MPLS-TE Introduction and Case Study

Liu Yong
Technical Support Engineer

2014/08/20

Agenda

- Understanding Traffic Engineering with MPLS
- Link Information Distribution
- Path Calculation
- Path Setup
- Forwarding Traffic Down Tunnels
- Basic Configuration Review
- Case Study

Understanding Traffic Engineering with MPLS

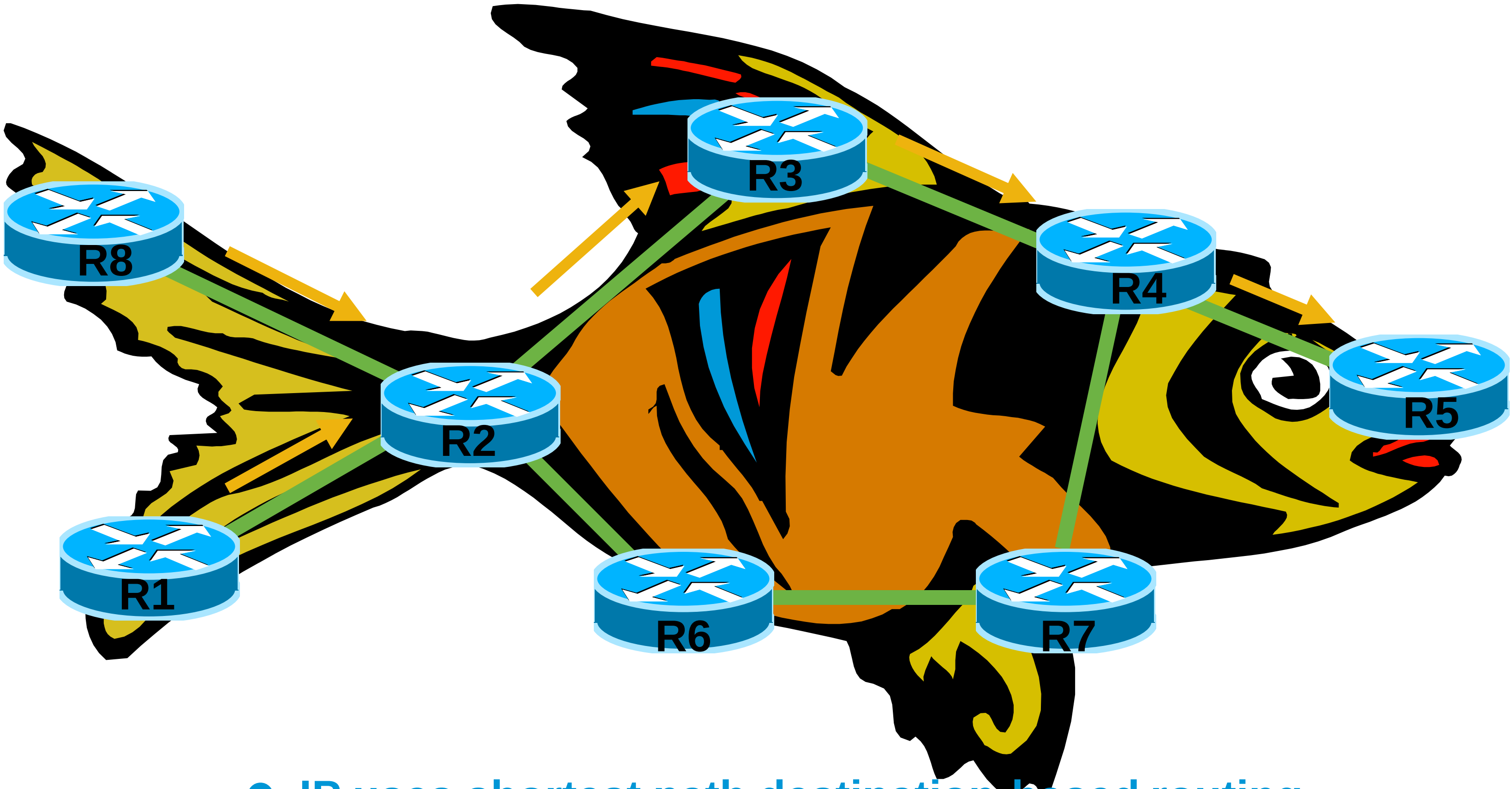
The meaning of Traffic Engineering

- Traffic Engineering is manipulating the traffic to fit the network
- Construct routes for traffic streams within a service provider in such a way, as to avoids causing some parts of the provider's network to be over-utilized, while others parts remain under-utilized

Motivation for Traffic Engineering

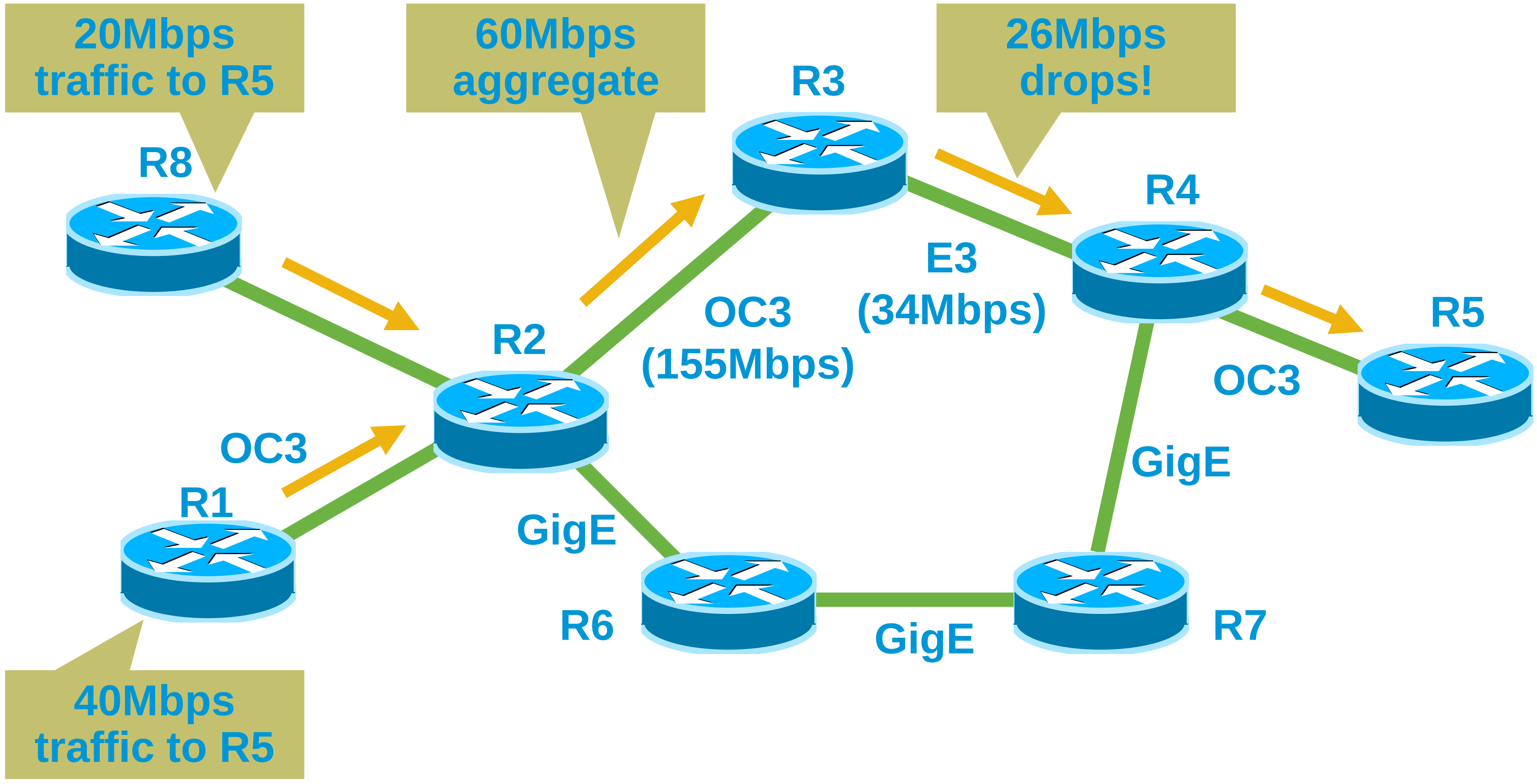
- Reduce the overall cost of operations by more efficient use of bandwidth resources
  - By preventing a situation where some parts of a service provider network are over-utilized (congested), while other parts under-utilized
- Ensures the most desirable/appropriate path for certain traffic types based on certain policies
  - Override the shortest path selected by the IGP (designed to move traffic along a path other than the IGP shortest path.)
- The ultimate goal is COST SAVING

The “Fish” Problem (Shortest Path)



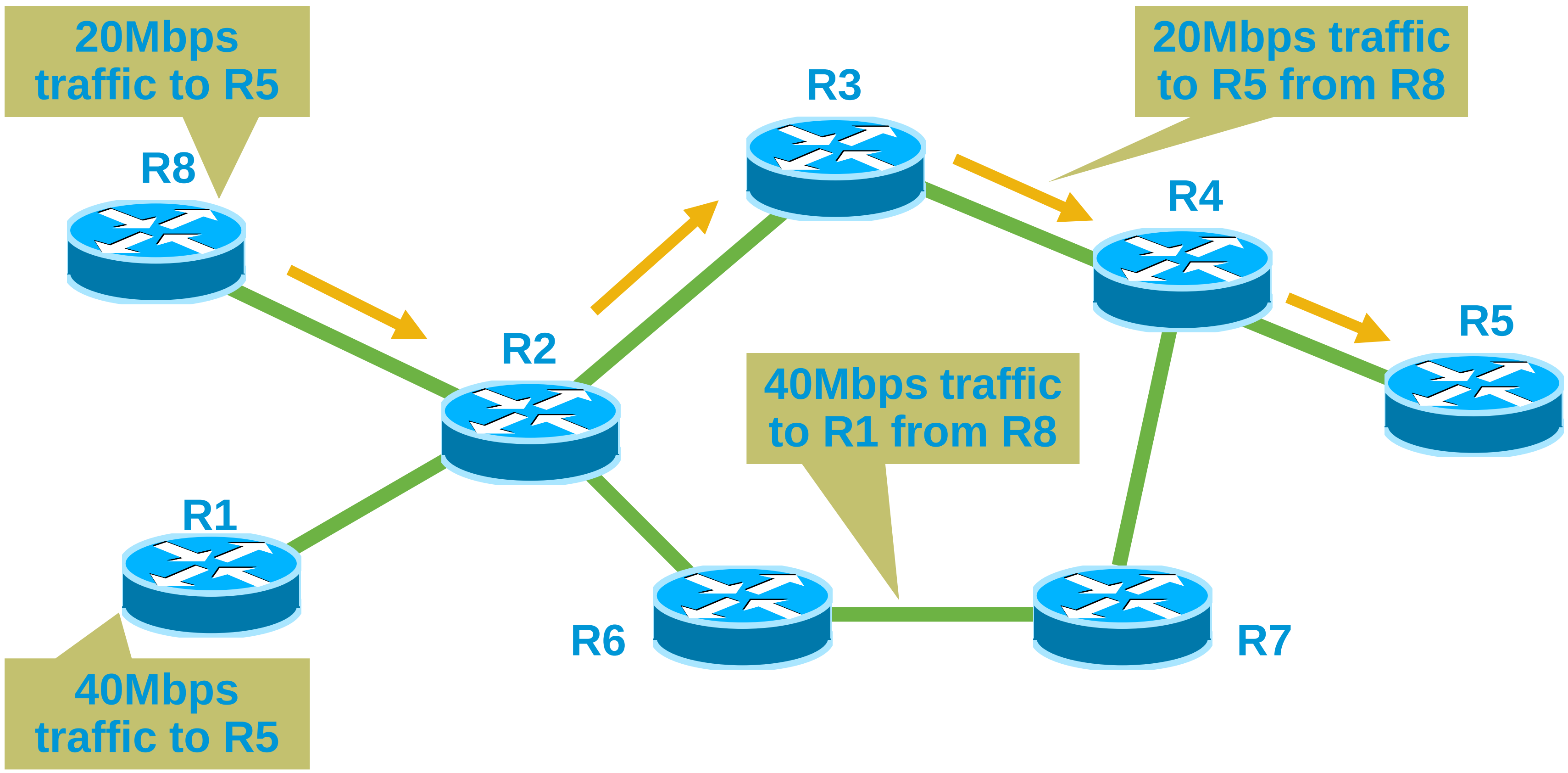
- IP uses shortest path destination-based routing
- Shortest path may not be the only path
- Alternate paths may be under-utilized
- Whilst the shortest path Is over-utilized

