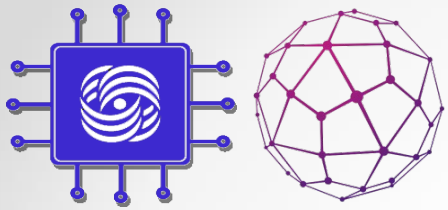
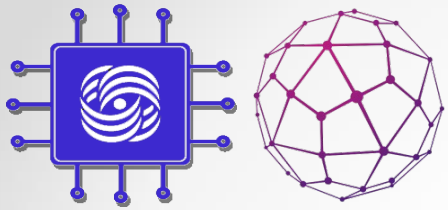


Применение MPLS в сетях связи. (Часть 3)

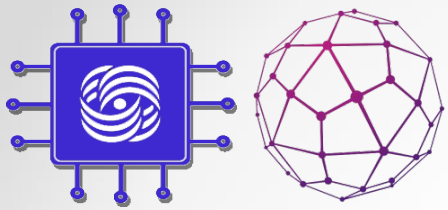
Дополнительные главы Компьютерных сетей и телекоммуникаций.

Васин В.В.
CCIE, ECE, CCSI



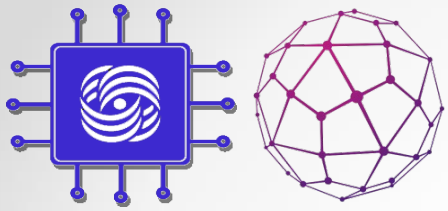
MPLS VPN Technology

Traffic Engineering Concepts



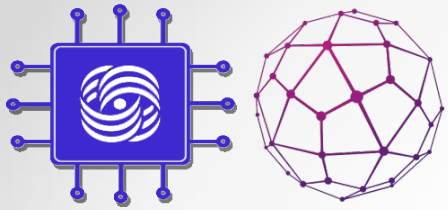
What Is Traffic Engineering?

- Term in common use in telephone voice network world
- Measures, models, and controls traffic to achieve various goals
- Provides an integrated approach to engineering traffic at Layer 3 (ISO/OSI)



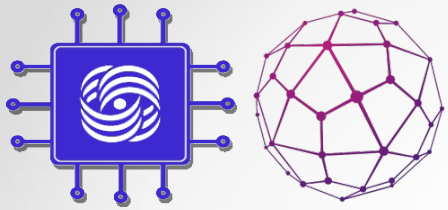
What Is Traffic Engineering? Traffic Engineering Motivations

- Reduce the overall cost of operations by more efficient use of bandwidth resources
- Prevent a situation where some parts of a service provider network are overutilized (congested), while other parts remain underutilized
- Implement traffic protection against failures
- Enhance SLA in combination with QoS



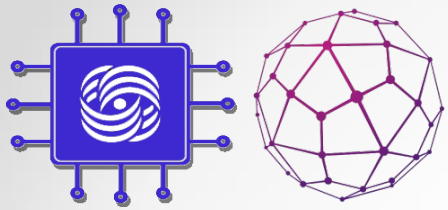
Business Drivers for Traffic Engineering

- Routers always forward traffic along the least-cost route as discovered by IGP.
- Network bandwidth may not be efficiently utilized:
  - The least-cost route may not be the only possible route.
  - The least-cost route may not have enough resources to carry all the traffic.



Business Drivers for Traffic Engineering (Cont.)

- Lack of resources results in congestion in two ways:
  - When network resources themselves are insufficient to accommodate offered load
  - When traffic streams are inefficiently mapped onto available resources
- Some resources are overutilized while others remain underutilized.



Congestion Avoidance and Traffic Engineering

- Network congestion can be addressed by either:
  - Expansion of capacity or classical congestion control techniques (queuing, rate limiting, etc.)
  - Traffic Engineering (TE), if the problems result from inefficient resource allocation

The focus of TE is on congestion problems that are prolonged, not on short-term bursts

Congestion Avoidance and Traffic Engineering

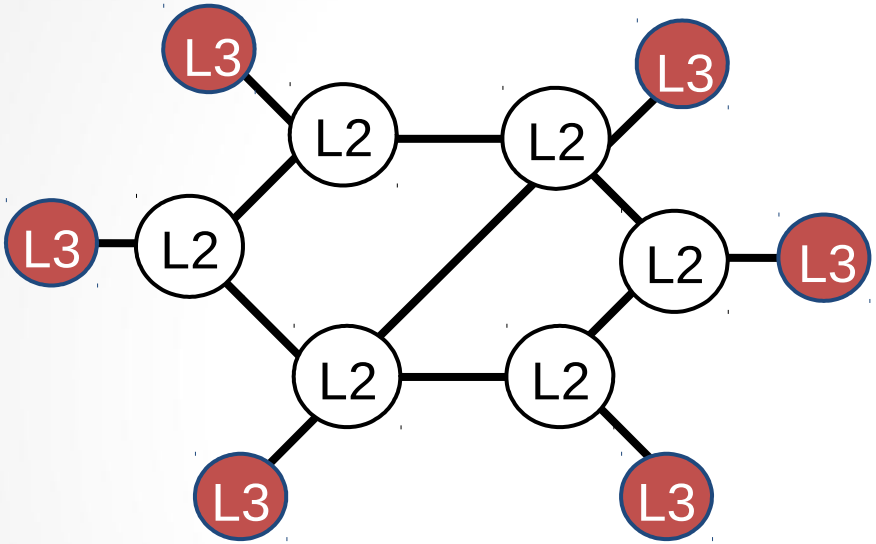
Without the use of TE, all traffic can be redirected to the route, where there is not enough bandwidth - drops will begin.

The convergence rate of OSPF or ISIS even when using BFD Bidirectional Forwarding Detection) is in the tens of «ms».

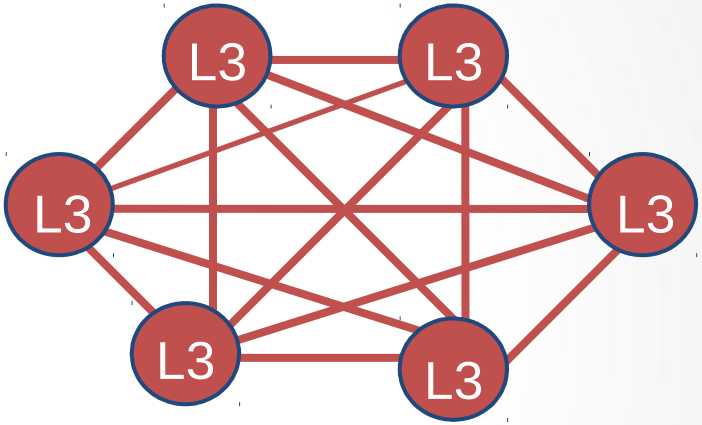
After that, the transport LSP must also be rebuilt.

It will not go unnoticed by subscribers.

Traffic Engineering with a Layer 2 Overlay Model



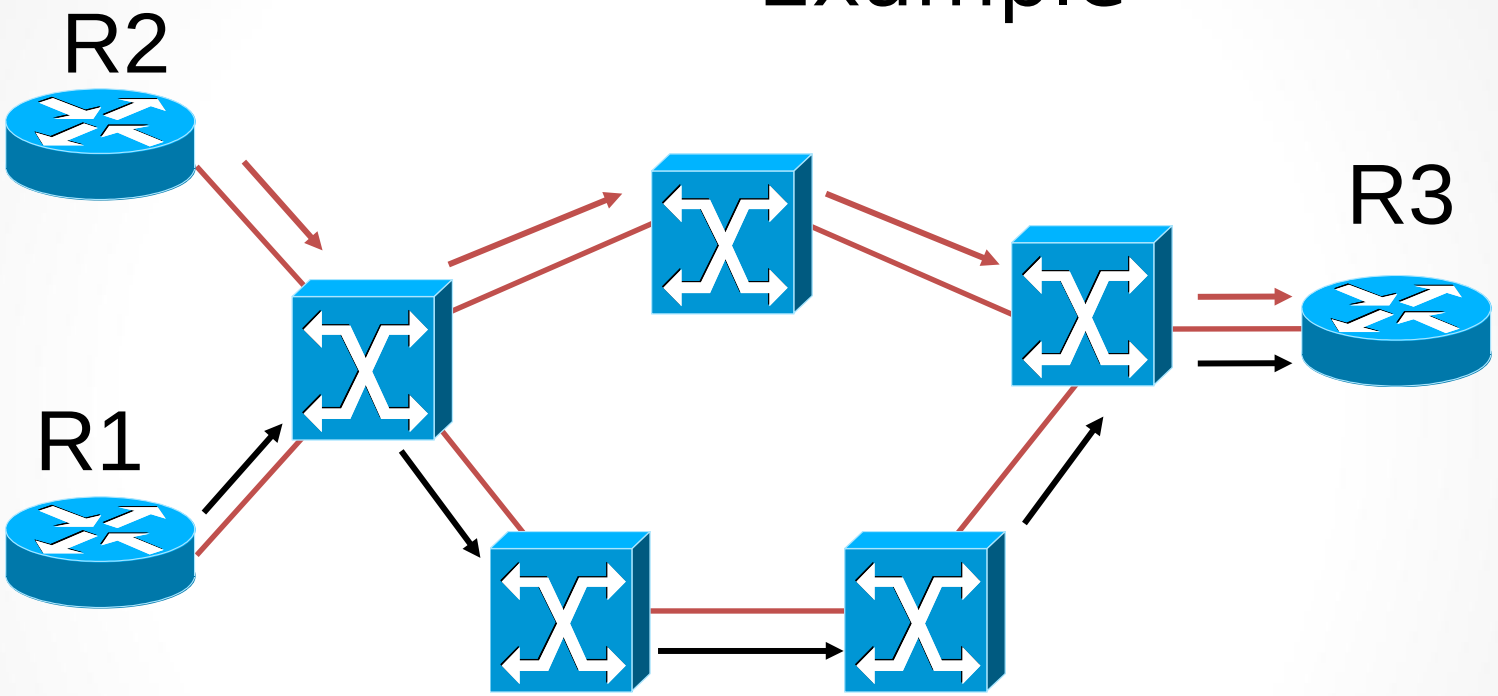
Physical

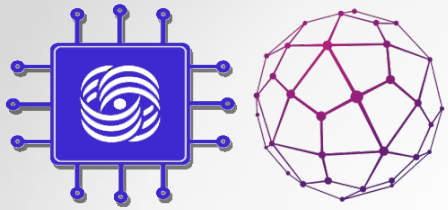


Logical

- The use of the explicit Layer 2 transit layer allows very exact control of how traffic uses the available bandwidth.
- Layer 3 at the edge sees a complete mesh.

Traffic Engineering with a Layer 2 Overlay Model: Example

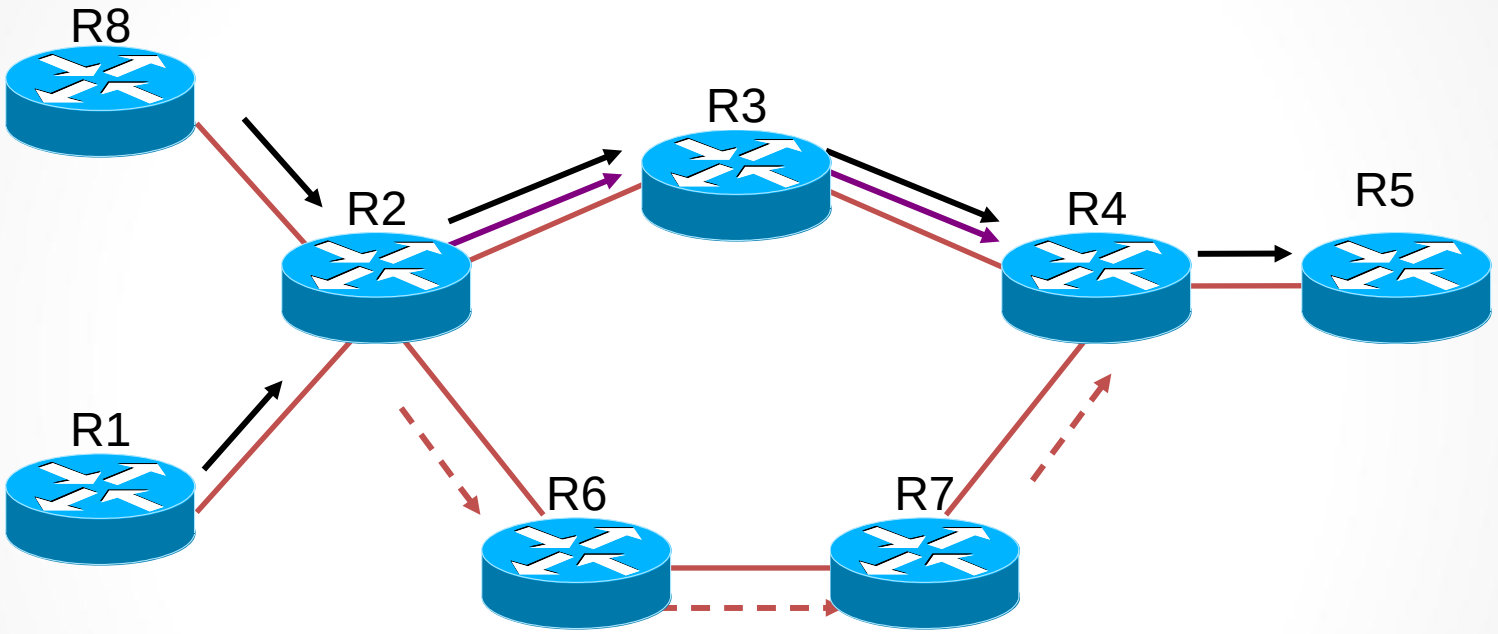




Traffic Engineering with a Layer 2 Overlay Model (Cont.)

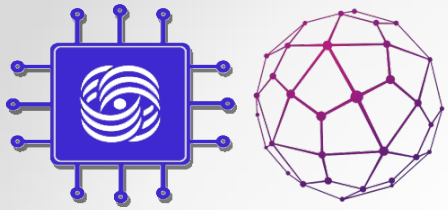
- Drawbacks of the Overlay Solution
  - Extra network devices
  - More complex network management:
    - Two-level network without integrated network management
    - Additional training, technical support, field engineering
  - IGP routing scalability issue for meshes
  - Additional bandwidth overhead (“cell tax”)
  - No differential service (class of service)

Traffic Engineering with a Layer 3 Model



IP (mostly) uses destination-based least-cost routing. Flows from R8 and R1 merge at R2. From R2, traffic to R3, R4, and R5 uses the upper route.

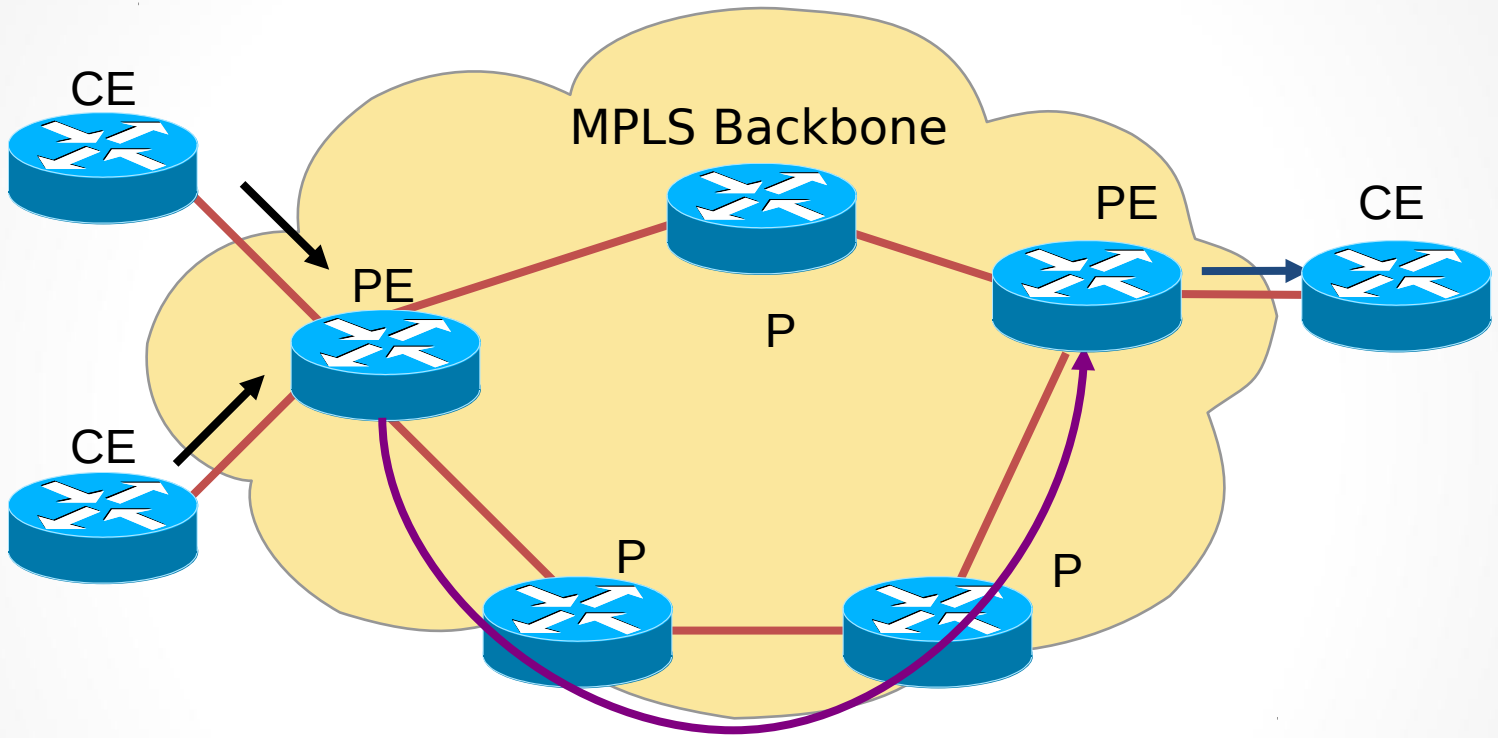
The dashed arrow denotes an underutilized alternative path.



