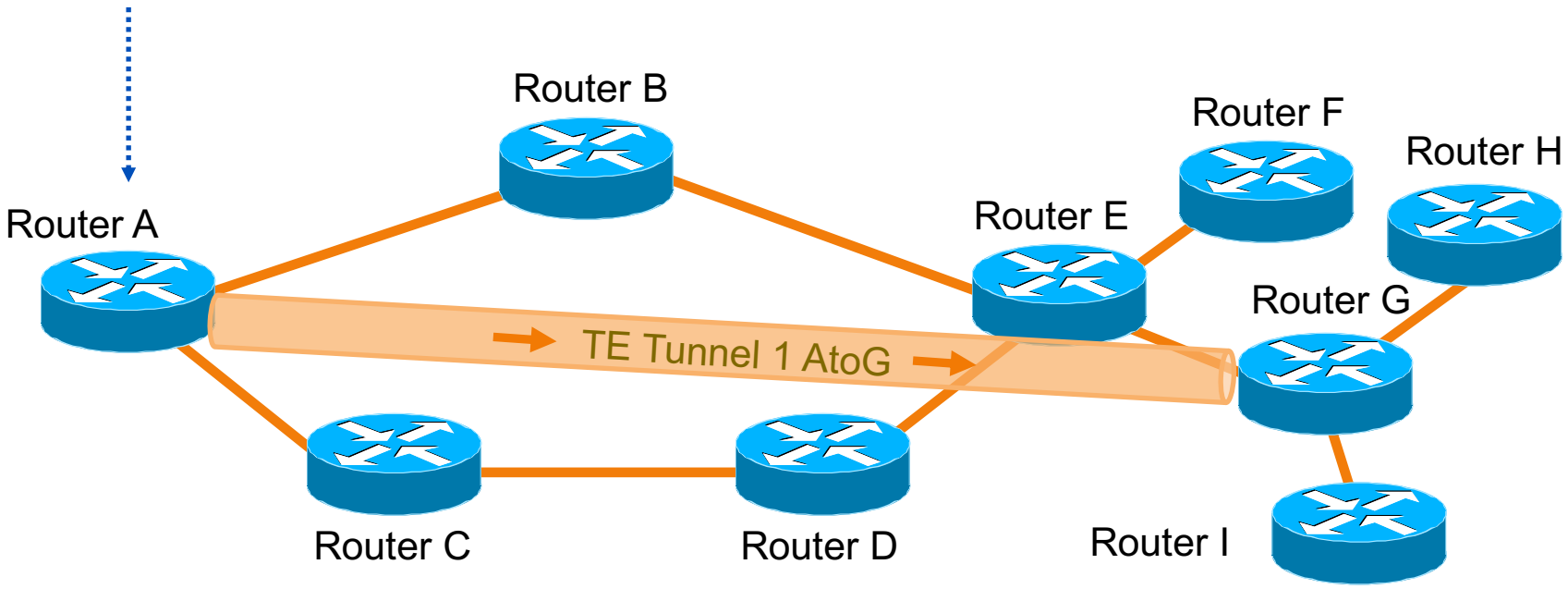
- In the following topology, cost=10 for each link.
- Tunnel 1 has been created on Router A from A to G.



Static Route

- Configure a static route pointing down the tunnel interface.

```
RtrA(config)# ip route H.H.H.H 255.255.255.255 Tunnel1
```