Shortest Path and congestion



EDCS-1171743

The MPLS TE solution



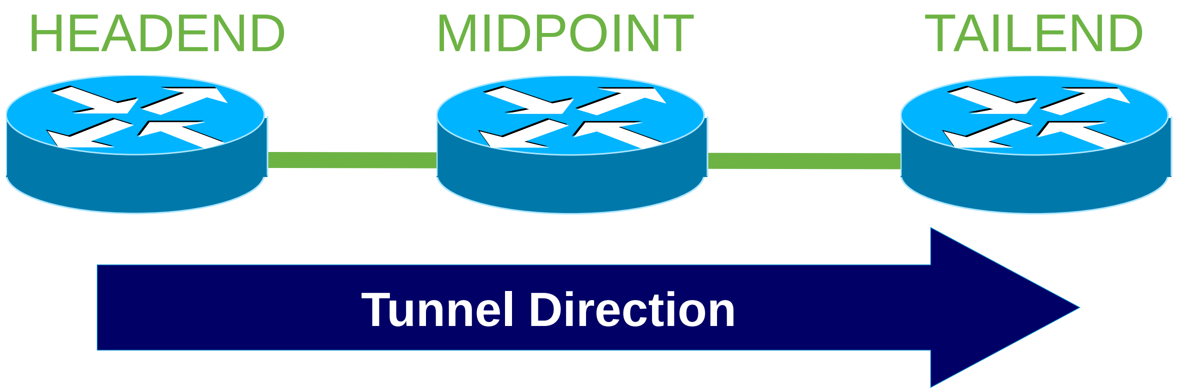
Terminology

- Constrained-Based Shortest Path First (CSPF)

  MPLS-TE uses CSPF to create a shortest path based on a series of constraints:

  - Bandwidth
  - Affinity/Link Attributes
  - Administrative weight

- Tunnels are UNI-DIRECTIONAL

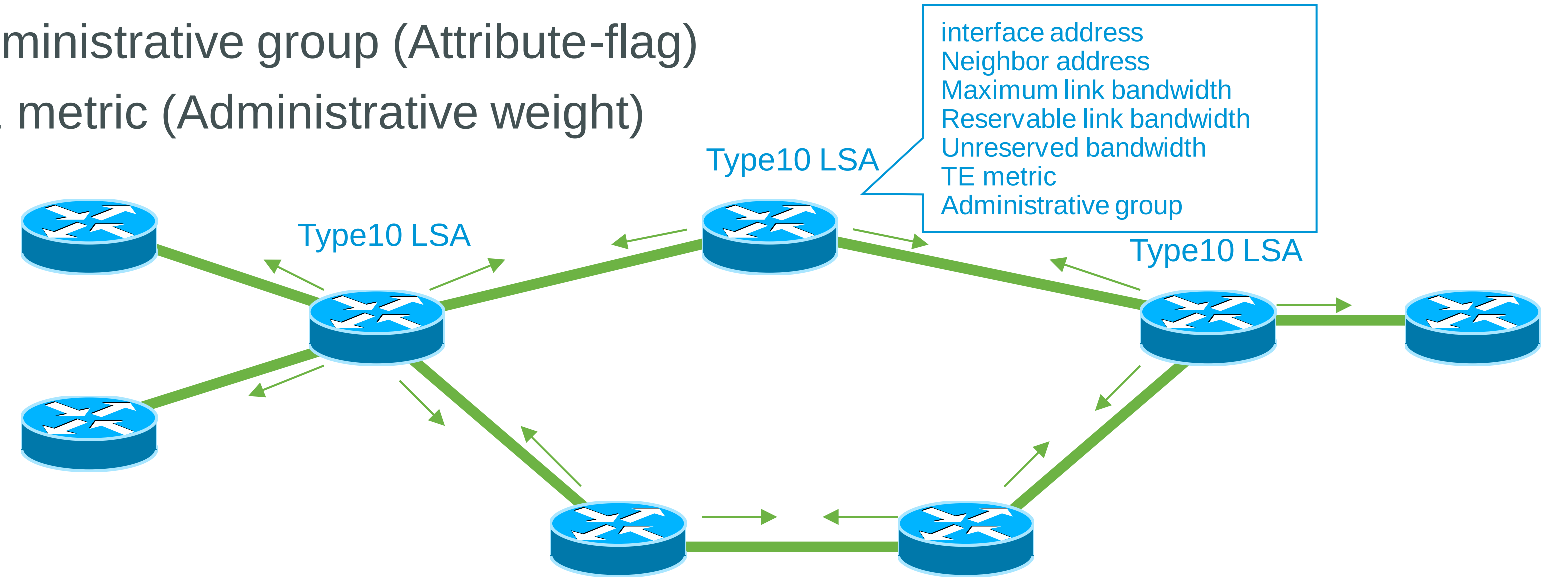


Basic Operation of MPLS TE

- Link Information Distribution
- Path Calculation
- Path Setup
- Forwarding Traffic Down Tunnels

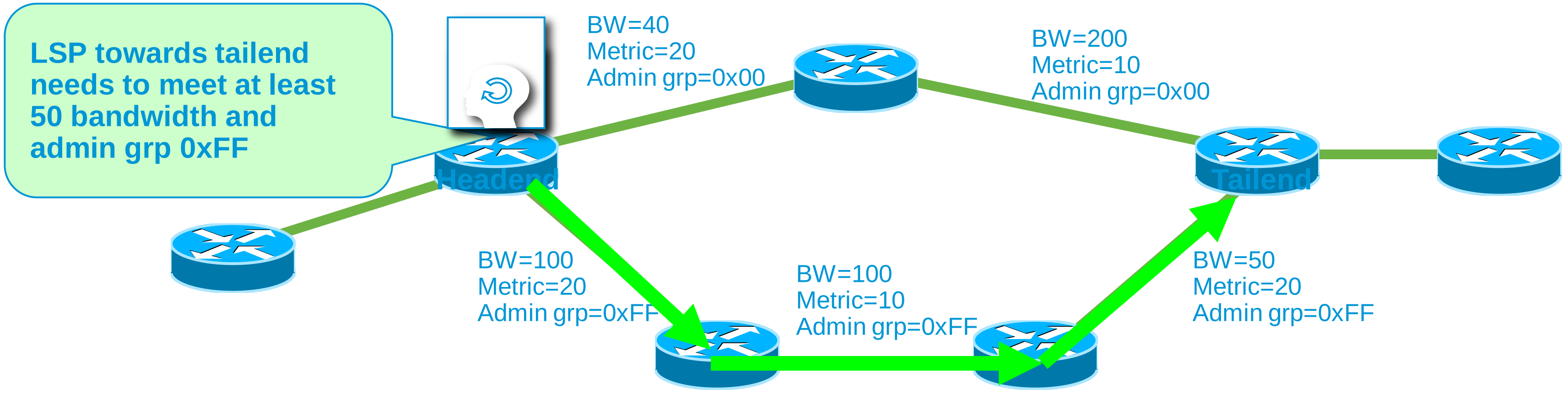
Link Information Distribution

- Extensions to existing IP link-state routing protocols (IS-IS and OSPF) for performing constraint-based calculation at LSRs
  - OSPF: type 10 opaque (area-local scope) LSA
  - IS-IS: TLV type 22
- Introducing new link attributes
  - Available bandwidth … meaningful only in control plane
  - Administrative group (Attribute-flag)
  - TE metric (Administrative weight)



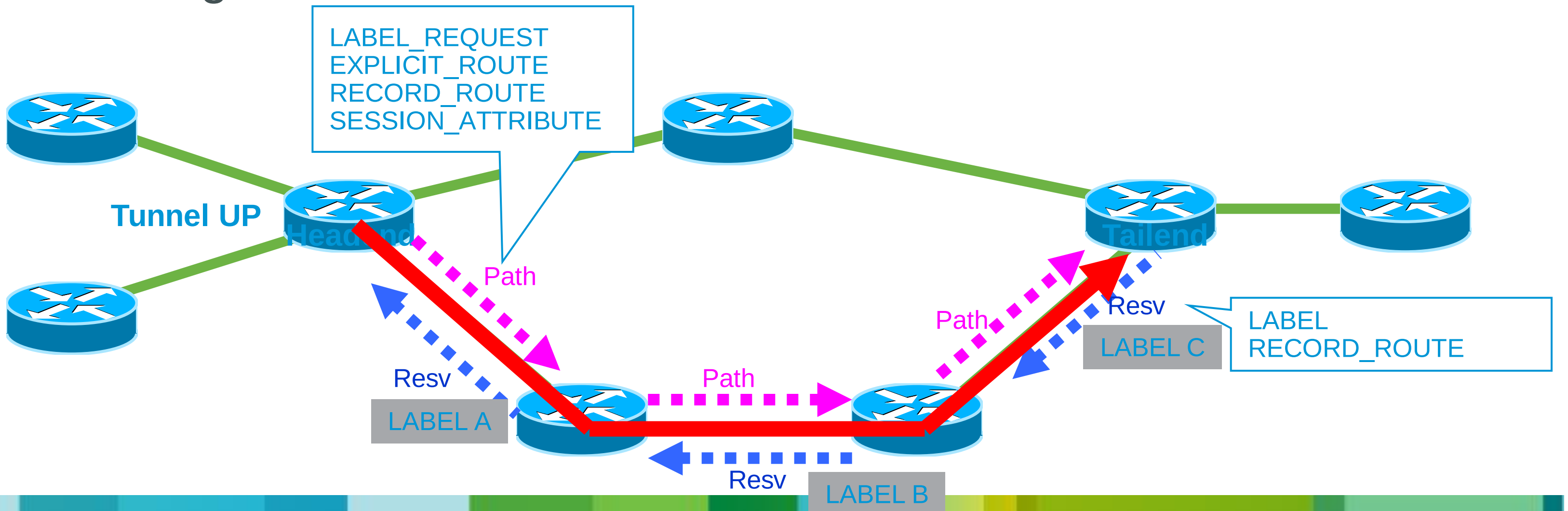
Path Calculation

- Once extended link information are flooded, LSRs can perform path calculation for the path from headend to tailend
  - Path may be explicitly configured by operator
- TE Headend does a “Constrained SPF” (CSPF) calculation to find the best path
- CSPF is just like regular IGP SPF, but it precludes links that do not meet the TE LSP constraints and uses TE metrics
  - e.g. required bandwidth, administrative group, and TE metric are taken into account
- Looks for the best path from a head to a single tail (unlike OSPF)



Path Setup

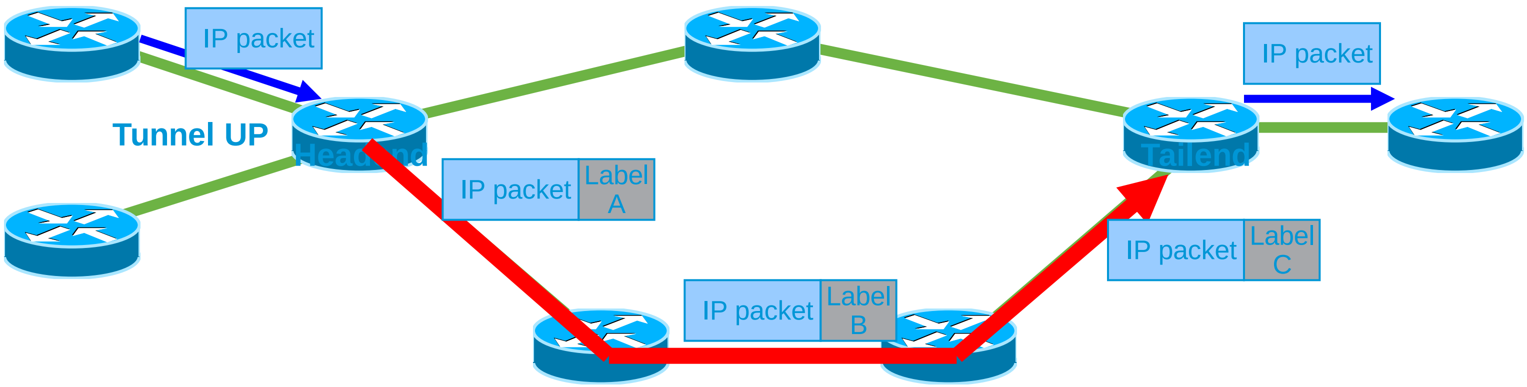
- Once the path is calculated, it must be signaled across the network
  - Reserve any bandwidth to avoid "double booking" from other TE reservations
  - Priority can be used to pre-empt low priority existing tunnels
- Extensions of RSVP used to set up TE LSP
  - Path messages (from head to tail) carries LABEL_REQUEST
  - Resv messages (from tail to head) carries LABEL
- When Resv reaches back to headend, tunnel interface = UP/UP
- RSVP messages are also used for LSP teardown and error notification



EDCS-1171743

Forwarding Traffic Down Tunnels

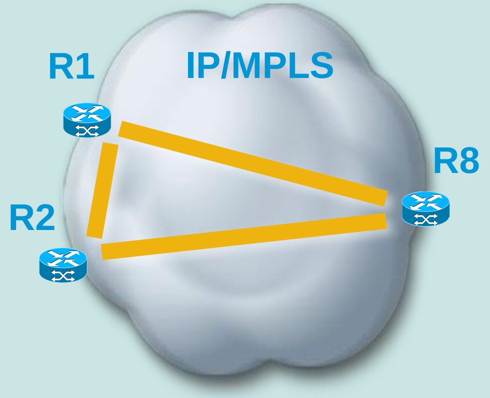
- Once TE LSP is established, TE tunnel has to be reflected into forwarding plane to flow the traffic on that LSP
- There are 4 options to forward traffic down TE tunnels;
  - Static route
  - Policy routing
  - Autoroute
    - Supporting Class-Based Tunnel Selection (CBTS) … (IOS)
    - Supporting Policy-Based Tunnel Selection (PBTS) … (IOS XR)
    - Supporting Multicast intact for preserving RPF lookup
  - Forwarding adjacency