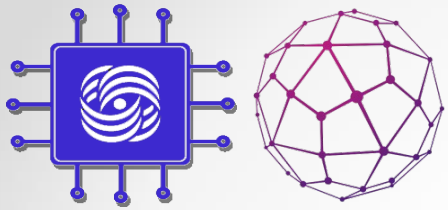
Traffic Engineering with a Layer 3 Model (Cont.)

- The current forwarding paradigm, centered around “destination-based,” is clearly inadequate:
  - Path computation based just on IGP metric is not enough.
  - Support for “explicit” routing (source routing) is not available.
  - Supported workarounds: static routes, policy routing.
  - Provide controlled backup and recovery.

Traffic Engineering with the MPLS-TE Model

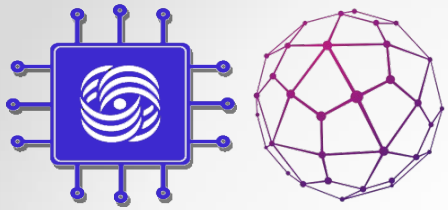


- Tunnel is assigned labels that represent the path (LSP) through the system.
- Forwarding within the MPLS network is based on labels (no Layer 3 lookup).



Traffic Engineering with the MPLS-TE Model (Cont.)

- The MPLS-TE LSPs are created by RSVP.
- The actual path can be specified:
  - Explicitly defined by the system administrator
  - Dynamically defined using the underlying IGP protocol



Traffic Engineering with the MPLS-TE Model

RSVP-Resource ReSerVation Protocol (RFC 2205) - 1993

The source node sends a special message in the RSVP Protocol format over the network before transmitting data that requires a certain non-standard quality of service (for example, constant bandwidth for video transmission). This message contains:

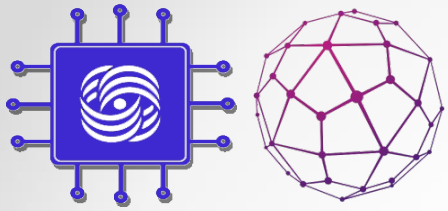
- type of information being transmitted
- bandwidth required.

It is transmitted between routers from the sending node to the destination address, and the sequence of routers in which you want to reserve a certain bandwidth is determined.

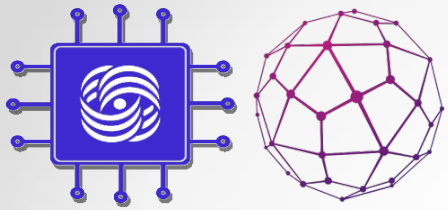
- When the router receives this message, it checks its resources.
- If the required bandwidth is achievable, the router configures the packet processing algorithm so that the specified bandwidth is always provided, and then sends the message to the next router along the path.
- In the absence, of the bandwidth the router rejects the request.

Traffic Engineering with the MPLS-TE Model

- The Path Packet reaches the recipient of the stream, who sends back a Resv message, confirming the allocation of resources throughout the path.
- The Original sender, having received Resv, understands that everything is ready for him, and he can send data.



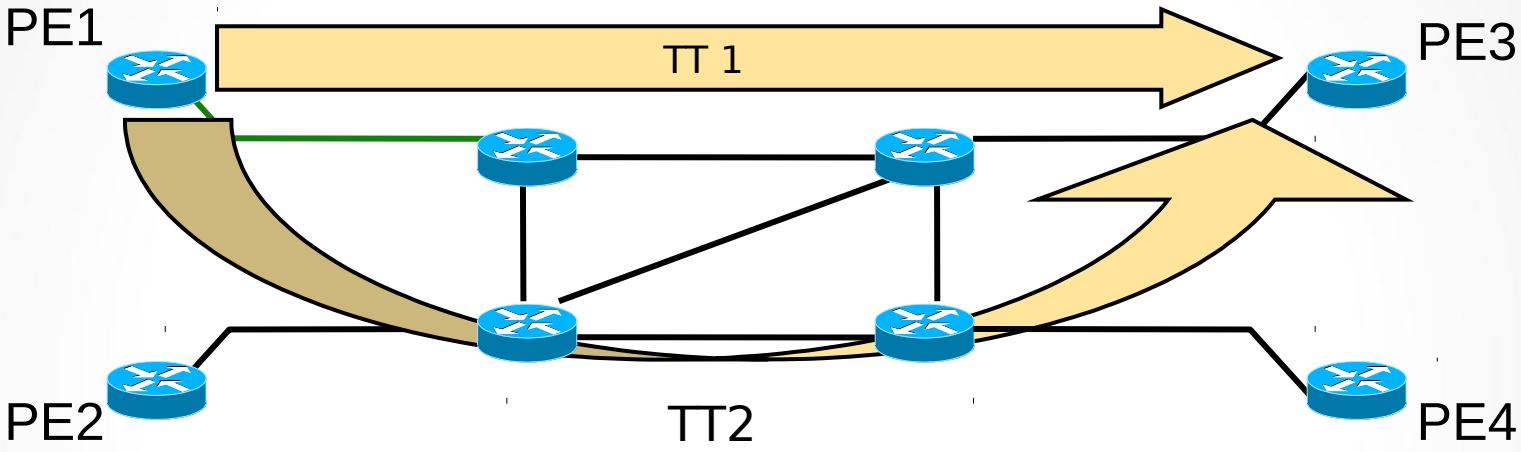
MPLS TE Components



Traffic Tunnels: Concepts

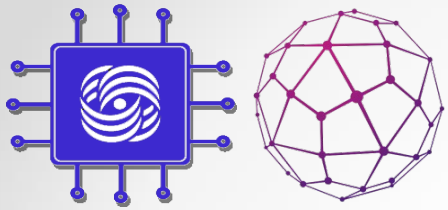
- The concept of traffic tunnels (MPLS-TE tunnels) was introduced to overcome the limitations of hop-by-hop IP routing:
  - A tunnel is an aggregation of traffic flows that are placed inside a common MPLS label switched path.
  - Flows are then forwarded along a common path within a service provider network.

Traffic Tunnels: Concepts (Cont.)



- Unidirectional single class of service model encapsulates all of the traffic between an ingress and an egress router.

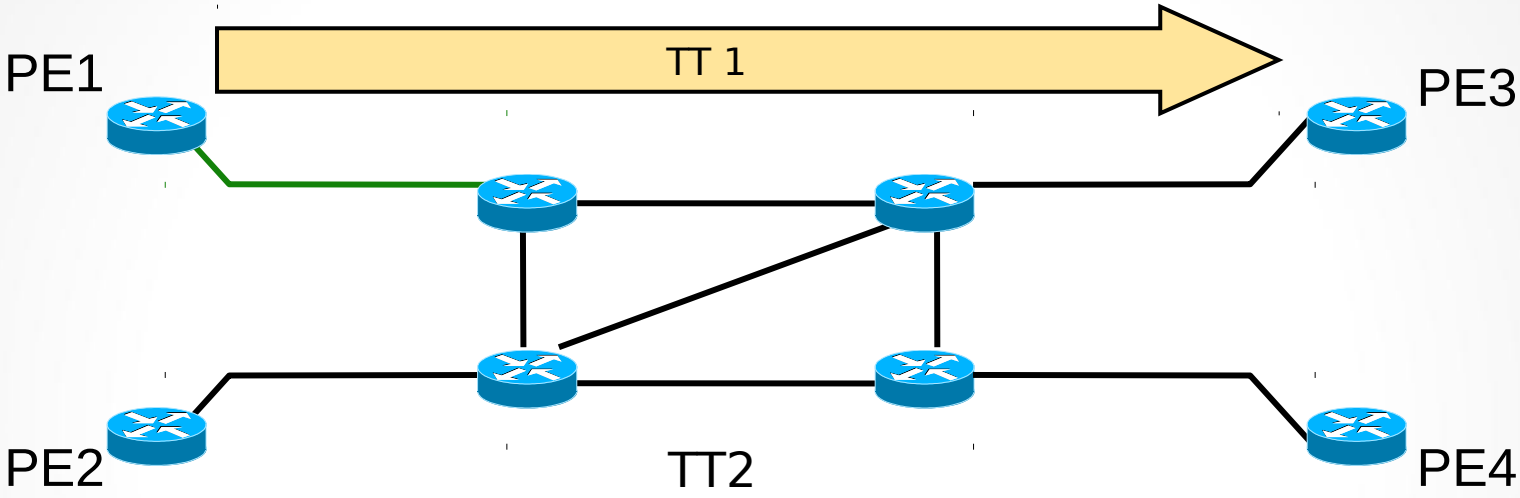
- **Different classes of service model assigns traffic into separate tunnels with different characteristics.**



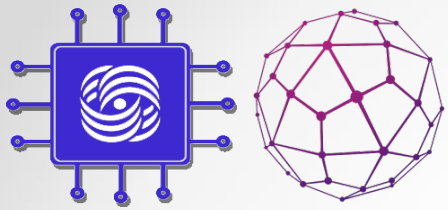
Traffic Tunnels – Characteristics

- Traffic tunnels are routable objects (similar to ATM VCs).
- A traffic tunnel is distinct from the MPLS LSP through which it traverses:
  - In operational contexts, a traffic tunnel can be moved from one path onto another
- A traffic tunnel is assigned attributes influencing its characteristics.

Traffic Tunnels – Attributes



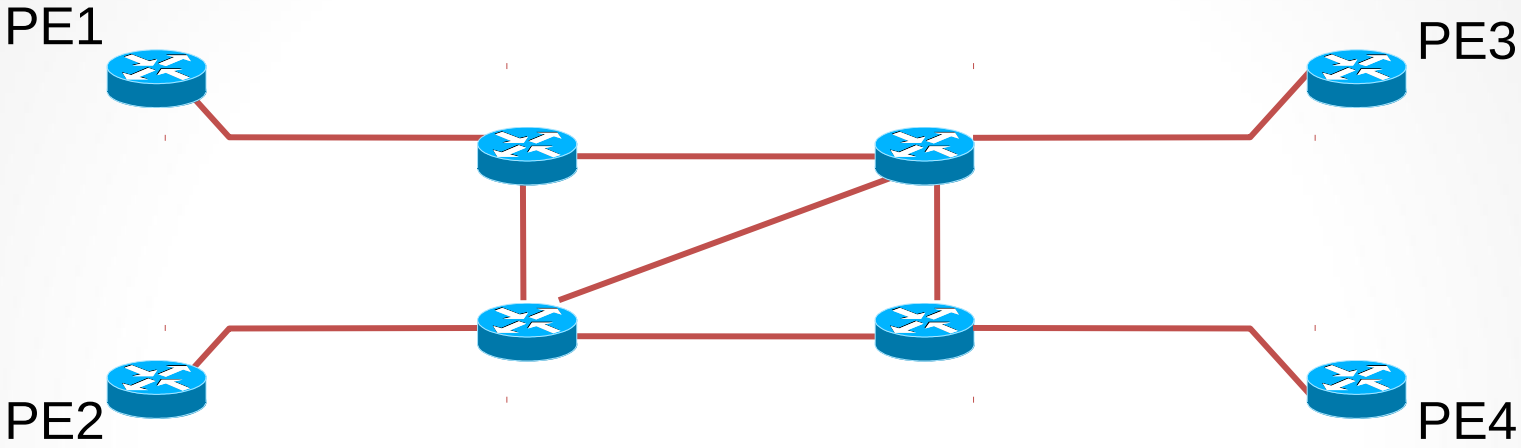
- Attributes are explicitly assigned to traffic tunnels through administrative action.
- A traffic tunnel is characterized by:
  - Its ingress and egress label switch routers
  - The forwarding equivalence class that is mapped onto it
  - A set of attributes that determine its characteristics



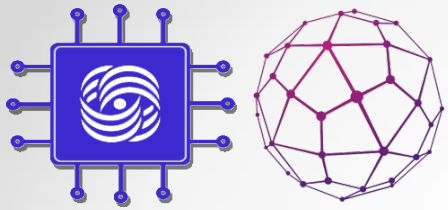
Traffic Tunnels – Attributes (Cont.)

- The administrator enters the relevant information (attributes) at the headend of the traffic tunnel:
  - Traffic parameter—resources required for tunnel (e.g., required bandwidth)
  - Generic path selection and management—path can be administratively specified or computed by the IGP
  - Resource class affinity—include or exclude certain links for certain traffic tunnels
  - Adaptability—should the traffic tunnel be reoptimized?
  - Priority and pre-emption—importance of a traffic tunnel and possibility for a pre-emption of another tunnel
  - Resilience—desired behavior under fault conditions

Network Links and Link Attributes



- Resource attributes (link availability) are configured locally on the router interfaces:
  - Maximum bandwidth
    - The amount of bandwidth available
  - Link affinity string
    - To allow the operator to administratively include or exclude links in path calculations
  - Constraint-based specific metric
    - Traffic engineering default metric



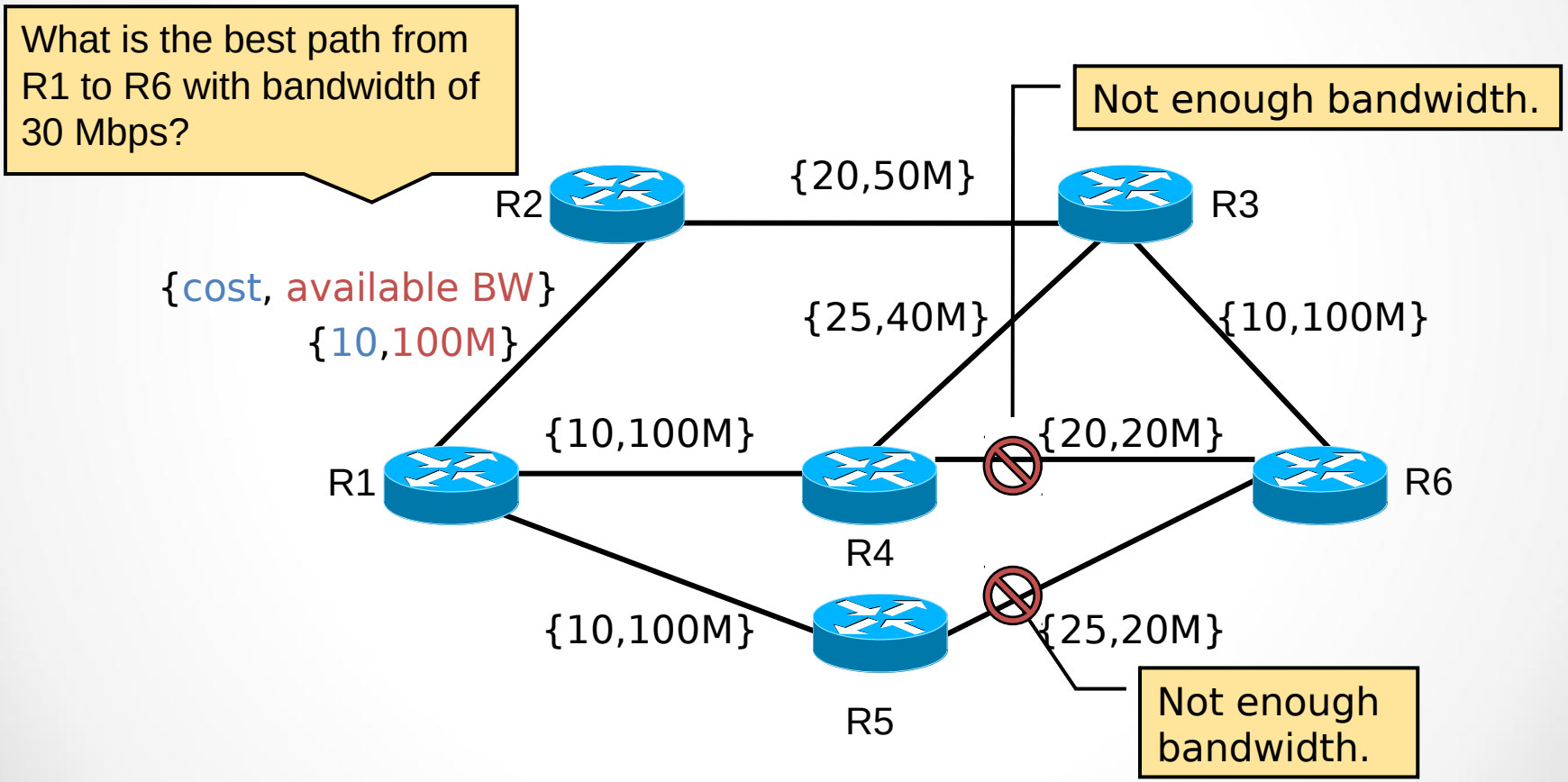
Constraint-Based Path Computation

- Constraint-based routing is demand-driven.
- Resource-reservation-aware routing paradigm:
  - Based on criteria including, but not limited to, network topology
  - Calculated at the edge of a network:
    - Modified Dijkstra's algorithm at tunnel headend (CSPF [constrained SPF] or PCALC [Path Calculation]).
    - Output is a sequence of IP interface addresses (next-hop routers) between tunnel endpoints.