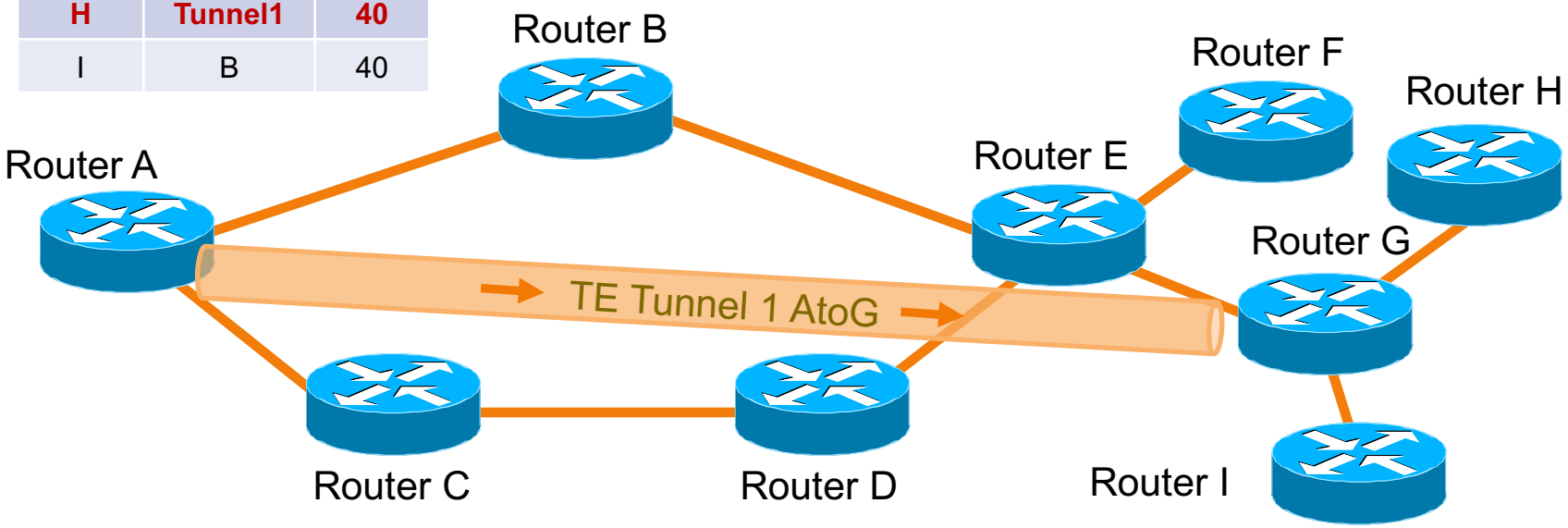


Static Route

Routing Table of RouterA

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	**B**	**30**
H	**Tunnel1**	**40**
I	B	40

- Router H is known via the tunnel
- Router G is **not** routed to over the tunnel, even though it's the tunnel tail!



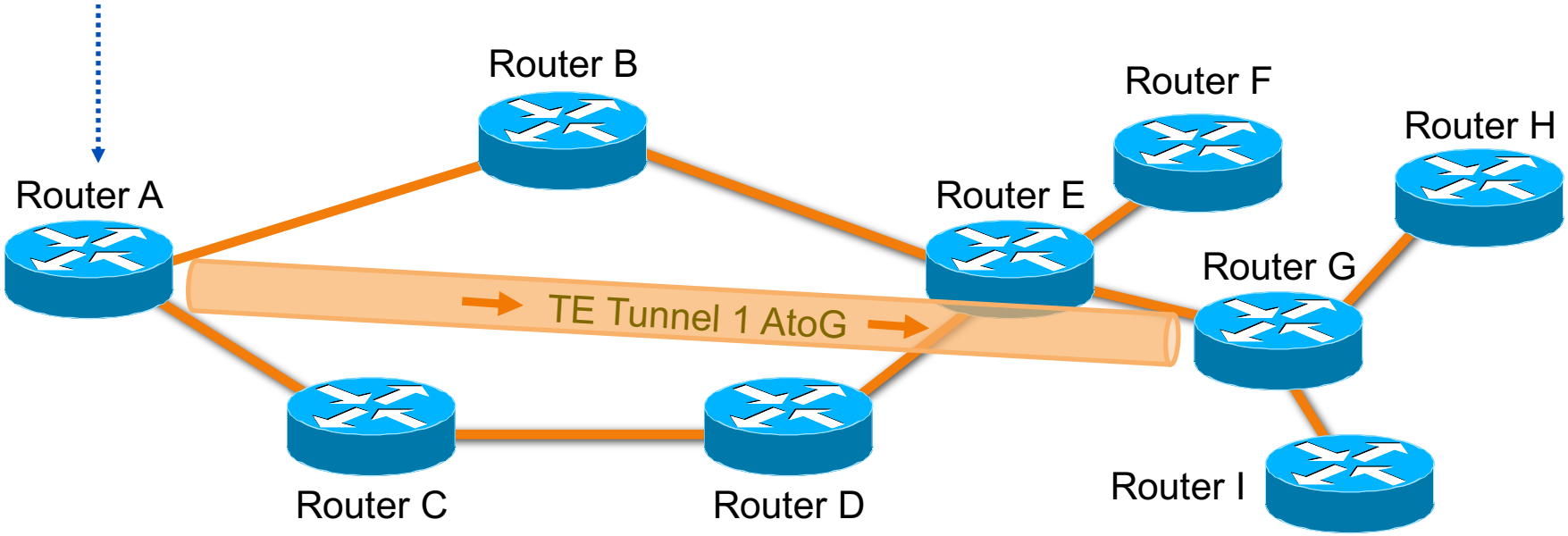
Auto Route

- Auto route allows a TE tunnel to participate in IGP route calculations as a logical link. The tunnel interface is used as the outbound interface of the route.
  - **IGP Shortcut**: The command is "**autoroute announce**" in Cisco IOS. Tunnel interface will be included in SPF tree on the **Head-end** router.
  - **Forwarding adjacency**: allows **all nodes** in IGP domains to see the TE tunnels and use them in the SPF calculation.