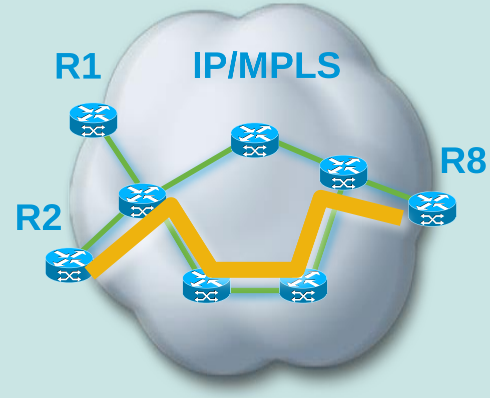
MPLS TE Deployment Models

Bandwidth Optimization

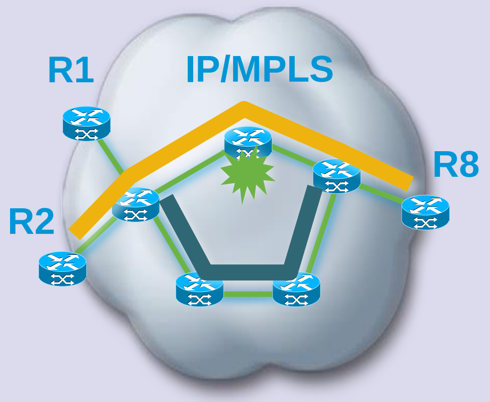
Strategic



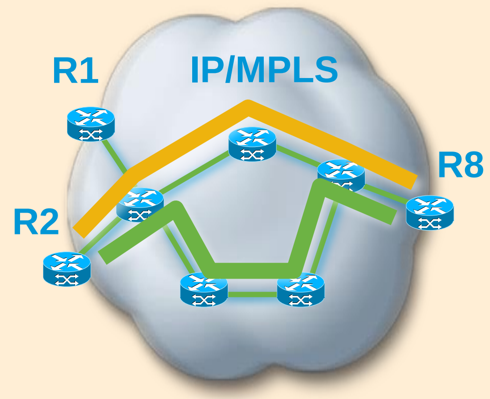
Tactical



Protection



Point-to-Point SLA



Link Information Distribution

MPLS TE prerequisites

- CEF is enabled on all of Label Switched Routers (LSRs)
- Using Link-state routing protocol (OSPF or IS-IS) as IGP
- A loopback interface to use as MPLS TE RID
  - Usually using the same loopback interface as is used at IGP RID
  - Address of loopback interface must be advertised via IGP
- Enabling MPLS TE

Enabling MPLS TE

IOS Configuration of Enabling MPLS TE

```
mpls traffic-eng tunnels
!
interface POS2/3
 ip address 10.1.3.1 255.255.255.0
 no ip directed-broadcast
 mpls traffic-eng tunnels
```

IOS XR Configuration of Enabling MPLS TE

```
mpls traffic-eng
 interface TenGigE0/1/0/0
 !
 interface TenGigE0/1/0/4
```

- All of LSRs which participate in MPLS TE must be enabled of MPLS TE explicitly

  `(config)#mpls traffic-eng tunnels` (IOS)

  `(config)#mpls traffic-eng` (IOS XR)

- All of interfaces on which TE tunnels go out and come in must be enabled of MPLS TE

  `(config-if)#mpls traffic-eng tunnels` (IOS)

  `(config-mpls-te)#interface {`*`interface`*`}` (IOS XR)

Enabling RSVP (IOS XR only)

IOS XR Configuration of Enabling RSVP

```
rsvp
 interface TenGigE0/1/0/0
 !
 interface TenGigE0/1/0/2
```

- When enabling MPLS TE on IOS XR, all of interfaces on which TE tunnels go out and come in must be enabled of MPLS TE

  `(config-rsvp)#interface {`*`interface`*`)` (IOS XR)

- For IOS, this step is not required since enabling MPLS TE for the interfaces also enables RSVP on those ones

Basic TE tunnel interface definition

IOS Configuration of TE tunnel interface

```
interface Tunnel13
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 3.3.3.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng path-option 90 dynamic
```

IOS XR Configuration of TE tunnel interface

```
interface tunnel-te1
 ipv4 unnumbered Loopback0
 destination 3.3.3.3
 path-option 90 dynamic
```

- Definition of tunnel interface in IOS is same with the other kind of tunnels e.g. L2TP or GRE, therefore ‘`tunnel mode mpls traffic-eng`’ is required and IOS XR identifies MPLS TE tunnel interfaces with the `tunnel-te` keyword
- Destination is RID of tailend router
- IP address for TE tunnel is have to be configured as unnumbered specifying loopback interface of its own TE RID

How Information is distributed

- Link information for TE LSP calculation is flooded using OSPF type 10 opaque-LSAs or IS-IS TLV type 22
  - *This document focus on the case using OSPF as IGP, therefore the description of IS-IS is omitted*
- Separate LSDB is created for CSPF TE LSP calculation

How Information is distributed
OSPF link information distribution

IOS Configuration of OSPF link information distribution

```
router ospf 100
 router-id 1.1.1.1
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng area 0
```

IOS XR Configuration of OSPF link information distribution

```
router ospf 100
 router-id 5.5.5.5
 area 0
  mpls traffic-eng
  interface Loopback0
   passive enable
  !
!
 mpls traffic-eng router-id Loopback0
```

- The command '`mpls traffic-eng router-id {`*`if-te_rid`*`}`' specifies TE RID, and it should be same with OSPF RID
- The configuration is also required for specifying which area the type 10 opaque LSAs should be flooded

  `(config-router)#mpls traffic-eng area {`*`area#`*`}` (IOS)

  `(config-ospf-ar)#mpls traffic-eng` (IOS XR)

What information is distributed

- Three major pieces of information are;
  - Available bandwidth
  - Attribute flags
  - TE metric (Administrative weight)

EDCS-1171743

Available Bandwidth

IOS Configuration of Max Reservable bandwidth

```
interface POS2/3
 ip address 10.1.3.1 255.255.255.0
 no ip directed-broadcast
 mpls traffic-eng tunnels
 ip rsvp bandwidth 100000
```

IOS XR Configuration of Max Reservable bandwidth

```
rsvp
 interface TenGigE0/1/0/0
  bandwidth
 !
 interface TenGigE0/1/0/2
  bandwidth 100000
```

- Default reservable bandwidth is 0
- If the bandwidth value is omitted at '`ip rsvp bandwidth`' ('`bandwidth`' with IOS XR), reservable bandwidth is configured as 75% of link bandwidth
- The configuration is not necessarily required if any of LSPs does not need bandwidth reservation
- If bandwidth reservation is required for LSP, the above configuration is required for each physical links which are possibly a part of LSP

Verification of Available Bandwidth

IOS Verification of Max Reservable bandwidth

```
P1_MF12410B#show ip rsvp interface

interface      allocated   i/f max   flow max sub max
PO0/0          0           116250K   116250K  0
PO0/1          0           116250K   116250K  0
PO0/2          0           116250K   116250K  0
PO1/0          0           116250K   116250K  0
PO2/3          50M         100M      100M     0
```

IOS XR Verification of Max Reservable bandwidth

```
RP/0/1/CPU0:MF12404C-XR_PE5#show rsvp interface
Interface    MaxBW (bps) MaxFlow (bps) Allocated (bps)      MaxSub (bps)
----------- ----------- ------------- -------------------- ------------
tt53146               0             0             0 (   0%)            0
Te0/2/0/0         7500M         7500M          100M (   1%)            0
```

Reserved Bandwidth

IOS Configuration of Reserved bandwidth for TE tunnel

```
interface Tunnel13
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 3.3.3.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng priority 7 7
 tunnel mpls traffic-eng bandwidth  50000
```

IOS XR Configuration of Reserved bandwidth for TE tunnel

```
interface tunnel-te1
 ipv4 unnumbered Loopback0
 signalled-bandwidth 50000
 destination 3.3.3.3
```

- LSP Bandwidth requested for reservation can be configured under tunnel interface
- Default reserved bandwidth is 0
- The configuration is not necessarily required when the LSP does not need bandwidth reservation
- Non zero bandwidth reservation is required if TE tunnel links have to be advertised with using forwarding adjacency in IOS XR

Verification of Reserved Bandwidth

IOS Verification of Reserved bandwidth for TE tunnel

```
P1_MF12410B#show mpls traffic-eng tunnels Tunnel13

Name: P1_MF12410B_t13                    (Tunnel13) Destination: 3.3.3.3
  Status:
    Admin: up         Oper: up     Path: valid       Signalling: connected

    path option 20, type dynamic (Basis for Setup, path weight 1)
    path option 10, type explicit t13-1

  Config Parameters:
    Bandwidth: 50000    kbps (Global)  Priority: 7  7   Affinity: 0x0/0xFFFF
    Metric Type: TE (default)
    AutoRoute:  enabled   LockDown: disabled  Loadshare: 50000    bw-based
    auto-bw: disabled

  InLabel  :  -
  OutLabel : POS2/3, implicit-null

P1_MF12410B#show mpls traffic-eng topology 1.1.1.1

link[2]: Point-to-Point, Nbr IGP Id: 3.3.3.3, nbr_node_id:11, gen:160
          frag_id 2, Intf Address:10.1.3.1, Nbr Intf Address:10.1.3.3
          TE metric:1, IGP metric:1, attribute_flags:0x0
          physical_bw: 2488000 (kbps), max_reservable_bw_global: 100000 (kbps)
          max_reservable_bw_sub: 0 (kbps)

                                  Global Pool       Sub Pool
               Total Allocated   Reservable        Reservable
               BW (kbps)         BW (kbps)         BW (kbps)
               ---------------   -----------       ----------
        bw[0]:            0         100000                 0
        bw[1]:            0         100000                 0
        bw[2]:            0         100000                 0
        bw[3]:            0         100000                 0
        bw[4]:            0         100000                 0
        bw[5]:            0         100000                 0
        bw[6]:            0         100000                 0
        bw[7]:        50000          50000                 0
```

Attribute Flags

IOS Configuration of Attribute flags for the link

```
interface POS2/3
 ip address 10.1.3.1 255.255.255.0
 no ip directed-broadcast
 mpls traffic-eng tunnels
 mpls traffic-eng attribute-flags 0xF
```

IOS XR Configuration of Attribute flags for the link

```
mpls traffic-eng
 interface TenGigE0/1/0/0
  attribute-flags 0xF
```

- Attribute flags are 32 bits attributes which can be freely configured for each links
- Default attribute flags value is 0x0
- With using affinity and mask configurations for TE tunnel, LSPs can be controlled as which links should be included or excluded before CSPF calculation
- The configuration is not necessarily required when the LSP does not need specific policy for link inclusion / exclusion

Verification of attribute flags

IOS Verification of attribute flags for the link

```
P1_MF12410B#show mpls traffic-eng topology 1.1.1.1

link[2]: Point-to-Point, Nbr IGP Id: 3.3.3.3, nbr_node_id:11, gen:160
          frag_id 2, Intf Address:10.1.3.1, Nbr Intf Address:10.1.3.3
          TE metric:1, IGP metric:1, attribute_flags:0x0
          physical_bw: 2488000 (kbps), max_reservable_bw_global: 100000 (kbps)
          max_reservable_bw_sub: 0 (kbps)

                                  Global Pool       Sub Pool
               Total Allocated    Reservable        Reservable
               BW (kbps)          BW (kbps)         BW (kbps)
               ---------------    -----------       ----------
        bw[0]:            0          100000                0
        bw[1]:            0          100000                0
        bw[2]:            0          100000                0
        bw[3]:            0          100000                0
        bw[4]:            0          100000                0
        bw[5]:            0          100000                0
        bw[6]:            0          100000                0
        bw[7]:        50000           50000                0
```

Attribute Flags - Affinity bits

IOS Configuration of affinity and mask for TE tunnel

```
interface Tunnel13
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 3.3.3.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng affinity 0x1 mask 0x1
```

IOS XR Configuration of affinity and mask for TE tunnel

```
interface tunnel-te1
 ipv4 unnumbered Loopback0
 destination 3.3.3.3
 affinity 1 mask 1
```

- With using attribute flags configurations for the links, LSPs can be controlled as which links should be included or excluded before CSPF calculation
- If (tunnel Affinity && Mask)==(link Attributes && Mask), the link is considered as match, so bit 1 in mask means care bit
- Default affinity/mask combination is 0x0/0xffff (0x0000ffff)
- The configuration is not necessarily required when the LSP does not need specific policy for link inclusion / exclusion

Verification of affinity and mask

IOS Verification of affinity and mask for TE tunnel

```
P1_MF12410B#show mpls traffic-eng tunnels Tunnel13

Name: P1_MF12410B_t13                     (Tunnel13) Destination: 3.3.3.3
  Status:
    Admin: up          Oper: up     Path: valid       Signalling: connected

    path option 20, type dynamic (Basis for Setup, path weight 1)
    path option 10, type explicit t13-1

  Config Parameters:
    Bandwidth: 50000    kbps (Global)  Priority: 7  7   Affinity: 0x0/0xFFFF
    Metric Type: TE (default)
    AutoRoute:  enabled   LockDown: disabled  Loadshare: 50000    bw-based
    auto-bw: disabled

  InLabel  :  -
  OutLabel : POS2/3, implicit-null

<SNIP>
```

Administrative Weight (TE metric)

IOS Configuration of administrative weight for the link

```
interface POS2/3
 ip address 10.1.3.1 255.255.255.0
 no ip directed-broadcast
 mpls traffic-eng tunnels
 mpls traffic-eng administrative-weight 10
```