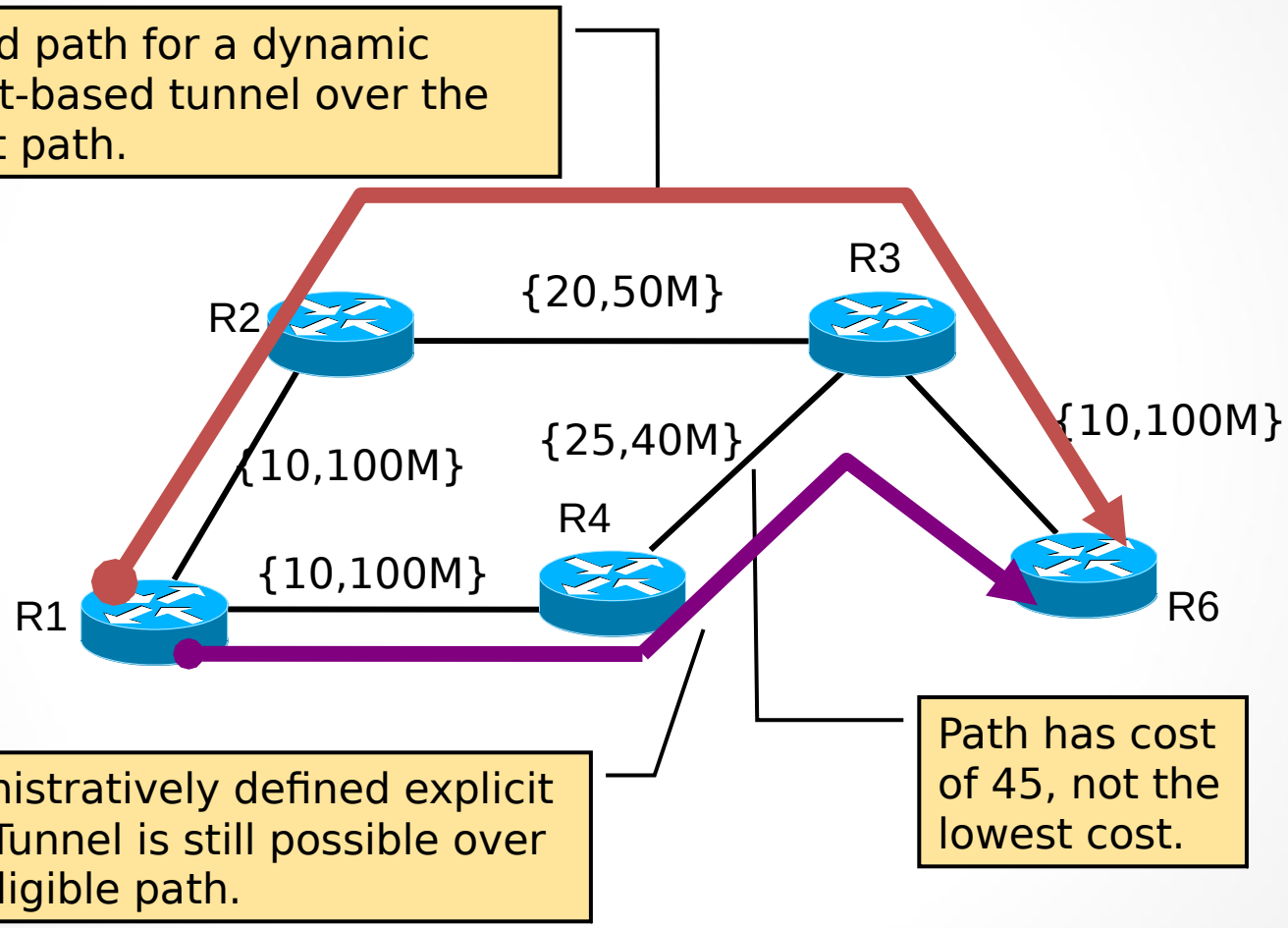
Constraint-Based Path Computation (Cont.)

- Constraint-based routing takes into account:
  - Policy constraints associated with the tunnel and physical links
  - Physical resource availability
  - Network topology state
- Two types of tunnels can be established across those links with matching attributes:
  - Dynamic—using the least-cost path computed by OSPF/IS-IS
  - Explicit—definition of a path by using OS configuration commands

Constraint-Based Path Computation (Cont.)

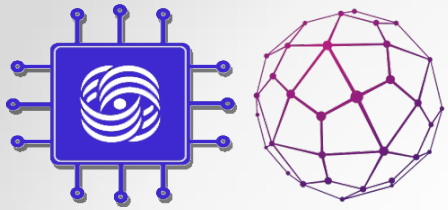


Constraint-Based Path Computation (Cont.)



Role of RSVP in Path Setup Procedures

- Once the path has been determined, a signaling protocol is needed:
  - To establish and maintain label switched paths (LSPs) for traffic tunnels
  - For creating and maintaining resource reservation states across a network (bandwidth allocation)
- The Resource Reservation Protocol (RSVP) was adopted by the MPLS workgroup of the IETF.



Role of RSVP in Path Setup Procedures

The goal of RSVP-TE is the same as that of LDP - to distribute the labels between the LSR and compile the resulting LSP from the recipient to the sender.

RSVP TE allows you:

- to build a primary and backup LSP,
- reserve resources on all nodes,
- detect network accidents,
- build pre-workarounds,
- do fast traffic redirection,
- avoid channels that physically pass through the same path.

LSP - unidirectional, resources will be reserved only in one direction.

Role of RSVP in Path Setup Procedures

The task is to create an LSP, that is to distribute the labels among the LSR.

In the simplest case:

- R1 needs LSP up to FEC 5.5.5.5/32. We will use the Tunnel interface on R1, which has a destination address of 5.5.5.5 and a Traffic Engineering type.
- It sends an RSVP path message in the direction of 5.5.5.5. A new object, Label Request, appears in this message. The Path message prompts the node to allocate a label for a given FEC.
- the Next node generates a new Path message and also sends it towards 5.5.5.5. Etc.

Role of RSVP in Path Setup Procedures

- the Path reaches the Egress LSR. It find that the packet is addressed to him, allocates a label and sends a Resv message to the source. The new object int the message is — Label. In it, Egress LSR passes its label for this FEC to the penultimate, …….
- Resv reaches the source. Thus, an LSP is created and the source is notified that everything is ready.

Labels are requested downstream (path message from sender to receiver) and transmitted upstream (Resv message from receiver to sender).

Role of RSVP in Path Setup Procedures

RSVP TE is very closely related to dynamic routing protocols

- We can use, only protocols based on link-state algorithms, i.e. OSPF and ISIS.
- OSPF and ISIS are expanding by introducing new elements. In OSPF - new type of LSA-Opaque LSA, in ISIS - new TLV IS Neighbor and IP Reachability.
- A special modification of the SPF — CSPF (Constrained Shortest Path First) algorithm is used to calculate the path between Ingress LSR and Egress LSR.

Role of RSVP in Path Setup Procedures

The Path message is transmitted by unicast address.

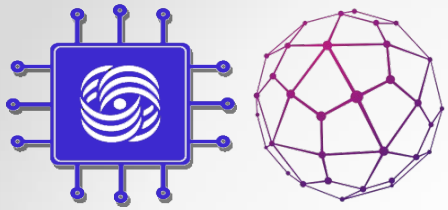
The sender's address is R1, and the recipient's is 5.5.5.5.

It could reach the recipient with the help of routing table.

But in fact the Path is transmitted over the network along a route that is defined on Ingress LSR

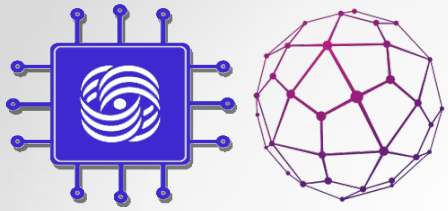
To build this path, RSVP TE needs to know the network topology that OSPF or ISIS receives

The CSPF algorithm is used to calculate the best path, the second-priority path, and, etc.

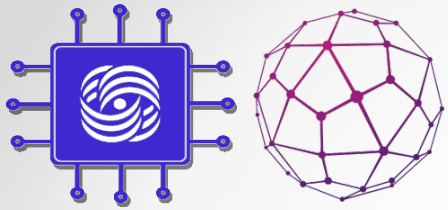


Forwarding Table Modifications

- IP routing is separate from LSP routing and does not see internal details of the LSP.
- The traffic has to be mapped to the tunnel:
  - Static routing—the static route in the IP routing table points to an LSP tunnel interface.
  - Policy routing—the next-hop interface is an LSP tunnel
  - Forwarding adjacency—the tunnel is announced as a point-to-point link to all other routers within an area
  - Autoroute—SPF enhancement:
    - The headend sees the tunnel as a directly connected interface (for modified SPF only).
    - The default cost of a tunnel is equal to the shortest IGP metric regardless of the used path.



Constraint-Based Path Computation



CSPF

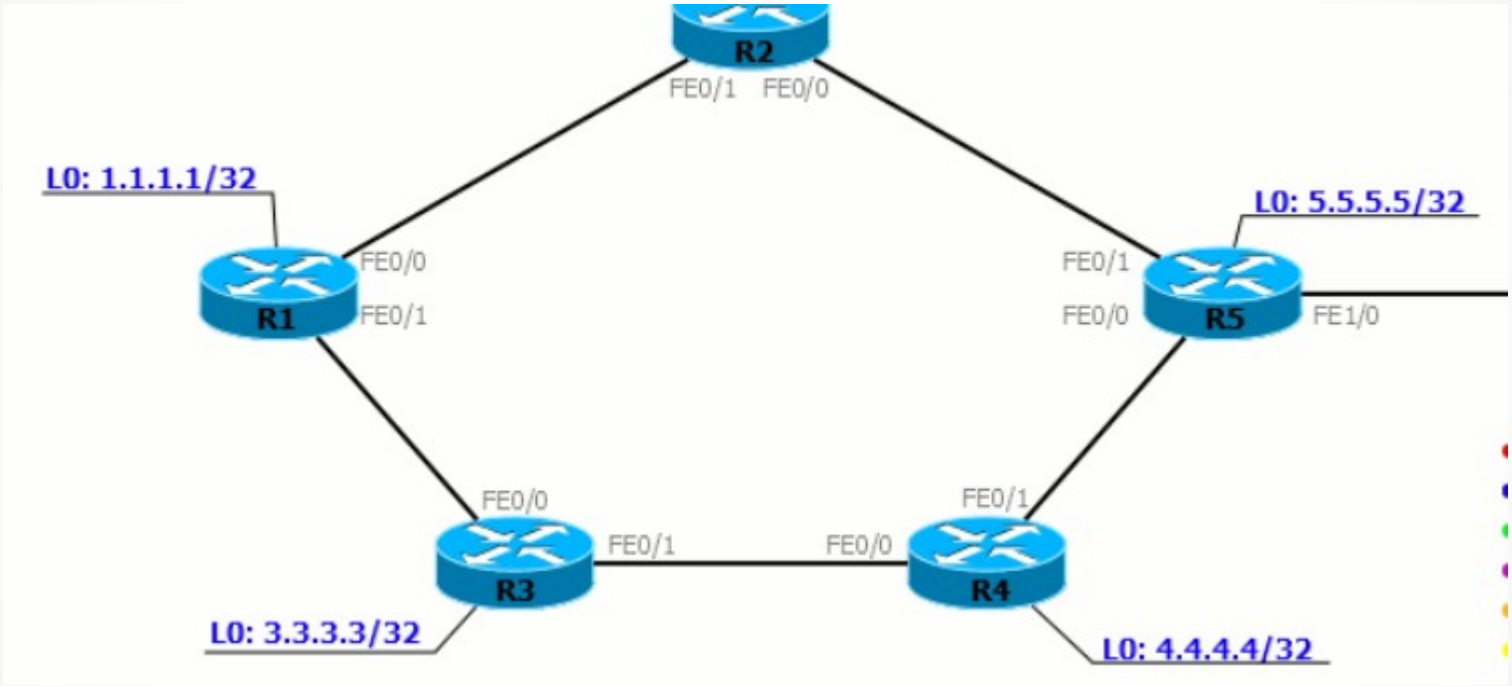
CSPF must be aware of constraints and of the available resources on the nodes of the entire network.
The input data - the restrictions specified in the tunnel and the network topology — (the topology contains information about available resources in addition to prefixes and metrics).
Routers communicate with each other through OSPF and ISIS messages not only basic information, but also characteristics of lines, interfaces, etc.

OSPF introduced 3 additional LSA types:

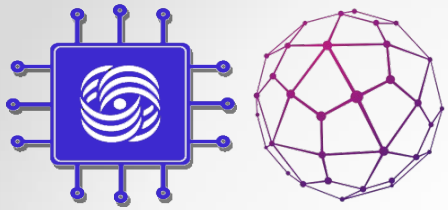
- Type 9 — link-local scope
- Type 10 — area-local scope
- Type 11 — AS scope

Opaque (for OSPF) - special types of LSA, not taken into account in the OSPF. They can be used by any other protocols for their needs. TE uses them to build its topology (it is called TED-Traffic Engineering Database).
ISIS works the same way. New messages: IS-IS TLV 22 (Extended IS Reach), 134 (Traffic Engineering router ID), 135 (Extended IP reach).

CSPF



On R1, we enabled TE MPLS and configured OSPF to transmit data to support TE. The routers exchanged information about available resources. In this step, TED is formed. RSVP is silent.

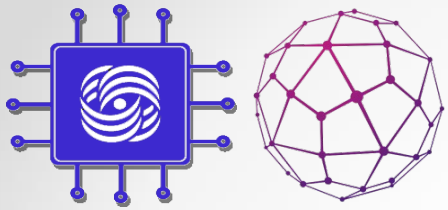


CSPF

- We created a tunnel interface, specified its type (Traffic Engineering), destination address (5.5.5.5) and necessary resource requirements.
- LSR starts CSPF: it is necessary to calculate the shortest path from R1 to 5.5.5.5 taking into account the imposed conditions. In this step we get the optimal path - a list of nodes from source to destination (R2, R5, R6).

- the Result of the previous step is passed to RSVP and transformed into an ERO object. R1 compiles the RSPV Path, and adds ERO. The destination of the package is 5.5.5.5. In addition, there is also a Label Request object that tells you that when you receive a package, you need to allocate a label for this FEC (5.5.5.5 / 32).
- ERO — Explicit Route Object — the special object of RSVP Path message. It contains a list of nodes through which this message is destined to pass.

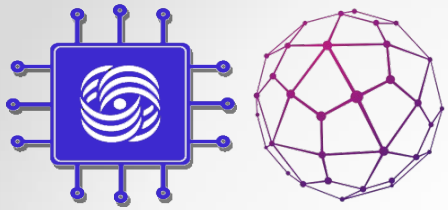
- the RSVP Path Message is transmitted in a special way- by the ERO list. In our case, the best IGP and ERO route are the same, so the packet is sent to R2.



CSPF

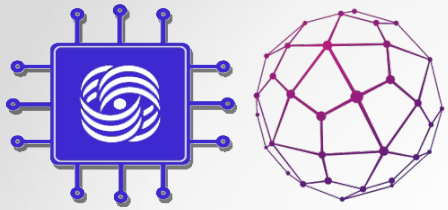
- R2, upon receiving the RSVP Path, checks for the required resources and, if any, allocates the MPLS label for FEC 5.5.5.5/32. The old Path package is destroyed and a new One is created, and the ERO object is changed — R2 itself is removed from it. This is done so that the next node does not attempt to return the packet to R2. That is, the new ERO is: (R5). R2 sends the updated Path to R5.