IGP Shortcut

- In the tunnel interface, configure *autoroute announce*

```
RtrA(config)# interface Tunnel1
RtrA(config-if)# tunnel mpls traffic-eng autoroute announce
```

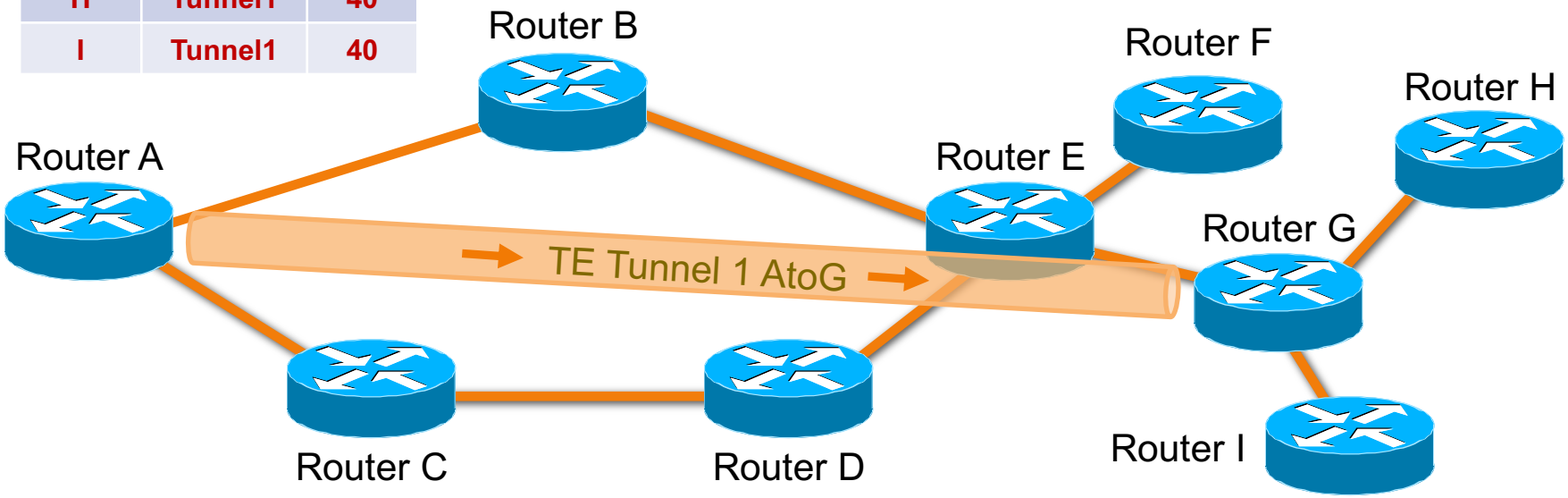


IGP Shortcut

Routing Table of RouterA

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	Tunnel1	30
H	Tunnel1	40
I	Tunnel1	40