IOS XR Configuration of administrative weight for the link

```
mpls traffic-eng
 interface TenGigE0/1/0/0
  admin-weight 10
```

- The default TE metric is same as the IGP metric but TE metric can be manually changed independent to IGP metric for each links
- Either of TE or IGP metric can be used for CSPF calculation

  Preferred metric for each TE tunnel can be configured with command;

  `(config-if)#tunnel mpls traffic-eng path-selection metric {te | igp}` (IOS)

  `(config-if)#path-selection metric {te | igp}` (IOS XR)

  Default preferred metric is TE

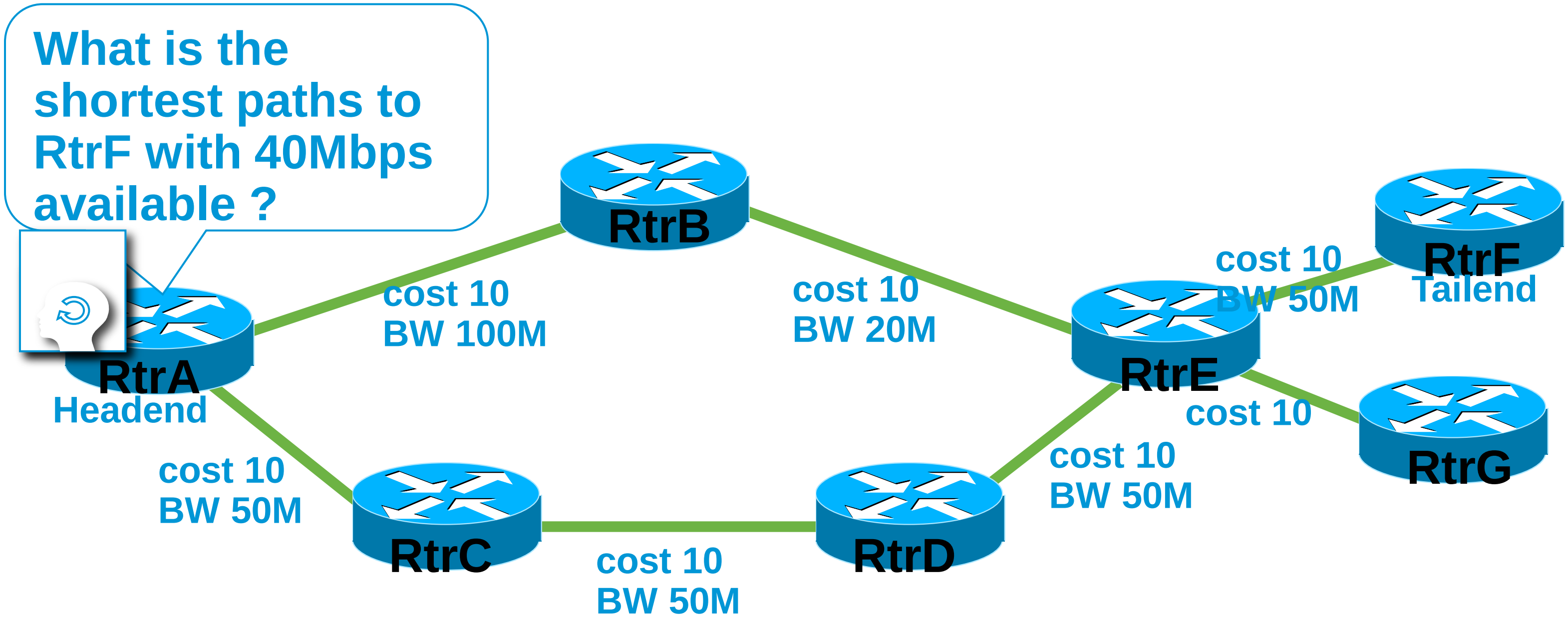
Verification of TE metric

IOS Verification of TE metric for link and preference of TE tunnel

```
P1_MF12410B#show mpls traffic-eng tunnels Tunnel13

Name: P1_MF12410B_t13                     (Tunnel13) Destination: 3.3.3.3
  Status:
    Admin: up         Oper: up     Path: valid       Signalling: connected

    path option 20, type dynamic (Basis for Setup, path weight 1)
    path option 10, type explicit t13-1

  Config Parameters:
    Bandwidth: 50000    kbps (Global)  Priority: 7  7   Affinity: 0x0/0xFFFF
    Metric Type: TE (default)
    AutoRoute:  enabled   LockDown: disabled  Loadshare: 50000    bw-based
    auto-bw: disabled

  InLabel  :  -
  OutLabel : POS2/3, implicit-null

P1_MF12410B#show mpls traffic-eng topology 1.1.1.1

link[2]: Point-to-Point, Nbr IGP Id: 3.3.3.3, nbr_node_id:11, gen:160
          frag_id 2, Intf Address:10.1.3.1, Nbr Intf Address:10.1.3.3
          TE metric:1, IGP metric:1, attribute_flags:0x0
          physical_bw: 2488000 (kbps), max_reservable_bw_global: 100000 (kbps)
          max_reservable_bw_sub: 0 (kbps)

                                  Global Pool       Sub Pool
               Total Allocated    Reservable        Reservable
               BW (kbps)          BW (kbps)         BW (kbps)
               ---------------    -----------       ----------
        bw[0]:             0          100000                 0
        bw[1]:             0          100000                 0
        bw[2]:             0          100000                 0
        bw[3]:             0          100000                 0
        bw[4]:             0          100000                 0
        bw[5]:             0          100000                 0
        bw[6]:             0          100000                 0
        bw[7]:         50000           50000                 0
```

Path Calculation

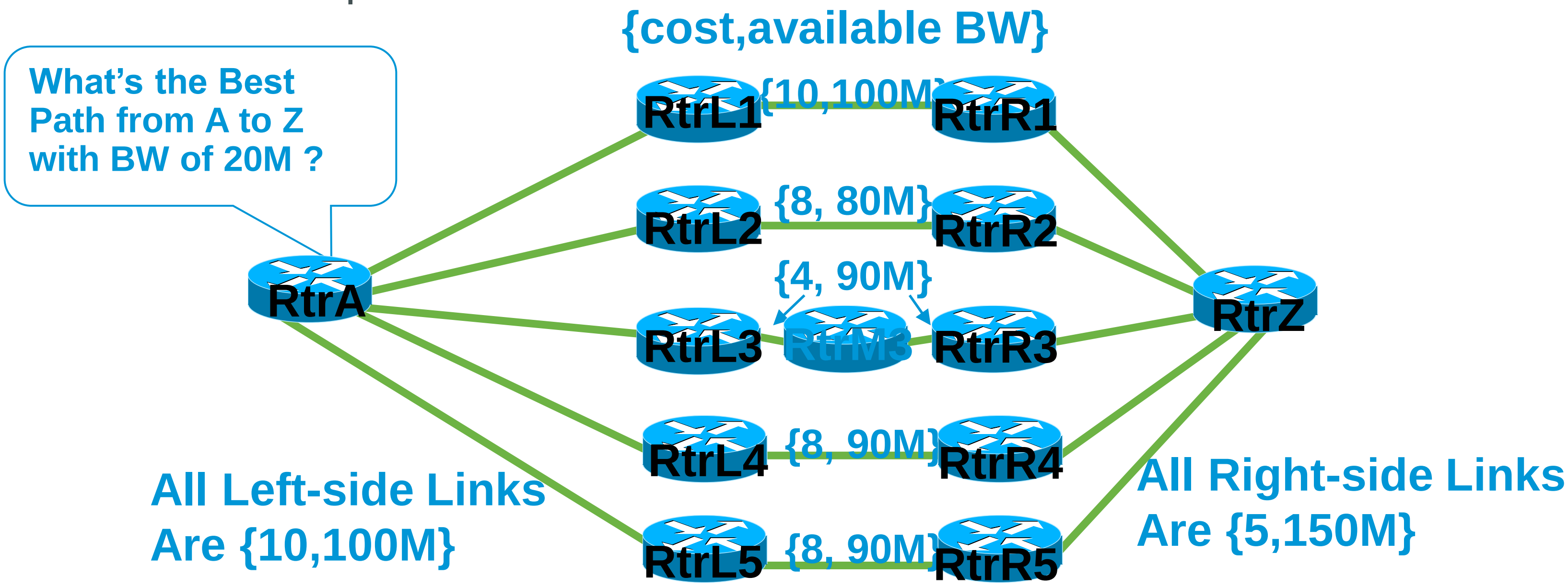
How CSPF (Constrained SPF) works



- CSPF also uses Dijkstra algorism
- CSPF is not designed to find the best route to all routers, but only to tunnel tailend
- CSPF considers link information e.g. reservable bandwidth, attribute flags, and TE metric, instead of single IGP cost
- CSPF is also called PCALC (Path calculation)

Tiebreakers in CSPF

- Best path is selected according to the following order;
  1. Find all paths with the lowest IGP cost
  2. Then pick the path with the highest minimum bandwidth along the path
  3. Then pick the path with the lowest hop count (number of routers in the path)
  4. Pick one path at random if there is still a tie



EDCS-1171743

Path configuration (CSPF Knobs)

IOS Configuration of path-option for TE tunnel

```
interface Tunnel13
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 3.3.3.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng path-option 10 explicit name t13-1
 tunnel mpls traffic-eng path-option 20 dynamic
```

IOS XR Configuration of path-option for TE tunnel

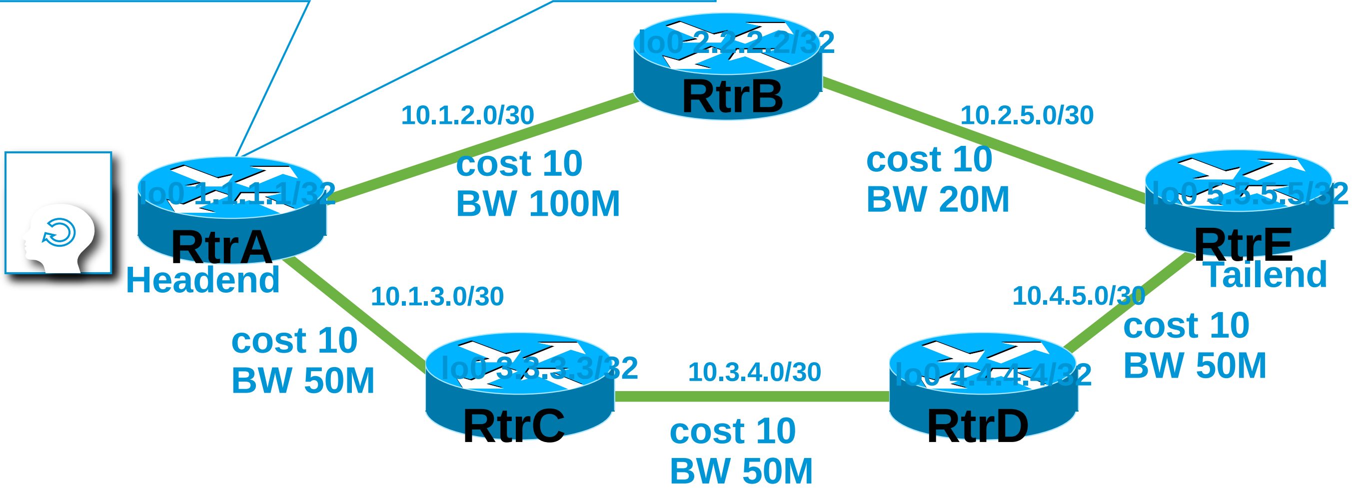
```
interface tunnel-te1
 ipv4 unnumbered Loopback0
 destination 3.3.3.3
 path-option 10 explicit name PATH1
 path-option 20 dynamic
```

- LSP can be controlled with configuring one or more path-options for TE tunnel interfaces
- CSPF is done at the headend according to the criteria configured path-options instruct
- Two kind of path-options can be configured; explicit and dynamic

Dynamic path-option

- CSPF is simply done for path to tailend with constraints; bandwidth, affinity, and TE metric

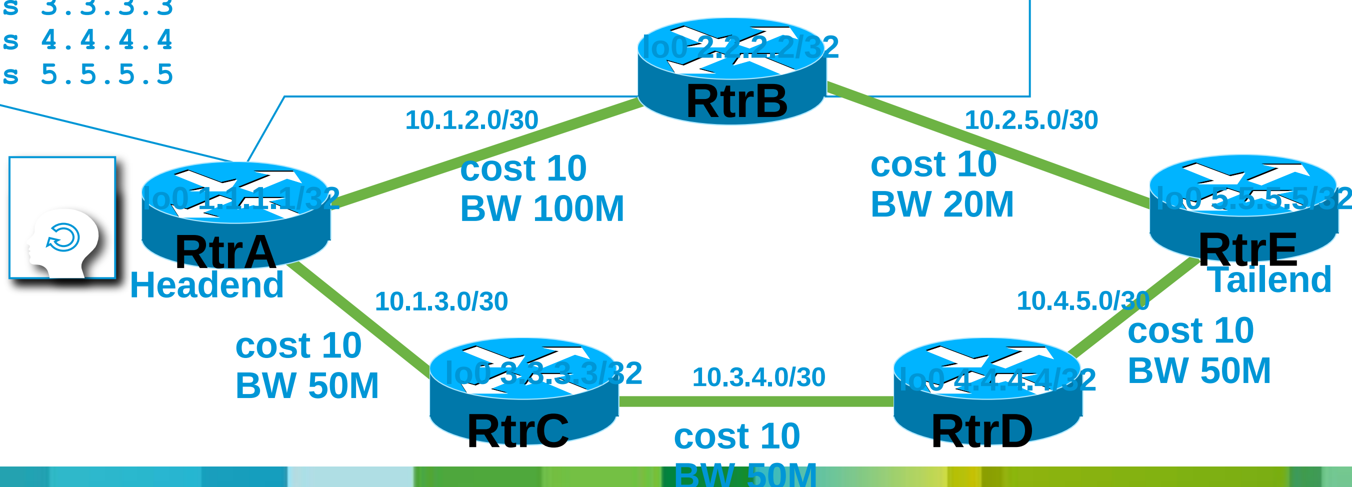
```
interface Tunnel15
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 5.5.5.5
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng bandwidth 40000
 tunnel mpls traffic-eng path-option 10 dynamic
```