- R5, having received the package. It destroys the Path, allocates a label for FEC 5.5.5.5, and inserts it as a Label object in the Resv response message.
- Note that before this step, labels were only allocated, but not distributed, now they begin to be announced by the LSR that requested them.
- the RESV Message advances to R1 (Ingress LSR), forming the LSP. Resv should go through the same nodes as Path, but in reverse order.
- an LSP from R1 to 5.5.5.5 is formed. Data on it can be transferred only from R1 to R6. To allow data transfer in the opposite direction, you need to create a tunnel interface on R6 with the destination address 1.1.1.1 — all actions will be the same.



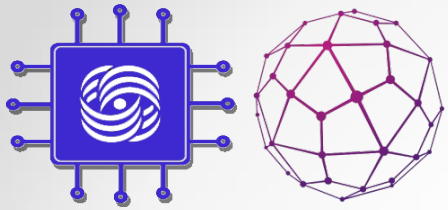
CSPF

- The ERO object may not actually contain all nodes from Ingress LSR to Egress LSR — some may be omitted.
- Therefore, each LSR must know where the message is eventually sent.
- The problem is in IGP zones (OSPF, ISIS)
- if you have a network divided into zones, CSPF can not calculate the whole path, because in its topology, the destination from another zone is a cloud.
- Explicit Path (not object ERO). The administrator can independently and explicitly specify the nodes through which the LSP needs to be routed. Ingress LSR must follow these instructions exactly.



Constraint-Based Path Computation

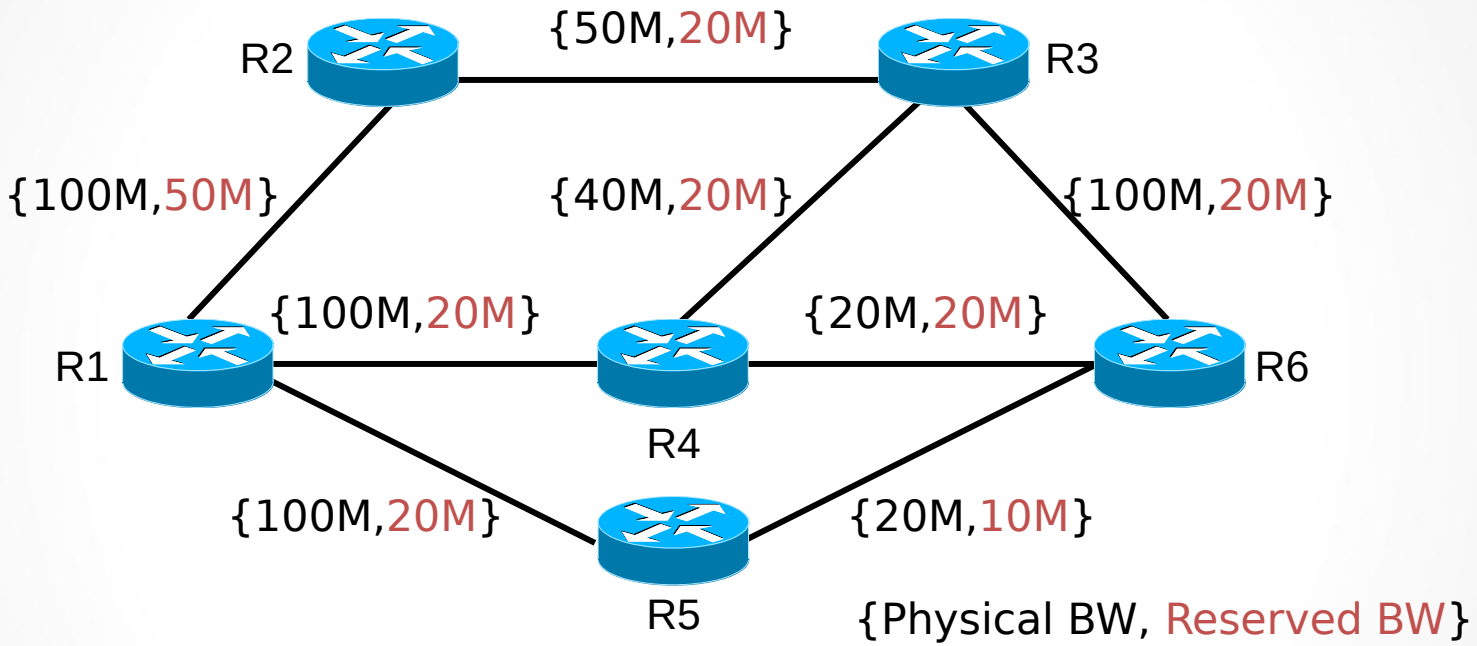
- Constraint-based path computation provides several resource attributes to control LSP path determination.
  - Link resource attributes that provide information on the resources of each link.
  - Traffic tunnel attributes characterize the traffic tunnel.



MPLS-TE Link Resource Attributes

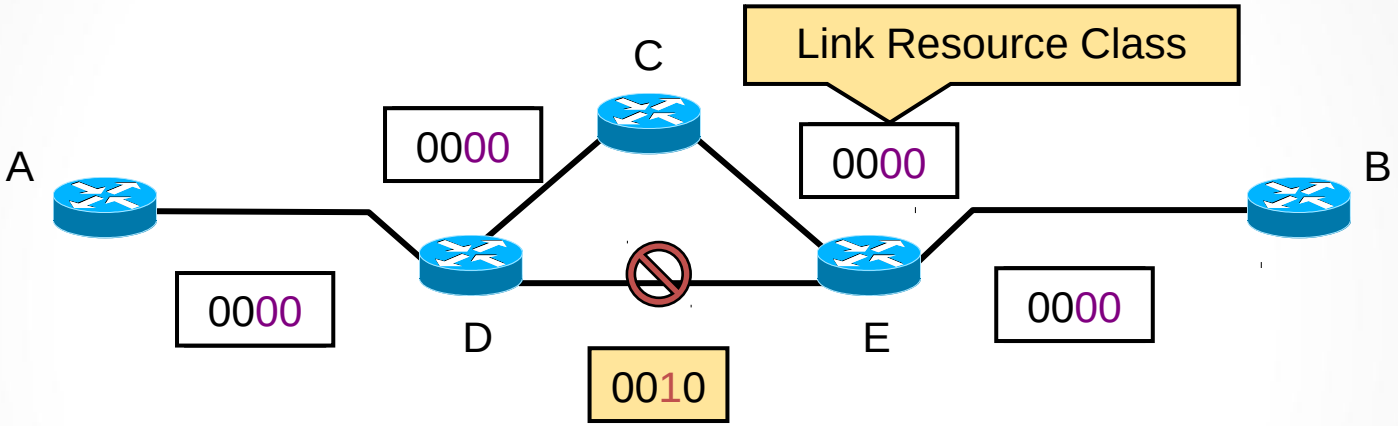
- Maximum bandwidth
- Maximum reservable bandwidth
- Link resource class
- Constraint-based specific link metric

MPLS-TE Link Resource Attributes: Maximum Allocation Bandwidth



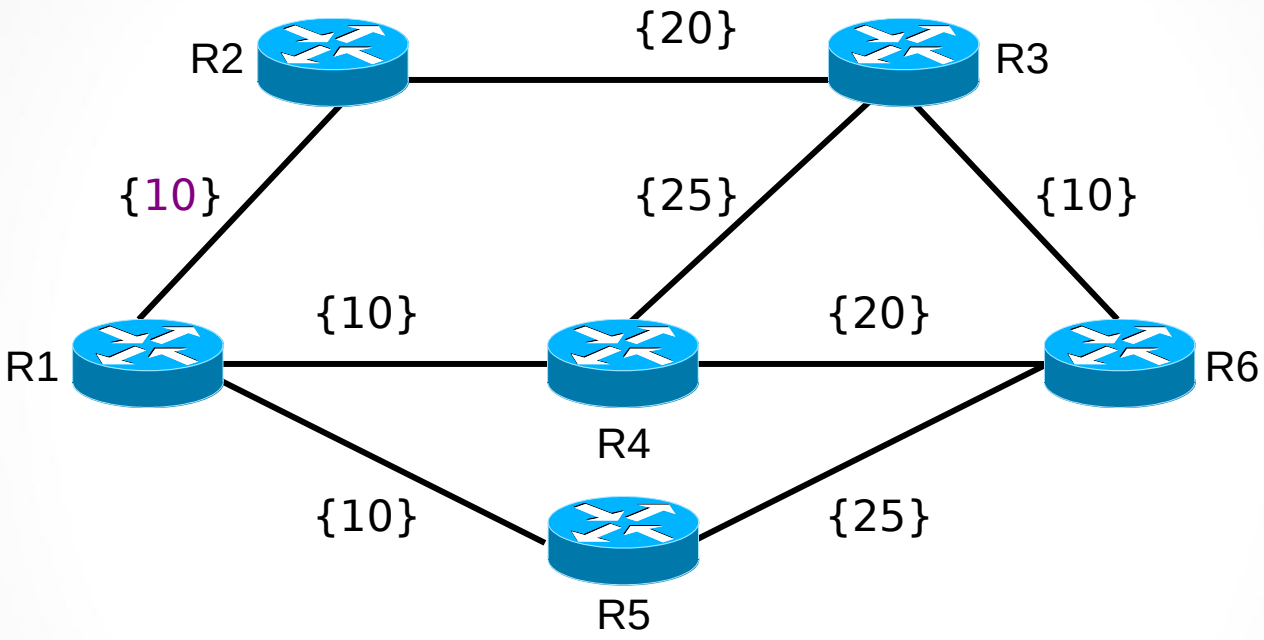
- Maximum bandwidth: the maximum bandwidth that can be used on this link in this direction (physical link)
- Maximum reservable bandwidth: The maximum amount of bandwidth that can be reserved in this direction on this link

MPLS-TE Link Resource Attributes: Link Resource Class

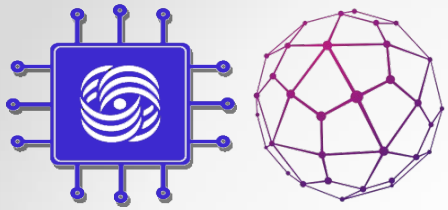


- Link is characterized by a 32-bit resource class attribute.
- Associated with a traffic tunnel in order to include or exclude certain links into or from the path of the traffic tunnel.

MPLS-TE Link Resource Attributes: Constraint-Based Specific Link Metric

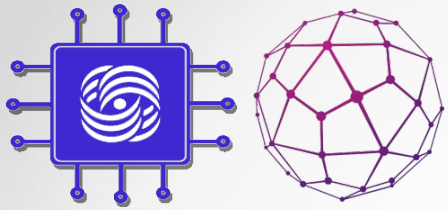


- This metric is administratively assigned to present a differently weighted topology to traffic engineering SPF calculations:
  - » Administrative weight (TE metric)



MPLS-TE Tunnel Attributes

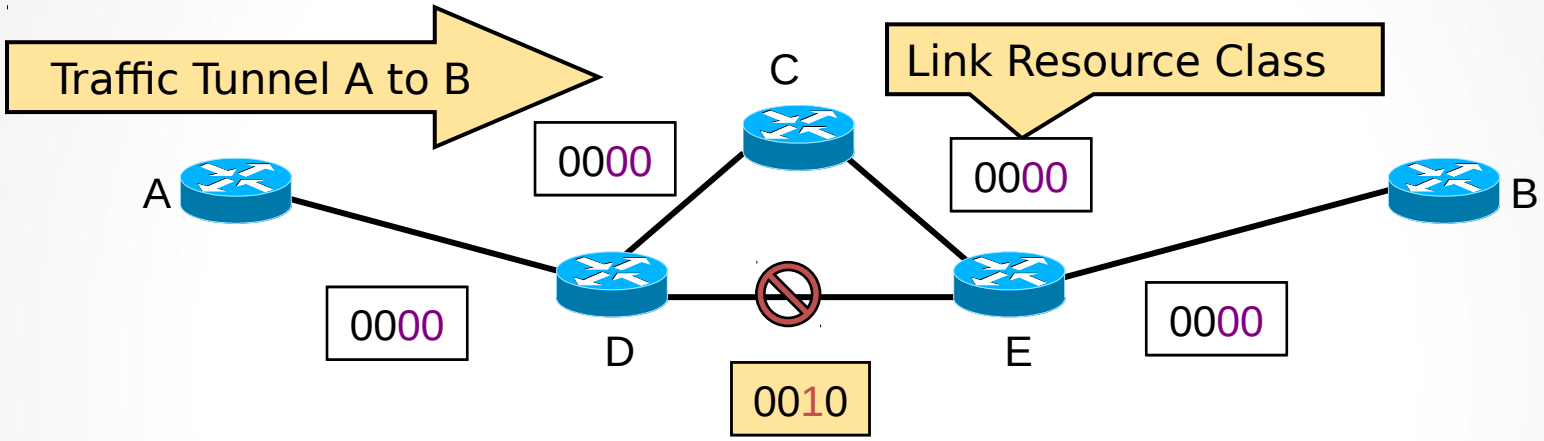
- Traffic parameter
- Generic path selection and management
- Tunnel resource class affinity
- Adaptability
- Priority
- Pre-emption
- Resilience



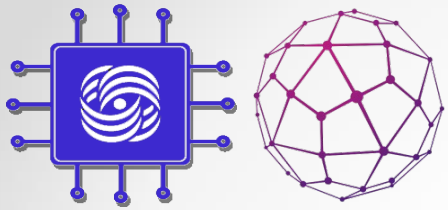
MPLS-TE Tunnel Attributes (Cont.)

- Traffic parameter:
  - Indicates the resource requirements (for example, bandwidth) of the traffic tunnel
- Generic path selection and management:
  - Specifies how the path for the tunnel is computed:
    - Dynamic LSP — Constraint-based computed paths based on a combination of bandwidth and policies
    - Explicit LSP — administratively specified off line (typically using CLI)

MPLS-TE Tunnel Attributes (Cont.)

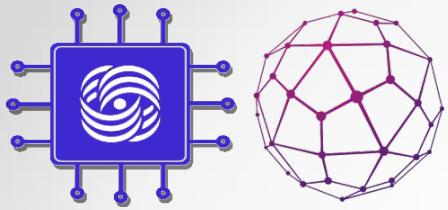


- Tunnel Resource Class Affinity:
  - The properties that the tunnel requires from internal links:
    - 32-bit resource class affinity bit string + 32-bit resource class mask
  - Link is included in the constraint-based LSP path when the tunnel resource affinity string or mask matches the link resource class attribute.



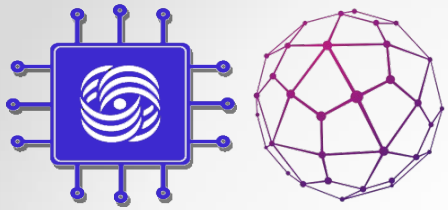
MPLS-TE Tunnel Attributes (Cont.)

- Adaptability:
  - If reoptimization is enabled, then a traffic tunnel can be rerouted through different paths by the underlying protocols:
    - Primarily due to changes in resource availability
- Priority:
  - Relative importance of traffic tunnels
  - Determines the order in which path selection is done for traffic tunnels at connection establishment and under fault scenarios:
    - Setup priority: Priority for taking a resource
- Pre-emption:
  - Determines whether another traffic tunnel can pre-empt a specific traffic tunnel:
    - Hold priority: Priority for holding a resource



MPLS-TE Tunnel Attributes (Cont.)