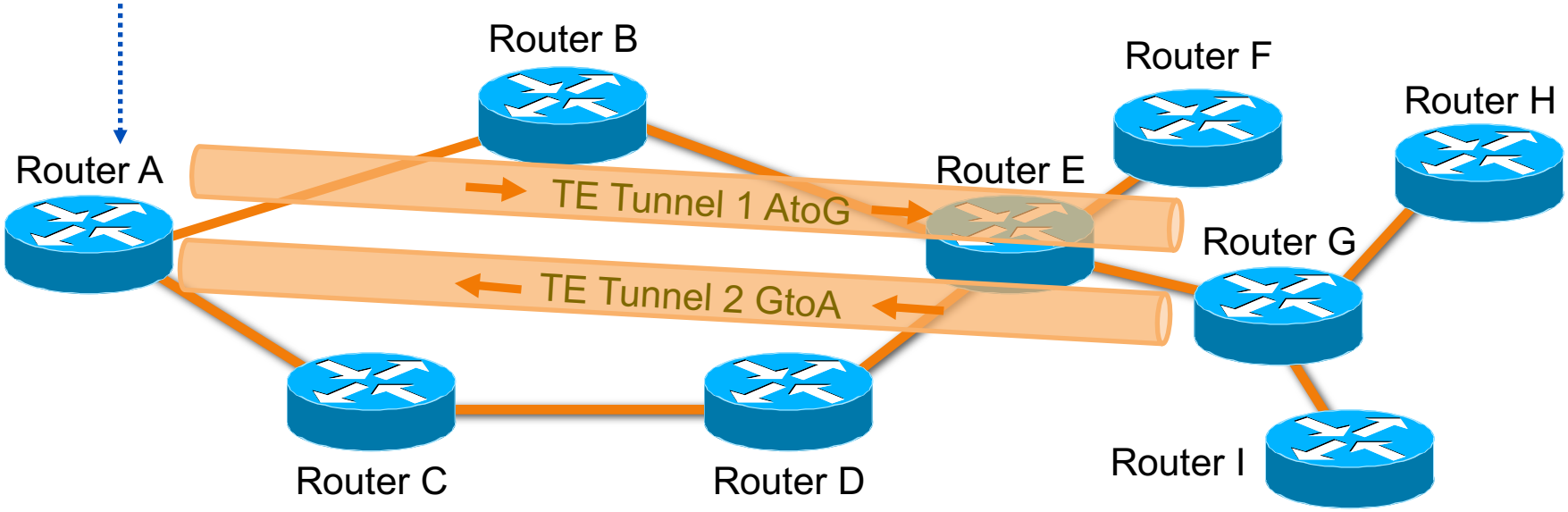
- Everything “behind” the tunnel is routed via the tunnel



Forwarding Adjacency

- Bidirectional tunnels: tunnel 1 AtoG, tunnel 2 GtoA
- Configure following commands on **both Router A and G**

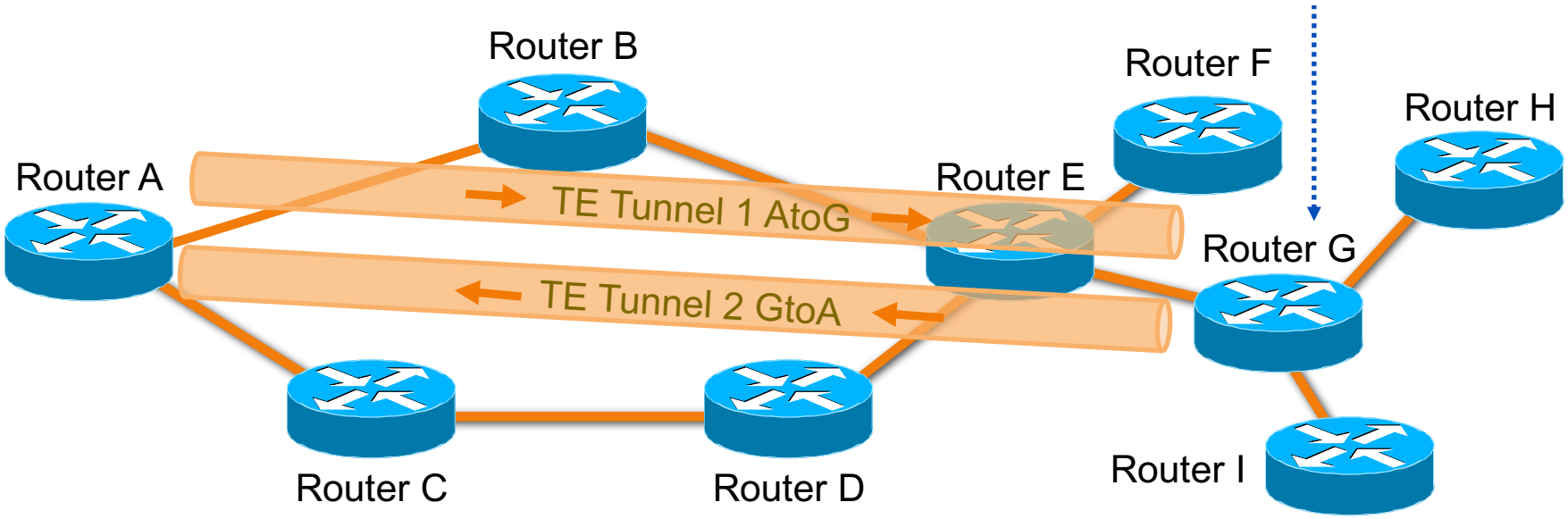
```
RtrA(config)# interface Tunnel1
RtrA(config-if)# tunnel mpls traffic-eng forwarding-adjacency
RtrA(config-if)# ip ospf cost 5
```