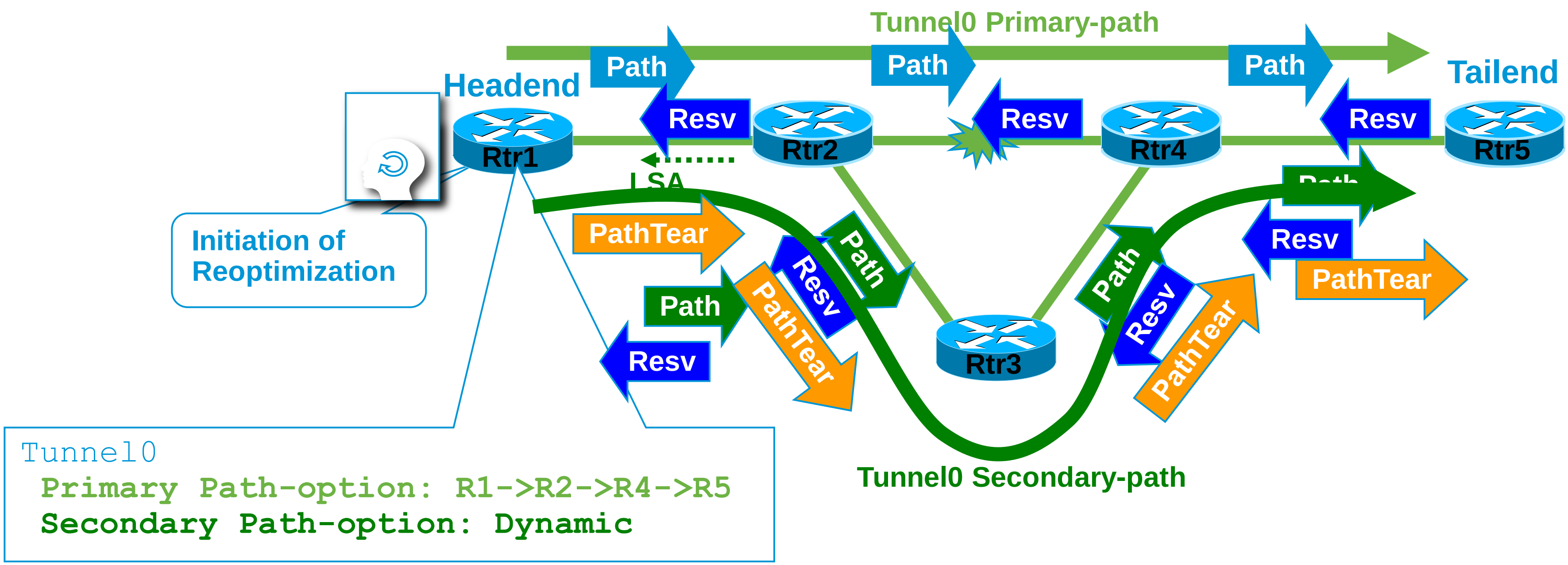
Explicit path-option

- Configured explicit path is checked before CSPF is done for path to tailend with constraints; bandwidth, affinity, and TE metric
- Specifying a sequence of RID of routers down to tailend is enough unless there are multiple links enabling TE between neighbors

```
interface Tunnel15
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 5.5.5.5
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng path-option 10 explicit name PATH-ACDE
!
ip explicit-path name PATH-ACDE enable
 next-address 3.3.3.3
 next-address 4.4.4.4
 next-address 5.5.5.5
```



Reoptimization (Headend reroute)



- Headend looks if there is a better path for the tunnel that are already up
- Headend behavior to switch tunnel LSP is called reoptimization or headend reroute
- Reoptimization is done with "**make-before-break**" manner

Path Reoptimization

- A headend can reoptimize automatically or manually existing TE LSPs in search of a better path
- New LSP is setup with better path in advance of tearing the existing path down … called **Make-before-break**
- Periodic reoptimization
  - 3600 sec by default, setting 0 is disabling the periodic reoptimization
  - `(config)#mpls traffic-eng reoptimize timers frequency {sec}` (IOS)
  - `(config-mpls-te)#reoptimize {sec}` (IOS XR)
- Manual reoptimization
  - Forcing reoptimization immediately by issuing exec command
  - `#mpls traffic-eng reoptimize [tunnel i/f]`
- Event-driven reoptimization
  - Forcing reoptimization immediately when physical interface comes up
    - Disabled by default
    - `(config)#mpls traffic-eng reoptimize events link-up` (IOS)
  - Fast reroute activation
    - When the headend receives PathErr with Error code 25, it immediately try to initiate reoptimzation

Verifying reoptimization timer

IOS Verification of reoptimization timer

```
MF7606G_P2#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 2809 seconds
    Periodic FRR Promotion:         Not Running
    Periodic auto-bw collection:    every 300 seconds, next in 109 seconds
TUNNEL NAME                      DESTINATION      UP IF     DOWN IF   STATE/PROT
Backup tunnel for link prote...  3.3.3.3          Te2/2     Te2/1     up/up
HalfQb_P3_t321                   1.1.1.1          Te2/1     Te2/2     up/up
HalfQb_P3_t324                   4.4.4.4          Te2/1     Gi1/1     up/up
Backup tunnel for node prote...  3.3.3.3          Gi1/1     Te2/1     up/up
Displayed 0 (of 0) heads, 4 (of 4) midpoints, 0 (of 0) tails
```

IOS XR Verification of reoptimization timer

```
RP/0/1/CPU0:MF12404C-XR_PE5#show mpls traffic-eng tunnels brief
Signalling Summary:
              LSP Tunnels Process:  running
                     RSVP Process:  running
                       Forwarding:  enabled
          Periodic reoptimization:  every 3600 seconds, next in 3012 seconds
           Periodic FRR Promotion:  every 300 seconds, next in 37 seconds
      Periodic auto-bw collection:  disabled
                     TUNNEL NAME         DESTINATION      STATUS  STATE
                  tunnel-te53146             6.6.6.6          up  up
Displayed 1 (of 1) heads, 0 (of 0) midpoints, 0 (of 0) tails
Displayed 1 up, 0 down, 0 recovering, 0 recovered heads
```

Path Setup

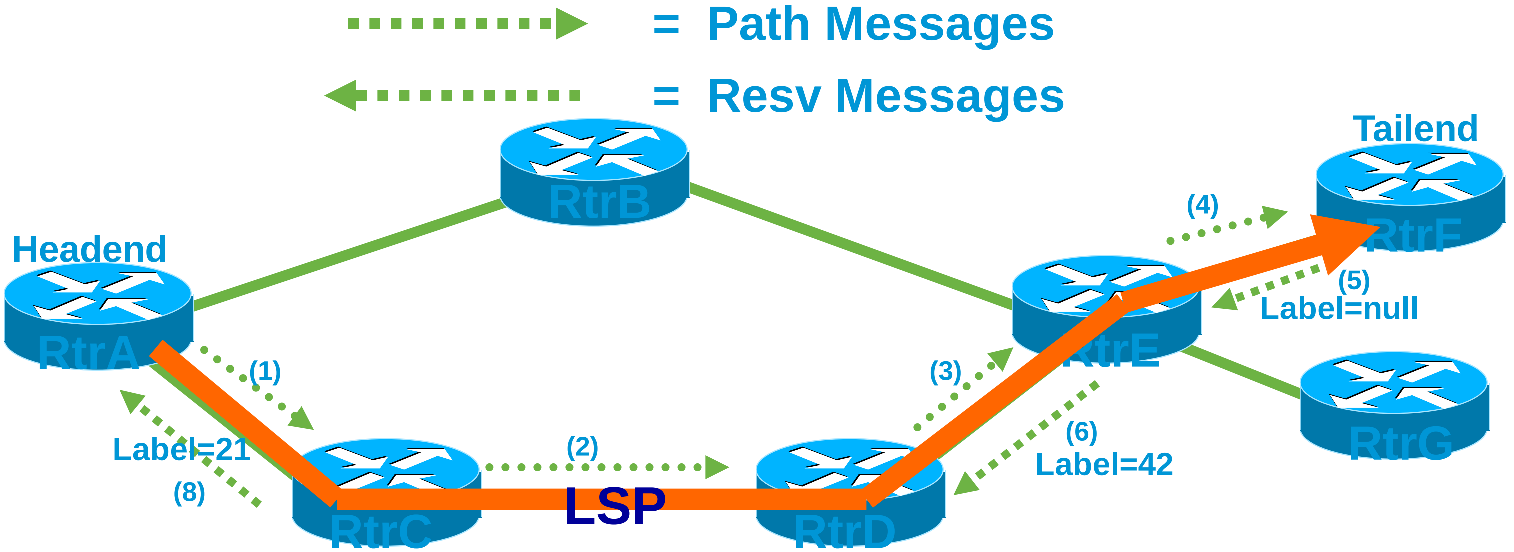
Path Setup

- Once the path is calculated with CSPF, that path needs to be signaled across the network:
  - To establish a hop-by-hop chain of labels that represent the path
  - To reserve any reservable resources (bandwidth) across the path
- RSVP is used for signaling protocol for MPLS TE in Cisco implementation
  - Original RSVP is specified in RFC 2205
  - RSVP extensions for MPLS TE is specified in RFC 3209
- RSVP uses Path and Resv messages to request an LSP along the calculated path

RSVP Basics

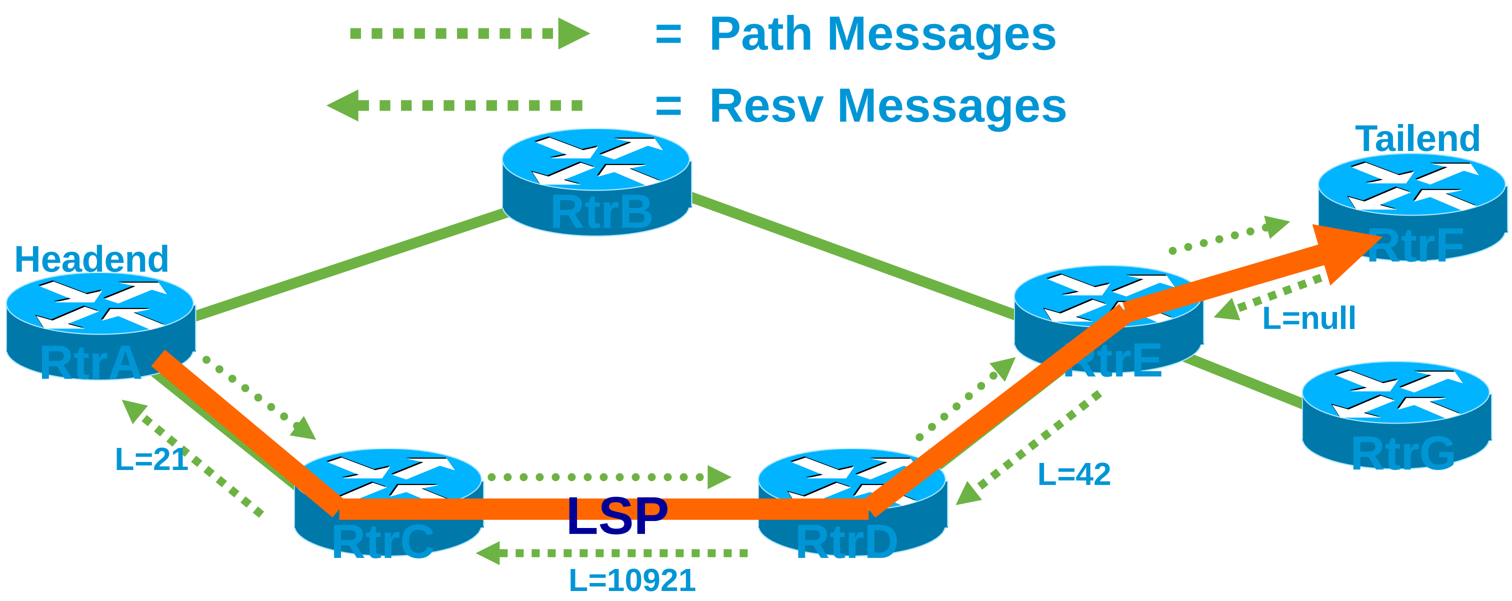
- RSVP originally specified as signaling protocol to maintain resource reservations across a network
- RSVP resource reservation is just at control plane, then it does not mean reservation at actual forwarding plane QoS
- Soft-state protocol
  - Periodical refresh is required by resignaling
  - Path setup is up till receiving the explicit signaling or reservation timeout