- Resilience:
  - Determines the behavior of a traffic tunnel under fault conditions:
    - Do not reroute the traffic tunnel
    - Reroute through a feasible path with enough resources
    - Reroute through any available path regardless of resource constraints



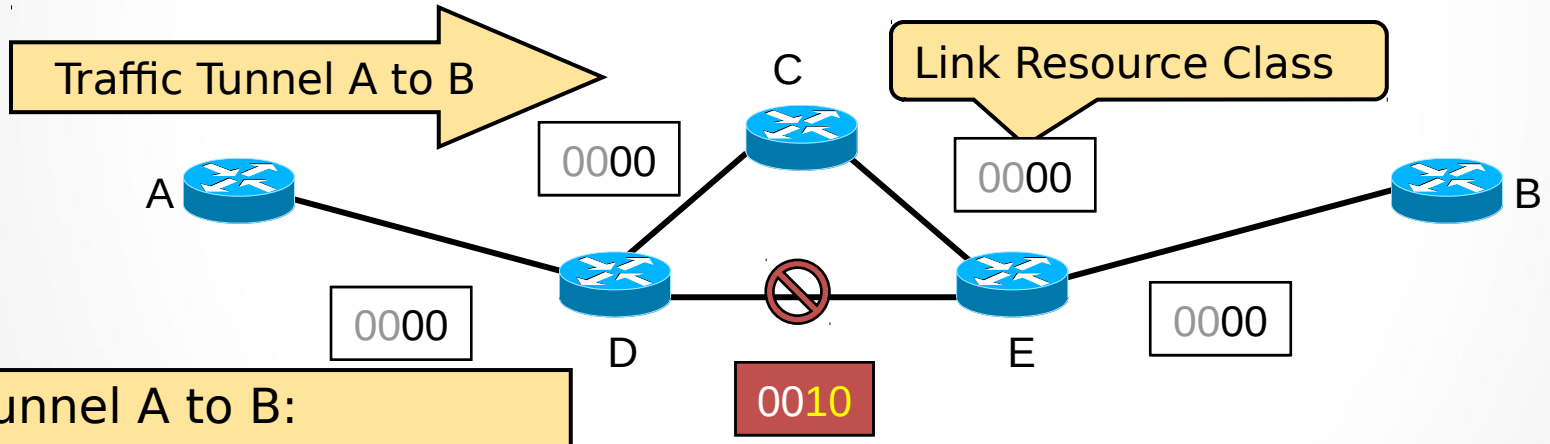
Implementing TE Policies with Affinity Bits

- Link is characterized by the link resource class
  - Default value of bits is 0
- Tunnel is characterized by:
  - Tunnel resource class affinity
    - Default value of bits is 0
  - Tunnel resource class affinity mask
    - (0=do not care, 1=care)
    - Default value of the tunnel mask is 0x0000FFFF

Implementing TE Policies with Affinity Bits (Cont.)

Setting a link bit in the lower half drives all tunnels off the link, except those specially configured.

Tunnel Affinity: bits = 0000, mask = 0011



Tunnel A to B:

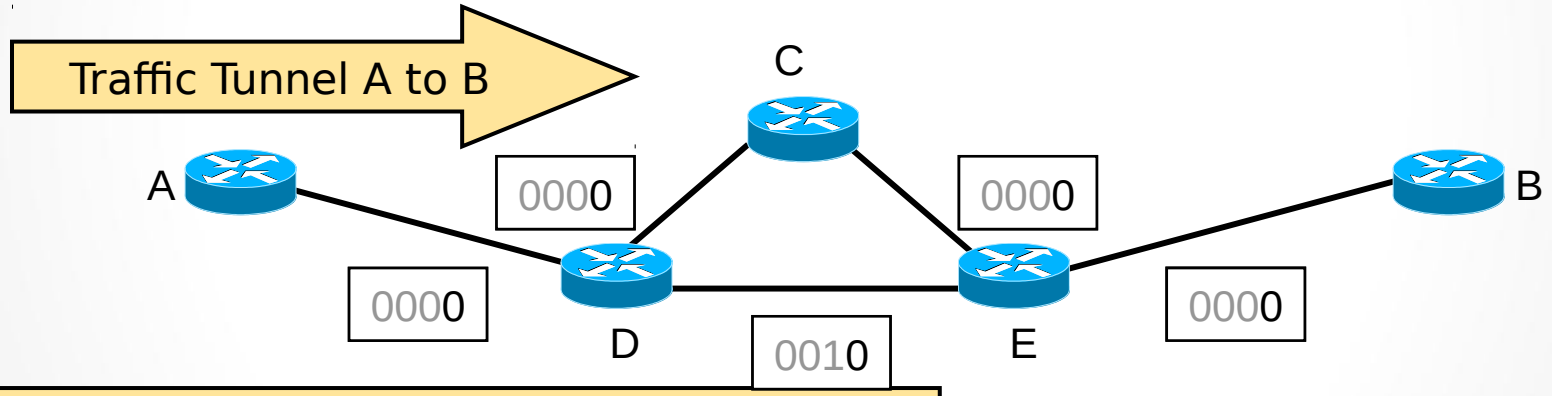
Only ADCEB is possible.

Using Affinity Bits to Avoid Specific Links

Implementing TE Policies with Affinity Bits (Cont.)

A specific tunnel can then be configured to allow all links by clearing the bit in its affinity attribute mask.

Tunnel Affinity: bits = 0000, mask = 0001



Tunnel A to B:

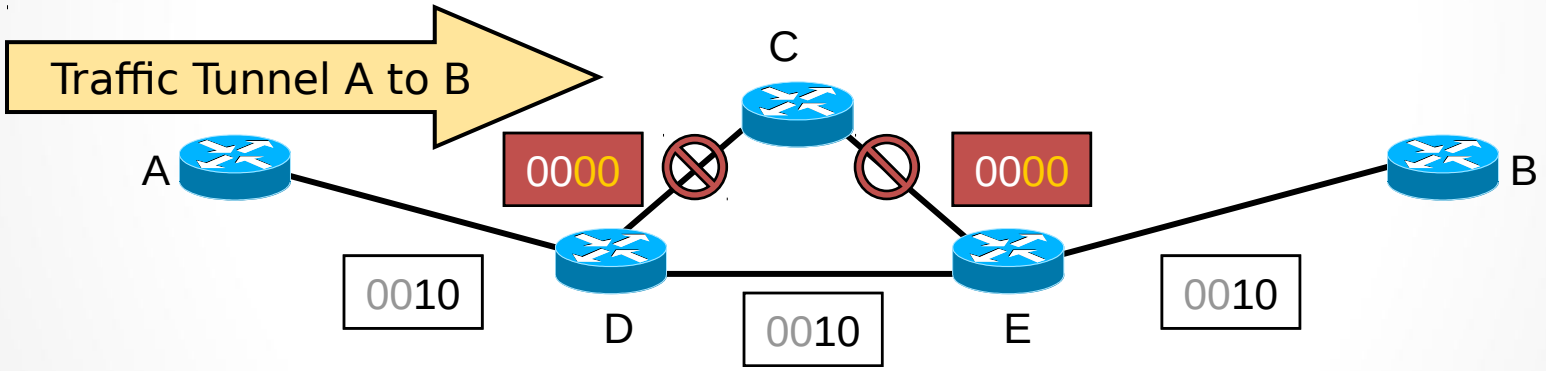
Again, ADEB and ADCEB are possible.

Using the Affinity Bit Mask to Allow all Links

Implementing TE Policies with Affinity Bits (Cont.)

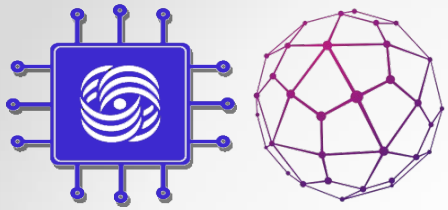
A specific tunnel can be restricted to only some links by turning on the bit in its affinity attribute bits.

Tunnel Affinity: bits = 0010, mask = 0011



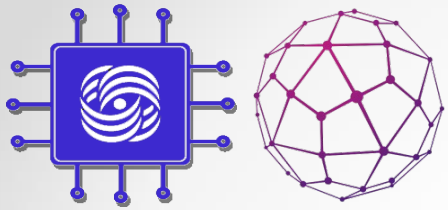
Tunnel A to B:
ADEB is possible.

Using Affinity Bits to Dedicate Links to Specific Purposes



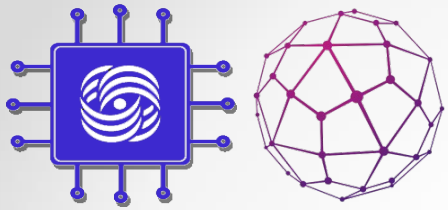
Propagating MPLS-TE Link Attributes with Link-State Routing

- IGP resource flooding takes place in the following situations:
  - Link-state changes
  - Resource class of a link changes:
    - Manual reconfiguration
    - Amount of available bandwidth crosses one of the preconfigured thresholds
  - Periodic (timer-based):
    - A node checks attributes; if they are different, it floods its update status
  - On LSP setup failure



Constraint-Based Path Computation

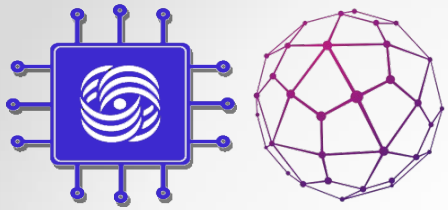
- When establishing a tunnel, the edge routers have knowledge of both network topology and link resources within its area:
  - Two methods for establishing traffic tunnels:
    - Static and dynamic path setup
  - In both cases the result is an explicit route expressed as a sequence of interface IP addresses (for numbered links) or TE router IDs (for unnumbered links) in the path from tunnel endpoints.
  - RSVP is used to establish and maintain constraint-based label switched paths for traffic tunnels along an explicit path.



Constraint-Based Path Computation (Cont.)

Constraint-Based Path Selection

- Path selection:
  - CBR uses its own metric (administrative weight, or TE cost; by default equal to the IGP cost)—used only during constraint-based computation
  - In case of a tie, select the path with:
    - The highest minimum bandwidth
    - The smallest hop count
    - If everything else fails, then pick a path at random
- LSP path setup—an explicit path is used by RSVP to reserve resources and establish LSP path
- Final result: unidirectional MLPS-TE tunnel, seen only at the headend router



Constraint-Based Path Computation (Cont.)

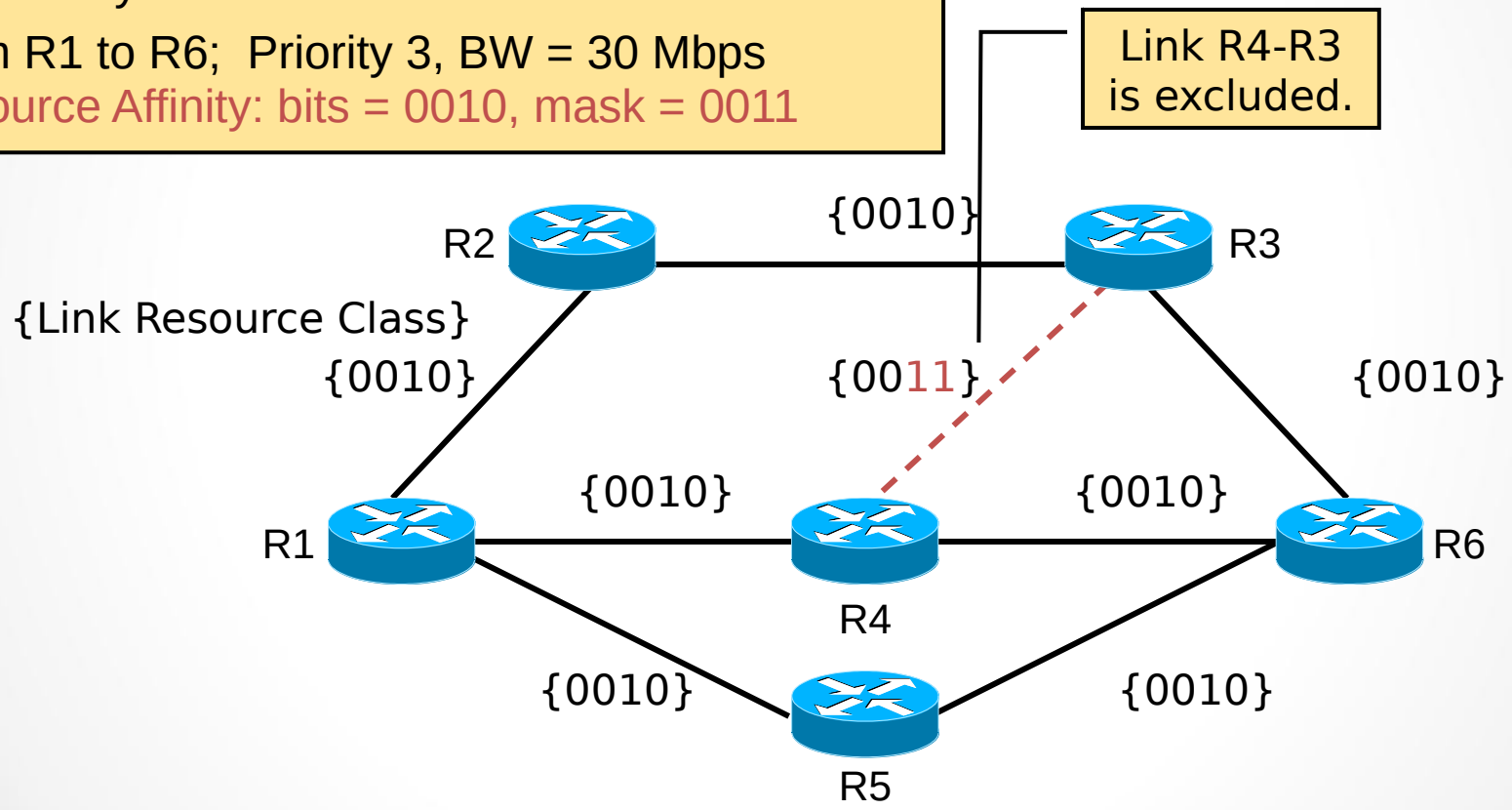
- MPLS-TE tunnel is not a link for link-state adjacency:
  - Establishment of a tunnel does not trigger any LSA announcements or a new SPF calculation (unless the forwarding adjacency feature is enabled).
  - Software uses tunnel interface for MPLS-TE tunnel creation and visualization, but behavior of MPLS-TE tunnels is fairly different from other tunnel protocols (for example, GRE).
- Only traffic entering at headend router will use tunnel
- IP cost: If autoroute is used, MPLS-TE tunnel in the IP routing table has a cost of the shortest IGP path to the tunnel destination (regardless of the LSP path)

Constraint-Based Path Computation (Cont.)

Request by tunnel:

From R1 to R6; Priority 3, BW = 30 Mbps

Resource Affinity: bits = 0010, mask = 0011



Path Selection Considering Policy Constraints

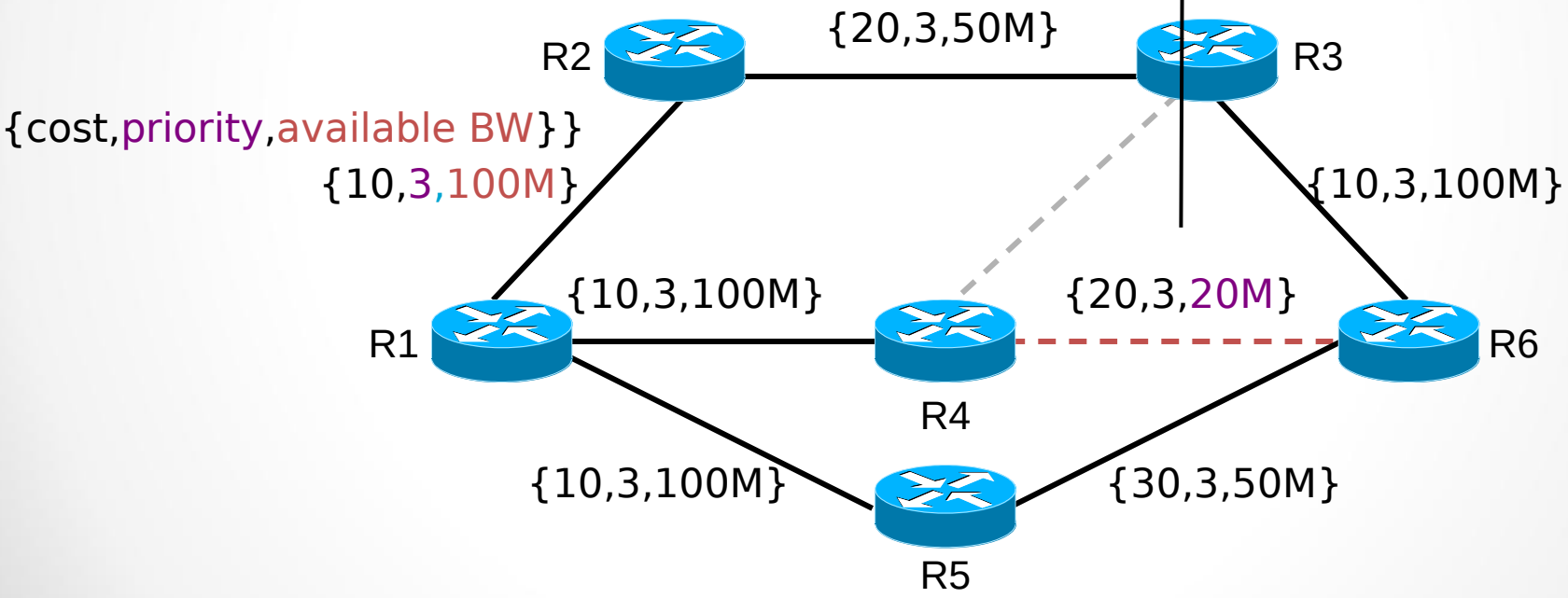
Constraint-Based Path Computation (Cont.)