Forwarding Adjacency

- Similar commands on **Router G**

```
RtrG(config)# interface Tunnel1
RtrG(config-if)# tunnel mpls traffic-eng forwarding-adjacency
RtrG(config-if)# ip ospf cost 5
```

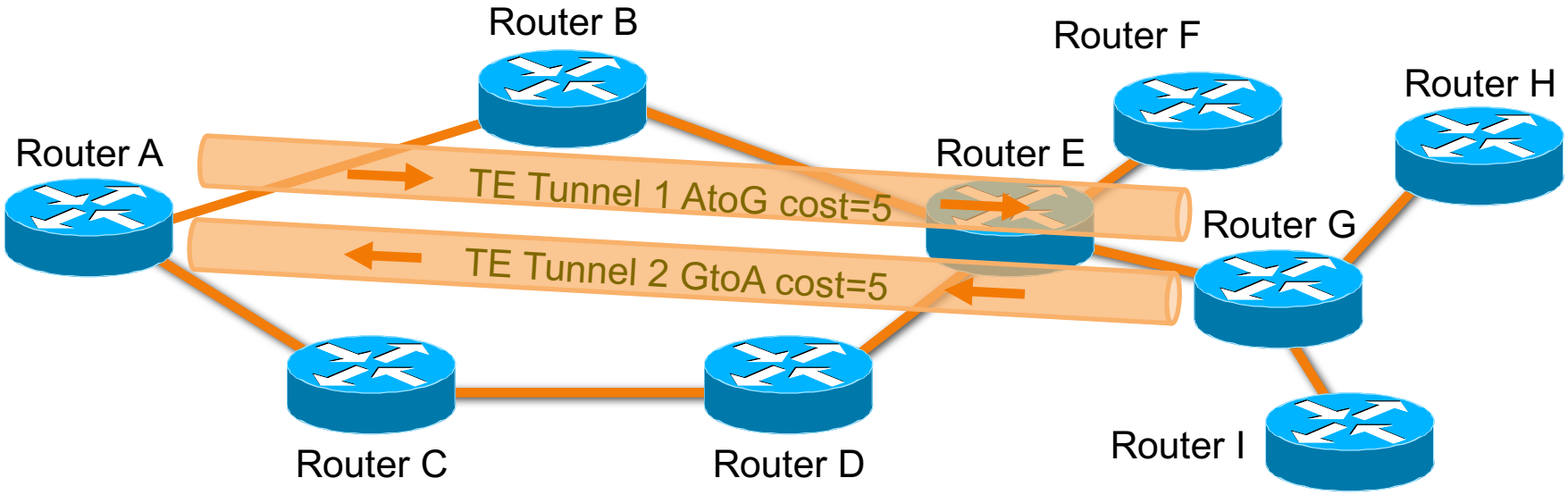


Before Configuring Forwarding Adjacency

- Assume cost of the tunnels is 5.
- Before FA configured, Router H does not include TE tunnel into its SPF calculation.