Path Setup



- Path message is sent from upstream to downstream router respectively
- Each routers which receive Path verify the message and requested resources (Call Admission Control)
- If CAC is successful, Resv message is sent from downstream to upstream router respectively
- Resv message contains incoming label, which the upstream router used to forward the traffic down the LSP

Path Maintenance



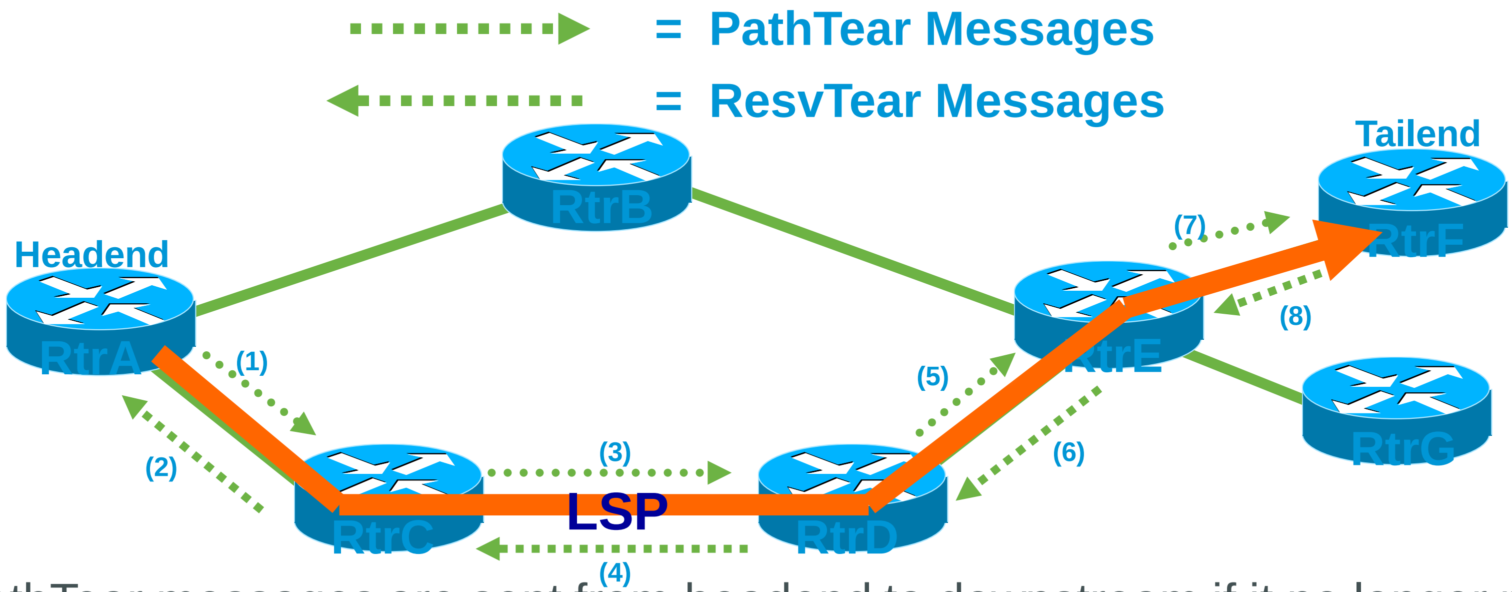
- Path and Resv messages are sent periodically to maintain the path
- The refresh timer (message interval) is 30 sec + 50 % variance (15 ~ 45 sec)

  Holdtime (L) is calculated

  L >= (K + 0.5) * 1.5 * R = 157.5 (R=30, K=3 in IOS implementation)

- Path and Resv messages are sent independently between neighbors

Path Teardown



- PathTear messages are sent from headend to downstream if it no longer need the existing path
- ResvTear messages are sent back in response to PathTear
- PathTear messages are also sent from midpoint to its downstream when it detect the error toward upstream or itself
- ResvTear messages are also sent from midpoint or tailend to upstream when detect error toward downstream or itself

Forwarding Traffic Down Tunnels

Methods for forwarding traffic down TE LSP

- For using TE LSP to forward the traffic, there are four options：
  - Autoroute
  - Forwarding adjacency
  - Static routes
  - Policy routing
- Static routes and Policy routing are straightforward since these can be configured with specifying tunnel interfaces as like physical interfaces, but these are rarely used since lacks of dynamic optimal routing and scaling due to the static nature

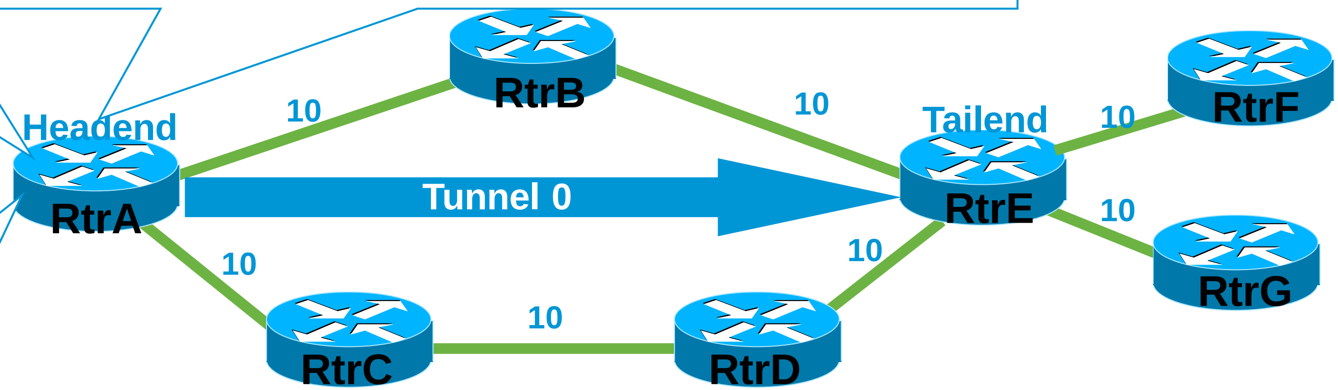
Autoroute - Concept

- IGP cannot run on TE tunnel
  - TE tunnel is unidirectional, thus the interface cannot receive the packet
- IGP does not need to run on TE tunnel
  - Topology data is not need to be flooded over TE tunnel since it is already flooded over physical topology (Intra-area TE)
- Autoroute concept “Treat this interface like the tunnel is a directly connected link to the tunnel tail, send any packets down the tunnel that are destined for either the tunnel’s tail or anything behind that tunnel tail”
- When SPF running, if IGP encounters a node that is either TE tunnel tail or located behind the TE tunnel tail, it installs TE tunnel to that node rather than to the IGP path in RIB

Autoroute

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	B	30

```
interface Tunnel0
 tunnel mpls traffic-eng autoroute announce
```



autoroute is applied

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	**Tunnel0**	**20**
F	**Tunnel0**	**30**
G	**Tunnel0**	**30**

- After autoroute is configured, the tunnel tail is always routed only through the tunnel
- Nodes behind the tunnel tail can generally be reached through the tunnel

TE tunnel metric with autoroute

- The metric assigned to a tunnel interface is shortest path cost the IGP calculated to tunnel tail by default
- Any routes behind TE tunnel have the cost of tunnel plus cost from tunnel tail to the target routes
- There are knobs to change the metric of tunnel which is autoroute announced
  - Fixed metric
  - Absolute metric (IS-IS only)
  - Relative metric

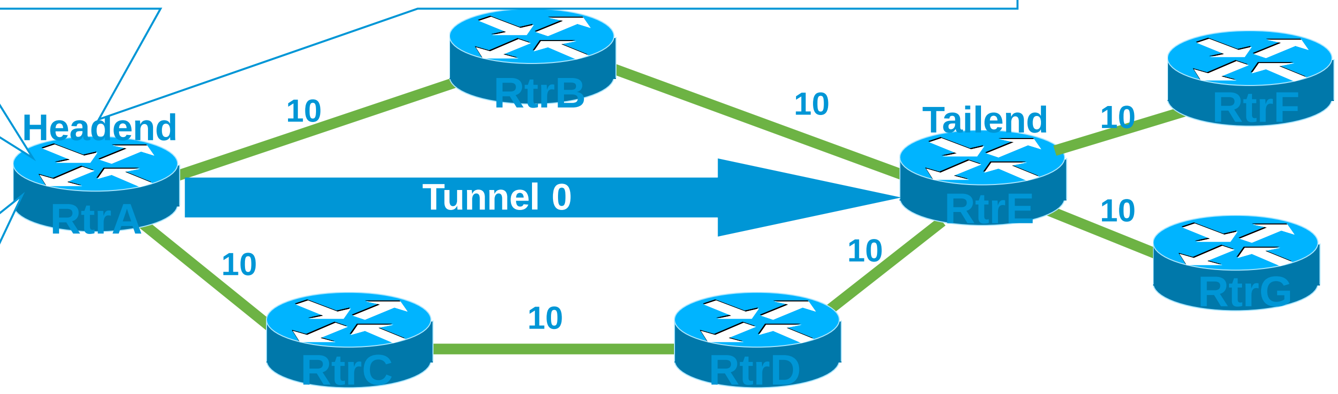
```
R1(config-if)#tunnel mpls traffic-eng autoroute metric ?
  <1-4294967295>  Set tunnel metric for autoroutes
  absolute        Set metric for all autoroutes over tunnel
  relative        Adjust tunnel metric for autoroutes relative to IGP
```

```
RP/0/1/CPU0:MF12404C-XR_PE5(config-if)#autoroute metric ?
  absolute  Set metric mode absolute
  relative  Set metric mode relative
```

Autoroute metric (Fixed metric)

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	Tunnel0	20
F	Tunnel0	30
G	Tunnel0	30

```
interface Tunnel0
 tunnel mpls traffic-eng autoroute metric 15
```



autoroute metric is applied

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	**Tunnel0**	**15**
F	**Tunnel0**	**25**
G	**Tunnel0**	**25**

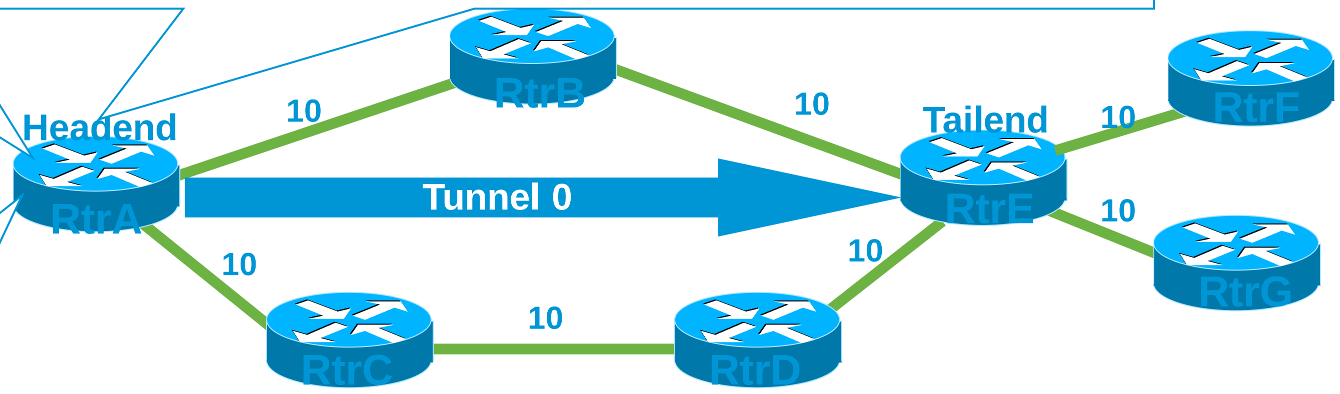
- The command ‘`tunnel mpls traffic-eng autoroute metric {`*`value`*`}`’ (IOS) overrides IGP shortest path cost
- It is usually configured when using TE tunnel if IGP shortest path is unavailable
  - If TE tunnel metric is configured as 25, IGP shortest path RtrA->RtrB-RtrC wins therefore tunnel is not used for forwarding unless this path become unavailable

EDCS-1171743

Autoroute metric (Absolute metric)

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	Tunnel0	20
F	Tunnel0	30
G	Tunnel0	30

```
interface Tunnel0
 tunnel mpls traffic-eng autoroute absolute 17
```



autoroute metric is applied

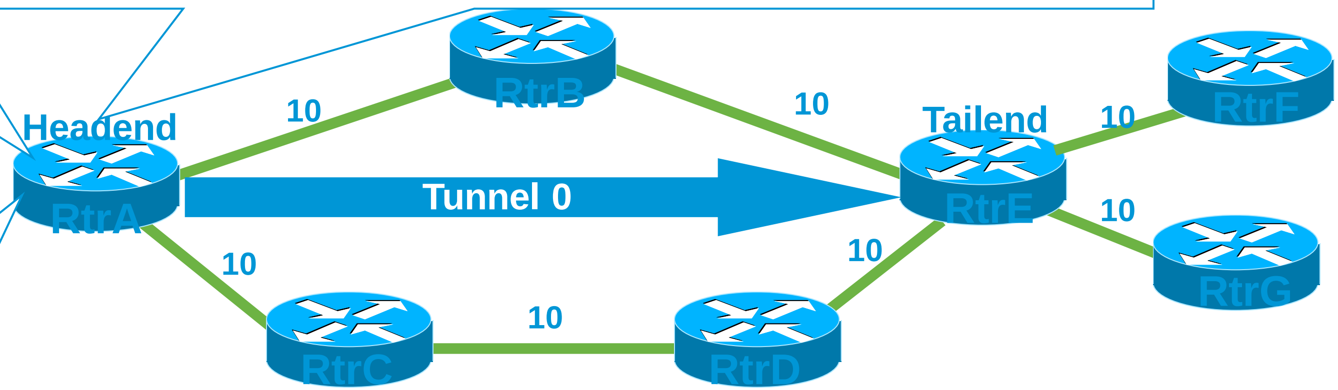
Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	**Tunnel0**	**17**
F	**Tunnel0**	**17**
G	**Tunnel0**	**17**