Request by tunnel:

From R1 to R6; Priority 3, BW = 30 Mbps

Resource Affinity: bits = 0010, mask = 0011

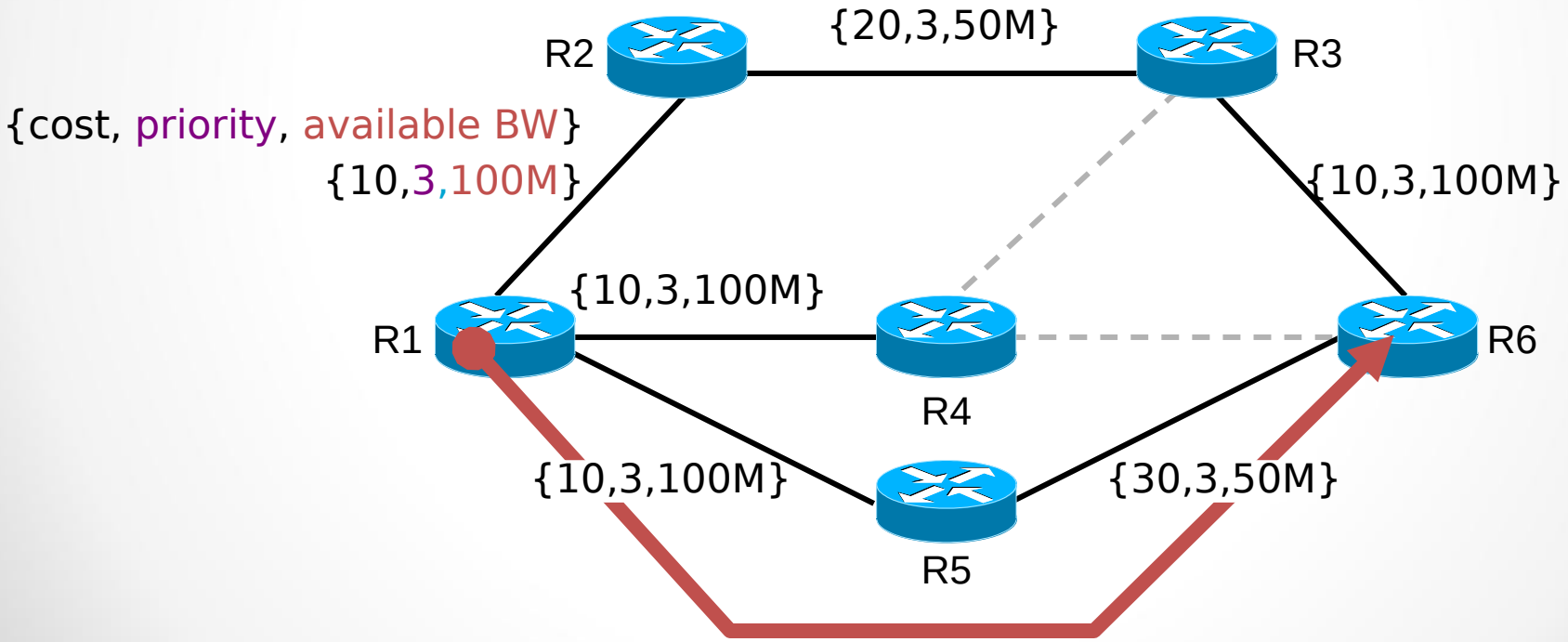
Not enough bandwidth



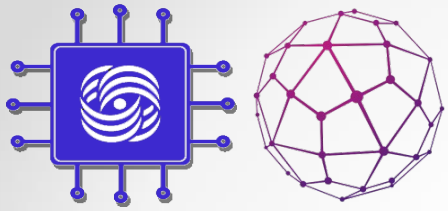
Path Selection Considering Available Resources

Constraint-Based Path Computation (Cont.)

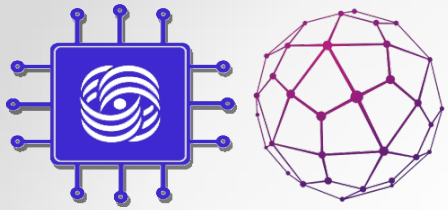
The headend router has two possible paths with a total cost of 40: R1 – R2 – R3 – R6 and R1 – R5 – R6, both offering at least 50 Mbps (minimim bandwidth). Because of the smaller hop count, R1 – R5 – R6 is selected.



Selecting the Best Path

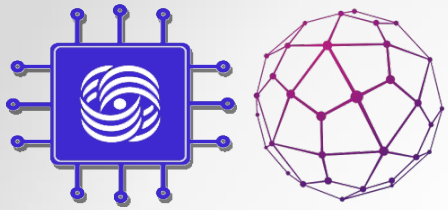


Path Setup and Maintenance



Path Setup

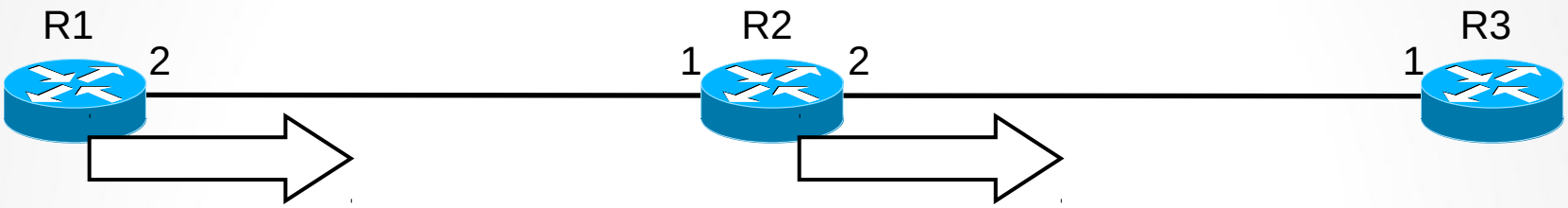
- LSP path setup is initiated at the headend of a tunnel.
- The route (list of next-hop routers) is either:
  - Statically defined
  - Computed by CBR
- The route is used by RSVP to:
  - Assign labels
  - Reserve bandwidth on each link
- Tunnel attributes that affect path setup:
    - Bandwidth
    - Priority
    - Affinity attributes



RSVP Usage in Path Setup

- RSVP makes resource reservations for both unicast and multicast applications:
  - RSVP provides support for dynamic membership changes and automatic adaptation to routing changes.
  - RSVP sends periodic refresh messages to maintain the state along the reserved path.
  - RSVP sessions are used between routers, not hosts.

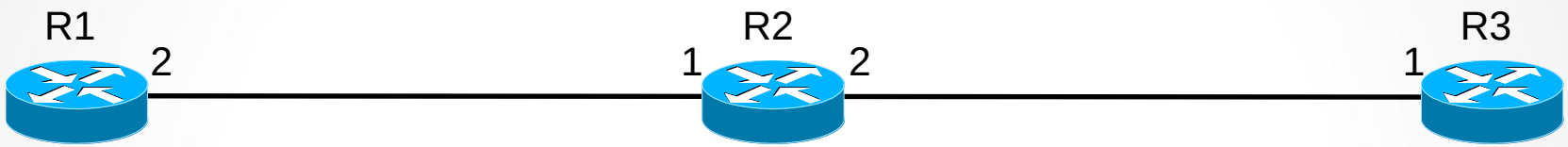
Hop-by-Hop Path Setup with RSVP



Path:
Common_Header
Session(R3-lo0, 0, R1-lo0)
PHOP(R1-2)
Label_Request(IP)
ERO (R2-1, R3-1)
Session_Attribute (...)
Sender_Template(R1-lo0, 00)
Record_Route(R1-2)

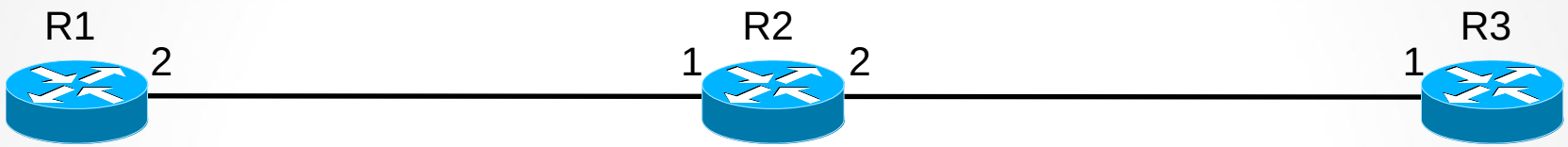
Path:
Common_Header
Session(R3-lo0, 0, R1-lo0)
PHOP(R2-2)
Label_Request(IP)
ERO (R3-1)
Session_Attribute (...)
Sender_Template(R1-lo0, 00)
Record_Route (R1-2, R2-2)

Hop-by-Hop Path Setup with RSVP (Cont.)



Path State:
Session(R3-lo0, 0, R1-lo0)
PHOP(R2-2)
Label_Request(IP)
ERO ()
Session_Attribute (...)
Sender_Template(R1-lo0, 00)
Record_Route (R1-2, R2-2, R3-1)

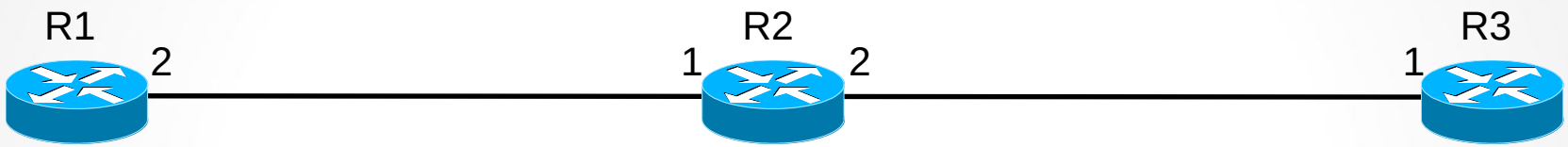
Hop-by-Hop Path Setup with RSVP (Cont.)



Resv:
Common_Header
Session(R3-lo0, 0, R1-lo0)
PHOP(R2-1)
Sender_Template(R1-lo0, 00)
Label=25
Record_Route(R2-1, R3-1)

Resv:
Common_Header
Sess[illegible]0, R1-lo0)
PHOP(R3-1)
Sender_Template(R1-lo0, 00)
Label=POP
Record_Route(R3-1)

Hop-by-Hop Path Setup with RSVP (Cont.)



Resv state:
Session(R3-lo0, 0, R1-lo0)
PHOP(R2-1)
Sender_Template(R1-lo0, 00)
Label=5
Record_Route(R1-2, R2-1, R3-1)

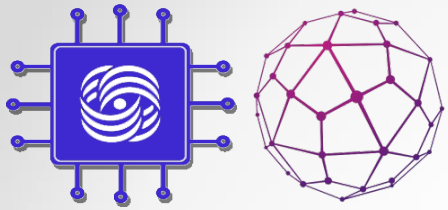
Tunnel and Link Admission Control (Cont.)

- Pre-emption
  - The process of LSP path setup may require the pre-emption of resources.
  - LCAC notifies RSVP of the pre-emption.
  - RSVP sends PathErr or ResvErr or both for the pre-empted tunnel.

Path Rerouting

Path rerouting may result from:

- Reoptimization due to a change in network resources
- Link failures that affect the LSP path

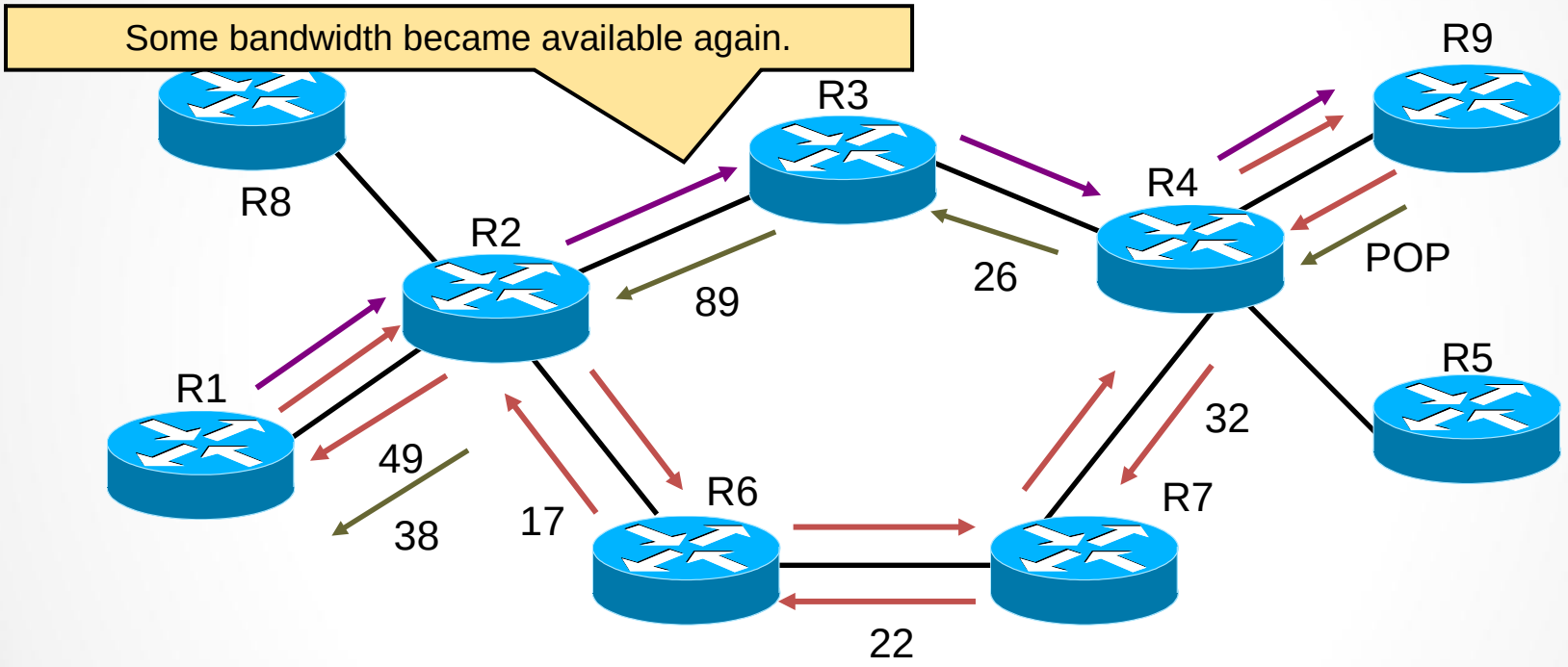


Path Reoptimization

- Problem: Some resources become available, which results in a nonoptimal path of traffic tunnels
- Solution: Reoptimization:
  - A periodic timer checks for the most optimal path
  - If a better LSP seems to be available:
    - The device attempts to signal the better LSP
    - If successful, replaces the old and inferior LSP with the new and better LSP

Path Reoptimization (Cont.)

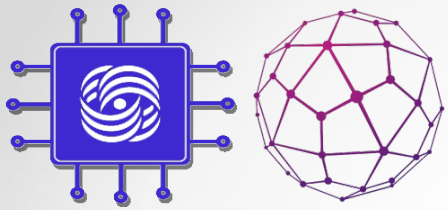
Nondisruptive Rerouting — Reoptimization



Current Path (ERO = R1 -> R2 -> R6 -> R7 -> R4 -> R9).

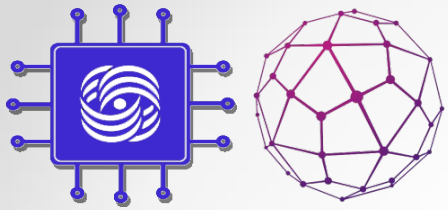
New Path (ERO = R1 -> R2 -> R3 -> R4 -> R9)—shared with current path and reserved for both paths.

Until R9 gets new Path message, current Resv is refreshed—PathTear can then be sent to remove old path (and release resources).



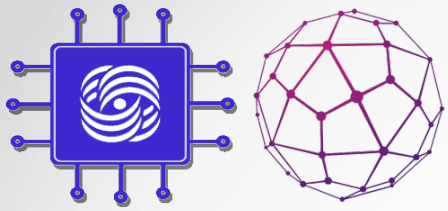
Path Rerouting: Link Failure

- The Goal
  - Repair at the headend of the tunnel in the event of failure of an existing LSP:
    - IGP or RSVP alarms the headend.
  - New path for LSP is computed, and eventually a new LSP is signaled.
  - Tunnel interface goes down if there is no LSP available for 10 sec.



Path Rerouting: Link Failure (Cont.)

- Link failure – What happens:
- (Example: One link along a dynamic tunnel LSP path goes down.)
  - RSVP PathTear causes the headend to flag LSP as dead
  - RSVP session is cleared
  - PCALC triggered:
    - No alternative path:
      - Headend sets the tunnel down
    - Alternative path found:
      - New LSP directly signaled
      - Adjacency table updated for the tunnel interface

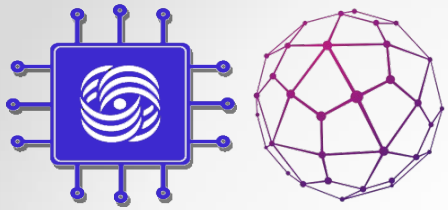


Assigning Traffic to Traffic Tunnels

Traffic Flow Modifications

- In contrast to LDP LSP, on which traffic runs by default, we need to direct traffic in TE-tunnels.