Routing Table of RouterH

Node	Next-Hop	Cost
A	G	40
…	…	…

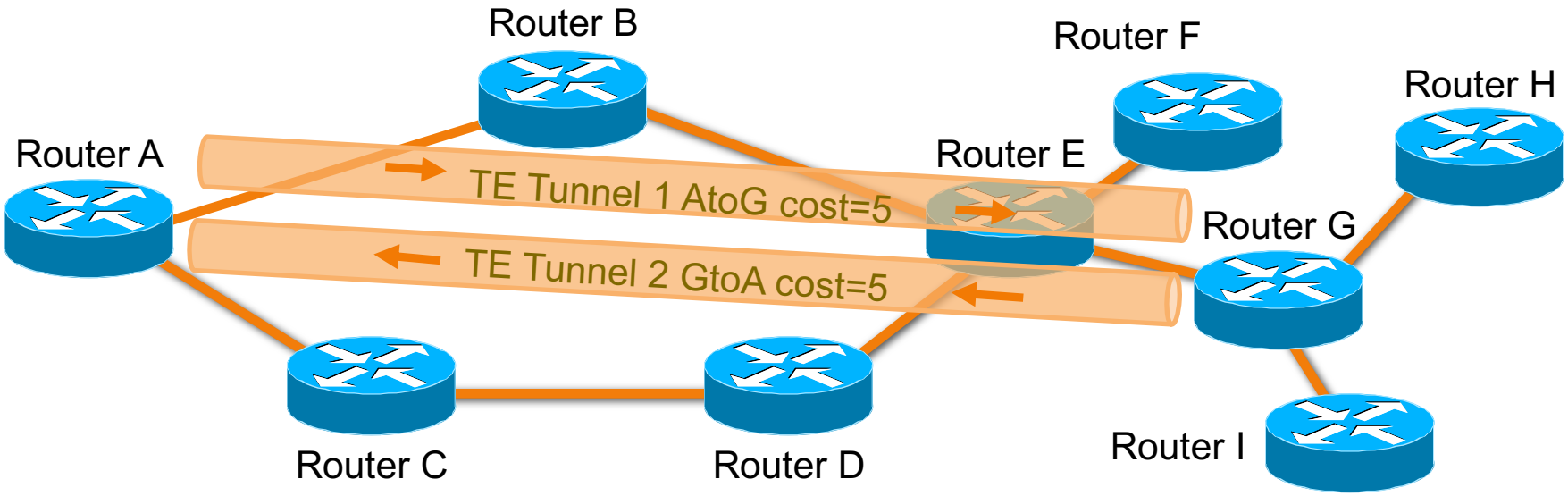


After Configuring Forwarding Adjacency

- After FA configured, Router G advertised as a link in an IGP network with the link's cost associated with it.
- Router H can see the TE tunnel and use it to compute the shortest path for routing traffic throughout the network.

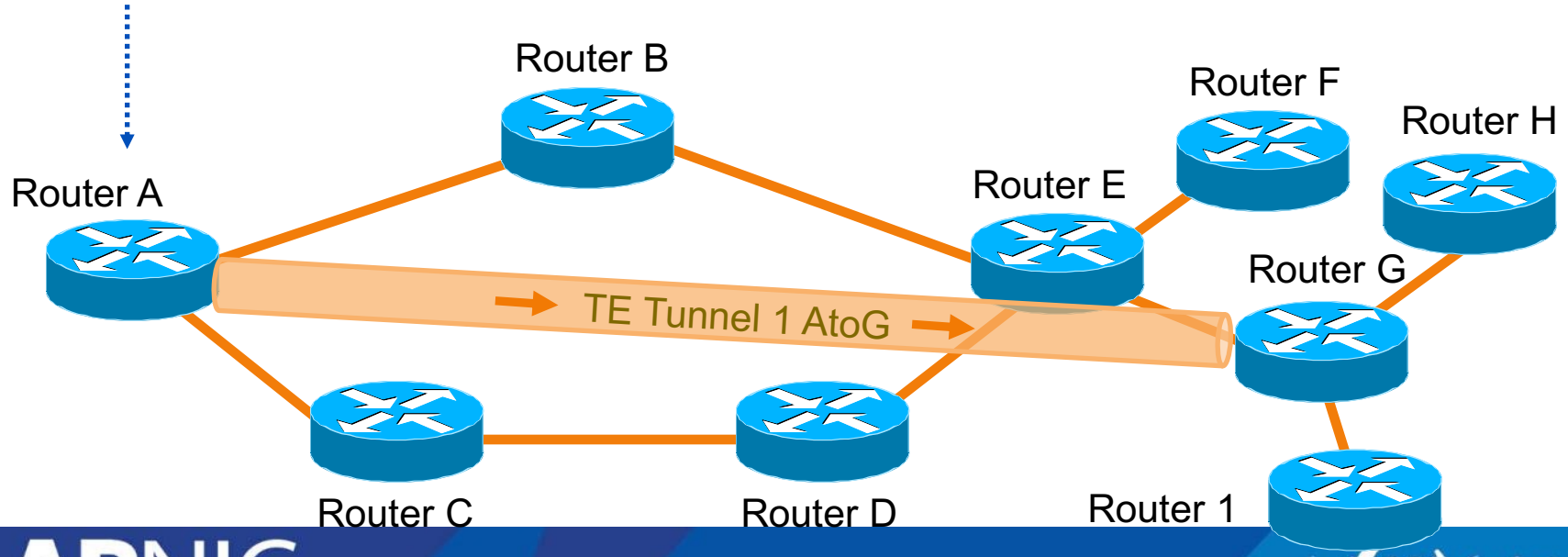
Routing Table of RouterH

Node	Next-Hop	Cost
A	G	15
…	…	...