- The following commands ‘overrides IGP shortest path cost to tunnel tail and all routes behind it
  - `tunnel mpls traffic-eng autoroute metric absolute {value}`’ (IOS)
  - `autoroute metric absolute {value}`’ (IOS XR)
- *It is not supported with OSPF*

Autoroute metric (Relative metric)

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	Tunnel0	20
F	Tunnel0	30
G	Tunnel0	30

```
interface Tunnel0
 tunnel mpls traffic-eng autoroute relative -6
```



autoroute metric is applied

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E	**Tunnel0**	**14**
F	**Tunnel0**	**24**
G	**Tunnel0**	**24**

- The following commands can be used for adjusting tunnel metric relative to the IGP shortest path cost

  `tunnel mpls traffic-eng autoroute metric relative {`*`-10 - 10`*`}` (IOS)

  `autoroute metric relative {`*`-10 - 10`*`}` (IOS XR)

- The tunnel metric dynamically changed along with the IGP topology and LSP changes

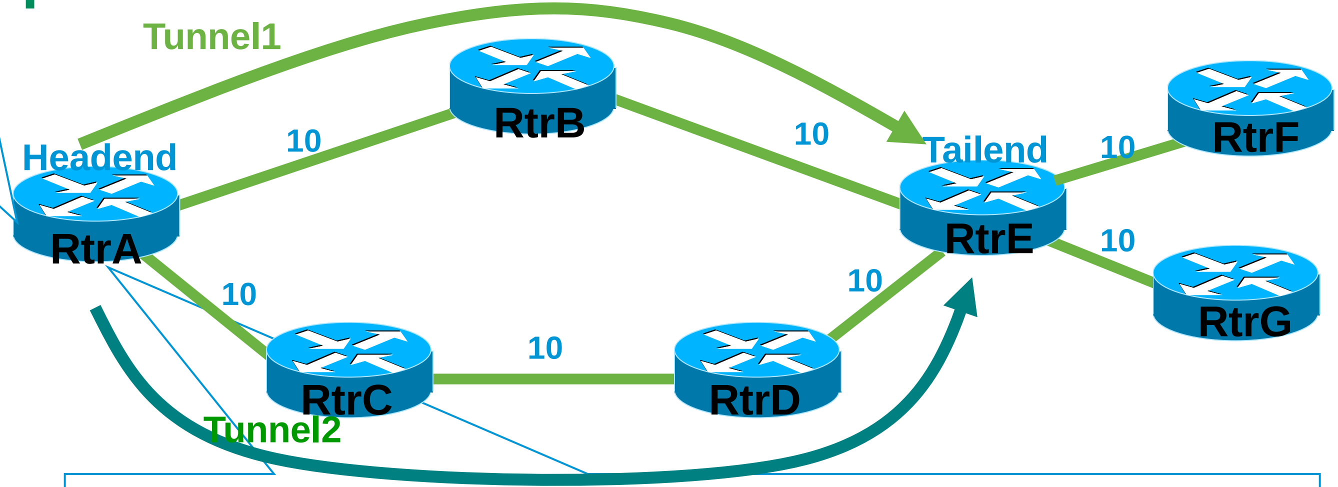
- *If positive value is configured, IGP shortest path always wins and tunnel is never used for traffic forwarding (No make sense)*

Class-Based Tunnel Selection (CBTS)

- Discriminating at the headend router as to which MPLS EXP values a node forwards through specific tunnels
  - Ingress non-labeled packets have to have been marked on IP TOS since selection is done before inbound policy
- The tunnel headend must have a group of tunnels with the same tailend to reach a particular destination
- This feature can be used for static routes or autoroute
  - The mix of static routes and autoroute within a group of tunnels will not work since static routes are always preferred over dynamic route
- One tunnel of a group can be configured as a default, or all of tunnels can be explicitly mapped with all possible MPLS EXP values
- The feature is available since 12.0(30)S depending on LC type, 12.2(33)SRA
- Policy-Based Tunnel Selection (PBTS) is the equivalent feature which will be supported on IOS XR with some differences

CBTS example

Node	Nhop	Cost
A	Self	0
B	B	10
C	C	10
D	C	20
E (EXP 5)	**Tunnel1**	**18**
E (default)	**Tunnel2**	**18**
F (EXP 5)	**Tunnel1**	**18**
F (default)	**Tunnel2**	**18**
G (EXP 5)	**Tunnel1**	**18**
G (default)	**Tunnel2**	**18**



Distribution setting can be checked with the command `'show mpls forwarding-table {*destination*} detail'`

```
interface Tunnel1
 ip unnumbered Loopback0
 tunnel destination 3.3.3.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng autoroute relative -2
 tunnel mpls traffic-eng priority 5 5
 tunnel mpls traffic-eng band width 10000
 tunnel mpls traffic-eng path-option explicit name PathA-B-C
 tunnel mpls traffic-eng exp 5
!
interface Tunnel2
 ip unnumbered Loopback0
 tunnel destination 3.3.3.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng autoroute relative -2
 tunnel mpls traffic-eng priority 7 7
 tunnel mpls traffic-eng band width 90000
 tunnel mpls traffic-eng path-option explicit name PathA-C-D-E
 tunnel mpls traffic-eng exp default
```

Basic Configuration Review

Step1 - Enabling MPLS TE on a Node

MPLS TE Node Configuration in Cisco IOS (*All LSRs*)

```
mpls traffic-eng tunnels
mpls traffic-eng logging lsp path-errors
mpls traffic-eng logging lsp setups
mpls traffic-eng logging lsp teardowns
mpls traffic-eng logging tunnel lsp-selection
mpls traffic-eng logging tunnel path change
```

MPLS TE Node Configuration in Cisco IOS XR (*All LSRs*)

```
mpls traffic-eng
```

Step2 - Enabling MPLS TE on an Interface

MPLS TE Interface Configuration in Cisco IOS (*All LSRs*)

```
interface GigabitEthernet1/0
 description Connection to P1 gi1/1
 ip address 192.168.14.4 255.255.255.0
 no ip directed-broadcast
 ip ospf network point-to-point
 mpls traffic-eng tunnels
 ip rsvp bandwidth
!
interface GigabitEthernet1/2
 description Connection to PE6 gi2/0
 ip address 192.168.46.4 255.255.255.0
 no ip directed-broadcast
 ip ospf network point-to-point
 mpls traffic-eng tunnels
 ip rsvp bandwidth
```

MPLS TE & RSVP Interface Configuration in Cisco IOS XR (*All LSRs*)

```
rsvp
 interface TenGigE0/1/0/0
 !
 interface TenGigE0/1/0/4
 !
!
mpls traffic-eng
 interface TenGigE0/1/0/0
 !
 interface TenGigE0/1/0/4
```

Step3 - Defining a TE Tunnel Interface

Definition of a TE LSP Headend in Cisco IOS (*Headend*)

```
interface Tunnel64135
 description Primary tunnel (Rtr6-4-1-3-5)
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 5.5.5.5
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng path-option 10 explicit name Path6-4-1-3-5
```

Definition of a TE LSP Headend in Cisco IOS XR (*Headend*)

```
interface tunnel-te53146
 description Primary tunnel (Rtr5-3-1-4-6)
 ipv4 unnumbered Loopback0
 destination 6.6.6.6
 path-option 10 explicit name Path5-3-1-4-6
```

 EDCS-1171743

Step4 - Defining an Explicit Path

Definition of Explicit path in Cisco IOS (*Headend*)

```
ip explicit-path name Path6-4-1-3-5 enable
 next-address 4.4.4.4
 next-address 1.1.1.1
 next-address 3.3.3.3
 next-address 5.5.5.5
```

Definition of Explicit path in Cisco IOS XR (*Headend*)

```
explicit-path name Path5-3-1-4-6
 index 1 next-address strict ipv4 unicast 3.3.3.3
 index 2 next-address strict ipv4 unicast 1.1.1.1
 index 3 next-address strict ipv4 unicast 4.4.4.4
 index 4 next-address strict ipv4 unicast 6.6.6.6
```

Step5 Link Information Distribution (for OSPF)

Configuration of TE Extension for OSPF in Cisco IOS (*All LSRs*)

```
router ospf 100
 router-id 4.4.4.4
 log-adjacency-changes
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng area 0
 passive-interface Loopback0
 network 0.0.0.0 255.255.255.255 area 0
```

Configuration of TE Extension for OSPF in Cisco IOS XR (*All LSRs*)

```
router ospf 100
 router-id 3.3.3.3
 area 0
  network point-to-point
  mpls traffic-eng
  interface Loopback0
   passive enable
  !
  interface TenGigE0/1/0/0
  !
  interface TenGigE0/1/0/4
  !
 !
 mpls traffic-eng router-id Loopback0
```

Step6 - Configuring Link Attributes (*Optional*)

Definition of Link Characteristics in Cisco IOS (*All LSRs*)

```
interface GigabitEthernet1/0
 description Connection to P1 gi1/1
 ip address 192.168.14.4 255.255.255.0
 no ip directed-broadcast
 ip ospf network point-to-point
 mpls traffic-eng tunnels
 mpls traffic-eng attribute-flags 0x2
 mpls traffic-eng administrative-weight 10
 ip rsvp bandwidth 500000
```

Definition of Link Characteristics in Cisco IOS XR (*All LSRs*)

```
rsvp
 interface TenGigE0/1/0/0
  bandwidth 5000000
 !
!
mpls traffic-eng
 interface TenGigE0/1/0/0
  admin-weight 5
  attribute-flags 0x1
```

Step7 - Configuring the TE LSP Constraints (*Optional*)

Definition of TE LSP Constraints in Cisco IOS (*Headend*)

```
interface Tunnel64135
 description Primary tunnel (Rtr6-4-1-3-5)
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 5.5.5.5
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng priority 5 5
 tunnel mpls traffic-eng bandwidth 400000
 tunnel mpls traffic-eng path-option 10 explicit name Path6-4-1-3-5
 tunnel mpls traffic-eng affinity 0x3 mask 0x2
```

Definition of TE LSP Constraints in Cisco IOS XR (*Headend*)

```
interface tunnel-te53146
 description Primary tunnel (Rtr5-3-1-4-6)
 ipv4 unnumbered Loopback0
 priority 5 5
 signalled-bandwidth 100000
 destination 6.6.6.6
 path-option 10 explicit name Path5-3-1-4-6
 affinity 3 mask 1
```

Step8 - Configuring RSVP (*Optional*)

Optional RSVP Configuration in Cisco IOS (*All LSRs*)

```
ip rsvp signalling refresh reduction
ip rsvp signalling hello
!
interface GigabitEthernet1/0
 description Connection to P1 gi1/1
 ip address 192.168.14.4 255.255.255.0
 no ip directed-broadcast
 ip ospf network point-to-point
 ip rsvp bandwidth 500000
 ip rsvp signalling hello
 ip rsvp signalling hello refresh interval 1000
```

Optional RSVP Configuration in Cisco IOS XR (*All LSRs*)

```
rsvp
 interface TenGigE0/1/0/0
  bandwidth 5000000
 !
 signalling greceful-restart
 signalling hello graceful-restart refresh interval 1000
```

Step10 Aligning Null label between IOS and IOS XR (Optional)

Configuration of implicit-null label signalling in Cisco IOS (*All LSRs*)

```
mpls traffic-eng signalling advertise implicit-null
```

Step11 - Installing TE LSP to forwarding plane (Case1 autoroute)

Configuration of autoroute in Cisco IOS (*Headend*)

```
interface Tunnel64135
 description Primary tunnel (Rtr6-4-1-3-5)
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 5.5.5.5
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng autoroute announce
 tunnel mpls traffic-eng autoroute metric relative -2
 tunnel mpls traffic-eng path-option 10 explicit name Path6-4-1-3-5
```

Configuration of autoroute in Cisco IOS XR (*Headend*)

```
interface tunnel-te53146
 description Primary tunnel (Rtr5-3-1-4-6)
 ipv4 unnumbered Loopback0
 autoroute announce
 autoroute metric relative -2
 destination 6.6.6.6
 path-option 10 explicit name Path5-3-1-4-6
```

Step11 - Installing TE LSP to forwarding plane (Case2 Forwarding adjacency)

Configuration of forwarding adjacency in Cisco IOS (*Headend*)