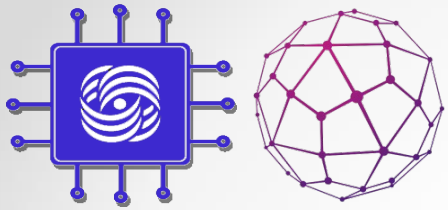
- Static route
- PBR
- IGP Shortcut
- Tunnel-policy



IGP Shortcut

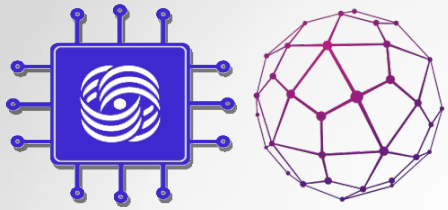
- This method is the most common and is supported by almost all manufacturers.
- The router treats the tunnel as a virtual interface. And through this interface, remote routers seem to be directly connected to the local.

- With the help of IGP Shortcut (AutoRoute Announce), we force the routing Protocol on Ingress LSR to consider the tunnel as a normal Egress LSR connected directly. All networks behind Egress LSR will be accessible through the tunnel and the packets for them will be sent to the tunnel, including VPN packages.



Metric of the tunnel (IGP Shortcuts)

- By default, the tunnel metric is the sum of the TE metrics of all lines from Ingress to Egress along the shortest IP path (not the one the tunnel is following). That is, the metrics of normal IP routes and routes through the tunnel will be the same, even if the tunnel is actually much longer.
- The tunnel metric should override the best IP path metric.
- If the metrics are equal, the router will choose the tunnel.
- If there are IP paths that do not share segments with the tunnel LSP, and their metrics are equal, balancing will take place.



Methods for managing the metric of the tunnel (IGP Shortcuts)

- Change the value of the TE metric on physical interfaces.
- Specify MPLS TE to use IGP metric instead of TE.
- Accordingly, change the IGP metric of the physical interface.
- To set directly the metric of the tunnel interface:
    - It can be specified statically.

  When calculating the route metric, the metrics of the sections will be summed.
    - Can specify absolute value. For all routes that are accessible through the tunnel, the metric will be the same.
    - Can set it relative to the IGP metric. For example, more on 5, or less on 7.

Tunnel management methods

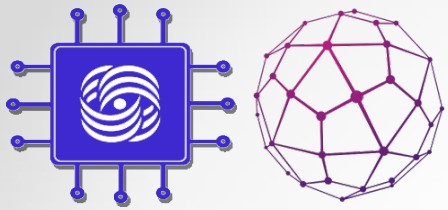
- MPLS te path metric
- The restriction on the bandwidth
- Explicit-Path
- The priorities of tunnels

Tunnel management methods

- The restriction on the bandwidth

- Offline Bandwidth - A method that uses a static setting of the required bandwidth value
- Auto-Bandwidth - This method involves tracking the tunnel load over a period of time and then adapting the reservation.

- Adjust Interval — the time during which the router monitors traffic and tracks peaks.
- Adjust Threshold-the threshold after which RSVP overwrites the reservation.



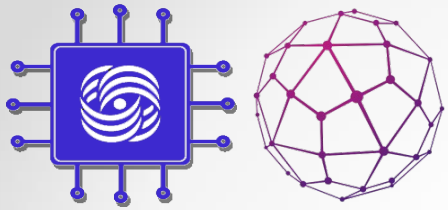
Tunnel management methods

- The priorities of tunnels

- **Setup Priority**-priority of RSVP LSP installation.
- **Hold Priority**-hold priority of RSVP LSP.

- If there is not enough bandwidth on the node and the new tunnel has a higher installation priority than the hold priority of the old one, the new one will be built — the old one is broken.

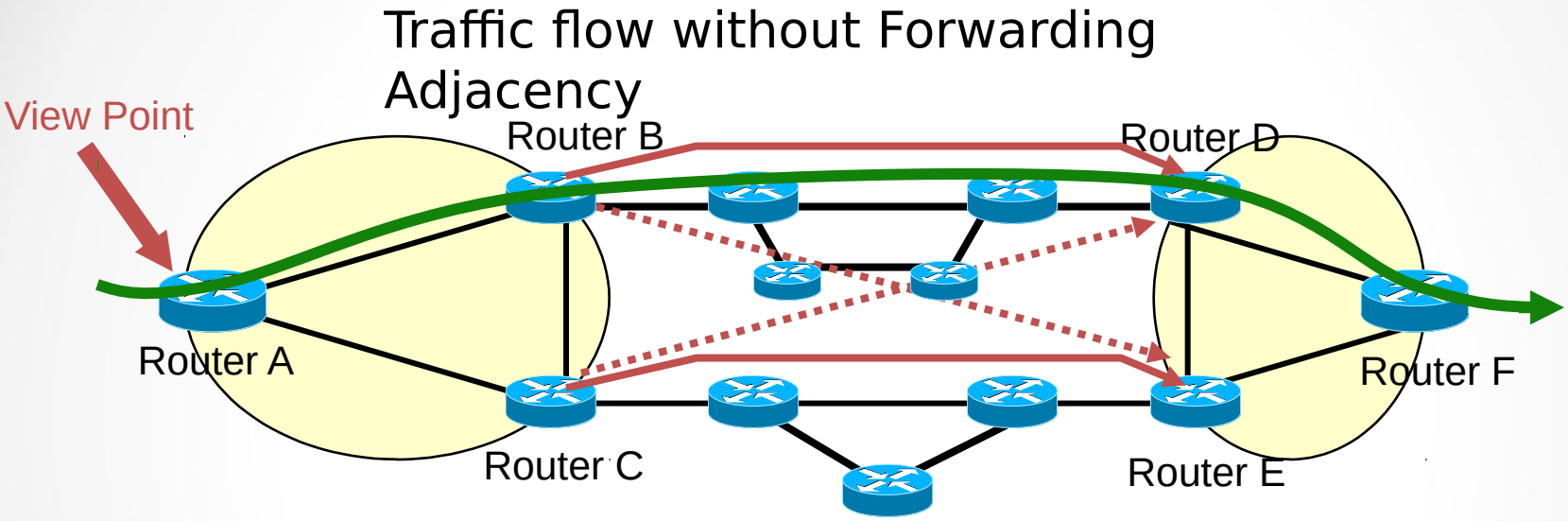
- Both have 8 values: 0 to 7. For both, 0 is the highest, 7 is the lowest.



Forwarding Adjacency

- Mechanism for:
  - Better intra- and inter-POP load balancing
  - Tunnel sizing independent of inner topology
- Allows the announcement of established tunnel via link-state (LSP) announcements

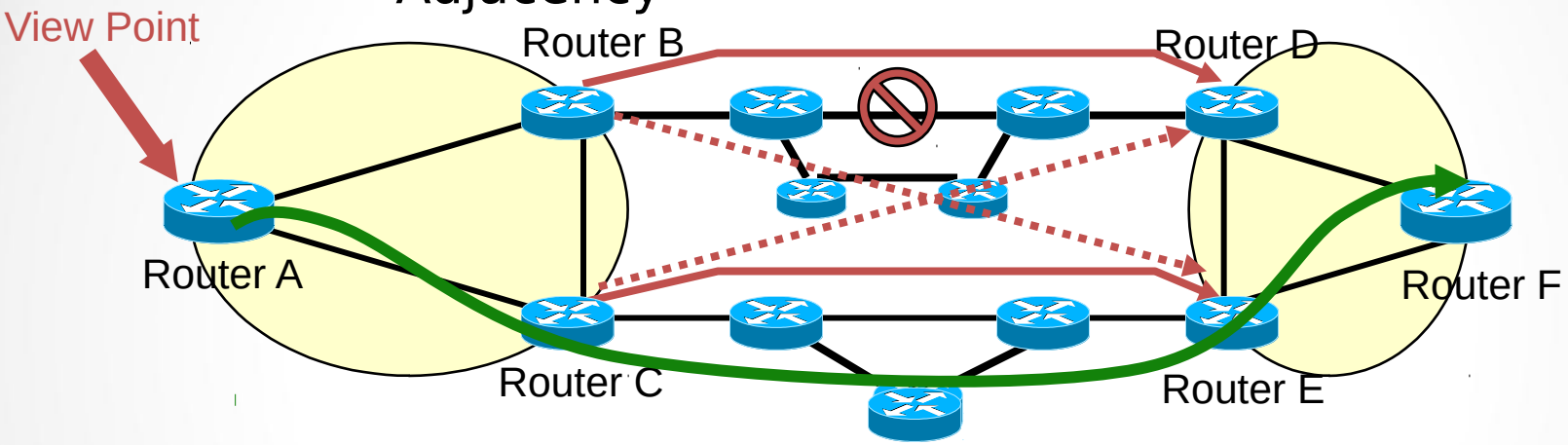
Forwarding Adjacency (Cont.)



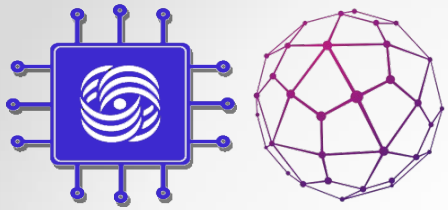
- **Tunnels created and announced to IP with autoroute with equal cost to load-balance.**
  - All the POP-to-POP traffic exits via the routers on the IGP shortest path:
    - No load balancing
    - All traffic flows on tunnel: A □ B □ D □ F

Forwarding Adjacency (Cont.)

Traffic flow without Forwarding Adjacency



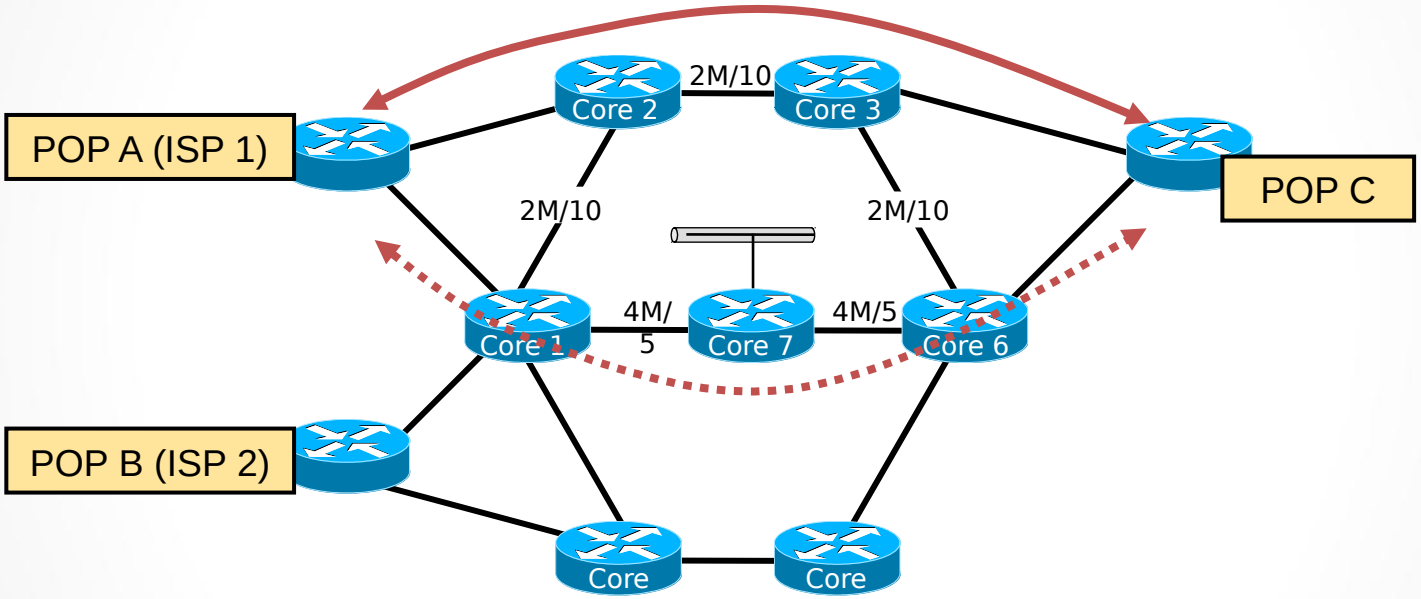
- All the POP-to-POP traffic exits via the routers on the IGP shortest path.
- Change in the core topology does affect the load balancing in the POP:
  - Normal state: All traffic flows A □ B □ D □ F
  - Link failure: All traffic flows A □ C □ E □ F



Advanced MPLS-TE Link Protection

Improving Convergence Time

The search for an alternative path and its signaling takes too long and has a negative impact on packet forwarding.

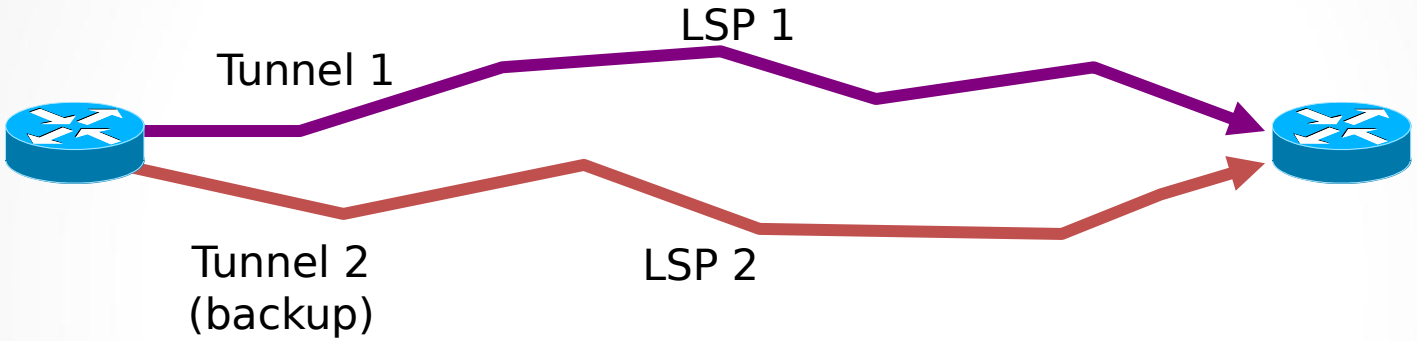


Solution with two pre-established tunnels to the same destination:

- One tunnel could be configured as a backup to another tunnel.
- LSP for the secondary tunnel is presignaled and available if the first tunnel fails.
- Double reservation can be avoided with a “make-before-break” mechanism.