Policy Routing

```
RtrA(config)# interface ethernet 1/0
RtrA(config-if)# ip policy route-map set-tunnel
RtrA(config)# route-map set-tunnel
RtrA(config-route-map)# match ip address 101
RtrA(config-route-map)# set interface Tunnel1
```

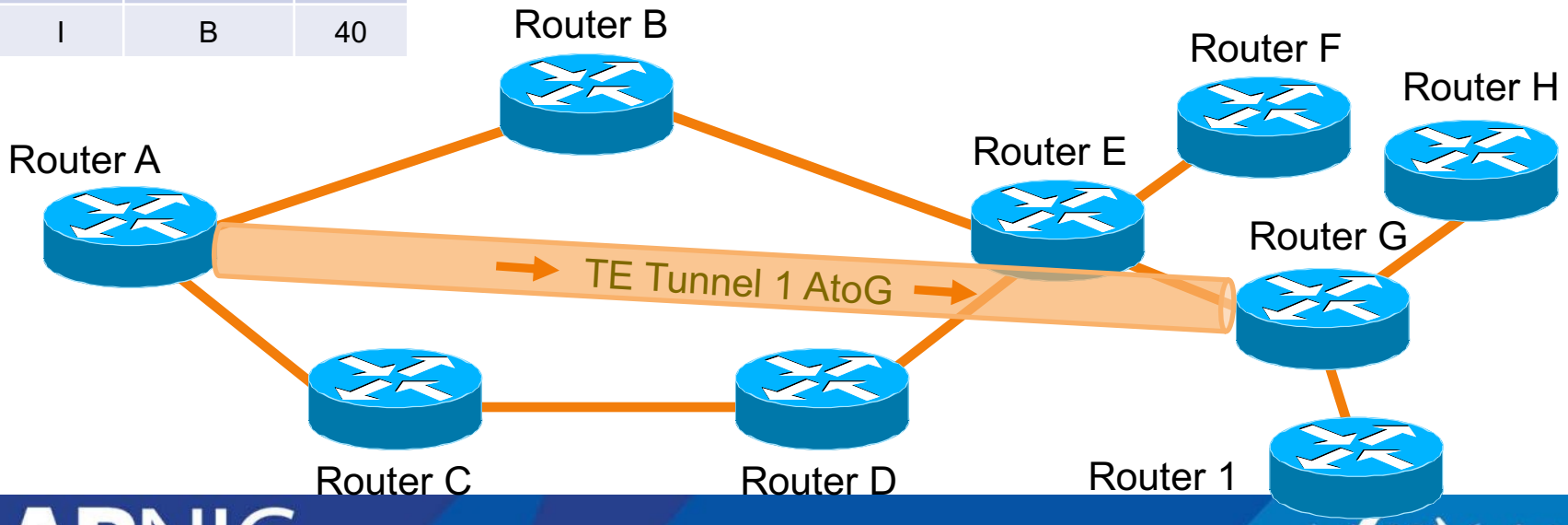


Policy Routing

Routing Table of RouterA

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	B	30
H	B	40
I	B	40

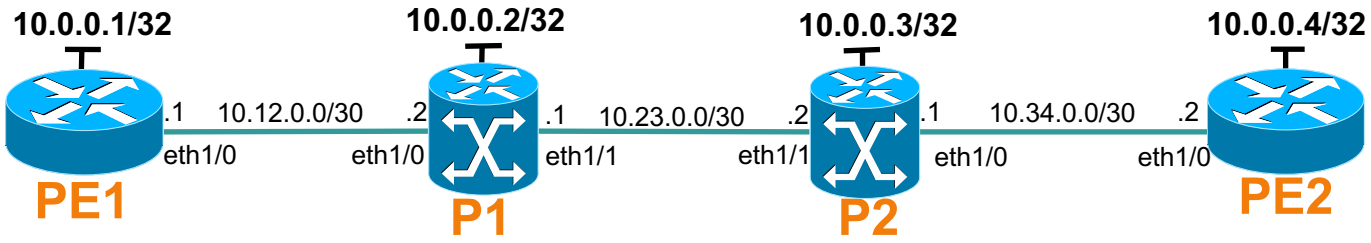
- Routing table isn't affected by policy routing
- Cisco IOS Need (12.0(23)S or 12.2T) or higher for 'set interface tunnel' to work



MPLS TE Configuration Example

Configuration Example

- Task: Bidirectional tunnels need to be set up between PE1 to PE2 (BW=500kbps), path will be selected automatically.
- Prerequisite configuration:
  - 1. IP address configuration on PE & P routers
  - 2. IGP configuration on PE & P routers
    - Make sure all the routers in public network can reach each other.
  - 3. Enable MPLS on the router and interfaces



Enable MPLS TE

- Configuration steps:
  - 1. Configure MPLS TE on all the routers and the reservable bandwidth on the router interface

```
mpls traffic-eng tunnels

interface ethernet1/0
mpls traffic-eng tunnels
ip rsvp bandwidth 2000

interface ethernet1/1
mpls traffic-eng tunnels
ip rsvp bandwidth 2000
```