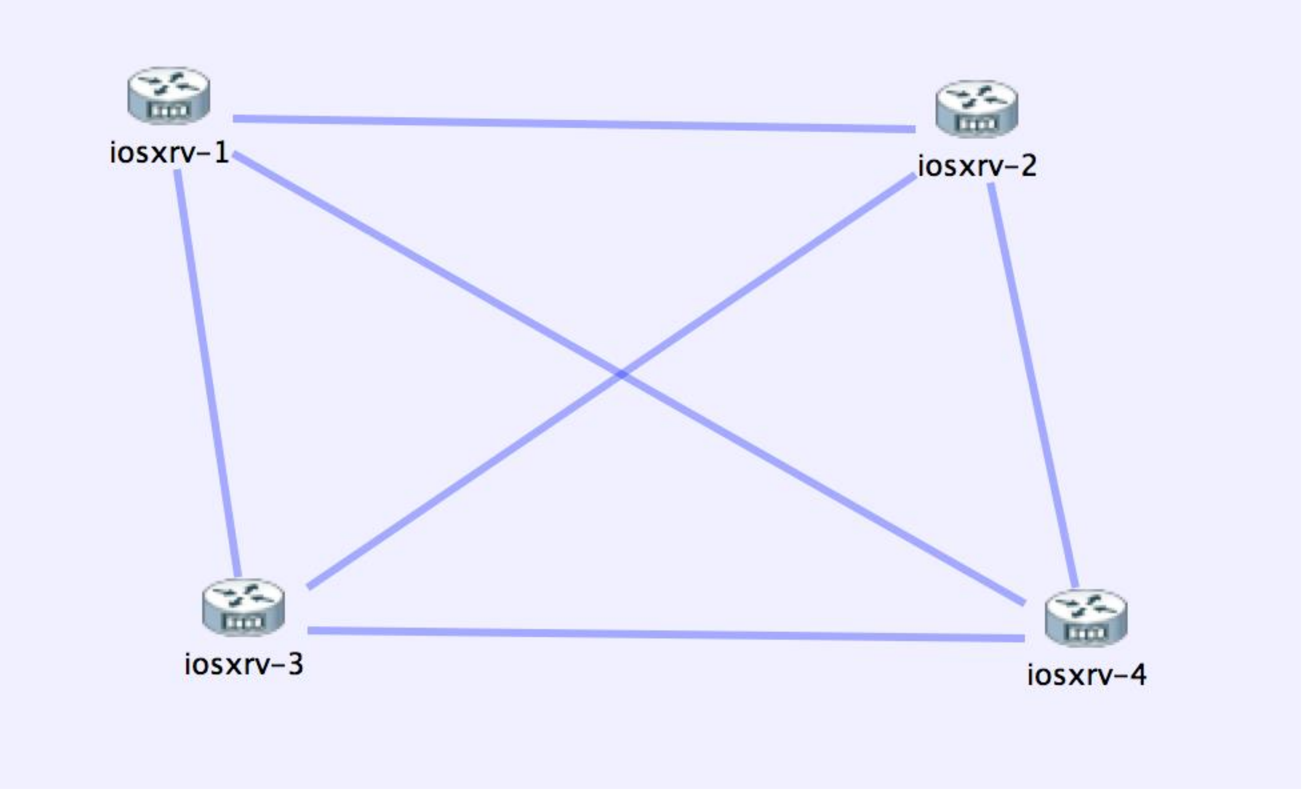
```
interface Tunnel64135
 description Primary tunnel (Rtr6-4-1-3-5)
 ip unnumbered Loopback0
 no ip directed-broadcast
 ip ospf cost 5
 tunnel destination 5.5.5.5
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng forwarding-adjacency
 tunnel mpls traffic-eng path-option 10 explicit name Path6-4-1-3-5
```

Configuration of forwarding adjacency in Cisco IOS XR (*Headend*)

```
interface tunnel-te53146
 description Primary tunnel (Rtr5-3-1-4-6)
 ipv4 unnumbered Loopback0
 signalled-bandwidth 100000
 destination 6.6.6.6
 forwarding-adjacency
 path-option 10 explicit name Path5-3-1-4-6
!
router ospf 100
 router-id 5.5.5.5
 area 0
  network point-to-point
  mpls traffic-eng
  interface Loopback0
   passive enable
  !
  interface tunnel-te53146
   cost 3
  !
  interface TenGigE0/2/0/0
```

Case Study

Topology and Questions



- Iosxrv-1的Tunnel-Te14 Down，在问题出现前，隧道正常工作超过1年。

Question Analysis

- IGP协议是否正常开启TE支持？
- 隧道目的地址是否可达？
- 隧道路径RSVP是否开启？
- 是否有足够剩余带宽建立隧道？
- 是否有其他影响隧道建立的参数改变？

EDCS-1171743

Question Analysis

- IGP正常启用TE支持，并且TED正常。

```
RP/0/0/CPU0:iosxrv-1#show run router isis
Mon Aug 18 14:04:19.859 UTC
router isis 1
 net 49.1921.6800.0001.00
 address-family ipv4 unicast
  metric-style wide
  mpls traffic-eng level-2-only
  mpls traffic-eng router-id Loopback0
 !
```

- 隧道目的地址可达。

```
RP/0/0/CPU0:iosxrv-1#ping 192.168.0.4
Mon Aug 18 14:05:19.025 UTC
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.0.4, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/9 ms
```

Question Analysis

- 经检查，所有设备RSVP正常开启。
- 各设备接口有足够剩余带宽。

```
RP/0/0/CPU0:iosxrv-1#show rsvp interface
Mon Aug 18 14:06:17.171 UTC

*: RDM: Default I/F B/W % : 75% [default] (max resv/bc0), 0% [default] (bc1)

Interface    MaxBW (bps)  MaxFlow (bps) Allocated (bps)      MaxSub (bps)
-----------  ------------ ------------- -------------------- -------------
Gi0/0/0/0             1G            1G          0 (   0%)              0
Gi0/0/0/1             1G            1G          0 (   0%)              0
Gi0/0/0/2             1G            1G          0 (   0%)              0
```

Question Analysis

- 检查发现TE异常由于没有到达目标的路径。

```
RP/0/0/CPU0:iosxrv-1#show mpls traffic-eng tunnels 14
Mon Aug 18 14:08:29.012 UTC


Name: tunnel-te14  Destination: 192.168.0.4
  Status:
    Admin:    up Oper: down   Path: not valid   Signalling: Down

    path option 10,  type dynamic
    Last PCALC Error: Mon Aug 18 14:08:21 2014
      Info: No path to destination, 192.168.0.4 (affinity)
    G-PID: 0x0800 (derived from egress interface properties)
    Bandwidth Requested: 20000 kbps  CT0
    Creation Time: Mon Aug 18 13:43:18 2014 (00:25:12 ago)
  Config Parameters:
    Bandwidth:    20000 kbps (CT0) Priority:  5  5 Affinity: 0x0/0xffff
    Metric Type: TE (default)
    Hop-limit: disabled
    AutoRoute: disabled  LockDown: disabled   Policy class: not set
    Forward class: 0 (default)
```

Question Analysis

- 检查TED数据库发现，链路配置有attribute(affinity)。

```
RP/0/0/CPU0:iosxrv-1#show mpls traffic-eng topology
Mon Aug 18 14:10:45.593 UTC
My_System_id: 1921.6800.0001.00 (IS-IS 1 level-2)
My_BC_Model_Type: RDM

Signalling error holddown: 10 sec Global Link Generation 584

IGP Id: 1921.6800.0001.00, MPLS TE Id: 192.168.0.1 Router Node  (IS-IS 1 level-2)

  Link[0]:Point-to-Point, Nbr IGP Id:1921.6800.0002.00, Nbr Node Id:2, gen:573
      Frag Id:0, Intf Address:10.0.0.1, Intf Id:0
      Nbr Intf Address:10.0.0.2, Nbr Intf Id:0
      TE Metric:1, IGP Metric:1
      Attribute Flags: 0x124
      Ext Admin Group:
          Length: 256 bits
          Value : 0x::124
      Attribute Names:
```

Analysis Summary

- 分析结果如下：
  - TE有关功能开启一切正常。
  - 检查数据库发现，多数链路配置有attribute flags。
- 由于Cisco设备TE隧道默认的affinity为0x0掩码为0xffff，所以如隧道所经的链路有着色情况，会导致因为affinity不匹配隧道Down的情况。
- 和客户进一步沟通发现，在隧道Down的时间点上，其他厂商设备有针对链路着色配置的修改。和数据库检查结果一致。

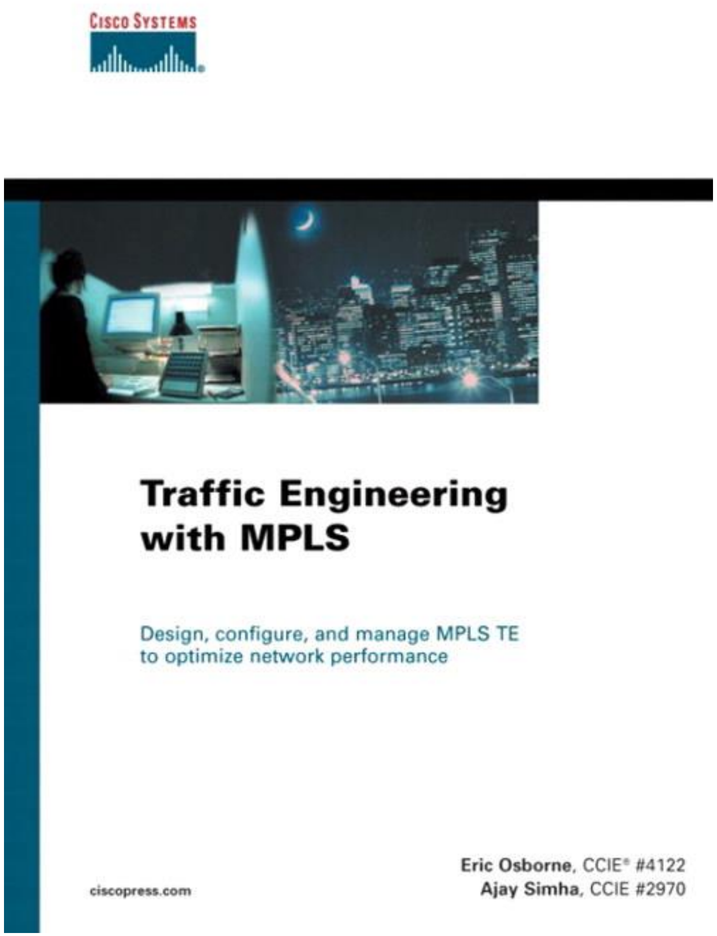
Action Plan

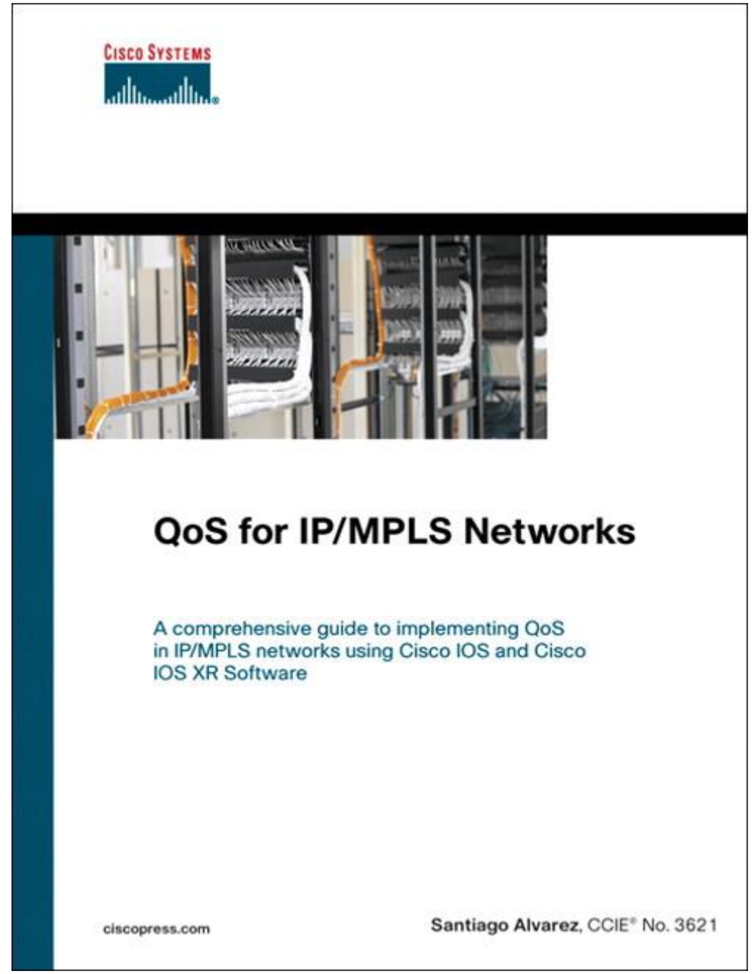
- 建议客户修改TE隧道下的affinity参数：
  - 隧道建立是忽略affinity的检查
  - 将隧道affinity与链路着色匹配。
- 最终，客户修改配置忽略affinity后隧道建立正常。

```
RP/0/0/CPU0:iosxrv-1#show mpls traffic-eng tunnels 14
Mon Aug 18 14:17:53.824 UTC


Name: tunnel-te14  Destination: 192.168.0.4
  Status:
    Admin:    up Oper:   up   Path:  valid   Signalling: connected

    path option 10,  type dynamic  (Basis for Setup, path weight 1)
    G-PID: 0x0800 (derived from egress interface properties)
    Bandwidth Requested: 20000 kbps  CT0
    Creation Time: Mon Aug 18 13:43:18 2014 (00:34:35 ago)
  Config Parameters:
    Bandwidth:    20000 kbps (CT0) Priority:  5  5
    Number of affinity constraints: 1
       Ignore all

    Metric Type: TE (default)
```


EDCS-1171743

Recommended Readings





- Traffic Engineering with MPLS
- QoS for IP/MPLS Networks

Thank you.