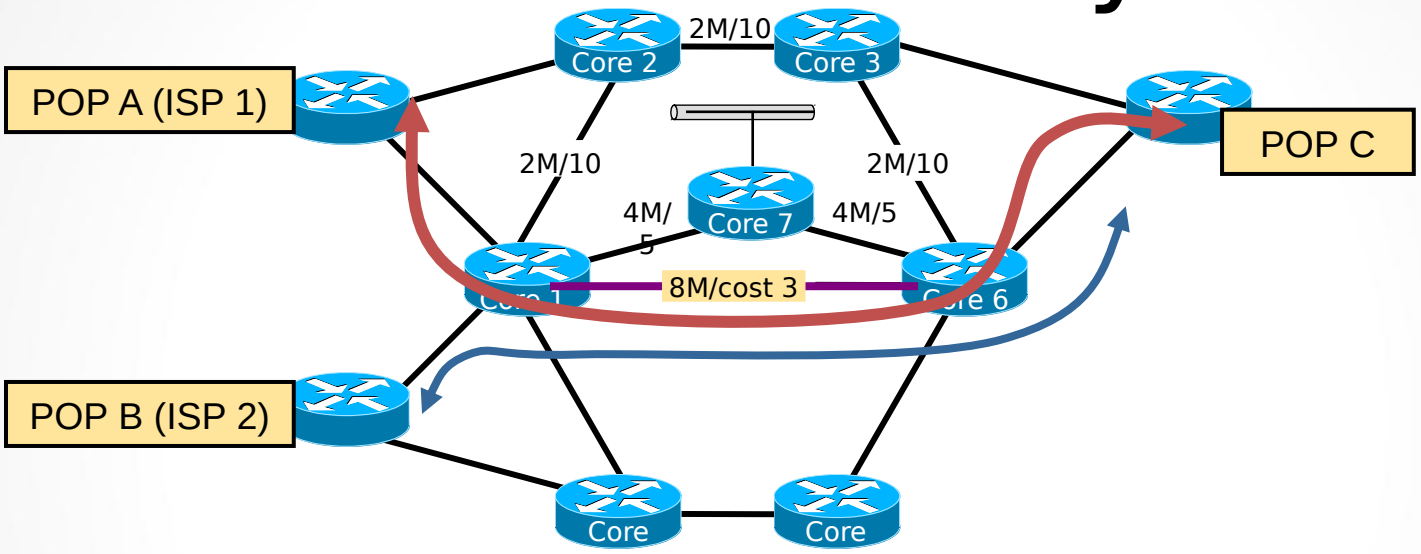
Drawbacks of Parallel Tunnels

Path Protection with Preconfigured Tunnels



- Preconfigured tunnels speed up recovery by moving the traffic on a preinstalled LSP as soon as the headend learns the primary LSP is down.
- Drawbacks:
  - Backup tunnel allocates labels and reserves bandwidth over the entire path
  - Double counting of reservations via RSVP over the entire path

Fast Reroute: Case Study



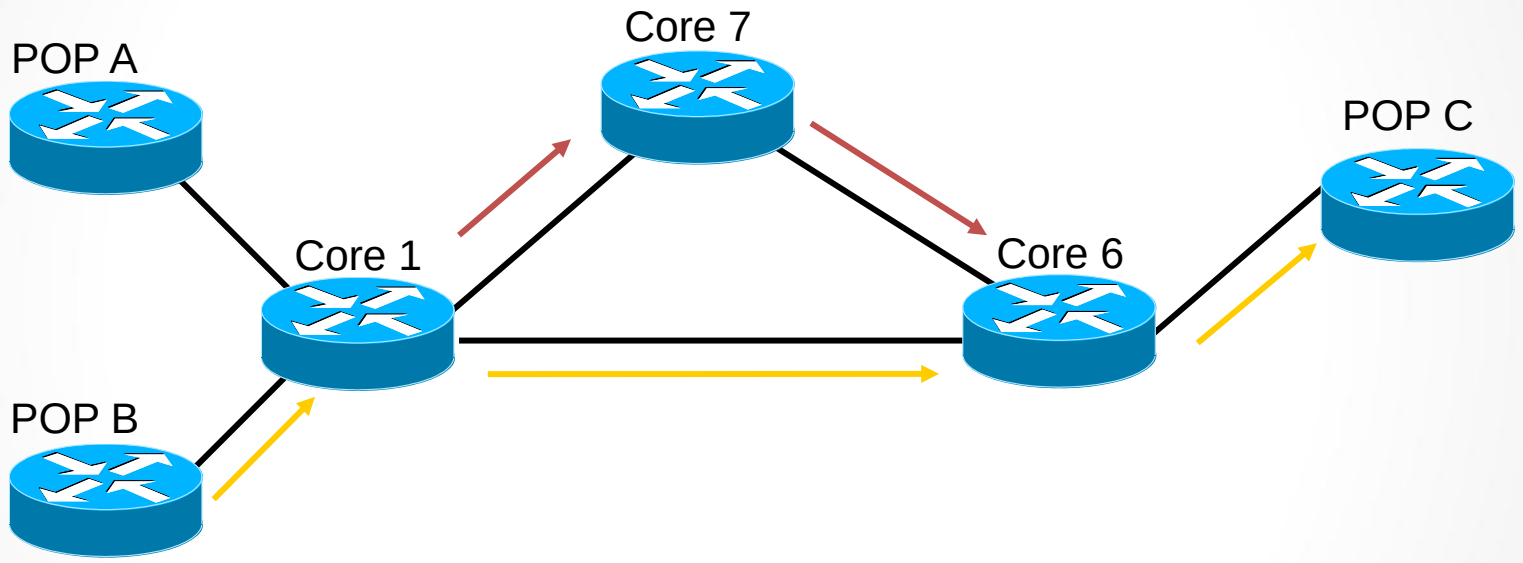
- The company decided to retain only dynamic tunnels. A new high- speed link was introduced between Core 1 and Core 6 to influence CBR and native path selection and speed up transport across the network.
- The new high-speed link is now heavily used by traffic tunnels and may cause a serious disruption.

Fast Reroute

- Fast Reroute allows for temporary routing around a failed link or a failed node while the headend is rerouting the LSP:
  - Controlled by the routers with preconfigured backup tunnels around the protected link or node (link or node protection).
  - The headend is notified of the failure through the IGP and through RSVP.
  - The headend then attempts to establish a new LSP that bypasses the failure (LSP rerouting).

Link Protection with FRR

Link Protection for Core 1 – Core 6 Link



End-to-end tunnel onto which data normally flows

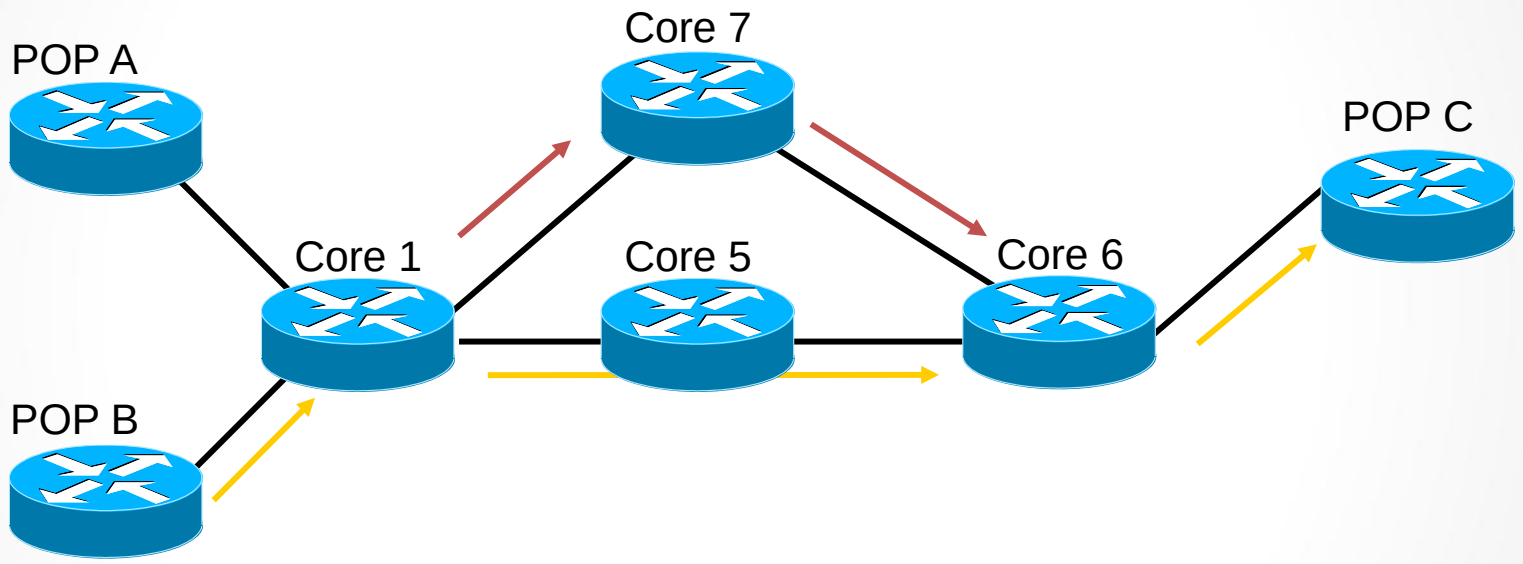
Bypass (backup) static tunnel to take in the event of a failure

Link Protection with FRR (Cont.)

- “Link Down” Event
  - The router, realizing the link is down:
    - Issues an IGP advertisement
    - Issues an RSVP message with session attribute flag 0x01=ON
      (do not break the tunnel; you may continue to forward packets during the reoptimization)
  - In the event of a failure, an LSP is intercepted and locally rerouted using a backup tunnel.
    - Original LSP nested within protection LSP
    - Minimum disruption of LSP flow
      (under 50 ms - time to detect and switch)
  - The headend is notified by RSVP PathErr and by IGP
    - Special flag in RSVP PathErr (reservation in place) indicates that the path states must not be destroyed, so the LSP flow is not interrupted.
    - The headend of the tunnel smoothly re-establishes the tunnel along a new route.

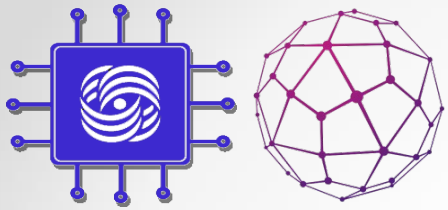
Node Protection with FRR

Node Protection for Core 5



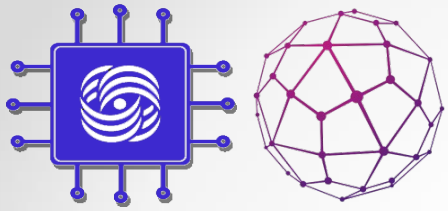
End-to-end tunnel onto which data normally flows

Bypass (backup) static tunnel to take in the event of a failure



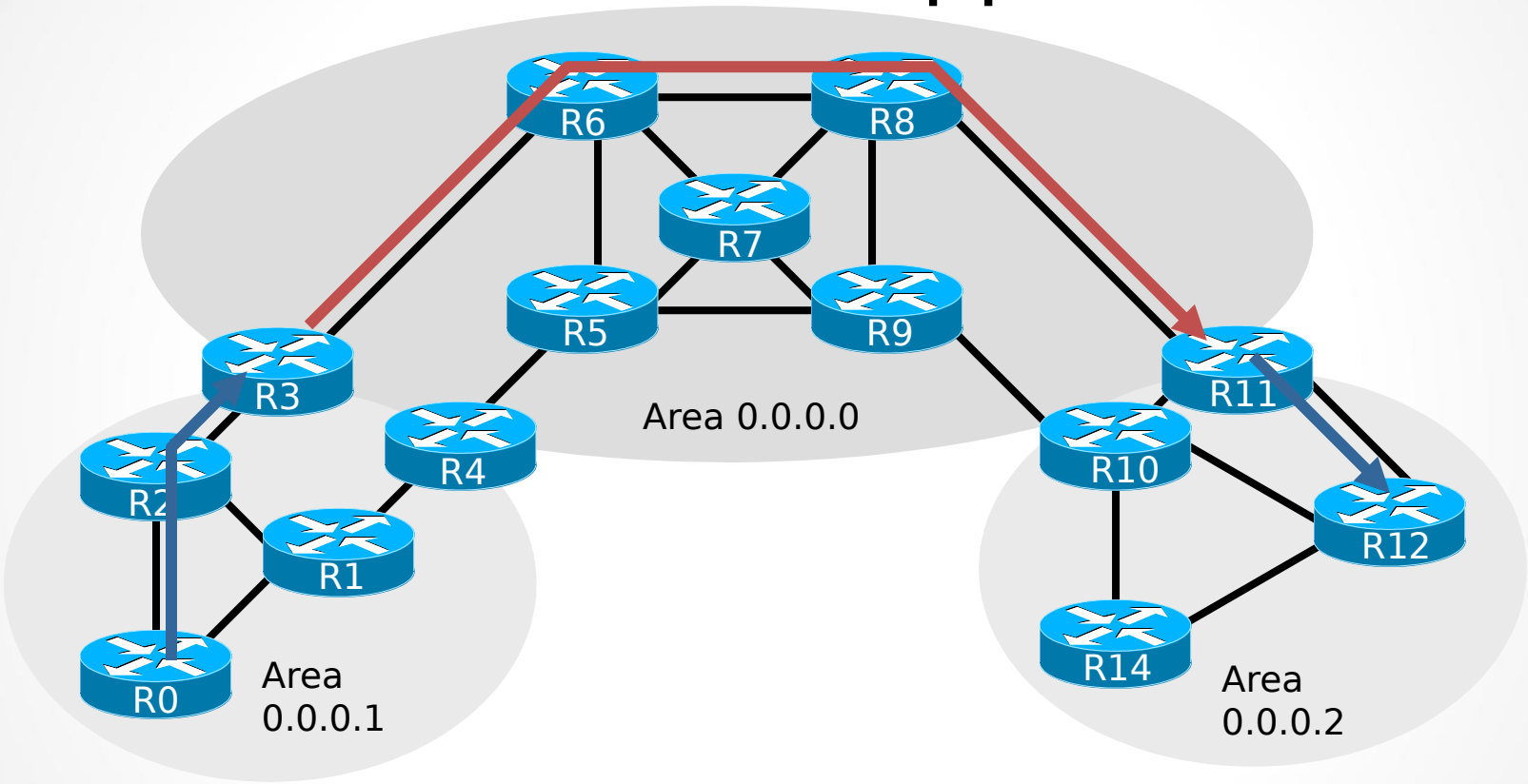
Node Protection with FRR (Cont.)

- Router node fails; the router detects this failure by an “interface down” notification.
  - It switches LSPs going out that interface onto their respective backup tunnels (if any).
- RSVP hellos can also be used to trigger Fast Reroute.
  - Messages are periodically sent to the neighboring router.
  - If no response is received, hellos declare that the neighbor is down.
  - Causes any LSPs going out that interface to be switched to their respective backup tunnels.



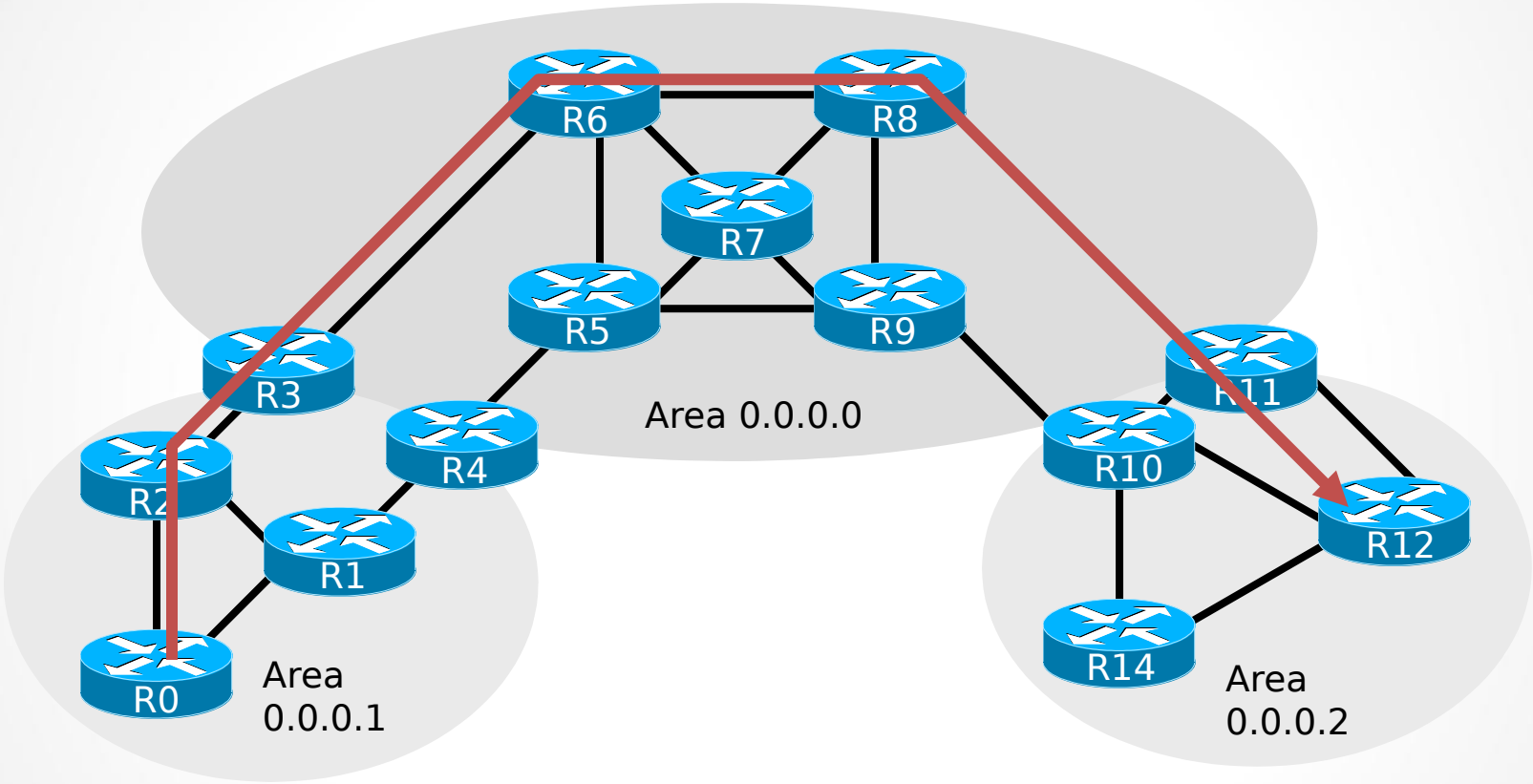
MPLS-TE Inter-Area Support

MPLS-TE Without Interarea Support



- One database exists per area.
- Three TE tunnels are needed to engineer the traffic from R0 to R12.

Interarea MPLS-TE



- With interarea TE a single TE tunnel can be set up dynamically, spanning areas from R0 to R12.