Enable OSPF to Support TE

– 2. Enable OSPF to support MPLS TE on every router:

```
router ospf 1
mpls traffic-eng router-id loopback 0
mpls traffic-eng area 0
```

Configure Tunnel Interface on Head-end Routers

– 3. Configure tunnel interface on both PE1 and PE2

```
interface tunnel 10
ip unnumbered loopback 0
mpls ip
tunnel mode mpls traffic-eng
tunnel destination 10.0.0.4
tunnel mpls traffic-eng path-option 10 dynamic
tunnel mpls traffic-eng bandwidth 500
tunnel mpls traffic-eng autoroute announce
tunnel mpls traffic-eng priority 0 0
```

Verify MPLS TE and RSVP Bandwidth on Interfaces

- Check MPLS TE configured on interfaces:

```
P1#show mpls interfaces
Interface              IP            Tunnel   BGP Static Operational
Ethernet1/1            Yes (ldp)     Yes      No  No     Yes
Ethernet1/0            Yes (ldp)     Yes      No  No     Yes
```

- Check RSVP bandwidth status for interfaces:

```
PE1#show ip rsvp interface
interface    rsvp       allocated  i/f max   flow max sub max   VRF
Et1/0        ena        500K       2000K      2000K     0
```

Verification of MPLS TE Support

- Check MPLS TE support on each router:

```
PE1#show mpls traffic-eng tunnels summary
Signalling  Summary:
    LSP Tunnels Process:            running
    Passive LSP Listener:           running
    RSVP Process:                   running
    Forwarding:                     enabled
    Periodic reoptimization:        every 3600 seconds, next in 2747 seconds
    Periodic FRR Promotion:         Not Running
    Periodic auto-bw collection:    disabled
    P2P:
      Head: 1 interfaces,   1 active signalling attempts, 1 established
            2 activations,  1 deactivations
            822 failed activations
            0 SSO recovery attempts, 0 SSO recovered
      Midpoints: 0, Tails: 1
    P2MP:
      Head: 0 interfaces,   0 active signalling attempts, 0 established
            0 sub-LSP activations,  0 sub-LSP deactivations
            0 LSP successful activations,  0 LSP deactivations
            0 SSO recovery attempts, LSP recovered: 0 full, 0 partial, 0 fail
      Midpoints: 0, Tails: 0
```

Verify Tunnel Interface Status

- Check tunnel interfaces on head-end routers:

```
PE1#show ip interface brief
Interface              IP-Address      OK? Method Status                Protocol
Ethernet1/0            10.12.0.1       YES NVRAM  up                    up
Tunnel10               10.0.0.1        YES TFTP   up                    up
```

```
PE1#show mpls traffic-eng tunnels brief
Signalling  Summary:
    LSP Tunnels Process:            running
    Passive LSP Listener:           running
    RSVP Process:                   running
    Forwarding:                     enabled
    Periodic reoptimization:        every 3600 seconds, next in 7 seconds
    Periodic FRR Promotion:         Not Running
    Periodic auto-bw collection:    disabled
P2P TUNNELS/LSPs:
TUNNEL NAME                     DESTINATION     UP IF      DOWN IF    STATE/PROT
Router1_t10                     10.0.0.4        -          Et1/0      up/up
Router4_t10                     10.0.0.1        Et1/0      -          up/up
Displayed 1 (of 1) heads, 0 (of 0) midpoints, 1 (of 1) tails
```

Verify Reachability

- Check ip routing table

```
PE1#show ip route
C        10.0.0.1/32 is directly connected, Loopback0
O        10.0.0.2/32 [110/11] via 10.12.0.2, 1d08h, Ethernet1/0
O        10.0.0.3/32 [110/21] via 10.12.0.2, 09:08:57, Ethernet1/0
O        10.0.0.4/32 [110/31] via 10.0.0.4, 00:15:34, Tunnel10
```

- Check LSP

```
PE1#traceroute mpls traffic-eng tunnel 10
Tracing MPLS TE Label Switched Path on Tunnel10, timeout is 2 seconds
Type escape sequence to abort.
  0 10.12.0.1 MRU 1500 [Labels: 16 Exp: 0]
L 1 10.12.0.2 MRU 1500 [Labels: 16 Exp: 0] 16 ms
L 2 10.23.0.2 MRU 1504 [Labels: implicit-null Exp: 0] 20 ms
! 3 10.34.0.2 20 ms
```

Questions?

APNIC